

Peptide-Based Modulators of α -Helix and β -Strand mediated Protein-Protein Interactions

Kristina Hetherington

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

University of Leeds
School of Chemistry

October 2020

Intellectual Property and Publications Statements

The candidate confirms that the work submitted is his/her own, except where work which has formed part of jointly-authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

Chapter 2 – The experimental alanine scanning work described in this chapter was included as part the following research paper: **Predicting and Experimentally Validating Hot-Spot Residues at Protein-Protein Interfaces**, A. A. Ibarra, G. J. Bartlett, Z. Hegedüs, S. Dutt, F. Hobor, K. A. Horner, K. Hetherington, K. Spence, A. Nelson, T. A Edwards, D. N. Woolfson, R. B. Sessions, A. J. Wilson, *ACS Chem. Biol.* 2019, 14, 10, 2252–2263.

With specific regard to Chapter 2, the computational alanine scan was performed by Gail Bartlet (University of Bristol) and Amaurys A. Ibarra (University of Bristol). Protein (SHANK1) production was done by Zsofia Hegedus (University of Leeds) as well as crystallisation and solving of L6F peptide variant to SHANK1. All other work was performed by the candidate unless otherwise stated throughout.

Chapter 3 – Computational work was done by Amaurys Avilla Ibarra (University of Bristol) and Richard Session (University of Bristol) who developed the Predictive Saturation Variation Scanning tool. Protein (MCL-1 and SUMO) production and purification were done by Pallavi Ramsahye (University of Leeds) and Fruzsina Hobor (University of Leeds). Lastly, two of the five SIM peptides (wt- SIM₂₇₀₅₋₂₇₁₇, SIM₂₇₀₅₋₂₇₁₇I2708D) were synthesised by Som Dutt (University of Leeds). All other work was performed by the candidate unless otherwise stated throughout.

Chapter 4 – **Stapled Peptides as HIF1- α /p300 inhibitors: Helicity Enhancement in the Bound State Increases Inhibitory Potency**, K. Hetherington, Z. Hegedus, T. A Edwards; R. B. Sessions; A. Nelson; A. J. Wilson.

Protein (p300) production was done by Zsofia Hegedus (University of Leeds). Molecular Modelling was done by Richard Sessions (University of Bristol) while Thomas Edwards, Adam Nelson and Andy Wilson supervised the project and assisted with manuscript preparation. All other work was performed by the candidate unless otherwise stated throughout.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

© 2020 The University of Leeds and Kristina Hetherington

Acknowledgments

Firstly, I would like to thank my supervisors, Prof. Andy J. Wilson and Prof. Adam Nelson for the opportunity to work on a this stimulating project. It was an honour to be part of such a collaborative project and hear so many ideas fly around the room during the PoPPI meetings, which will not be forgotten – they were insightful and at times provided comic relief. Many thanks to the Leeds University for awarding me a Leeds Anniversary Scholarship to fund my project and to the School of Chemistry and Steve Marsden for financing my final year.

Thank you to those who have supported the project in a scientific capacity, listened to my scientific and personal woes and also helped build my confidence up: Mark Howard your presence in the department helped lift my spirits but more importantly lifted the department's ability to perform NMR experiments to the highest calibre; Martin Hushcroft thank you for helping with the HPLC and LCMS issues; Steve Marsden thank you for the chats and jokes; Stuart Warriner you've always had a minute to spare and advice to give, thank you; Mike Webb it was a pleasure having you as my internal assessor, thank you for ensuring my meetings were engaging; Terry Kee thank you for making sure we are all spiritually well. As well, thank you to Ian Manfield (ITC) and Nasir Khan (CD) for technical help and always sparring a minute for a chat. Last but no least, thank you to Jennie Dickinson, Mandy Clarkson, Louise Hawksworth, Shelly Parker and Helen Linyard for always saying hello, sharing their stories with me, asking how things are going and giving a general interest in my well-being.

To the collaborators I had the pleasure of sharing my PhD journey – thank you for your insights and support and all your hard work. Sesh, from the moment we talked movies you were doomed with my friendship. Thank you, Richard Sessions, Amaurys Ibarra, Gail Bartlett and Debbie Shoemark for providing the computational work that flows throughout this project.

Past and present Wilson group members thank you for making the group a family and fostering a championing environment – Irene I wouldn't be here without you, you've supported me in more ways than I can count so I've been keeping all the labels on my new clothes so I can return them and buy you a thank you gift; Silvia you were the best mentor; Heather thank you for making my accent sound Yorkshire and telling us all something for nothing; Zsofia you imparted knowledge with every conversation and lots of tequila shots; speaking of shots, thanks Emma for your sass; Martin you are a fountain of knowledge and support, could not imagine the group without your leadership; Sonja you were my partner in crime; Suzanne "Kool" you kept my sanity with your stories and friendship; Chris thank you for being the best student I could ever dream of working with, you renegade; Rob better luck next prank, and thank you for The Office; Matt thank you for your witty jokes; Kirstin who knew you leaving meant a stronger friendship. These are just a few cherished memories I will take away, thank you all again.

To my family, thank you for caring and trying to understand my project. Most importantly, to my wife, you deserve an award for dealing with my mood swings and late nights in the lab missing dinners you carefully planned. This would not have been possible without your constant support and love; without you by my side.

Abstract

Protein-protein interactions (PPIs) are physical contacts between two or more protein molecules that control, virtually, all biological responses in a cell. Consequently, any interference, such as overexpression or under-expression of a protein, affects the natural progression of PPIs and can contribute to the development and continuation of disease. PPIs have emerged as therapeutic targets with significant potential in chemical biology, medicinal chemistry and therapeutics.

PPIs were once considered impossible targets for small molecule modulation due to the large protein surfaces coming together and the complexity of the binding compared to the simple lock and key model observed in traditional drug targets e.g. enzymes. Recent efforts in medicinal chemistry accompanied with advances in computational analysis and fragment based screening show that development of PPI modulators is feasible, albeit still challenging.

Structured-based design and synthesis of peptide-based inhibitors of both α -helix and β -strand mediated PPIs, through the combination of computational design and experimental validation, was attempted. Through biophysical studies (Fluorescence Anisotropy, Isothermal Titration Calorimetry), the elements driving the interaction of four clinically relevant PPI systems – GKAP/SHANK1 linked to neurodegenerative diseases; NOXA B/MCL-1 relevant in oncology; SIM/SUMO relevant in a number of response pathways including inflammation and oncology; HIF-1/p300 relevant in oncology – were determined and harnessed in the design of peptide-based inhibitors which were then further tested for their ability to inhibit the PPIs using biophysical and structural methods (Fluorescence Anisotropy, Isothermal Titration Calorimetry, Circular Dichroism, Nuclear Magnetic Resonance).

Chapter 2 focused on the GKAP/SHANK1 PPI, mediated through an extended β -strand conformation adopted by the hexapeptide GKAP (E₁A₂Q₃T₄R₅L₆-COOH) upon binding. The goals were to determine the GKAP residues important for

binding to SHANK1, probe this interaction further using selected peptide variants and finally use the knowledge gained through the first two experiments to develop a modulator for this interaction. Assay development aided the experimental alanine scan of the GKAP/SHANK1 system and confirmed the residues of the GKAP ligand important for binding (T₄, L₆ and carboxylate). Furthermore, the GKAP ligand sequence was altered to include point variations, rationally chosen either based on other SHANK1 binding sequences or to observe steric effects of bulky side chains, and determine if these variations have any effect on GKAP binding to SHANK1. Interestingly, variations at the A₂ and R₅ positions in the GKAP sequence showed no effect on binding to SHANK1, eliminating these two residue positions as facilitators to binding. Perhaps more remarkable was the variation of the L₆ residue to Phe which showed similar binding affinity as the wild-type GKAP sequence. Where once binding was lost with the L₆A variation during the alanine scan, the L₆F variation recovered the binding. The knowledge gained from these variations in the GKAP sequence along with the knowledge gained from the experimental alanine scan were harnessed in the design and synthesis of two, an acyclic and cyclic, peptide-based modulators for this β -strand mediated PPI. However, the difficulty encountered with the design and synthesis of a macrocyclic peptide-based inhibitor to target this β -strand interaction affirmed the need for continued, rigorous efforts to be placed when targeting β -strand mediated PPIs.

Chapter 3 was inspired by the design and testing of select GKAP sequences containing point variations. The goal of Chapter 3 was to validate computationally generated peptide interacting motifs (PIMs) of the α -helix mediated NOXA B/MCL-1 and β -strand mediated SIM/SUMO systems. Previously in the group Kat Horner and Som Dutt performed the experimental alanine scan for NOXA B/MCL-1 and SIM/SUMO respectively, making these systems ideal model systems. As such, these were subjected to a Predictive Saturation Variation Scan (PreSaVS) tool developed by the researchers at the University of Bristol, Amaury A. Ibarra and Richard Sessions. The PreSaVS tool computationally varied the sequence of each ligand peptide, sequentially, to every other amino acid generating novel peptide sequence binders, PIMs. Experimental validation of the PIMs established this methodology as an effective and fast approach

towards generating PIMs with binding affinities comparable to the wild-type ligand sequences, applied on both an α -helix and a β -strand mediated PPI. Extending the application of computational approaches to develop tools such as the PreSaVS tool, ultimately aid in the fast and efficient development of PPI modulators.

Chapter 4 shifted focus on dibromomaleimide peptide stapling used in targeting α -helix mediated PPIs. The goal was to expand the application of this stapling technique to a challenging PPI target, specifically through the design and synthesis of stapled helices of the HIF-1 α ligand for targeting p300. Typically peptide stapling has been shown to improve binding affinity of peptide ligands to their target by conformationally restricting the peptide in its bioactive, helical conformation. Although the stapled helix of HIF-1 α showed improved inhibitory potency against p300, contrary to expectations, the peptide displayed only a modest increase in helicity while in solution as a consequence of stapling; 20% helicity when compared to the 11% helicity of the wild-type peptide. Further investigation using CD difference spectra and MD simulations suggested the interesting conclusion that the staple enhanced the helical content of the bound state peptide, hence improving the inhibitory potency. This finding suggests stringent analysis of peptides in the bound state to their target, rather than in isolation, could result in more catered and successful stapled peptide binders.

The thesis highlights the importance of experimental studies still required when studying PPIs. Although the computational predictions done by researchers at the University of Bristol were harnessed and helped in the design of peptide-based inhibitors as well as leading to the development a new tools for studying PPIs, these are a long way from being reliable without experimental validation.

Table of Contents

Intellectual Property and Publications Statements	iii
Acknowledgments	v
Abstract	viii
Table of Contents	xi
List of Tables	xiv
List of Figures	xv
List of Schemes	xxv
List of Abbreviations and Symbols	xxvi
Amino Acids Abbreviations	xxx
Chapter 1: Protein-Protein Interactions	1
1.1 Protein-Protein Interactions	1
1.2 Key Features of PPI Affinity: Computational Analysis	2
1.3 Structural Features of PPI Interfaces.....	4
1.3.1 α -Helix Mediated PPIs.....	5
1.3.2 β -Strand Mediated PPIs	8
1.4 PPIs and Disease	11
1.5 Strategies for Targeting PPIs	12
1.5.1 α -Helix Mimetics: Stapled Peptides.....	15
1.5.2 β -Strand Mimetics: Macrocycles	18
1.5.3 Peptide Phage Display Libraries: Complementary Tool for Ligand Discovery.....	22
1.6 Small-Molecule Inhibitors of PPIs.....	23
1.7 Project Aims	26
1.7.1 Targeting α -Helix and β -Strand Mediated PPIs.....	26
Chapter 2: Targeting a β-strand mediated PPI: GKAP/SHANK1	30
2.1 PDZ Domains	30
2.1.1 GKAP/SHANK1 PPI.....	32
2.1.2 Inhibitors of the PDZ domain of SHANK Proteins	33
2.2 Objectives.....	34
2.3 Assay Development and Experimental Alanine Scan.....	35
2.3.1 Peptide Synthesis	35
2.3.2 Fluorescence Anisotropy.....	36
2.3.3 Assay Development for GKAP/SHANK1	37

2.3.4	Experimental Alanine Scan of GKAP Ligand	38
2.4	Computational Alanine Scan of GKAP Ligand.....	39
2.5	Design, Synthesis and Evaluation of Selected Peptide Variants	41
2.6	Design, Synthesis and Testing of a Peptide-Based Inhibitor	45
2.7	Conclusions.....	56
Chapter 3: Design of Novel Peptide Interacting Motifs Using Predictive Saturation Variation Scanning (PreSaVS)		58
3.1	Introduction.....	58
3.2	Objectives.....	60
3.3	Targets of Interest	61
3.3.1	α -Helix Mediated PPI: NOXA/MCL-1.....	61
3.3.2	β -Strand Mediated PPI: SIM/SUMO.....	62
3.4	Results and Discussion	63
3.4.1	Computational PreSaVS Scan and Design of Novel Peptide Binders.....	64
3.4.2	Experimental Validation of the Novel Peptide Sequences ...	66
3.4.3	NOXA/MCL-1	67
3.4.4	SIM/SUMO	70
3.5	Conclusions and Future Work	72
Chapter 4: Design, Synthesis and Evaluation of Stapled peptide inhibitors of the HIF-1α/p300 PPI		74
4.1	HIF1- α /p300 Introduction.....	74
4.1.1	HIF-1 α /p300 Binding Interaction	75
4.1.2	Inhibitors of HIF-1 α /p300.....	76
4.2	Objectives.....	79
4.3	Results and Discussion	80
4.3.1	Design and Synthesis of HIF-1 α Peptides.....	80
4.3.2	Binding Assays of HIF-1 α Peptides.....	82
4.3.3	Structural Analysis of HIF-1 α Peptides: Investigation into Structure to Explain the Observed Binding Behaviours.....	86
4.4	Conclusions.....	94
Thesis Summary and Future Work.....		97
Chapter 5: Experimental.....		101
5.1	General Experimental Remarks	101
5.2	PreSaVS Information and Full Scan Results	102
6.1	Peptide Synthesis Procedure	104

6.2 Peptide Biophysical Binding Assays.....	108
6.3 Synthesis of Hao Building Block.....	113
References.....	120
Appendix	133
A.1 Chapter 2 Peptide Analytical Data.....	133
A.2 Chapter 3 Peptide Analytical Data.....	139
A.3 Chapter 4 Peptide Analytical Data.....	144

List of Tables

Table 1. IC ₅₀ values for alanine and C-terminal amide variants.....	39
Table 2. Further investigation of the binding interaction with different GKAP variants.	42
Table 3. IC ₅₀ values for peptide variants designed to probe the binding interaction of GKAP to SHANK1.....	43
Table 4. Summary of data of wt-NOXA and variants.	69
Table 5. IC ₅₀ values for the SIM peptides.	71
Table 6. Summary of inhibitory potencies and helicity of peptide variants.	87
Table 7. Percentage Helicity based on MD simulations over the first 50 ns timeframe for each peptide.	93

List of Figures

Figure 1. Illustration of why PPIs can be challenging to target: a) A simple cartoon displaying the enzyme/substrate complex where the inhibitor design followed the substrate closely; b) Protein-protein interaction which occurs over a larger surface area; c) MCL-1/NOXA B interaction (PDB: 2JM6¹⁰) displayed as surfaces; followed by MCL-1 shown as surface and NOXA a cartoon with all the amino acid side chains projecting outwards to display the number of possible interactions that could be responsible for driving the binding interaction. Finally, the interface surface on MCL-1 shown as mesh to display the large area over which this PPI occurs.1

Figure 2. Computational techniques used in deciphering PPIs: a) Native binding interface; b) Alanine scanning where each residue is varied for alanine in order to identify hotspot residues; c) Hydrophile scanning where each residues is varied for charged residues.3

Figure 3. Typical structural features found at the interface of PPIs (PDBs left to right: HIF-1 α /p300 (1L8C³⁴), GKAP/SHANK1 (1Q3P³⁵), Ib1 dimer (2FPF³⁶), HIF-1 α /p300 (1L8C³⁴).4

Figure 4. α -Helix structural elements: a) the α -helix shown in stick mode with the backbone hydrogen bonding that stabilises the structure shown as dotted yellow lines; b) top-view of the helical conformation showing three main faces where residues stack on top of each other and interactions with the side chains can occur; c) Crystal structure of NOXA B peptide in cyan, bound to the surface of MCL-1 in green (PDB 2JM6¹⁰) showing residues at i , $i+4$ and $i+7$, positioned at the recognition surface of the continuous groove of MCL-1.....5

Figure 5. Helix mediated PPIs where residues on different faces of the helices contribute towards binding with a – c as examples of protein complexes with hotspot residues on one, two and three faces (PDBs from left to right 2JM6¹⁰,

1XIU⁴⁷, 1OR7⁴⁸s of NOXA B/MCL-1, LBD/RXR, *Escherichia coli* ECF/RseA respectively).....7

Figure 6. NMR structure of p300 (green) and HIF-1 α (cyan) (PDB: 1L8C): a) Cartoon representation of the NMR structure with zinc ions in grey. b) HIF-1 α cartoon representation wrapped around the surface of p300.....8

Figure 7. β -Strand structural information: a) Idealised β -Strand conformation of a poly-alanine sequence in its extended conformation; b) cartoon representation of two antiparallel oriented β -strands interacting.....9

Figure 8. β -Strand interactions showing: a) the p65 dimer PPI (PDB: 1OY3⁶⁴) where only side chains are involved in binding; b) The GKAP/SHANK1 PPI (PDB: 1Q3P³⁵) where both side-chains and backbone interactions drive the binding interaction with polar contacts shown by yellow dotted lines.10

Figure 9. Structure based PPI inhibitor design: i) The target PPI is studied using computational methods where topology at the interface and hotspot residues can be ascertained; ii) Constrained peptides or structural mimetics can now be designed based on the information gained from the computational studies. Small molecules can also be identified through *in silico* docking and structure similarity search; iii) The designed or discovered ligands are then tested in biophysical assays and the best binders are taken further through rounds of optimisation. 14

Figure 10. Initial peptide constraint approaches: a) The arrangement of atoms in an α -helix compared to that of a helix stapled at the residues using a hydrocarbon constraint and head to tail backbone nitrogen hydrogen bond surrogate (HBS); b) Crystal structure of the p53 hydrocarbon stapled peptide, where the staple drapes over the hDM2 protein partner (PDB: 3V3B¹⁰⁸); c) Crystal structure of the co-activator peptide bound to ER α where the hydrocarbon staple drapes over the protein partner (PDB: 2YJD¹⁰⁹).15

Figure 11. Helix enforcing techniques: a) Stapling techniques through side chain residues: i) Hydrocarbon staple achieved through the insertion of a non-natural amino acid and ring closing metathesis; ii) Lactam bridge obtained from using natural residues with special protecting groups on both the acid and the amine side-chain residues; iii) Disulfide constraint formed by the oxidation of cysteine residues; iv) Cu(I) catalysed azide-alkyne cycloaddition; b) Dibromomaleimide stapling design with reversibility accessed via TCEP reducing agent.....17

Figure 12. Macrocycles designed to target β -strand interactions: a) HIV-1 protease 1 inhibitor; b) Visual overlay and comparison of the binding conformation of the HIV-1 protease natural peptidic substrate (cyan, PDB: 1MT7¹²⁷) and the macrocyclic inhibitor (orange, PDB: 1D4L); c) Nowick designed macrocycle 2 with the **Hao** (hydrazine, 5-amino-2-methoxybenzoic acid, oxalic acid) aromatic building block incorporated, which fosters hydrogen bonding (blue bonds) between the strands. The rigidity of the aromatic ring ensures that the molecule forms and intermolecular hydrogen bond which also aids the molecule in adapting a conformation that maximises the hydrogen bonds between strands.....19

Figure 13. Protein epitope mimetics based on constrained beta-hairpin: a) Cartoon representation of how a loop interaction can help influence the design of the PEM structure which consists of three main variables: loop, building blocks and template regions; b) beta-hairpin mimic of the p53 peptide bound to *hDM2* (PDB: 2AXI) showing in red (middle frame) the three hotspot residues mimicked and in an overlap with the p53 peptide (PDB: 1YCR).21

Figure 14. Small molecule inhibitors of the p53/*hDM2* PPI: a) Crystal structure (PDB: 4J3E) of Nutlin 3a bound to *hDM2* discovered through HTS; b) Chemical structure of Nutlin 3a and the second generation compounds optimised by Roche.24

Figure 15. Small molecule inhibitors of BCL-x: a) Crystal structure of ABT-737 molecule bound to the protein (PDB: 2YXJ); b) Chemical structure of ABT-based developed inhibitors.25

Figure 16. Summary of experimental systems investigated and chapters which cover each of these.....	26
Figure 17. Cartoon representation of a canonical PDZ domain structure (forest green) made up of 2 α -helical regions and 6 β -strands, with a typical C-terminus ligand (cyan) binding between the α B and the β B structural elements. The terminal ϕ hydrophobic residue is accommodated within a GLGF pocket in the PDZ domain.....	30
Figure 18. Cartoon representation of the scaffold SHANK1 protein and its many PPIs in the post-synaptic density. GKAP is represented in green, interacting with SHANK1 through the PDZ domain while other protein interactions are mediated through other SHANK1 domains. GKAP is observed to form a further interaction with PSD95 which in turn interacts with the receptor.	31
Figure 19. Crystal structure of the GKAP (cyan) and SHANK1 (green) PPI (PDB: 1Q3P): a) GKAP/SHANK1 with GKAP in an extended β -strand conformation showing the binding interactions from both side-chains and backbone; b) An enlarged image of the Thr residue forming a hydrogen bond on the opposing side of the PDZ binding interface when compared to the backbone hydrogen bonds. Leucine shown magnified, packed in a hydrophobic pocket.	32
Figure 20. Peptide-based SHANK3 and SHANK1 inhibitor based on the PIX sequence 3 and SHANK3 inhibitors based on the tetrahydroquinoline carboxylate 4-7.....	33
Figure 21. Buffer screen for the assay development for the GKAP/SHANK1 PPI all between pH 7.4-7.5 using 100 nM tracer read after a 45 minute incubation period: a) Buffer conditions of 50 mM Tris, 150 mM NaCl; b) Triton 0.01% was added to the initial 50 mM Tris, 150 mM NaCl buffer; c) Buffer system contained 50mM Tris, 150 mM NaCl and 0.02 mg/mL BSA; d) Buffer contained all 4 components: 50 mM Tris, 150 mM NaCl, 0.01% Triton, 0.02 m/mL BSA.	38

Figure 22. FA data for the GKAP/SHANK1 PPI; a) Direct titration assay (50 mM Tris, 150 mM NaCl, using 100 nM Fitc-Ahx-GKA); b) Competition assays for the GKAP alanine variants and the C-terminal amide variant (50 mM Tris, 150 mM NaCl, 50 nM tracer, 1 μ M SHANK1).39

Figure 23. Alanine scanning data for GKAP/SHANK1 (PDB: 1Q3P); a) The computational BUDE AlaScan results denoted by red dots compared with the experimental data represented as grey bars; b) BUDE AlaScan $\Delta\Delta G$ data shown by the red dots and BUDE predictions incorporating a 1 ms MD data for the mutants depicted by the blue dots with a dotted line depicting the hot-spot cut-off.....40

Figure 24. Competition assay data of specific residues of the GKAP sequence all done in 50 mM Tris, 150 mM NaCl using 50 nM tracer and 1 μ M SHANK1; a) Competition assay curves for GKAP₁₋₆A2D and GKAP₁₋₆A2F compared to wt-GKAP₁₋₆; b) Competition assay for GKAP₁₋₆R2W compared to GKAP₁₋₆; c) Competition assay for GKAP₁₋₆T4S compared to GKAP₁₋₆; d) Competition assay for GKAP₁₋₆L2F compared to wt-GKAP₁₋₆.42

Figure 25. ITC and crystal data for GKAP₁₋₆L6F variant compared to wt-GKAP₁₋₆; a) ITC data for wt-GKAP₁₋₆; in 50 mM Tris, 150 mM NaCl, pH 7.5, at 25 $^{\circ}$ C; b) ITC data for GKAP₁₋₆L6F; in 50 mM Tris, 150 mM NaCl, pH 7.5, at 25 $^{\circ}$ C; c) GKAP/SHANK1 (PDB: 1Q3P) crystal structure displaying the binding of the GKAP sequence; d) for GKAP₁₋₆L6F/SHANK1 crystal structure; e) Overlay of wt-GKAP₁₋₆ (cyan) with GKAP₁₋₆L6F (orange).44

Figure 26. Peptidomimetic design based on the GKAP peptide sequence and incorporating the Hao molecular building block.....46

Figure 27. Analysis of compound 1 against wt-GKAP₁₋₆; a) Competition assay in 50 mM Tris, 150 mM NaCl using 50 nM tracer and 1 μ M SHANK1; b) ITC data in 50 mM Tris, 150 mM NaCl, pH 7.5, at 25 $^{\circ}$ C.54

Figure 28. Structure of 1 with red arrows displaying the potential key interactions to look for in NMR as signs of pre-organisation.	54
Figure 29. NOESY NMR analysis of 8.	55
Figure 30. PreSaVS workflow: following target selection the native ligand peptide sequence is computationally placed through the PreSaVS software and new peptide sequences with point variations suggested that require experimental validation.....	60
Figure 31. Structural, computational and experimental data for NOXA/MCL-1; a) NMR structure of NOXA/MCL-1 (PDB 2JM6 ¹⁰) showing NOXA in its α -helical conformation upon binding, with the hot-spot residues highlighted in red; b) The computational BUDE AlaScan results compared with experimental data ¹⁹	62
Figure 32. SIM/SUMO; a) NMR structure of SIM/SUMO (PDB 2LAS ²⁰³) showing SIM in an extended β -strand conformation upon binding, with the hot-spot residues highlighted in red; b) The computational BUDE AlaScan results compared with the experimental data.	63
Figure 33. Design of variant peptides from the computational PreSaVS; a) Excerpt of the computational PreSaVS results of the NOXA peptide; b) NOXA variant peptide sequences selected for synthesis and further testing; c) Excerpt of the computational PreSaVS results of the SIM peptide; d) SIM variant peptide sequences selected for synthesis and further testing.	65
Figure 34. FA direct titration curves; a) Direct titration assay for NOXA/MCL-1 in 50 mM Tris, 150 mM NaCl buffer using 25 nM tracer concentration; b) Direct titration assay for SIM/SUMO in 50 mM Tris, 150 mM NaCl buffer using 50 nM tracer concentration.	66

Figure 35. Data set for wt-NOXA₇₅₋₉₃ and the variant sequences presented in Table 4; a) Competition FA data for all NOXA peptides in 50 mM Tris, 150 mM NaCl buffer using 25 nM tracer concentration and 150 nM MCL-1; b) ITC data for the peptide variants in 50 mM Tris, 150 mM NaCl buffer using 500 μM peptide concentration and 20 μM MCL-1 followed by plotted thermodynamic profiles from the ITC data for the NOXA peptide variants and wild type NOXA; c) Competition assays of the NOXA variant peptides and wt-NOXA peptide against the BCL-xL protein in in 50 mM Tris, 150 mM NaCl buffer using 25 nM tracer concentration and 150 nM BCL-xL protein.....68

Figure 36. CD data for the wt-NOXA and binding peptide variants; a) CD in 20 mM potassium phosphate buffer, 150 μM peptide concentration; b) CD in 20 mM potassium phosphate buffer with 30% TFE, 150 μM peptide concentration.70

Figure 37. FA data for the SIM peptide variants; a) Competition assay data for the wt-SIM₂₇₀₀₅₋₂₇₁₇ and SIM peptide variants, 50 mM Tris, 150 mM NaCl buffer, 150 μM competitor concentration, 25 nM tracer, 100 nM SUMO; b) SIM/SUMO (PDB: 2LAS²⁰³) with I2708, V2710, I2711, V2713 residues in red c) I2708 proximity to SUMO K residues, I2711 proximity to SUMO residues Y21 and K37, V2713 pointing towards solvent.71

Figure 38. Condensed summary cartoon of the Hif-1α regulation pathway: with HIF-1 α degradation taking place under normal oxygen levels (left) and HIF-1α activity under hypoxia (right).74

Figure 39. Structural features of the HIF-1α/p300 interaction (HIF-1 in cyan and p300 in forest green); PPI determined through solution NMR experiments (PDB: 1L8C): a) p300 in a very compact 4 helix bundle structurally supported through 3 zinc atoms; b) HIF-1α wrapped around p300; c) both proteins shown as surfaces to emphasize the large surface area which a small molecule would not be able to cover; d) Helix 2 with Cys800 and Asn803 and e) Helix 3 with Leu818, Leu822 and Val825 residues which are believed to be important for binding p300.75

Figure 40. Small molecules thought to target the HIF-1 α /p300 PPI.	77
Figure 41. HBS structures derived from helix 2 and helix 3 of HIF-1 α to target the HIF-1/p300 for inhibition with K_d values determined using ITC. The best inhibitors from each helix were used to treat mice xenografts models and showed tumour reduction when compared to the control mice.....	78
Figure 42. Peptidomimetic approaches targeting the HIF-1/p300 PPI: a) The best oligoamide structure targeting the helix 3 interface on p300; b) Comparison of binding affinities for the full length HIF-1 α ₇₇₆₋₈₂₆ , HIF-1 α ₇₉₄₋₈₂₆ and the bionic protein; c) The general structure of oxopiperazines designed from a linear peptide are displayed on the left panel, with blue squiggly lines denoting where the linkage is present; A list of the OHM structures designed and tested by the Arora group on the right hand side of the figure.	79
Figure 43. Project aim: design and test stapled peptides based on helix 3 of HIF-1 α for binding to p300.	80
Figure 44. Peptide design and synthesis based on helix 3 of HIF-1 α : a) The two main lengths of HIF-1 α that were used in this work; b) Residues chosen to be replaced by cysteines and stapled following a methodology previously reported; c) Stapling reaction which was improved upon by lowering the overall reaction time from 16 hours to 30 minutes; d) All variant peptides synthesised and tested for binding to p300.	81
Figure 45. FA data for HIF-1 α ₇₈₆₋₈₂₆ : a) Direct titration assay of 25 nM FITC-Ahx-HIF-1 α ₇₈₆₋₈₂₆ (tracer) in 20 mM Tris, 100 mM NaCl, 0.1 mM DTT, pH 7.46; b) Competition assay using the acetylated HIF-1 α ₇₈₆₋₈₂₆ as competitor using 50 nM tracer, 100 nM p300 in 20 mM Tris, 100 mM NaCl, 0.1 mM DTT, pH 7.46.	83

Figure 46. FA competition data for the HIF-1 $\alpha_{812-826}$ E817C-A821C peptide series: a) Competition assay for the peptide variants (25 nM) FITC-Ahx-HIF-1 $\alpha_{776-826}$, (100 nM) p300 in 20 mM Tris, 100 mM NaCl, 0.1 mM DTT, pH 7.46; b) Peptide sequences.84

Figure 47. FA data for the HIF-1 $\alpha_{812-826}$ E816C-R820C peptide series: a) Competition assay for the peptide variants using (25 nM) FITC-Ahx-HIF-1 $\alpha_{776-826}$, (100 nM) p300, in 20 mM Tris, 100 mM NaCl, 0.1 mM DTT, pH 7.46; b) Peptide sequences.85

Figure 48. HRMS data of the assay wells following the 45 minute incubation period to ensure peptides are still in their oxidised and reduces states: top two panels show HIF-1 $\alpha_{812-826}$ E817C-A821C series with the reduced peptide mass showing the two extra hydrogen atoms and oxidised peptide; similarly the bottom two panels show the HIF-1 $\alpha_{812-826}$ E816C-R820C reduced and oxidised peptides maintaining their respective states.85

Figure 49. CD spectra for wt-HIF-1 $\alpha_{812-826}$, HIF-1 $\alpha_{812-826}$ oxE816C-R820C, HIF-1 $\alpha_{812-826}$ sE816C-R820C and HIF-1 $\alpha_{812-826}$ redE816C-R820C at 250 μ M in 20 mM sodium phosphate, pH 7.55.87

Figure 50. CD thermal melting experiments from 25 to 90°C data, labelled with a black arrow in the direction of unfolding with increasing temperature, of peptides at 250 M, in 20 MM sodium phosphate, pH 7.55: a) Thermal unfolding for wt-HIF-1 $\alpha_{812-826}$; b) Thermal unfolding for HIF-1 $\alpha_{812-826}$ oxE816C-R820C; c) Thermal unfolding for HIF-1 $\alpha_{812-826}$ redE816C-R820C; d) Thermal unfolding HIF-1 $\alpha_{812-826}$ sE816C-R820C.88

Figure 51. Thermal denaturation experiments in 20 mM sodium phosphate, pH 7.55: a) Thermal CD of p300 at 20 μ M concentration showing a smooth and consistent transition between folding and unfolding CD traces from 20-90°C; b) Thermal CD of p300/HIF-1 $\alpha_{812-826}$ sE816C-R820C, 1:1 ratio at 20 μ M concentration exhibiting a moderate increase in thermal stability; c) Thermal CD

of p300/HIF-1 $\alpha_{812-826}$ redE816C-R820C, 1:1 ratio at 20 μ M where a lack of structural folding is observed over the temperature range; d) CD thermal denaturation curve showing structural effects of HIF-1 $\alpha_{812-826}$ sE816C-R820C and HIF-1 $\alpha_{812-826}$ redE816C-R820C on p300: the hollow points show p300 denaturation curve, with defined folded and unfolded states; the orange trace of p300 in a 1:1 ratio with HIF-1 $\alpha_{812-826}$ redE816C-R820C, showing severe loss of structure indicated by the lack of transition; lastly, the blue trace of p300 in a 1:1 ratio with HIF-1 $\alpha_{812-826}$ sE816C-R820C with a reduced T_m but with transition between folding and unfolding still evident.....89

Figure 52. Conformational analyses of HIF-1 $\alpha_{812-826}$ E816C-R820C: a) 100 ns MD simulation of wt- HIF-1 $\alpha_{812-826}$ peptide in the absence (top left panel) and presence (bottom left panel) of p300; b) 100 ns MD simulation of HIF-1 $\alpha_{812-826}$ oxE816C-R820C peptide in the absence (top right panel) and presence (bottom right panel) of p300; c) MD simulations of HIF-1 $\alpha_{812-826}$ sE816C-R820C peptide in isolation and in complex with p300 in the two dominant conformations adopted; d) MD snapshots of two stable p300 bound conformations of HIF-1 $\alpha_{812-826}$ oxE816C-R820C.....92

Figure 53. CD analysis: a) Cartoon representation of the method for acquiring difference CD data; b) CD data for the additive (grey), combined (black) and difference spectra (blue) for the HIF-1 $\alpha_{812-826}$ sE816C-R820C variant peptide in the presence of p300; c) Difference CD: spectra for binding of wt-HIF-1 $\alpha_{812-826}$, (black) stapled peptide HIF-1 $\alpha_{812-826}$ sE816C-R820C (blue) and HIF-1 $\alpha_{812-826}$ oxE816C-R820C (green) to p300.94

List of Schemes

Scheme 1. Solid phase peptide synthesis cycles and manual couplings that followed.....	35
Scheme 2. Synthetic route towards obtaining Fmoc protected Hao.....	47
Scheme 3. Failed Fmoc deprotection and acetylation of Hao building block after its manual coupling to the GKAP peptide.....	48
Scheme 4. Synthetic route to obtain the acetylated Hao building block.....	48
Scheme 5. Final step, standard peptide cleavage off resin, to obtain the 1st Generation Peptidomimetic.....	49
Scheme 6. Synthetic efforts to deprotect and cyclize 15 to obtain the 2 nd Generation peptidomimetic 2: a) Linear peptide design with the Asp replacing Arg at the 5 position in the GKAP peptide sequence; b) Initial deprotection and cyclisation of 15 which was unsuccessful; c) Attempted Fmoc deprotection and acetylation to test if this was the location of the synthetic issues encountered.....	50
Scheme 7. New design for the 2 nd Generation peptidomimetic: a) the GKAP peptide with the insertion of a Hao building block with a ornithine linker attached; b) Synthetic route for ornithine containing Hao building block.	51
Scheme 8. Synthetic route for building block 26.	52

List of Abbreviations and Symbols

Å	Angstrom
βPIX	p21-Activated Protein Kinase Exchange Factor Alpha
aa	Amino acid
BAK	Bcl-2 homologous Antagonist Killer protein
BAX	Bcl-2 Associated X-Protein
BLC-2	B-cell lymphoma 2
BCL-xL	B-cell lymphoma extralarge
BID	BH3-Interacting Domain death agonist
BIM	Bcl-2 Interacting Mediator of cell death
BSA	Bovin Serum Albumin
^t Boc ₂ O	<i>tertiary</i> -Butoxycarbonyl anhydride
BUDE	Bristol Universal Docking Engine
CAS	Computational Alanine Scan
CBP	CREB-binding protein
CD	Circular Dichroism
DIPEA	<i>N,N'</i> -Diisopropylethylamine
DTT	Dithiothreitol
EDT	1,2-Ethanedithiol
ERα	Estrogen receptor alpha
FA	Fluorescence Anisotropy
FDA	Food and Drug Administration
FITC	Fluorescein-isothiocyanate
Fmoc	Fluorenylmethyloxycarbonyl
GKAP	Guanylate-Kinase-Associated Postsynaptic-Density Protein
HATU	1-[Bis(dimethylamino)methylene]-1 <i>H</i> -1,2,3-triazolo[4,5- <i>b</i>]pyridinium 3-oxid hexafluorophosphate
Hao	hydrazine, 5-amino-2-methoxybenzoic acid, oxalic acid

HBS	Hydrogen Bond Surrogate
HCTU	2-(6-Chloro-1H-benzotriazole-1-yl)-1,1,3,3-tetramethylammonium hexafluorophosphate
<i>hDM2</i>	Human Double Minute two
HIF-1 α	Hypoxia-Inducible Factor 1 alpha
HIV-1	Human Immunodeficiency Virus type 1
HPLC	High-Pressure Liquid Chromatography
HRMS	High Resolution Mass Spectrometry
HTS	High-Throughput Screening
Ib1 dimer	Islet-brain 1
IC ₅₀	Half Maximal Inhibitory Concentration
IgG	Immunoglobulin G
ITC	Isothermal Titration Calorimetry
K _d	Dissociation constant
K _i	Inhibition constant
LBD	Ligand Binding Domain
LC-MS	Liquid Chromatography–Mass Spectrometry
MBHA	4-Methylbenzhydrylamine
Mcl-1	Induced Myeloid Leukaemia cell differentiation protein
MRE	Mean Residual Ellipticity
NMR	Nuclear Magnetic Resonance
NMP	<i>N</i> -Methylpyrrolidinone
NOESY	Nuclear Overhauser Effect Spectroscopy
NOXA B	Phorbol-12-myristate-13-acetate-induced protein 1
NTD	N-Terminal Domain
p53	Tumour Protein 53
p65	Nuclear Factor NF-kappa-B
p300	E1A binding Protein 300
Pbf	2,2,4,6,7-pentamethyldihydrobenzofurane

PDB	Protein Data Bank
PDZ	Post Synaptic Density Protein 95 (PSD-95), Drosophila Disc Large Tumor Suppressor (Dig1), Zona Occludens-1 Protein (ZO-1)
PEG	Polyethyleneglycol
PEM	Protein Epitope Mimetic
PIMs	Peptide Interacting Motifs
PPI	Protein-protein interaction
PSD-95	Post Synaptic Density Protein 95
PSD	Postsynaptic Density
QSAR	Quantitative Structure Activity Relationship
RCM	Ring-Closing Metathesis
RMSD	Root-Mean-Square Deviation
RXR	Retinoid X Receptors
SAR	Structure-Activity Relationship
SBD	Structure-Based Design
SHANK	SH3 and multiple ankyrin repeat domains also known as proline-rich synapse-associated protein 2 (ProSAP2)
SICLOPPS	Split-Intein Circular Ligation of Peptides and Proteins
SPPS	Solid Phase Peptide Synthesis
TCEP	Tris(2-carboxyethyl)phosphine
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
TIS	Triisopropylsilane
TLC	Thin Layer Chromatography
TNF	Tumour Necrosis Factor
UV	Ultraviolet
WT	Wild Type

Amino Acids Abbreviations

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

Chapter 1: Protein-Protein Interactions

1.1 Protein-Protein Interactions

Protein-protein interactions (PPIs) regulate protein function and co-ordinate signalling pathways to control biological functions, with an estimated 650 000 PPIs in existence¹⁻³. PPIs are at the heart of virtually all biological mechanisms and mis-regulation in PPI pathways can lead to disease⁴. Accordingly, PPIs have emerged as targets for drug discovery campaigns and for understanding of biological systems. However, initially PPIs were considered undruggable⁵⁻⁷ and although success has now been achieved in bringing inhibitors to the clinic and understanding signalling pathways involving PPIs⁸ these still remain quite challenging to target. Conventional enzyme/substrate complexes were tractable targets for drug discovery because the substrate served as a template for design of the inhibitor (Figure 1a). However, the large surface area over which PPIs occur and the driving forces governing these interactions remain as hurdles when designing modulators (Figure 1b, c)^{2,9}.

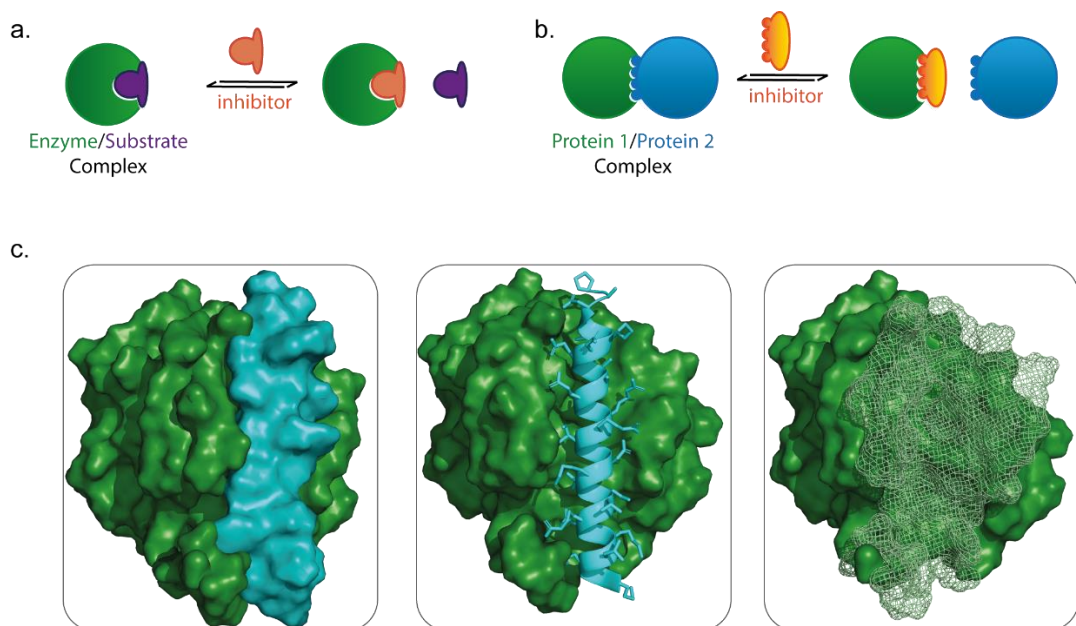


Figure 1. Illustration of why PPIs can be challenging to target: a) A simple cartoon displaying the enzyme/substrate complex where the inhibitor design followed the substrate closely; b) Protein-protein interaction which occurs over a larger surface area; c) MCL-1/NOXA B interaction (PDB: 2JM6¹⁰) displayed as surfaces; followed by MCL-1 shown as surface and NOXA a cartoon with all the amino acid side chains projecting outwards to display the number of possible interactions that could be responsible for driving the binding interaction. Finally, the interface surface on MCL-1 shown as mesh to display the large area over which this PPI occurs.

Furthermore, PPIs generally bind rapidly and reversibly and range in binding affinities between low nanomolar to high micromolar¹¹, making the recapitulation of the necessary intermolecular interactions into a specific and relatively potent inhibitor difficult^{12,13}. Ultimately, inhibitors must be designed such that they bind their target selectively, and binding affinity is one key challenge that directs the pharmacological effect of the drug along with potency, cellular permeability, plasma stability, clearance and side effects¹⁴. The degree of plasticity, flexibility and dynamics of binding of PPIs are factors which affect the mechanisms of interactions and could help in the type of modulator to design but that are yet to be understood fully^{15,16}.

1.2 Key Features of PPI Affinity: Computational Analysis

Usually, the relatively large surface area of PPIs (1500 to 3000 Å²) poses significant challenges in the design of inhibitors. The experimental technique known as alanine-scanning mutagenesis was developed to identify key residues important for binding, hot-spots¹⁷, through the sequential variation of each residue to alanine at the PPI interface¹⁸. When a loss in binding was observed due to the variation to alanine, that native residue would be identified as a hot-spot, a residue which contributed significantly towards binding affinity. This technique allowed for the drug discovery efforts to focus on a particular area within a PPI interface. Alongside experimental methods, advances in computational tools have also made the analysis of PPIs more accessible by finding smaller, more localised regions within the binding interface as important for binding. Computational alanine-scanning (CAS) is an *in silico* approach which identifies such regions made up of hotspot residues¹⁷ (Figure 2b). Each amino acid in a peptide sequence is varied to alanine (Ala) and the binding free energy difference ($\Delta\Delta G$) between binding of the native and variant peptides to the target protein is determined^{17,19–21}. A 'hot-spot' is a residue which contributes significantly to binding within a PPI evidenced after its computational variation to alanine shows a subsequent loss in binding^{19,20}. Thus, a loss in binding energy by $\Delta\Delta G \geq 2.5$ kJ/mol would be interpreted as a hotspot residue¹⁹. Often, Short Linear Interacting Motifs or interface peptides are sufficient to recapitulate a

binding site²². However, both small molecules and peptidomimetics have been designed to target hotspot residue clusters^{19,23–25}. For example the well-known three hotspot residues of p53, from the p53/*hDM2* interaction, (Phe19, Trp23 and Leu26) were used to inform the design of small molecule inhibitors^{25–28} (further discussed in section 1.5.4).

Complementary to CAS, but not as widely used, computational hydrophile-scanning was developed to explore the systematic mutation of residues in the BIM peptide sequence to charged residues and the effect on its binding with Bcl-x_L and MCL-1²⁹ (Figure 2c). This technique differs from CAS as each residue in a sequence is swapped for charged residues, as opposed to alanine, offering insight to potential charged interactions which may reinforce or disadvantage the binding.

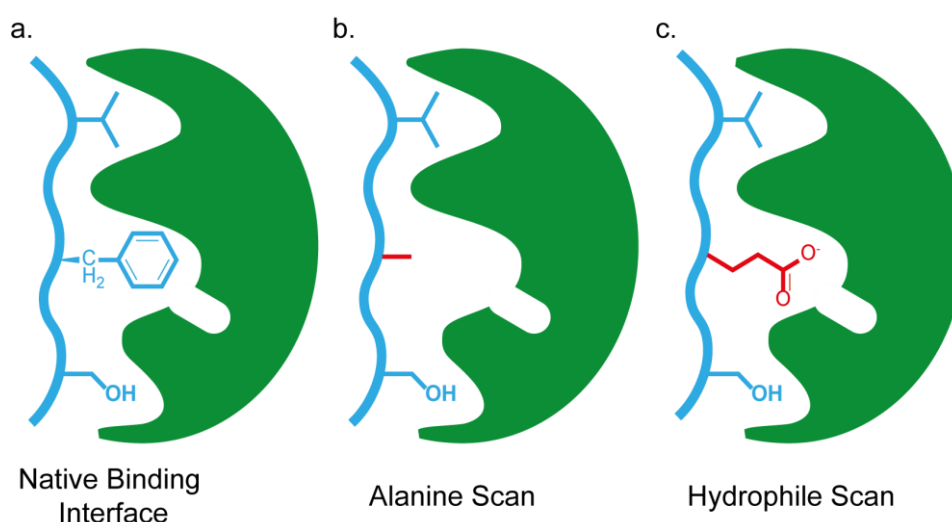


Figure 2. Computational techniques used in deciphering PPIs: a) Native binding interface; b) Alanine scanning where each residue is varied for alanine in order to identify hotspot residues; c) Hydrophile scanning where each residues is varied for charged residues.

Combining the results of the hydrophile scan with the information gathered from CAS, it was possible to design variant BIM peptides with minimal sequence variations that achieved selectivity towards either BCL-x_L or MCL-1²⁹. Thus, integrated computational methods, such as CAS or hydrophile scanning, and experimental characterization can be successfully used in the design of protein interacting motifs (PIMs) with enhanced properties.

1.3 Structural Features of PPI Interfaces

A PPI can involve two globular proteins, a protein domain and a linear peptide motif or between two peptides⁹. More importantly, the physical features at the interface of these interactions range from structured to non-regular, devoid of a secondary structure^{9,30}. The more prevalent structural motifs mediating PPIs are the α -helix and β -strand followed by loops which can also be further classified³¹, and unstructured or extended chain peptide motifs where no secondary structured arrangement is adopted at the PPI binding interface^{32,33} (Figure 3).

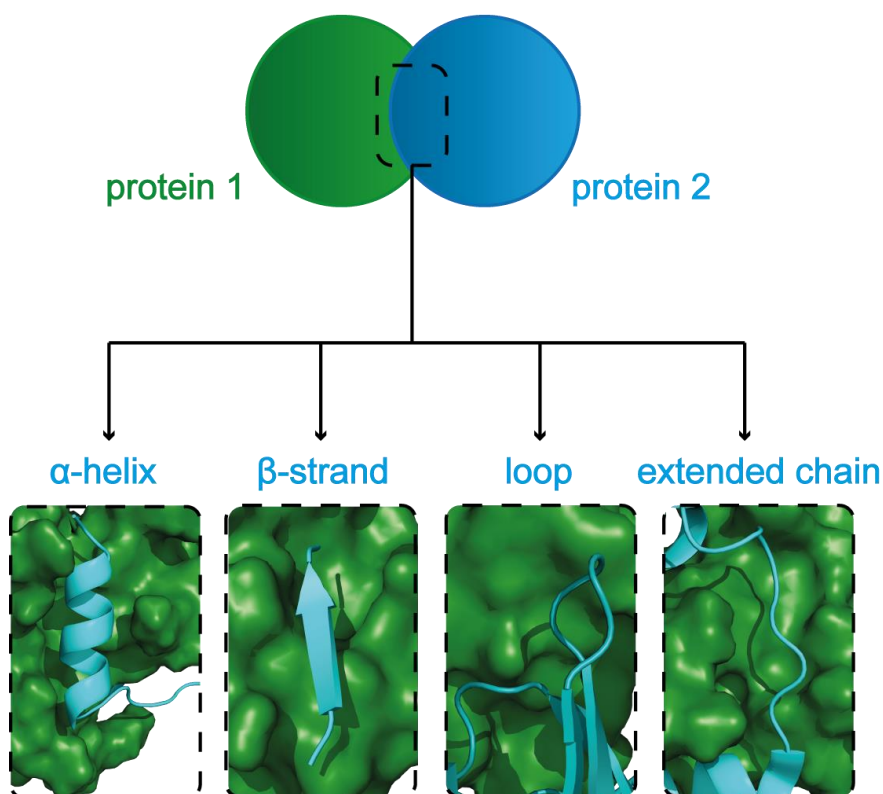


Figure 3. Typical structural features found at the interface of PPIs (PDBs left to right: HIF-1 α /p300 (1L8C³⁴), GKAP/SHANK1 (1Q3P³⁵), Ib1 dimer (2FPF³⁶), HIF-1 α /p300 (1L8C³⁴).

The identification and structural classification of PPI interfaces revealed new terrain for targeting PPIs³⁷. Identifying and dissecting PPIs based on the structural motifs present at the interface helps aid in the rational design of inhibitors. For example stapled peptides targeting α -helix mediated PPIs and peptide based macrocycles targeting β -strand and loop mediated interactions

were developed based on the native structural motifs found at PPI interfaces and will be discussed in depth section 1.5.1 and 1.5.2 respectively. Both peptide-based and non-peptidic based inhibitors were designed to mimic secondary structures and recapitulate key hotspots of binding and shown to inhibit PPIs^{1,8,27,38–40}.

1.3.1 α -Helix Mediated PPIs

The α -helix is a common motif in protein secondary, however, this secondary structural motif also mediates PPIs. With a turn at every 3.6 residues, the α -helix forms through hydrogen bonding between the carbonyl oxygen and the amide proton of every third residue in the turn^{38,41,42} (Figure 4a). As such, the side chains of the α -helix project outward and those located every 3-4 residues are spatially adjacent to one another (Figure 3b). Overall, the α -helix presents up to three faces, each endowed with side chain residues, available for molecular recognition (Figure 4b). Generally the PPI takes place when the helical motif of one protein docks into the interface of its binding partner (Figure 3c).

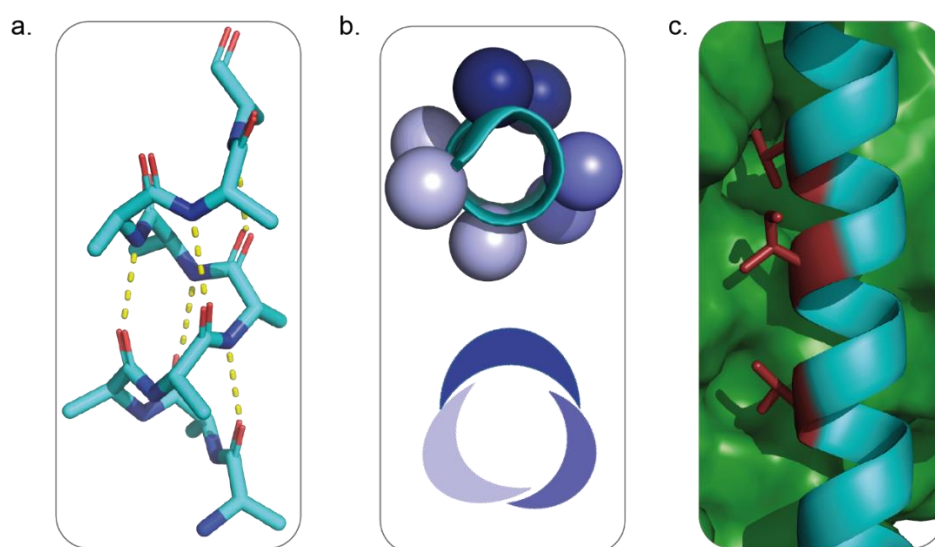


Figure 4. α -Helix structural elements: a) the α -helix shown in stick mode with the backbone hydrogen bonding that stabilises the structure shown as dotted yellow lines; b) top-view of the helical conformation showing three main faces where residues stack on top of each other and interactions with the side chains can occur; c) Crystal structure of NOXA B peptide in cyan, bound to the surface of MCL-1 in green (PDB 2JM6¹⁰) showing residues at i , $i+4$ and $i+7$, positioned at the recognition surface of the continuous groove of MCL-1.

1.3.1.1 PPIs Mediated by a Single Helix

Protein complexes in the Protein Data Bank (PDB) that feature one helix at the interaction interface have previously been identified computationally⁴³. The goal of the study was to aid elegant, systematic design of helix mimetics by bridging the gap between the irregular types of helix mimetics used and the interfaces that have been or could be targeted. Firstly, examination of all PDB entries showed that 15% of the databank contained multiprotein complexes of which 62% featured a helix at the PPI interface⁴³. Through CAS the interactions where the helix was significant to binding were identified if the helix contained hotspot residues. Overall, 480 helix mediating PPIs were identified, of which 60% featured hotspot residues on a single face of the helix, 30% featured hotspots on two faces of the helix and 10% required all three faces of the helix for recognition⁴³. Since the study was done in 2009, it is reasonable to assume the number of PPIs deposited since has increased and so have the PPIs mediated through α -helices. More recently, as well as screening for single helix mediated PPIs in the PDB, non-helical mediated PPIs were also virtually screened and assessed as potentially suitable for accommodating helical constructs at the interface²³. Targeting PPIs which are not helix-mediated could become more accessible if these can accommodate modulators or small molecule scaffolds designed based on the α -helix motif. In this study, 17 PPIs from the whole PDB database were identified and predicted as being amenable to helix based ligands even though the PPIs were not mediated through this secondary structure motif²³.

Among the single helix mediated PPIs, the B-cell Lymphoma 2 (BCL-2) protein MCL-1, pro-survival signalling protein, interacts with the pro-apoptotic member NOXA B⁴⁴. Intrinsically unstructured in solution, upon binding, NOXA B adopts a helical conformation which docks into a groove/cleft on MCL-1 occupying a site of about 1000 Å⁴⁵. The binding affinity of NOXA B₆₈₋₈₇, 0.70 μ M⁴⁶, arises predominantly from the side chain residues of one face of the NOXA B helix at $i+4$ and $i+7$ (Figure 5a).

In contrast to the NOXA B/MCL-1 PPI where a single helical face is responsible for containing the side chains which drive the binding, complexity of interaction could be observed when two or three faces of the single helix are involved in

molecular recognition⁴³. For example, 4 hotspot residues for the LBD/RXR PPI were identified through CAS study as L690, H691, L693 and L694 which span over two faces of the interacting helix at i , $i+1$, $i+3$ and $i+4$ respectively (Figure 5b). All three helical faces are involved in molecular recognition for the *Escherichia coli* ECF/RseA as CAS predicted 5 hotspot residues of one of the helical regions of *Escherichia coli* ECF at i , $i+4$, $i+5$, $i+6$ and $i+10$ as W33, H37, L38, I39 and M43 respectively (Figure 5c). A functionalised molecular scaffold, would have to reproduce or mimic the composition and orientation of the side chains in order to exploit the interactions necessary for binding affinity and specificity. However, small molecule scaffolds that project side chains to target such complex PPIs have yet to be developed.

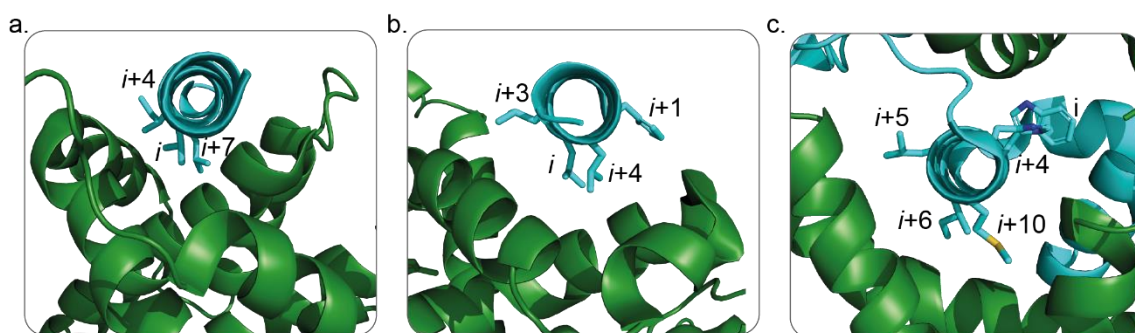


Figure 5. Helix mediated PPIs where residues on different faces of the helices contribute towards binding with a – c as examples of protein complexes with hotspot residues on one, two and three faces (PDBs from left to right 2JM6¹⁰, 1XIU⁴⁷, 1OR7⁴⁸s of NOXA B/MCL-1, LBD/RXR, *Escherichia coli* ECF/RseA respectively)

1.3.1.2 PPIs Mediated by Multiple Helices

Although PPIs mediated through a single helix can show increasing complexity, depending on how many faces of the helix are contributing to binding, arguably an even more challenging PPI can be that mediated through multiple helices. In the HIF-1 α /p300 interaction, HIF-1 α wraps around the CH1 domain of p300 forming three helical regions (Figure 6). Despite HIF-1 α being a single polypeptide chain, the three helical regions mixed with unstructured segments result in a very large surface area devoid of a particular region for targeting when compared to a single helix mediated PPI. This PPI, within the hypoxic response pathway, is very attractive for potential developments in cancer therapeutics^{49–52}. Significant efforts to understand this PPI^{53–55} have been undertaken.

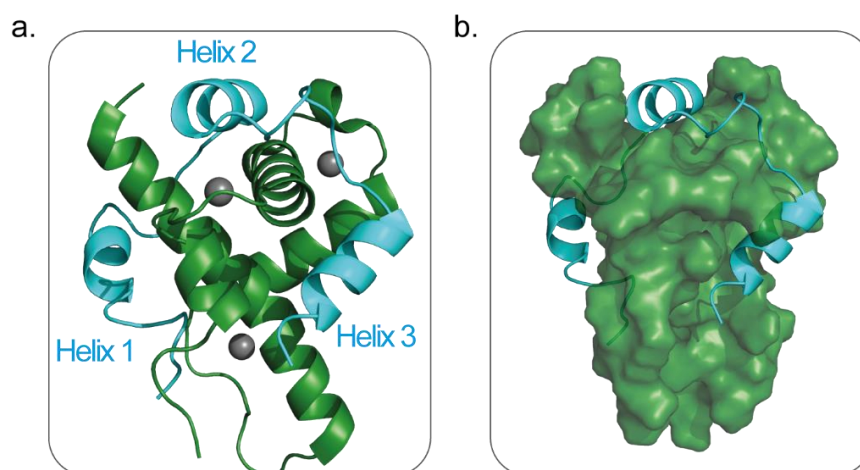


Figure 6. NMR structure of p300 (green) and HIF-1 α (cyan) (PDB: 1L8C): a) Cartoon representation of the NMR structure with zinc ions in grey. b) HIF-1 α cartoon representation wrapped around the surface of p300.

Saturation mutagenesis studies as well as CAS studies have provided little understanding of which residues on which helices are important for binding^{53,56,57} (discussed in Chapter 4, section 4.1.1). However, HIF-1 α truncation studies have identified regions of the protein, which are necessary for interaction with p300. Truncation of the shortest helical segment, helix 1, resulted in a moderate effect on activity⁵⁸. However, truncation of the largest helical segment, helix 3, resulted in complete loss of HIF-1 α activity⁵⁹. Although several approaches have been applied to discover or design inhibitors for this PPI (discussed in Chapter 4), the current knowledge has been insufficient in aiding discovery and development of successful clinical modulators.

1.3.2 β -Strand Mediated PPIs

The β -strand also has a fundamental role in many PPIs¹. The secondary structure can provide structural support or can act as a recognition motif at PPI interfaces⁶⁰. For example the β -strand structural motif is often recognized by proteolytic enzymes⁶¹ and by scaffold proteins such as the SHANK proteins³⁵.

In a β -strand conformation the side chains of each residue project on alternating sides of the backbone, with maximum distances from each other minimizing steric clashes and exposing the backbone atoms to hydrogen bond with the receptor (Figure 6a). Analysis of the structures of PPIs in the PDB mediated through β -

strands revealed that this structural motif interacts with partner proteins, either in a strand or through multiple strands coming together in a sheet (Figure 1b), *via* side chain recognition, with or without hydrogen bonding from its backbone⁶². Furthermore, antiparallel β -sheet orientation was observed to be three times more prevalent across the PDB in the formation of the interaction between protein partners⁶². This could be due to the fact that antiparallel orientation of sheets is thought to provide better orientation for the formation of backbone hydrogen bonds⁶³.

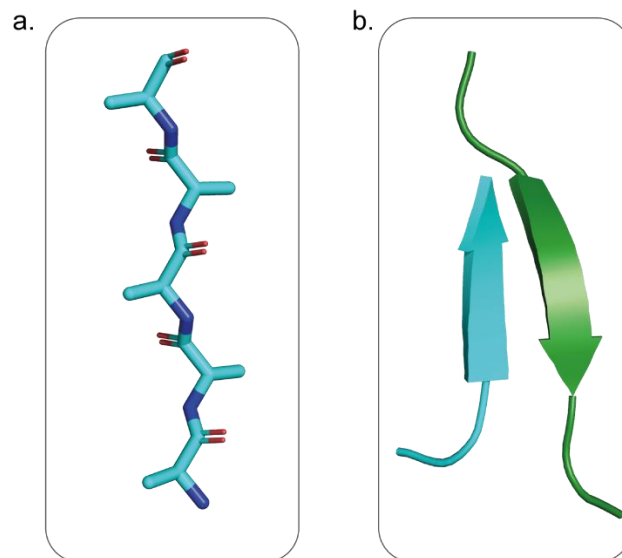


Figure 7. β -Strand structural information: a) Idealised β -Strand conformation of a poly-alanine sequence in its extended conformation; b) cartoon representation of two antiparallel oriented β -strands interacting.

1.3.2.1 β -Strand: Side-Chain Driven Interactions

The linear arrangement of amino-acids in the β -strand projects side chains of residues above and below the plane of the peptide backbone. Thus, one of the faces can exclusively bind to the interface of its ligand: only the side chains are involved in the affinity (Figure 8a). This is the case for the p65 dimer where the interaction is entirely driven by hydrophobic side-chains from one monomer interacting with the other and a glutamate-arginine charged interaction⁶⁴ (Figure 8a light blue residues).

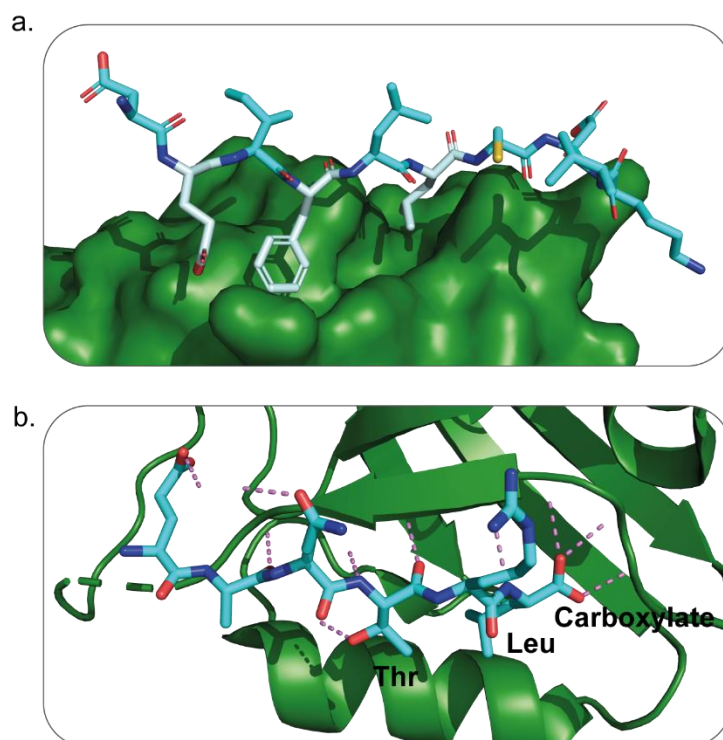


Figure 8. β -Strand interactions showing: a) the p65 dimer PPI (PDB: 1OY3⁶⁴) where only side chains are involved in binding; b) The GKAP/SHANK1 PPI (PDB: 1Q3P³⁵) where both side-chains and backbone interactions drive the binding interaction with polar contacts shown by yellow dotted lines.

1.3.2.2 β -Strand: Side-Chain and Backbone Interactions

Both faces of the β -strand and the backbone can contribute significantly towards recognition and binding events in PPIs (Figure 8b). In this case, the design of inhibitors becomes more difficult since targeting such a dispersed arrangement of polar contacts, an average of 10 hydrogen bond acceptors and donors, with a small molecule would pose considerable challenges. In particular, small molecules would need to be designed to recreate some of the hydrogen bond acceptors and donors displayed by the ligand⁶² as well as any other interactions from side chain contributions.

The guanylate-kinase-associated proteins (GKAP) are postsynaptic proteins, shown to be crucial for the recruitment and build-up of the SHANK scaffold proteins at the synapse excitation site^{65,66}. The GKAP/SHANK interaction is a prime example of a β -strand mediated PPI where GKAP adopts an extended conformation driving the binding through both side chains and backbone

hydrogen bonding to the SHANK receptor (Figure 8b)³⁵. The GKAP hexapeptide GKAP ligand forms seven backbone hydrogen bond contacts, 3 stemming from the carboxylate, oriented towards one face of the binding interface while the side chain hydrogen bond from Thr extends to the opposing side of the interface^{35,67}. Lastly, it is known that the Leu side chain is accommodated within a hydrophobic pocket of the SHANK1 PDZ domain^{35,67}. Although the Gln and Arg residues form salt bridges, there is little evidence to suggest these are imperative for binding¹⁹.

1.4 PPIs and Disease

PPI do not work in isolation but rather in networks, where protein complexes are formed in a cascade with a myriad of other partner proteins in order to elicit a biological response. Modulation of a PPI network, such as a mutation, over or under expression of a protein, can initiate or perpetuate disease states. For example, in specific cases overexpression of α -helix mediated proteins, involved in PPIs that regulate cell apoptosis, contributes to cancer proliferation and progression^{4,68–70}. Specifically, the p53/*hDM2* PPI has been studied extensively leading to several potent inhibitors⁷¹. A large number of PPIs targeted in anti-cancer research are related to cell apoptosis, however a number of other PPIs are also relevant to cancer such as the previously mentioned HIF-1 α /p300 PPI. Furthermore, aggregation of β -sheet structures in some proteins was thought to be responsible for a number of neurological disorders³⁹. The study of beta amyloid aggregation found in Alzheimer's disease is one example⁷². In addition, neurological synapses depend upon PPIs for amplification and localisation of neural transmissions. The SHANK family of proteins mediate such interactions through recognition of the β -strand structural motif. Malfunction of these PPIs are thought to result in a number of neurological diseases^{73–76}.

1.5 Strategies for Targeting PPIs

Three main strategies have been pursued which have enabled the discovery of modulators of PPIs: i. phenotypic drug discovery, ii. target-based screening and iii. structure-based design (SBD). Phenotypic drug discovery involves screening of drug-like compound libraries, with the identity of the target protein elucidated only after the desired biological effect is observed^{1,77}. For example, lenalidomide was discovered through a phenotypic screen where its ability to downregulate tumour necrosis factor (TNF) was observed⁷⁸. The target protein, E3 ubiquitin ligase cereblon, was elucidated after the FDA approval of lenalidomide in the treatment for multiple myeloma⁷⁹.

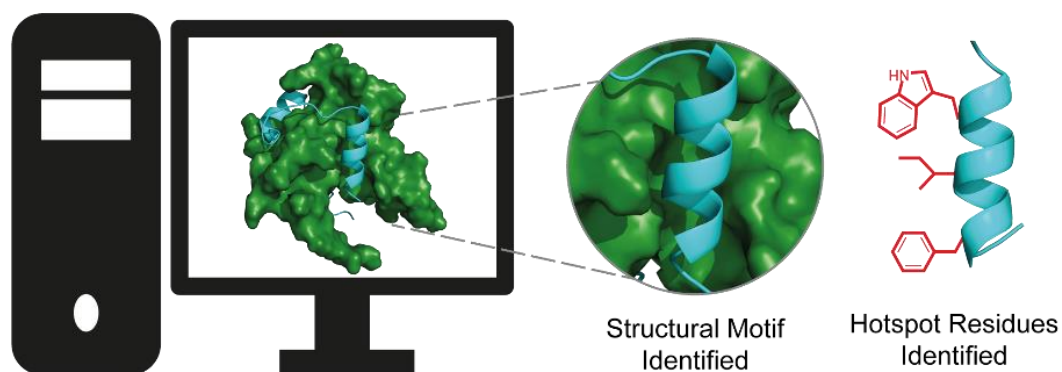
In contrast, the role of the target protein is known in target-based screening where high-throughput screen (HTS) or fragment-based discovery can be applied, with a high-resolution X-ray/NMR/cryo-EM structure favourable but not essential^{1,80}. HTS allows screening of compounds effectively and economically and has been adopted by both academia and industry. The HTS campaign success rate is dependent on the quality of compound in the library being screened to offer ideal hits with drug-like properties, low toxicity and high target specificity⁸¹. Nutlin compounds, inhibitors of the p53/*hDM2* PPI were discovered using HTS^{25,38} (section 1.6). However, as well as relying heavily on the quality of compound library, HTS also suffers from low hit rate and high numbers of false-positives and false-negatives. These disadvantages warrant the development of peptide-based inhibitors where the number of false positives or false negatives can be kept low, inhibitor design can recapitulate key hotspots to maintain superior target specificity. Similar to HTS, fragment-based drug discovery relies on screening libraries, however these contain fragments with much lower molecular mass. The idea is to discover several core motifs which can then be elaborated further as was the case with the ABT-737, a potent inhibitor against the BCL-xL synthesised through linking of two fragment hits⁸² (section 1.6). Due to their smaller size, low heavy atom count, fragments tend to bind with low affinity and require the combination of sensitive biophysical techniques such as NMR, SPR, X-Ray or ITC to validate the hits⁸³.

Although the first two mentioned strategies for targeting PPIs are valuable, the main focus will be given to SBD where high-resolution X-ray or NMR structures of the target are required, and inhibitors are designed based on key structural elements found at PPI interfaces¹ (Figure 9). A number of approaches can be applied when intimate knowledge of the target protein is available: peptide/peptidomimetic design, small molecule discovery, *in silico* structure based design⁸⁴, virtual library screening⁸⁵ and fragment-based discovery⁸³ to name a few. Following the selection of a target, steps are then taken to identify the driving forces of binding through computational methods applied to the crystal structure or NMR ensemble¹ (Figure 9i). Although the interface of the selected PPI can be large, computational assessment can uncover binding regions critical for the interaction. First the topology at the PPI interface is identified (helical, sheet etc.). These regions can be further studied through the previously mentioned CAS strategy where hotspot residues are identified (Figure 9i). Experimental methods are then used to validate the computational predictions and inhibitor design that can engage with the target surface through mimicry of the important residues of binding can begin (Figure 9ii). Design approaches can be categorised into four classes: A) peptide mimetics with small modifications made in order to optimise properties i.e. stabilised/stapled helices; B) modifications to backbone and side chains are introduced in this class (α/β peptides⁸⁶, peptoids^{87,88}, or insertion of a small molecule-like building blocks^{89,90}) that also extends to foldamers, oligomeric structures capable of adopting a desired conformation in solution; C) structural mimics which consist of a non-peptidic backbone, a small molecule scaffold that projects key residue side chains; D) molecules with no link to the original binding peptide structure^{14,91,92}. Although there are examples of foldamers^{93–95} and other topographical mimics^{14,91,96,97}, often the simplest of strategy in targeting a difficult PPI is to begin with a peptide based inhibitor.

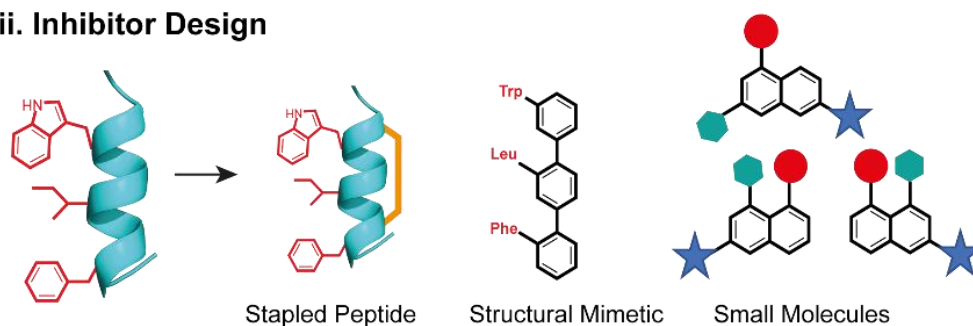
As well, high-throughput *in silico* docking can be used to identify potential ligands that target the same binding site using the apo-structure for docking. Following experimental binding studies of these compounds the best compounds are selected and further undergo several rounds of optimisation. Explicit examples of this methodology applied for the design of both peptide-based inhibitors and small molecules are further focused and expanded on in future subsections (1.5.1 –

1.5.2) and mention of screening based methods showcasing as complimentary tools to SBD also mentioned (1.5.3).

i. Computational Analysis of Target



ii. Inhibitor Design



iii. Binding and Biophysical Analysis

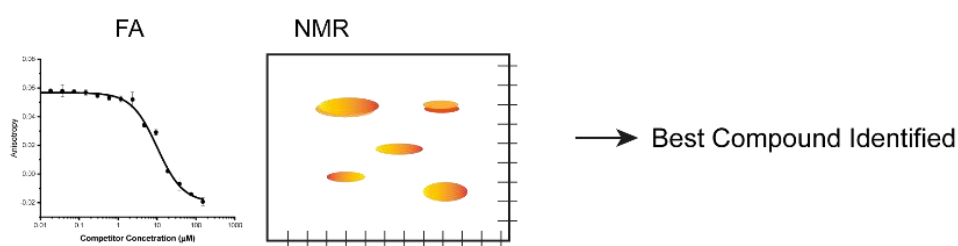


Figure 9. Structure based PPI inhibitor design: i) The target PPI is studied using computational methods where topology at the interface and hotspot residues can be ascertained; ii) Constrained peptides or structural mimetics can now be designed based on the information gained from the computational studies. Small molecules can also be identified through *in silico* docking and structure similarity search; iii) The designed or discovered ligands are then tested in biophysical assays and the best binders are taken further through rounds of optimisation.

1.5.1 α -Helix Mimetics: Stapled Peptides

Computational filtering of multi-protein complexes available in the PDB showed that roughly 60% were mediated via an α -helix, and often with hotspot residues lying on a single face of the helix⁴³. Thus, inhibitor design based on the conserved helix structure, has made targeting this type of mediated PPI a success story^{38,92,98}.

Peptides are attractive starting points in targeting PPIs as they already retain surface recognition properties and as they are derived from native proteins, less likely to cause unwanted side effects. However, in solution, peptides typically do not adopt a stable folded conformation and suffer from poor cell permeability. In particular, constrained helices have been developed as versatile peptidomimetics for targeting PPIs^{99,100}. Constraining a peptide in a helical conformation increases peptide proteolytic stability^{101,102}, improves cellular uptake^{102,103} and can pre-organise the peptide into a bioactive conformation resulting in less unfavourable change in binding entropy and thus enhanced target affinity^{104,105}. As well, depending on the helix stabilisation technique used, the mimetic could be endowed to target its partner through one, two or all three faces of the helix^{106,107}.

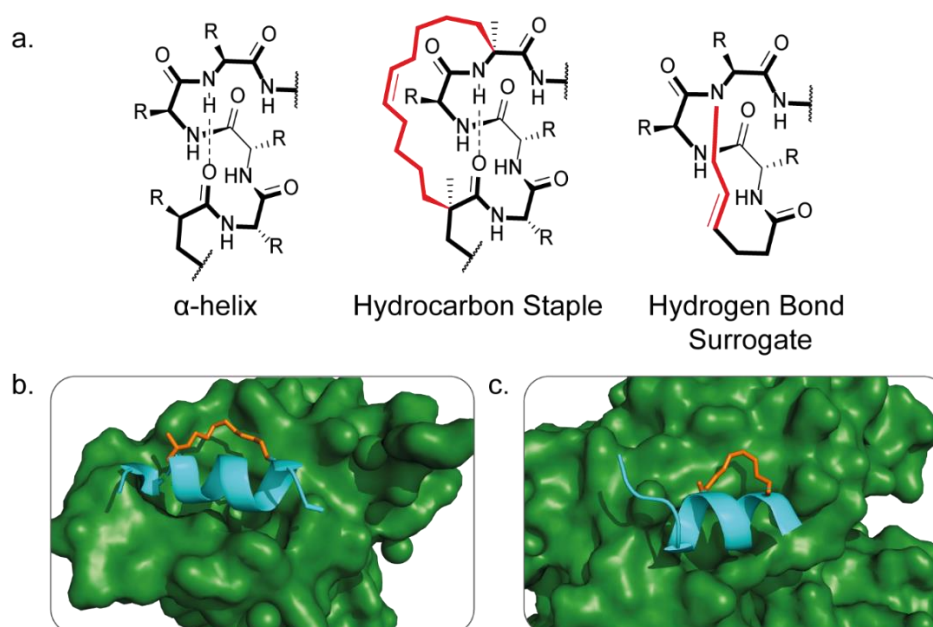


Figure 10. Initial peptide constraint approaches: a) The arrangement of atoms in an α -helix compared to that of a helix stapled at the residues using a hydrocarbon constraint and head to tail backbone nitrogen hydrogen bond surrogate (HBS); b) Crystal structure of the p53 hydrocarbon stapled peptide, where the staple drapes over the *hDM2* protein partner (PDB: 3V3B¹⁰⁸); c) Crystal structure of the co-activator peptide bound to ER α where the hydrocarbon staple drapes over the protein partner (PDB: 2YJD¹⁰⁹).

The first reported stapled peptides were obtained through the introduction of non-natural α,α' -disubstituted amino acids with olefin tethers, which afforded cross-linked hydrocarbon stapled peptides accessed *via* ring-closing metathesis (Figure 10). Considerable efforts were placed in understanding the optimal positioning of the staple, the stereochemistry and length of the linker. These studies generated a breakthrough with ligands based on the BID BH3 sequence which showed the stapled sequence upheld helical attributes, increased binding affinity as well as improved cell-permeability and proteolytic stability¹⁰⁴. The enhanced helical conformation of stapled peptides conceals the peptide bonds preventing them from recognition by proteolytic enzymes which hydrolyse extended conformations *in vivo* and *in vitro* while maintaining biological affinity and reducing tumors in leukaemia xenografts^{12,110}.

Hydrocarbon stapling has been extensively expanded with stapled peptides of Bcl-2 family members and stapled peptides designed against HIV-1/gp41 studied and potent inhibitors for these PPIs identified^{100,101}. Furthermore, crystal structures of the p53/hDM2¹¹¹ and ER α /co-activators¹⁰⁹ both showed the hydrocarbon staple as an active participant in the binding, draping over the protein partners and forming additional hydrophobic interactions (Figure 10b,c). This discovery highlighted the importance of staple placement and the need for careful analysis to avoid clashes or to enhance potency through staple interactions depending on the PPI.

Hydrogen bond surrogates (HBS Figure 10a) designed by the Arora group have also been successful, studied and applied in both *in vitro* and *in vivo* set-ups. The HBS helices based on HIF-1 α showed improved helicities by CD and modest potencies to p300 by ITC and are discussed in depth in Chapter 4. Mouse xenographs models, treated with one of the designed HBS, exhibited 53% tumour reductions when compared to mice in control groups. In depth discussion of the HIF-1 α /p300 PPI will be captured in Chapter 4.

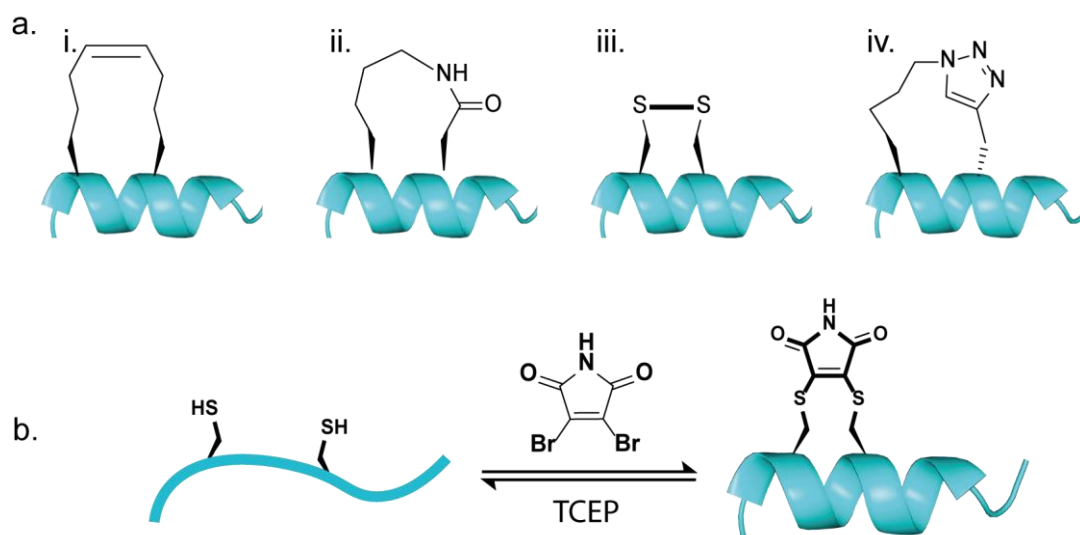


Figure 11. Helix enforcing techniques: a) Stapling techniques through side chain residues: i) Hydrocarbon staple achieved through the insertion of a non-natural amino acid and ring closing metathesis; ii) Lactam bridge obtained from using natural residues with special protecting groups on both the acid and the amine side-chain residues; iii) Disulfide constraint formed by the oxidation of cysteine residues; iv) Cu(I) catalysed azide-alkyne cycloaddition; b) Dibromomaleimide stapling design with reversibility accessed via TCEP reducing agent.

Similar success has been achieved with many types of peptide constraints. The toolbox for constraining peptides includes, but is not limited to: hydrocarbon constraint^{112,113}, lactam bridges¹¹⁴, disulfide bridges¹¹⁵, thiol crosslinked systems^{116,117}, Cu(I) catalysed azide-alkyne cycloaddition¹¹⁸ and hydrogen-bond surrogates^{99,113,119} (Figure 11a). Adding to this toolbox, the use of simple disulfide bridges, crosslinking of (homo)cysteine residues and other modification of thiols have been described^{100,117,120–123}. As an original example, the disulfide bridge was used to constrain a short peptide sequence which, due to containing the LXXLL motif, was shown to bind the estrogen receptor α ¹²⁰. The *i* and *i*+3 disulfide-constrained peptide showed better binding activity towards the receptor compared with a *i* and *i*+4 amide stapled analogue, establishing this technique as a viable avenue for constrained peptides¹²⁰. Recently, within the Wilson group, a staple protocol that exploits the reaction of dibromomaleimide with two (homo)cysteine residues judiciously placed at *i* and *i*+4 in the sequence of peptides for BID and RNase S has been reported¹¹⁷. The approach was shown to be accessible, proceeding in buffer on unprotected peptides, reversible, and did not require the insertion of non-natural amino acids in the peptide sequences (Figure 11b). However, the technique should continue to undergo further

investigation to extend its application to other PPIs and achieve success; to become established among the hydrocarbon and HBS structures in the stapled peptides toolbox.

1.5.2 β -Strand Mimetics: Macrocycles

Despite its important role at the interfaces of PPIs, applications of β -strand mimics as modulators are limited when compared to the α -helix¹. β -Strand based design is challenging because it needs to account not only for the role of side chains in protein recognition, but also the hydrogen-bond interactions from the peptide backbone. Often strand mimetics can form insoluble aggregates because they are inherently designed to form hydrogen bonds with other strands they encounter^{1,62}.

Macrocycles are frequently encountered as bioactive molecules in nature such as mycobacillin, a cyclic peptide with antifungal activity¹²⁴. Furthermore, macrocycles have been identified as structures that can target β -sheets or β -strand mediated PPIs⁸. HIV-1 protease accommodates the β -sheet structural motif of the substrate in its binding cavity and has been successfully targeted with macrocyclic compounds^{39,90,125}. The macrocycle compound targeting HIV-1 protease contains a para-substituted aromatic ring within the macrocyclic portion, which replaces two of the side chain residues responsible for binding within the PPI interface^{126,127} (Figure 12a,b). Although this is not a PPI inhibitor, proteases recognise their substrates in an extended conformation and the learnings from protease inhibitor design could be built upon and applied towards targeting PPIs similarly mediated. This type of macrocycle is devoid of intramolecular hydrogen bonding. In contrast, macrocycles such as **2** designed by the Nowick group, to disrupt amyloid formation, relied on stabilisation of the β -strand through the insertion of the non-natural building block Hao (hydrazine, 5-amino-2-methoxybenzoic acid, oxalic acid) and formation of intramolecular hydrogen bonding⁸⁹ (Figure 12c).

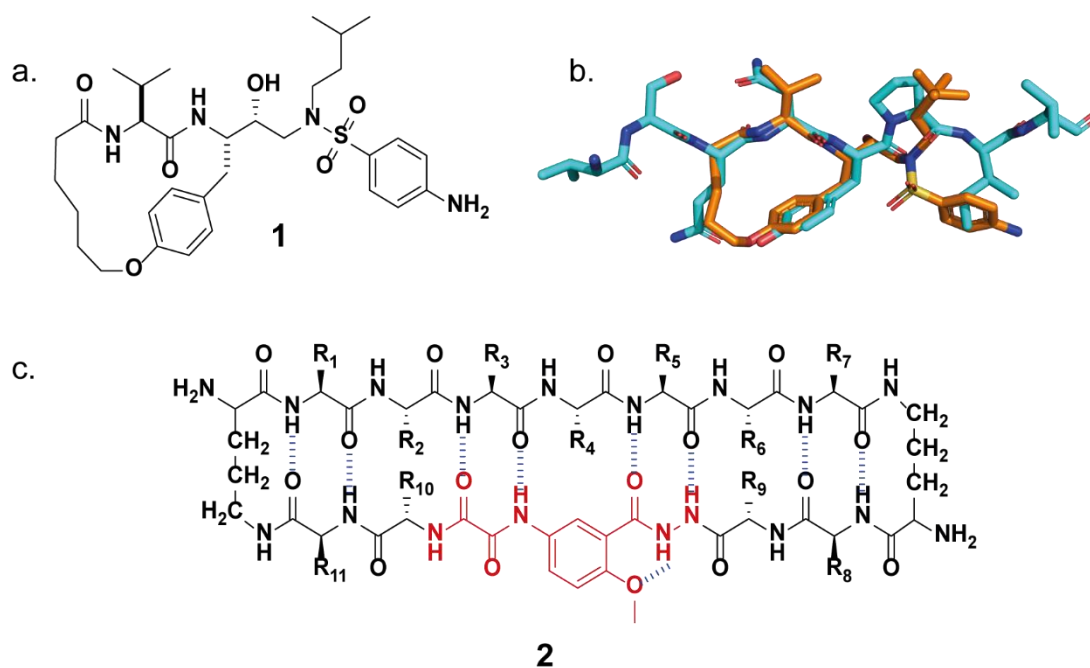


Figure 12. Macrocycles designed to target β -strand interactions: a) HIV-1 protease 1 inhibitor; b) Visual overlay and comparison of the binding conformation of the HIV-1 protease natural peptidic substrate (cyan, PDB: 1MT7¹²⁷) and the macrocyclic inhibitor (orange, PDB: 1D4L); c) Nowick designed macrocycle 2 with the Hao (hydrazine, 5-amino-2-methoxybenzoic acid, oxalic acid) aromatic building block incorporated, which fosters hydrogen bonding (blue bonds) between the strands. The rigidity of the aromatic ring ensures that the molecule forms and intermolecular hydrogen bond which also aids the molecule in adapting a conformation that maximises the hydrogen bonds between strands.

The macrocycle confers a degree of conformational pre-organisation permitting side chains to interact with partner proteins in an extended conformation and form important interactions. Equally, the macrocycle is not completely rigid, preserving sufficient flexibility in order to mould in the appropriate conformation. Thus, the macrocyclization restricts peptide freedom and affords similar optimisation results as helix-constrained mimetics, with increased proteolytic stability by fending amide bonds, lower entropic cost upon binding while still potent^{39,128}. The Hao building block was used as a platform for successful rationally designed macrocyclic inhibitors of seven amyloidogenic peptide sequences, among which A β and Tau were used, to control amyloid aggregation⁸⁹. The ability of the Hao building block to template the seven different peptide sequences into β -strands was proven using X-Ray crystallography and ¹H NMR studies, while the ability of the macrocycle to slow down aggregation was monitored using Thioflavin T fluorescence and electron microscopy experiments⁸⁹. Although the Hao building

block has been studied extensively for its ability to template β -sheets in order to prevent aggregation pathways, it has not been applied in the design and study as a β -strand mediated PPI inhibitor.

Protein epitope mimetics (PEM) constitute another successful approach in targeting secondary structure mediated PPIs (Figure 13a). The biologically relevant amino acid sequence involved in binding is transferred onto a synthetic scaffold, for example a semi-rigid hairpin stabilised template to afford a β -hairpin macrocyclic PEM¹²⁹. Initially, loops of varying sizes were cyclised through the D-Pro-L-Pro template, which adopts a β -turn, in the design of cationic antimicrobial peptides based on protegrin I¹³⁰. Moreover, as opposed to using the native peptide sequences phage display technology has been exploited to afford PEMs. The discovery of peptide sequence (DCAWHLGELVWCT) that binds the human IgG antibody which displayed a beta-hairpin structure in crystallography studies¹³¹ is a prime example where the introduction of the D-Pro-L-Pro template to this sequence produced a macrocyclic PEM inhibitor with higher affinity to IgG than the native sequence¹³². However, it can be envisioned that phage display could be harnessed in the design of novel ligands for any target protein even without any knowledge of the native ligand sequence. Lastly, PEM technology is very versatile since it was shown to be able to mimic a helix and disrupt a helix mediated PPI. An PEM inhibitor based on p53 was designed in a octapeptide beta-hairpin containing the three hotspot residues (Phe19, Trp23 and Leu26) further optimised to contain a chlorotryptophan and showed a K_d of 25 nM to *hDM2*¹³³ (Figure 13b).

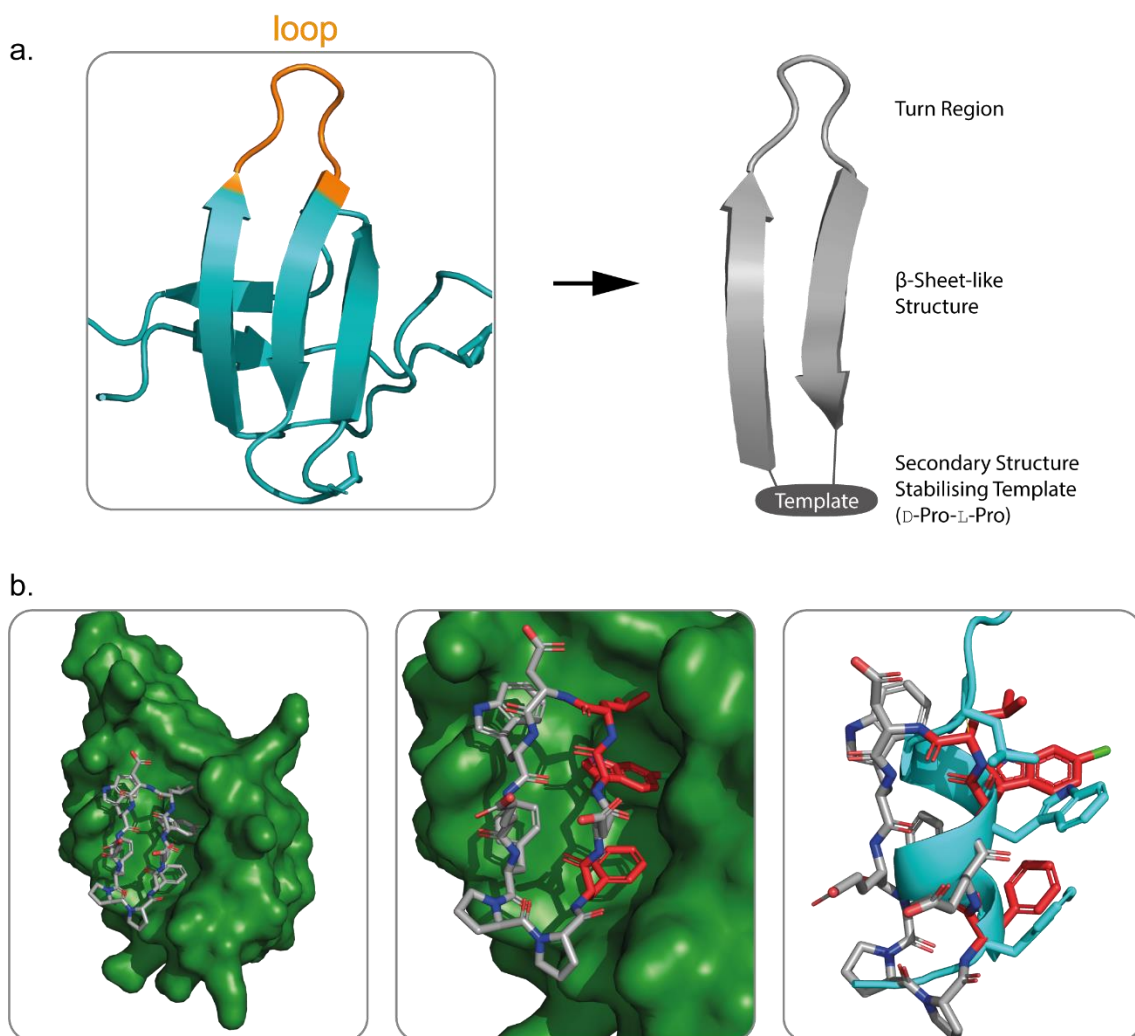


Figure 13. Protein epitope mimetics based on constrained beta-hairpin: a) Cartoon representation of how a loop interaction can help influence the design of the PEM structure which consists of three main variables: loop, building blocks and template regions; b) beta-hairpin mimic of the p53 peptide bound to *hDM2* (PDB: 2AXI) showing in red (middle frame) the three hotspot residues mimicked and in an overlap with the p53 peptide (PDB: 1YCR).

Due to their fairly large molecular mass, macrocycles do not abide by the ‘rule of 5’ applied to small molecule inhibitor discovery and have been under-explored in the past^{128,134}. However, macrocycles have undergone a resurgence. It has become evident that the macrocyclic structure could be designed based on any linear peptide sequence of interest that binds based on an extended conformation resulting in a potent and selective inhibitor, as seen with macrocycles developed to target proteases⁹⁰. Synthetically, a number of approaches can afford macrocycles: head to tail, side chain to tail, side chain to side chain cyclisation with or without incorporation of non-natural building blocks¹³⁵.

1.5.3 Peptide Phage Display Libraries: Complementary Tool for Ligand Discovery

Although not a structure-based approach, phage display technology has become a powerful combinatorial tool in PPI ligand discovery that is worth mentioning when discussing approaches towards targeting PPIs, as we have seen in the discussion of PEM design. Exploiting the phage particles ability to display numerous peptide structures on their coat proteins, large and diverse peptide libraries can be expressed using molecular biology methods^{136–138}. In a phage screening campaign, the target protein is immobilised, incubated with the phage library, and any non-binding phage is washed away. Following several rounds of panning (screen, wash, elute, amplify), the binding phage is eluted and amplified in order to be sequenced and analysed to offer insight into driving forces of binding to the target protein^{138,139}.

Binders with high affinity and specificity are identified through this technology^{139–141}. Notably, the peptide sequences can be novel, and as such can be taken forward in rational design as a new starting point when designing mimetics. This technique can be used as a complementary tool alongside other structure-based approaches mentioned thus far. For example, the application of phage display based on affimer template, uncovered selective ligands for several members of the same family of proteins, BCL-2¹⁴⁰. The binding ligands, discovered for MCL-1, BCL-xL, BCL-2, BAK and BAX, were analysed using several biophysical characterisation methods to validate binding interactions and found to inhibit these PPIs and elicit selective and potent binding to their respective target. Although a number of inhibitors were identified, 11 for MCL-1 and 11 for BCL-xL, only 3 were found by X-Ray crystallography to bind within the BH3 cleft of BCL-xL observed to be narrower than when bound to native BIM ligand¹⁴⁰. This new binding mode accessed for BCL-xL is significant, and, along with the new binders could offer inspiration in future structure-based design approaches for this family of proteins. These types of investigations can help elaborate the understanding of molecular recognition elements, particular side chains, needed for binding for a family of proteins which usually bind some of the same partner proteins.

Phage display technology has proven to be very versatile, adapted to also generate cyclic peptides¹³⁸ and bi-cyclic¹⁴² peptide binders. Generating novel binders of mimetic quality in one technology. However, the technique can be time consuming due to the several rounds of panning and tends to show bias towards hydrophobic residues¹¹. Based on recent advancements in computational chemistry, a more rapid *in silico* alternative to phage displays could help shorten the time span and potentially generate peptide sequences similar to those obtained experimentally.

Similarly to phage display technology, split-intein circular ligation of peptides and proteins (SICLOPPS) enables the generation of cyclic peptides for screening of PPIs¹⁴³. The screening method takes place *in vitro*, with active peptides disrupting a PPI formed as a natural complex, however, the methodology is limited to screening peptides made of natural amino acids and prior information about the PPI complex is favourable.

1.6 Small-Molecule Inhibitors of PPIs

Irrespective of a particular strategy for targeting PPIs it is worth mentioning some examples of successful PPI modulation achieved with completely non-peptidic molecules. These small molecules inhibit PPIs without necessarily mimicking secondary structure motifs. The following examples stand as evidence that PPIs are tractable molecular targets.

Perhaps the most renowned examples are the tetra-substituted imidazoline (Nutlin)²⁵ and the ABT compounds⁸². Identified by high-throughput screening (HTS), the Nutlin series of compounds target the p53/*hDM2* PPI. Several rounds of synthetic optimization yielded Nutlin-3a with an IC_{50} of 90 nM, activity largely due to the ability of the molecule to recapitulate interactions of p53 similar to those exhibited by the hotspot residues Phe19, Trp23 and Leu26²⁵ (Figure 14a). Further work by Roche yielded second generation compounds RG7112¹⁴⁴ and RG7388¹⁴⁵ (Idasanutlin) with 18 nM and 6 nM binding affinities, respectively, which also mimic the hot spot constellation (Figure 14b). Although RG7112 failed to pass Phase I clinical trials, Idasanutlin is currently in Phase III clinical trials^{145,146}.

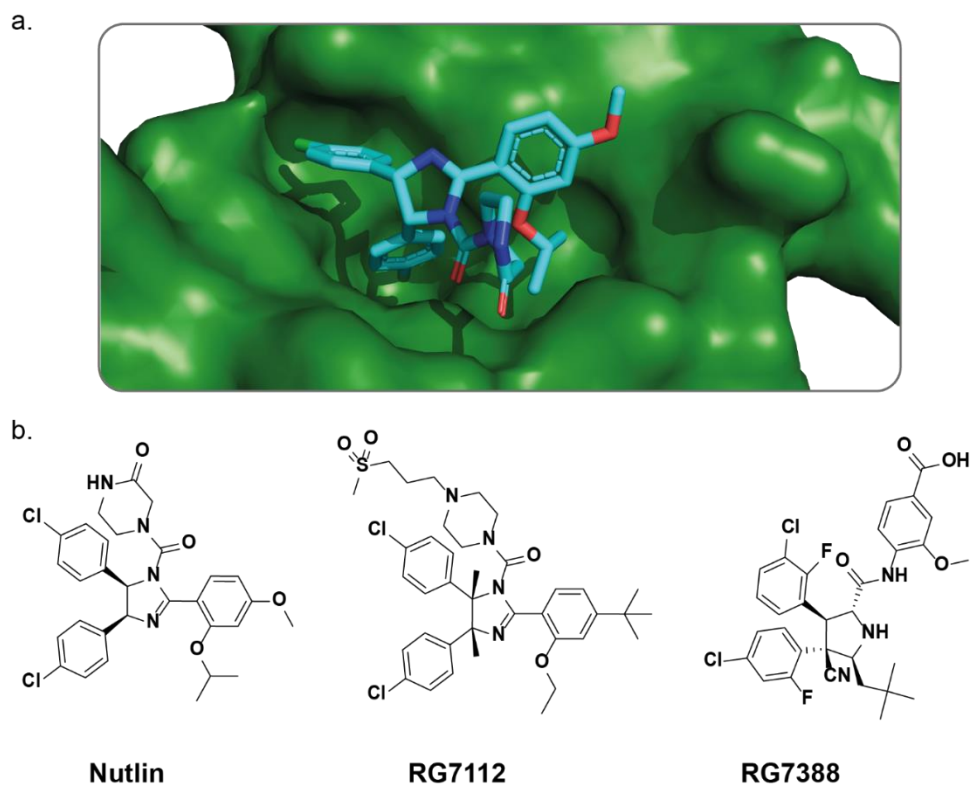


Figure 14. Small molecule inhibitors of the p53/*hDM2* PPI: a) Crystal structure (PDB: 4J3E) of Nutlin 3a bound to *hDM2* discovered through HTS; b) Chemical structure of Nutlin 3a and the second generation compounds optimised by Roche.

Abbot Laboratories used an NMR-based fragment screen and discovered two fragments in two adjacent sites which were then linked together and elaborated through several rounds to generate ABT-737, a potent inhibitor against the BCL-xL with a K_i of 0.6 nM⁸². Although ABT-737 failed in clinical trials due to poor bioavailability, optimisation resulted in ABT-263 (Navitoclax),¹⁴⁷ which was further optimised to the clinically successful ABT-199 (Venetoclax) with a $K_i < 0.01$ nM. Venetoclax has been successfully used against chronic lymphocyte leukaemia¹⁴⁸. The size of these compounds is relatively large when held against Lipinski's rules, and may be a consequence of the challenging PPI target.

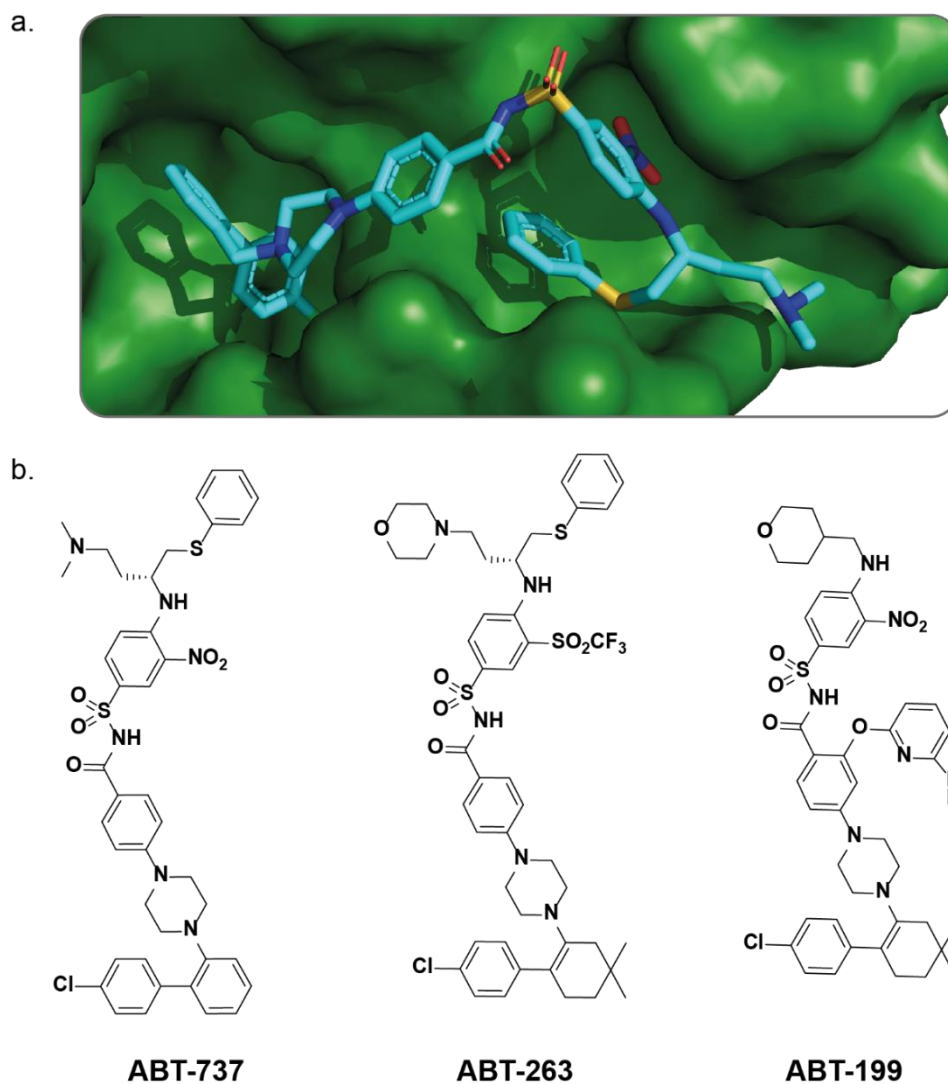


Figure 15. Small molecule inhibitors of BCL-x: a) Crystal structure of ABT-737 molecule bound to the protein (PDB: 2YXJ); b) Chemical structure of ABT-based developed inhibitors.

Although no common methodology has been achieved in identifying small molecule modulators of PPIs, more examples of inhibitors are emerging^{8,149,150}. The large number and structural diversity of PPIs still make these challenging to design inhibitors for.

1.7 Project Aims

1.7.1 Targeting α -Helix and β -Strand Mediated PPIs

Designing suitable small molecules to modulate PPIs is desirable. However, targeting PPIs needs further scientific attention to explore some of the more challenging secondary structure mediated PPIs. Peptide based modulators of α -helix and β -strand mediated PPIs are fundamental for understanding the driving forces and binding affinities of PPIs and could help inform the structure based design approach and make it more efficient. The goal of the project was to study both α -helix and β -strand mediated PPIs, to understand the driving forces for each interaction and harness structural information in the design of peptide-based inhibitors.

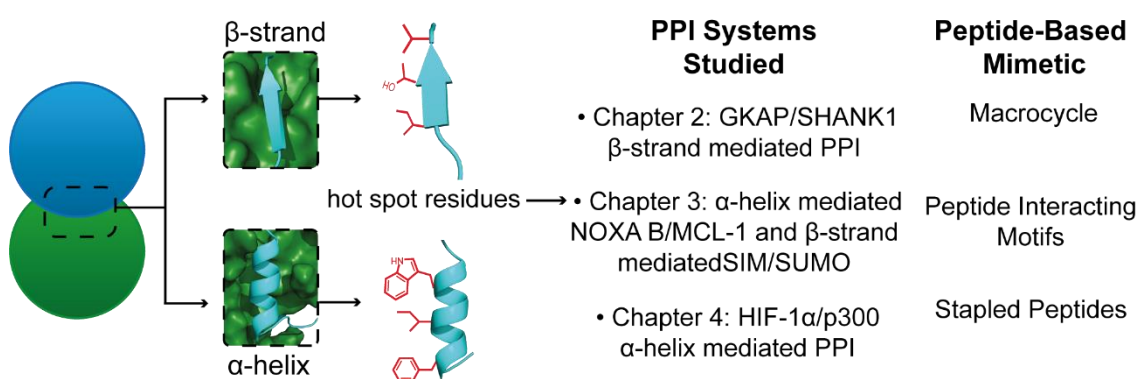


Figure 16. Summary of experimental systems investigated and chapters which cover each of these.

- Chapter 2: Probe the driving forces of the β -strand mediated GKAP/SHANK 1 to aid the rational design of a β -strand mimetic to target this interaction.

The relatively short GKAP sequence that binds SHANK1 could also be an ideal starting point towards studying and designing mimetics for a β -strand mediated PPI. Assay development will enable experimental alanine scanning and the results of the alanine scan will be compared with the anticipated GKAP key hotspot residues within the literature. The outcome would further be compared with the predicted hotspot residues generated by the University of Bristol

researchers in their endeavour to improve their alanine-scanning tool and comment on the necessity of experimental alanine scan. Furthermore, probing of the GKAP/SHANK1 binding would be tested using designed peptide variants, where point-variations to the GKAP ligand based on other PDZ domain binders, would be attempted. The ultimate goal was to use the results from the experimental alanine scan and the designed peptide variants to design a peptidomimetic inhibitor of this interaction.

- Chapter 3: Synthesis of Peptide Motifs of α -helix mediated NOXA B/MCL-1 and β -strand mediated SIM/SUMO, to validate a new computational tool that could predict new peptide motifs rapidly and efficiently

NOXA B/MCL-1 and SIM/SUMO are relatively well studied PPIs and were chosen as model systems for the development of peptide interacting motifs (PIMs). In collaboration with the computational team in Bristol, the ligands of these model PPIs were subjected to a rapid computational mutagenesis scan and novel peptide binders akin to those generated through phage displays were to be tested for binding to their respective proteins. The hot-spot residues were previously studied for these two systems and with that information at hand the experimental validation of *in silico* generated PIMs was at the forefront of this branch of the project. Ideally, the computational PIMs would maintain binding affinity to the target proteins. These could then be used to further understand the driving forces of binding for these two PPI systems. Furthermore, the new PIMs could point towards the development of alternative templates for inhibitor design later down the line.

- Chapter 4: Expand the application of the dibromomaleimide stapling to target the HIF-1/p300 PPI.

HIF-1 α /p300 remains an attractive target for cancer therapeutics. Previously, the Wilson group investigated the driving force of the HIF-1 α /p300 interaction by analysing each helical region separately and later combined to understand the affinity to p300. Helix 3 was identified as the region with the highest affinity for p300. A constrained helix was to be designed and synthesized to mimic the largest helical region of the HIF-1 α , Helix 3 using the dibromomaleimide stapling

technology. Ultimately the goal was to expand the application of the dibromomaleimide staple and further understand advantages or disadvantages of this approach. Ideally more insight as to whether this region could be amenable to small molecule modulation would also ensue from this line of inquiry.

Chapter 2: Targeting a β -strand mediated PPI: GKAP/SHANK1

2.1 PDZ Domains

PDZ domains have been the focus of intense research as they facilitate a large number of PPIs and could be potential drug targets in neurological diseases and cancers¹⁵¹. Canonical PDZ domains consist of approximately 90 amino acids and are composed of six β -strands forming an antiparallel barrel and two α -helices. The binding groove, located between one of the α -helix (α B) and β -strand (β B) structural elements of the PDZ domain, recognises the C-terminus consensus sequence X-(Ser/Thr)-X- ϕ -COOH where X can be any amino acid and the ϕ hydrophobic residue^{151–154} (Figure 17). However, it is worth noting some PDZ domains also recognise internal peptide motifs of some target proteins^{152,155}.

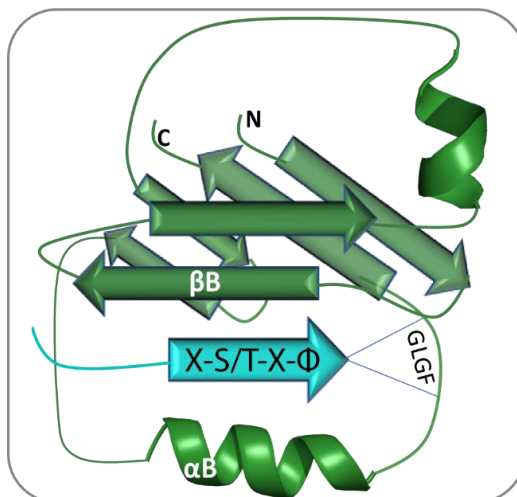


Figure 17. Cartoon representation of a canonical PDZ domain structure (forest green) made up of 2 α -helical regions and 6 β -strands, with a typical C-terminus ligand (cyan) binding between the α B and the β B structural elements. The terminal ϕ hydrophobic residue is accommodated within a GLGF pocket in the PDZ domain.

These domains are found in hundreds of signalling proteins. In particular, the PDZ domain is one of many domains found in scaffold proteins - large multidomain structures that facilitate the formation of large complex structures through protein-protein interactions¹⁵⁶. The formation of such complexes facilitate multiple functions among which is the localization and enhanced efficiency of signalling^{75,157–159}. The PDZ domain, thus, plays a crucial role in the formation

and stability of protein complexes with many biological functions among which signal transduction is one of the chief roles¹⁵².

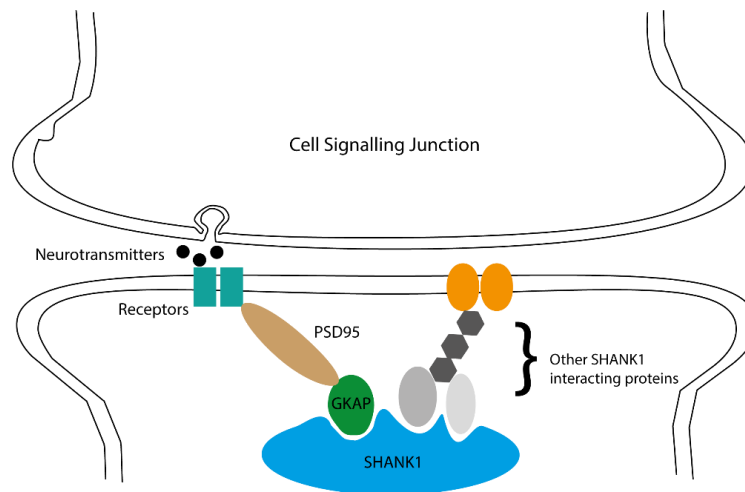


Figure 18. Cartoon representation of the scaffold SHANK1 protein and its many PPIs in the post-synaptic density. GKAP is represented in green, interacting with SHANK1 through the PDZ domain while other protein interactions are mediated through other SHANK1 domains. GKAP is observed to form a further interaction with PSD95 which in turn interacts with the receptor.

Among the PDZ domain containing scaffold proteins are SHANK proteins. These are found at the post-synaptic synapse and link neurotransmitter proteins to intracellular signalling pathways¹⁶⁰. They are known as ‘master’ scaffolding proteins – large multidomain structures that facilitate the formation of large complex structures through protein-protein interactions¹⁶¹. Crucially, the SHANK proteins are enriched at the postsynaptic density (PSD), where they aid proper synaptic development and function through organizing postsynaptic complexes^{158,162,163}. The PDZ domain of SHANK proteins interact with partner proteins such as guanylate-kinase-associated postsynaptic-density protein (GKAP)^{35,164,165} and G-protein coupled receptors¹⁵⁸ (Figure 18). However, precise understanding of driving forces and specificity of binding between PDZ domains and their ligands is lacking^{154,166,167}. There are three known family members: SHANK1-3. In mice, SHANK1 gene mutations presented impaired dendritic spine development and weaker synaptic transmissions⁷⁴. These mutations were associated with increased anxiety related behaviours, impaired long-term memory and fear learning reaffirming the relevance of SHANK1 in cognitive processes⁷⁴. The importance of the SHANK family of proteins is further underscored by research which shows mutations in SHANK3 are associated with

autism spectrum disorders^{76,163}. Further understanding of the SHANK1 family member, is pivotal in understanding neurological mechanisms and the implications of their malfunction in disease⁷⁵.

2.1.1 GKAP/SHANK1 PPI

One PPI between the PDZ domains of SHANK proteins is with the GKAP family. The PDZ domain of SHANK recognizes the sequence X-Thr-X-Leu-COOH found at the C-terminus of GKAP¹⁶². The 6 residues of the GKAP C-terminal sequence adopt an extended β -strand conformation upon binding, using back-bone hydrogen bonding and side chain groups for recognition observed in the crystal structure³⁵ (Figure 19). A total of 7 backbone hydrogen bond interactions on one face, a hydrogen bond through the Thr side chain on the opposing side of the binding interaction and the Leu residue packed within a hydrophobic pocket in the PDZ surface can be observed (Figure 19). Crystal structures of the SHANK1 PDZ domain with and without the GKAP binding ligand were solved in 2003³⁵. Once the GKAP/SHANK connection is made, GKAP in turn interacts with another guanylate kinase domain protein PSD-95¹⁶⁸ (Figure 18). It is thought that these interactions are organizational and are involved in synaptic targeting of receptor membrane proteins, working as linkers in the synaptic pathway.¹⁶² Of the four GKAP family members known thus far, all are recognised except splice variants of GKAP1 which terminate in GQSK and does not interact with PDZ¹⁶².

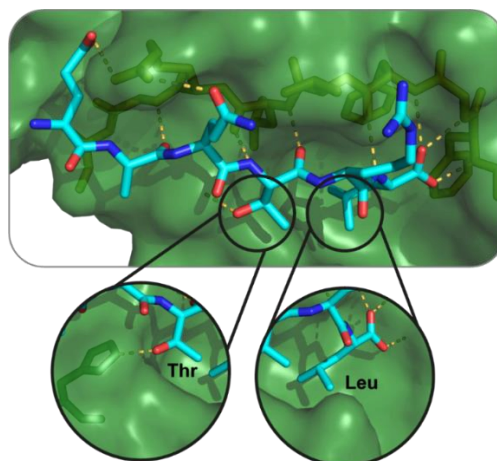


Figure 19. Crystal structure of the GKAP (cyan) and SHANK1 (green) PPI (PDB: 1Q3P): a) GKAP/SHANK1 with GKAP in an extended β -strand conformation showing the binding interactions from both side-chains and backbone; b) An enlarged image of the Thr residue forming a hydrogen bond on the opposing side of the PDZ binding interface when compared to the backbone hydrogen bonds. Leucine shown magnified, packed in a hydrophobic pocket.

2.1.2 Inhibitors of the PDZ domain of SHANK Proteins

Although targeting the PDZ domains of all types of proteins has been experiencing intense research, using both peptidomimetic and small-molecule inhibitors, very few inhibitors of the PDZ domain of the SHANK proteins are reported^{151,169,170}.

The peptide inhibitor **3** was designed based on another SHANK PDZ domain binder β PIX (p21-activated kinase interacting exchange factor) (Figure 20)^{171,172}. The β PIX actually displays 3 PDZ binding sites clustered by a parallel coiled-coil and yet the interaction takes place in a 1:1 stoichiometry¹⁷².

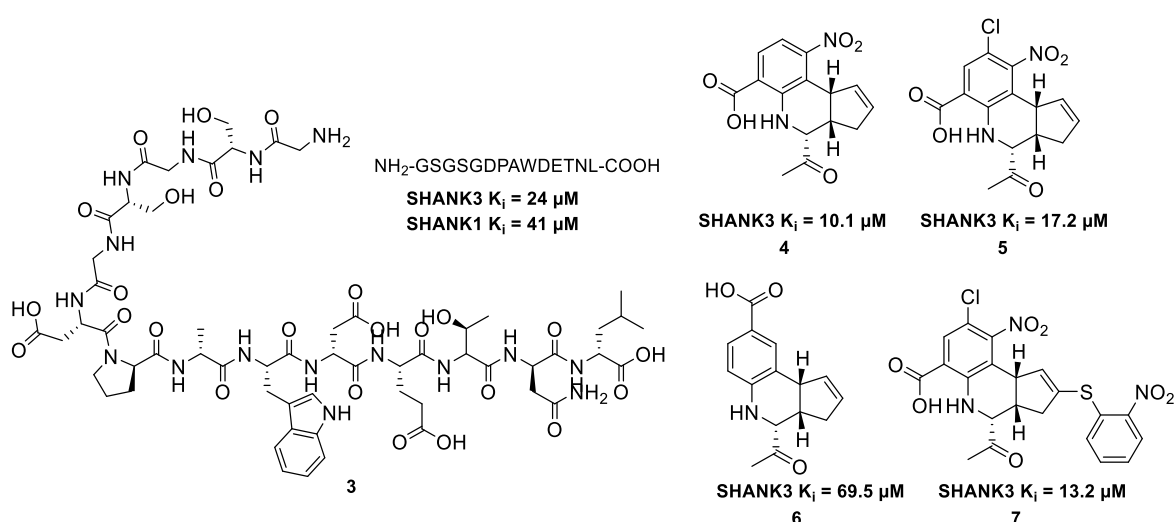


Figure 20. Peptide-based SHANK3 and SHANK1 inhibitor based on the PIX sequence **3** and SHANK3 inhibitors based on the tetrahydroquinoline carboxylate **4-7**.

A high-throughput screen identified tetrahydroquinoline carboxylate as an inhibitor of the PDZ domain of SHANK3 with an IC₅₀ below 250 μ M¹⁵⁵ (Figure 20). Subsequent SAR studies resulted in the tetrahydroquinoline carboxylate-based molecules **4-7**. Using ¹H,¹⁵N HSQC-NMR binding studies the compounds displayed modest activity with a range of 10-70 μ M affinity, displacing a hexapeptide derived from GKAP, a known binding partner of as well SHANK3^{151,173}.

2.2 Objectives

The main goal of this work was the rational design of a peptide-based β -strand mimetic. The β -sheet structural motif is a crucial recognition element in many PPIs however inhibitor design is relatively unexplored in comparison to α -helix mimetics³⁹. This is due to the extended strand conformation which allows for binding interactions to take place both from side chains and backbone, making inhibitor design difficult.

To rationally design an appropriate peptide-based inhibitor, three main milestones were identified:

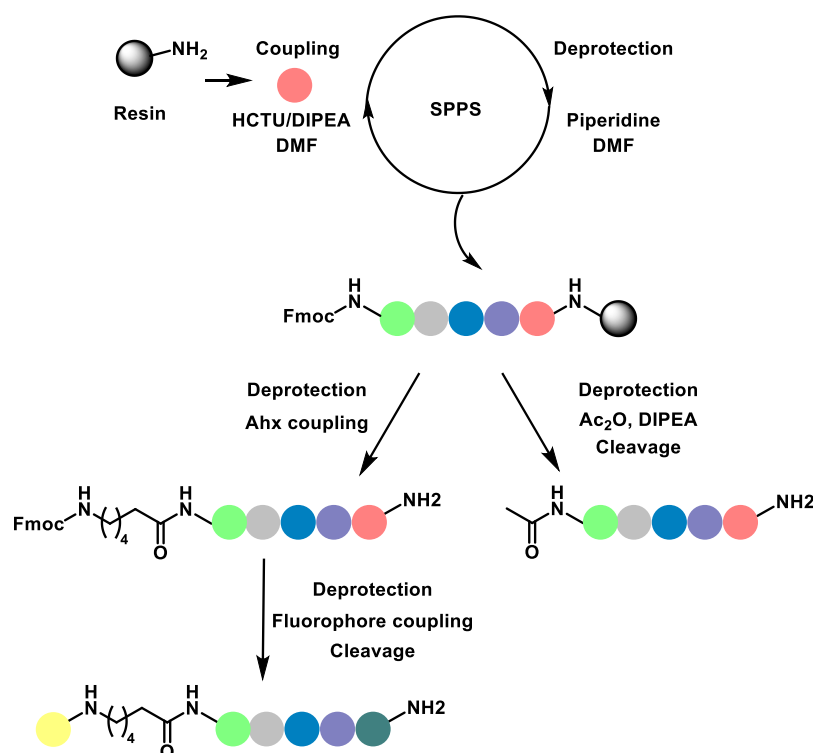
- Assay development Fluorescence Anisotropy (FA) for the GKAP/SHANK1 PPI followed by synthesis of alanine variants of the GKAP ligand for testing and identification of hotspot residues
- Comparison of experimental alanine scan with computational predictions
- Design and testing of select peptide sequences based on other known PDZ binders as well as variation of certain residues to bulkier side chain residues to probe the space around the binding interface

The GKAP ligand is small, 6 residues long, and the known binding motif for the PDZ domain X-S/T-X- ϕ -COOH, makes the GKAP/SHANK1 PPI a good model system for assay development. Experimental validation of the hotspot residues would help understand the forces driving the binding. Hot spot residues are deemed as important for binding and would thus need to be either kept or mimicked closely in the effective design of a GKAP based mimetic. Furthermore, the results of the experimental alanine scan would be compared with *in silico* predictions, done by Amaury A. Ibarra and Richard Sessions (University of Bristol) using their in-house designed BUDE alanine scanning tool, to help comment on the effectiveness of their computational tool in predicting hotspots as well as gauge the necessity of experimental alanine scanning. Next, the GKAP/SHANK1 PPI would undergo further probing through rational, point-variations of the GKAP sequence to other amino acid residues selected based on other PDZ domain binding ligands in order to further understand the PPI.

2.3 Assay Development and Experimental Alanine Scan

2.3.1 Peptide Synthesis

All peptides were synthesized on preloaded Wang resin as solid support to obtain C-terminal carboxylate peptides or Rink Amide MBHA resin in order to obtain C-terminal amidated peptides. All amino acids were *N*-Fmoc protected and side chains protected with Boc (Lys); *O**t*Bu (Asp, Ser, Thr); Trt (Asn, Gln); Pbf (Arg). Peptide elongation was done *via* solid phase peptide synthesis (SPPS) with repetitive coupling cycles of Fmoc deprotection in 20% piperidine in DMF, followed by amide coupling under standard conditions using HCTU as activating reagent and DIPEA as base (Scheme 1). Once automated peptide elongation was complete, the C-terminal was manually acetylated or reacted with 6-(Fmoc-amino)hexanoic acid linker (Ahx) followed by fluorophore, fluorescein isothiocyanate (FITC), coupling. Peptide cleavage from resin, and simultaneous side chain deprotection, was done using a standard cleavage cocktail of TFA/TIPS/H₂O (95:2.5:2.5). Due to the short peptide lengths and relatively simple sequences no issues were encountered and all peptides were obtained in good yield.



Scheme 1. Solid phase peptide synthesis cycles and manual couplings that followed.

2.3.2 Fluorescence Anisotropy

FA was selected as the biophysical tool to determine binding affinity between target protein and fluorescently labelled ligand peptide, in this case FITC labelled peptide. During FA polarised light is used to excite the FITC on the fluorescently labelled peptide. The fluorophore has the ability to emit light in the plane of polarised light however this depends on the rotational diffusion of the molecule; the fluorescently labelled peptide is tumbling quickly in solution resulting in unpolarised light being emitted thus low anisotropy detected. When the relatively small fluorescently labelled peptide binds the larger target protein, it's rotational diffusion slows and higher anisotropy can be detected (Figure 21 left hand side).

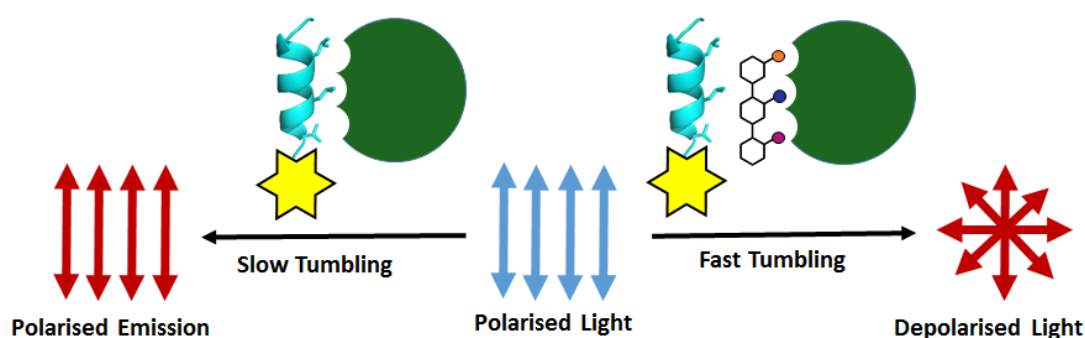


Figure 21. Cartoon depicting FA in a direct titration (left hand side) or under competition conditions (rights hand side).

Thus a fluorescently labelled ligand can be utilised to identify inhibitors of a PPI in competition mode to obtain an IC_{50} value, half maximal inhibitory potency, as opposed to a direct binding K_d . Designed inhibitors can be titrated in a solution containing the protein of interest and the FITC labelled peptide. If the inhibitor displaces the labelled peptide a decrease in anisotropy would be observed (Figure 21 right hand side). Although direct binding is not obtained, IC_{50} values can be used to compare inhibitors to each other and in relationship to the native peptide.

2.3.3 Assay Development for GKAP/SHANK1

Fluorescently-labelled GKAP (FITC-Ahx-GKAP) was prepared using microwave assisted Fmoc solid phase synthesis on Rink Amide MBHA resin for the peptide sequence, followed by manual coupling of Ahx linker and FITC to afford the tracer molecule for FA testing. Tracer preparation followed routine procedures and no challenges were encountered. GST tagged SHANK1 PDZ domain (656-762) production was done by Zsofia Hegedus and provided in a 50 mM Tris, 150 mM NaCl buffer at pH 7.5.

To establish optimal assay conditions, the starting protein concentration was varied (50 to 100 μ M) whilst the tracer concentration was kept constant (100 nM) in order to obtain a K_d value for binding of the tracer to SHANK1. A total of four buffer systems was trialled. The first buffer condition to be tested was the same as the buffer in which the SHANK1 protein is purified and stored in. This buffer gave a good result, with a curve which passed through the data points smoothly when fitted to a 1:1 model (Figure 22a). The following buffers had additions to the standard (50 mM Tris, 150 mM NaCl). First, Triton X (0.01%) was added to the standard conditions, a detergent that helps prevent non-specific binding interactions (Figure 22b). Bovine Serum Albumin (BSA) protein, prevents adhesion of other proteins to vessels such as pipette tips, and was the next addition to the standard conditions (Figure 22c). Finally, a buffer system containing all the components was screened (Figure 22d). Overall, all buffers gave good results (K_d values ranging from 1.1-1.9 μ M); however the standard conditions (50 mM Tris, 150 mM NaCl) were taken forward as this matched the buffer of the protein and could be kept consistent in further analysis. Overall the tracer molecule was observed to have a K_d of 1.5 ± 0.4 μ M in the direct titration assay (Figure 23a).

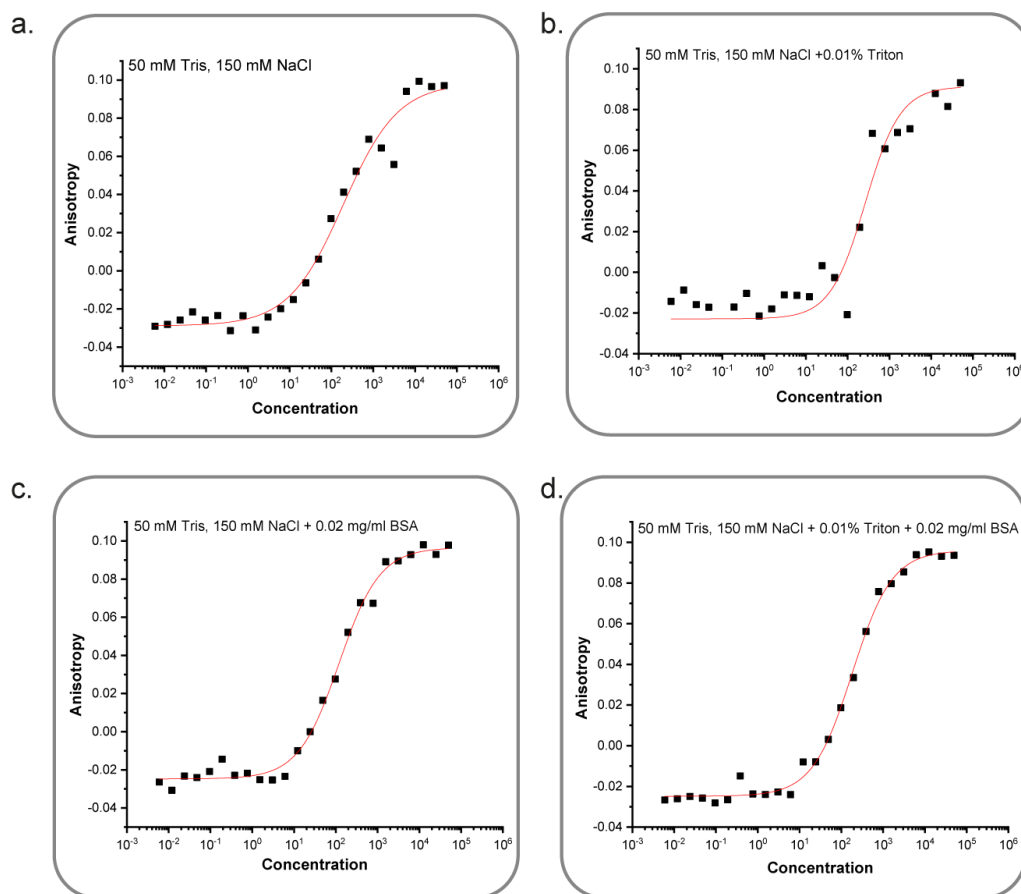


Figure 22. Buffer screen for the assay development for the GKAP/SHANK1 PPI all between pH 7.4-7.5 using 100 nM tracer read after a 45 minute incubation period: a) Buffer conditions of 50 mM Tris, 150 mM NaCl; b) Triton 0.01% was added to the initial 50 mM Tris, 150 mM NaCl buffer; c) Buffer system contained 50mM Tris, 150 mM NaCl and 0.02 mg/mL BSA; d) Buffer contained all 4 components: 50 mM Tris, 150 mM NaCl, 0.01% Triton, 0.02 m/mL BSA.

2.3.4 Experimental Alanine Scan of GKAP Ligand

Alanine variants were prepared using microwave assisted Fmoc solid phase synthesis on Rink Amide MBHA resin as described above. These were all tested in competition mode (Figure 23b) and the resulting IC_{50} s are shown in Table 1. The experimental results for the GKAP₁₋₆R5A variant gave an $IC_{50} = 11 \pm 0.2 \mu M$ in the competition assay, similar to that of the wt-GKAP₁₋₆, $IC_{50} = 10 \pm 1 \mu M$. Some loss in binding was observed for GKAP₁₋₆E1A, however more importantly, GKAP₁₋₆T4A, GKAP₁₋₆L6A and the carboxylate to amide variant were unable to inhibit the tracer/protein interaction. Thus, this data supported T4, L6 and carboxylate to amide as critical requirements for binding, hotspots, following the X-(Ser/Thr)-X- ϕ -COOH consensus from literature.

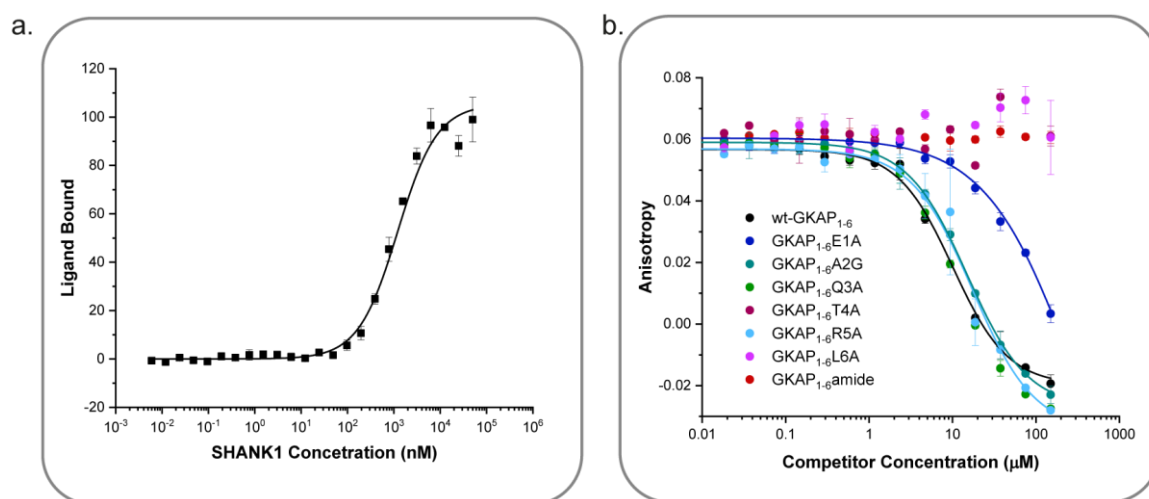


Figure 23. FA data for the GKAP/SHANK1 PPI; a) Direct titration assay (50 mM Tris, 150 mM NaCl, using 100 nM FITC-Ahx-GKA); b) Competition assays for the GKAP alanine variants and the C-terminal amide variant (50 mM Tris, 150 mM NaCl, 50 nM tracer, 1 μ M SHANK1).

Table 1. IC₅₀ values for alanine and C-terminal amide variants.

GKAP	Wt	E1A	A2G	Q3A	T4A	R5A	L6A	amide
IC ₅₀ (μ M)	10 \pm 1	< 300	15 \pm 1	1.7 \pm 1	n.d.	11 \pm 0	n.d.	n.d.

The experimental alanine scan was successfully completed and the results aligned with literature³⁵. The GKAP/SHANK1 PPI proved to be an ideal model system, well behaved in the assay with the GKAP ligand amenable to the systematic alanine variation. The PPI was taken through a computational alanine scan by Amaury Ibarra (Bristol University) for hotspot predictions in order to compare the competence of their tool against experimental data.

2.4 Computational Alanine Scan of GKAP Ligand

In silico alanine predictions of the GKAP sequence were undertaken by Amaury Ibarra (University of Bristol), using the Bristol Universal Docking Engine (BUDE) package^{19,174}. The BUDE computational tool was previously developed for rapid small-molecule protein docking and was modified by the researchers at the University of Bristol for computational alanine scanning¹⁷⁵. The tool is an empirical free energy approach that uses standard force fields; where energy

parameters are used to calculate the energy potential of a system of atoms¹⁷⁶. Subsequently the web interface BudeAlaScan was developed to make the capability user-friendly¹⁷⁴.

The GKAP hexapeptide was subjected to BUDE analysis. Since the GKAP sequence contains an alanine residue, A2, this was computationally modified to a glycine. A threshold of $\Delta\Delta G$ 4.5 kJ/mol was selected and any residue varied to Ala that gave a value at or above this threshold was considered a hot spot.

BUDE computational alanine scan predicted R5 marginally ($\Delta\Delta G < 4.2$ kJ/mol), and L6 ($\Delta\Delta G < 20$ kJ/mol) and carboxylate ($\Delta\Delta G < 19$ kJ/mol) as important for binding, however, failed to pick the T4 residue as a hotspot (Figure 24a). The Bristol team ran MD simulations, a technique that could offer insight into the dynamic behaviour of the peptide. This result was then combined with the BUDE predictions to correct the T4 prediction with little success (Figure 24b). Overall, the experimental FA results mirrored the BUDE computational prediction for the L6 and carboxylate variants and for the residues not predicted as hotspots, bar R5, as well.

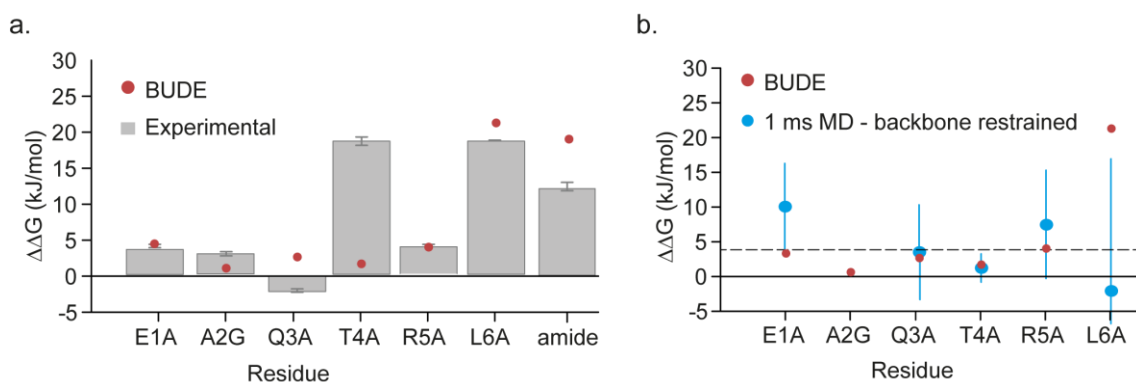


Figure 24. Alanine scanning data for GKAP/SHANK1 (PDB: 1Q3P); a) The computational BUDE AlaScan results denoted by red dots compared with the experimental data represented as grey bars; b) BUDE AlaScan $\Delta\Delta G$ data shown by the red dots and BUDE predictions incorporating a 1 ms MD data for the mutants depicted by the blue dots with a dotted line depicting the hot-spot cut-off.

BUDE failed to predict T4 as a hot spot residue and showed a slight tendency to overestimate charged residues such as R5 as important for binding. As such, the Bristol team endeavoured to use 5 CAS methods on 3 diverse PPIs, among which GKAP/SHANK1 was investigated, as a comparative analysis between the tools

but also compared with experimental analyses in order to discern best practices when predicting hotspots. The Bristol group found that averaging the predictions of multiple *in silico* methods led to better prediction rather than using a single computational tool¹⁹. Overall experimental alanine scanning is still necessary and computational methods cannot be solely relied on for identifying hotspot residues.

2.5 Design, Synthesis and Evaluation of Selected Peptide Variants

Following the alanine scan, variations in the GKAP sequence to other amino acid residues were undertaken. Although the alanine scan helped identify GKAP hotspots, the extent of involvement of all other residues was not clear, ie. are any of the non-hotspot residues involved in specificity or could they be varied to further probe the binding interface of SHANK1. Selected peptide variants were designed based on other known PDZ domain binding ligands or simply to probe the interaction for any steric effects when variations to bulkier side-chains were selected. All peptide variants were prepared using microwave assisted Fmoc solid phase synthesis on Rink Amide MBHA resin.

The A2, T4, R5 and L6 positions were probed further (Table 2) in order to understand how residues other than alanine could affect the binding, would these be tolerated and would any improvement in binding be observed. The A2 and R5 position were picked in order to further verify that these residues do not contribute in any way to the binding interaction towards SHANK1. Based on a different known SHANK1 PDZ binder β -PIX, mentioned earlier with the sequence WDETNL-COOH¹⁷⁷, GKAP₁₋₆A2D was synthesised to check if this variation could potentially enhance binding. In addition GKAP₁₋₆A2F was also synthesised to probe scan and assess for any steric clash that could ensue due to such a bulky side chain. The recognition consensus suggests Ser at position 4 ought to be tolerated in the GKAP sequence. Thus, the GKAP₁₋₆T4S variant was synthesised to test this hypothesis. Next, the R5 position was also probed with GKAP₁₋₆R5W; W located at position 5 in the sequence of a different PDZ domain binding ligand was shown to expel a water molecule from a cavity on SHANK1 and potentially aid in binding affinity¹⁷⁸. Finally, GKAP₁₋₆L6F was designed since

it is a preferred residue for internal peptide motifs recognised by PDZ domains^{22,139}.

Table 2. Further investigation of the binding interaction with different GKAP variants.

Sequence	Rationale
EAQTRL-COOH	native sequence
EDQTRL-COOH	inspired by β -PIX sequence which also binds PDZ
EFQTRL-COOH	would the bulkier side chain cause steric clashes?
EAQSRL-COOH	test if S at the 4 position tolerated as mentioned in literature ¹⁷⁹
EAQTWL-COOH	steric probing + would W expel a water molecule ¹⁷⁹
EAQTRF-COOH	Inspired by internal peptide motif binders of the PDZ domain

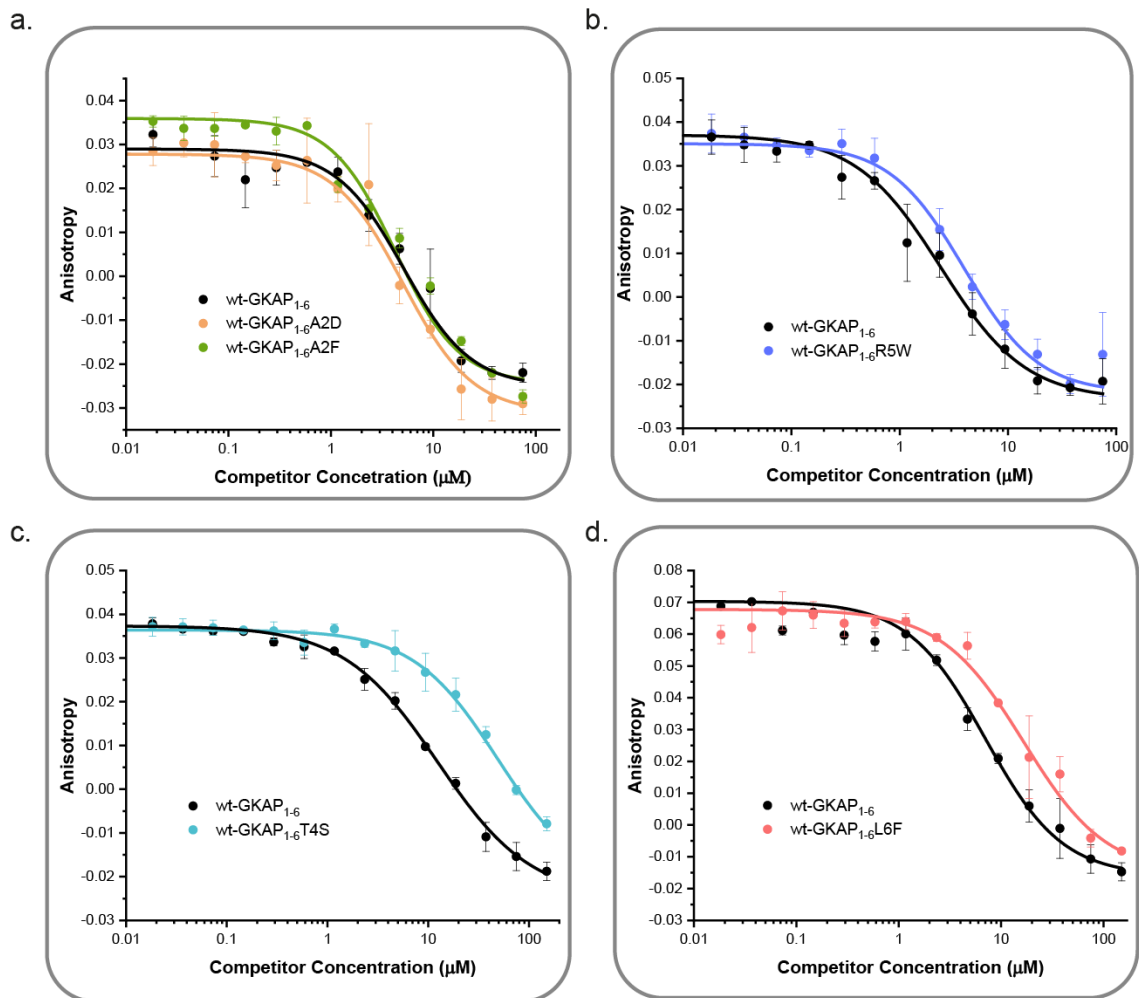


Figure 25. Competition assay data of specific residues of the GKAP sequence all done in 50 mM Tris, 150 mM NaCl using 50 nM tracer and 1 μ M SHANK1; a) Competition assay curves for GKAP₁₋₆A2D and GKAP₁₋₆A2F compared to wt-GKAP₁₋₆; b) Competition

assay for GKAP₁₋₆R2W compared to GKAP₁₋₆; c) Competition assay for GKAP₁₋₆T4S compared to GKAP₁₋₆; d) Competition assay for GKAP₁₋₆L2F compared to wt-GKAP₁₋₆.

Table 3. IC₅₀ values for peptide variants designed to probe the binding interaction of GKAP to SHANK1.

GKAP	Wt	A2D	A2F	R5W	T4S	L6F
IC ₅₀ (μM)	7.1 ± 2.3	5.0 ± 0.6	5.1 ± 0.7	4.0 ± 0.3	40 ± 7	16 ± 8

Based on the competition assays the variations at the A2 and R5 positions seemed to suggest that these two residues have little impact on the binding affinity to the SHANK1 PDZ domain, correlating well with the computational and experimental alanine scanning were these were not found to be hotspots (Figure 25a, b). All of the aforementioned variations showed similar IC₅₀ values to the wt-GKAP₁₋₆ peptide (Table 3). However GKAP₁₋₆T4S showed slight loss in ability to compete the wt-GKAP tracer, IC₅₀ of 40 ± 7 μM, lower than that of the wt-GKAP₁₋₆. This would suggest that the T4 position holds a potential role for specificity in binding through the conformational restriction of the branched side chain. Lastly, L6 was confirmed as a hot-spot residue during the alanine scan. The ability to recover binding to SHANK1 using the considerably larger F residue was successful (Figure 25d). The GKAP₁₋₆L6F variant showed similar binding to SHANK1 as the wild-type peptide. To gain confidence in this result the variant was further investigated using ITC. Furthermore, a co-crystal with the PDZ domain of SHANK1 was obtained by Zsofia Hegedus (Figure 26). As can be observed in the crystal structures the binding of GKAP₁₋₆L6F resembled the binding of the native sequence, with the F residue very well accommodated in the binding pocket. An overlay of the two peptide ligands, wt-GKAP₁₋₆ with GKAP₁₋₆L6F shows the variant peptide is almost perfectly aligned.

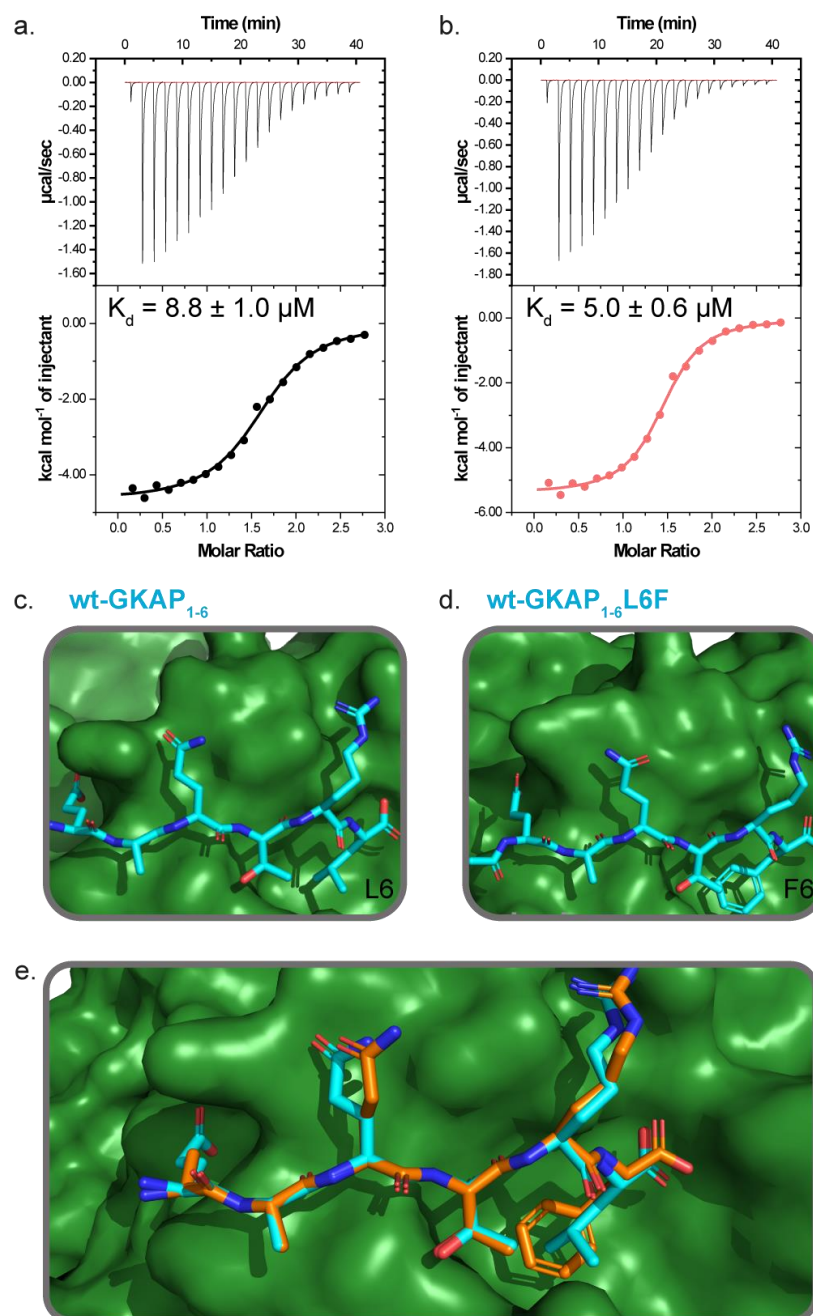


Figure 26. ITC and crystal data for GKAP1-6L6F variant compared to wt-GKAP₁₋₆; a) ITC data for wt-GKAP1-6; in 50 mM Tris, 150 mM NaCl, pH 7.5, at 25 °C; b) ITC data for GKAP₁₋₆L6F; in 50 mM Tris, 150 mM NaCl, pH 7.5, at 25 °C; c) GKAP/SHANK1 (PDB: 1Q3P) crystal structure displaying the binding of the GKAP sequence; d) for GKAP1-6L6F/SHANK1 crystal structure; e) Overlay of wt-GKAP₁₋₆ (cyan) with GKAP₁₋₆L6F (orange).

Overall the selected peptide variants helped satisfy us that the A2 and R5 residues have no influence on binding affinity to SHANK1, a conclusion which matched well with the alanine study. Additionally, variation of these two residues to bulkier amino acids did not elicit any steric clash or subsequent detriment to

binding. Furthermore, the T4 variation to Ser seemed to implicate this position as a potential driver for specificity. In future, the study of a known ligand of the PDZ domain containing a Ser at the 4 position could help elucidate this theory. Lastly, the analysis of the selected peptide variants confirmed Phe can easily replace the L6 residue, as is observed in the extended binding motifs of PDZ domains. In combination with the alanine study it is evident that the A2 or R5 positions would be the ideal locations for any changes in the design of a peptide-based inhibitor of this PPI.

2.6 Design, Synthesis and Testing of a Peptide-Based Inhibitor

Based on the knowledge acquired from the alanine study and the selected peptide variants, development of a peptidomimetic structure was pursued. The short wt-GKAP₁₋₆ peptide ligand, only 6 residues long, had 3 elements important for binding to SHANK1; the threonine at the 4 position (T4), the terminal leucine (L6) and the carboxylate. The decision was made to build the peptidomimetic around the full length of the GKAP ligand in order to maintain the 3 important elements of recognition intact with a residue change at the R5 position. Inspiration was gained from the work conducted by the Nowick group. It was envisioned that the motif in red (Hao), previously shown by Nowick and co-workers⁶⁰ to stabilise β -sheet fold without increasing aggregation, would form hydrogen bonds with the peptide backbone and the entropic cost upon binding would be lowered by pre-organizing the GKAP peptide into an extended β -strand conformation in solution (Figure 27). Specifically, the Hao building block would hydrogen bond with the peptide backbone of the opposing SHANK1 recognition surface, pre-organising the peptidomimetic for recognition.

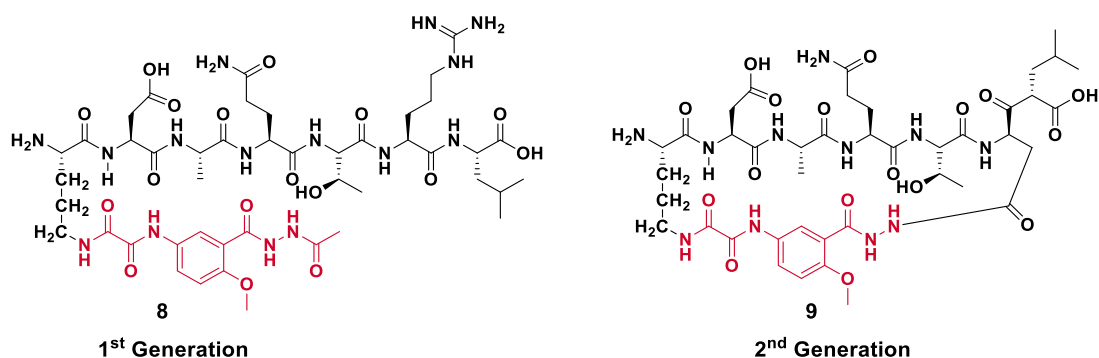
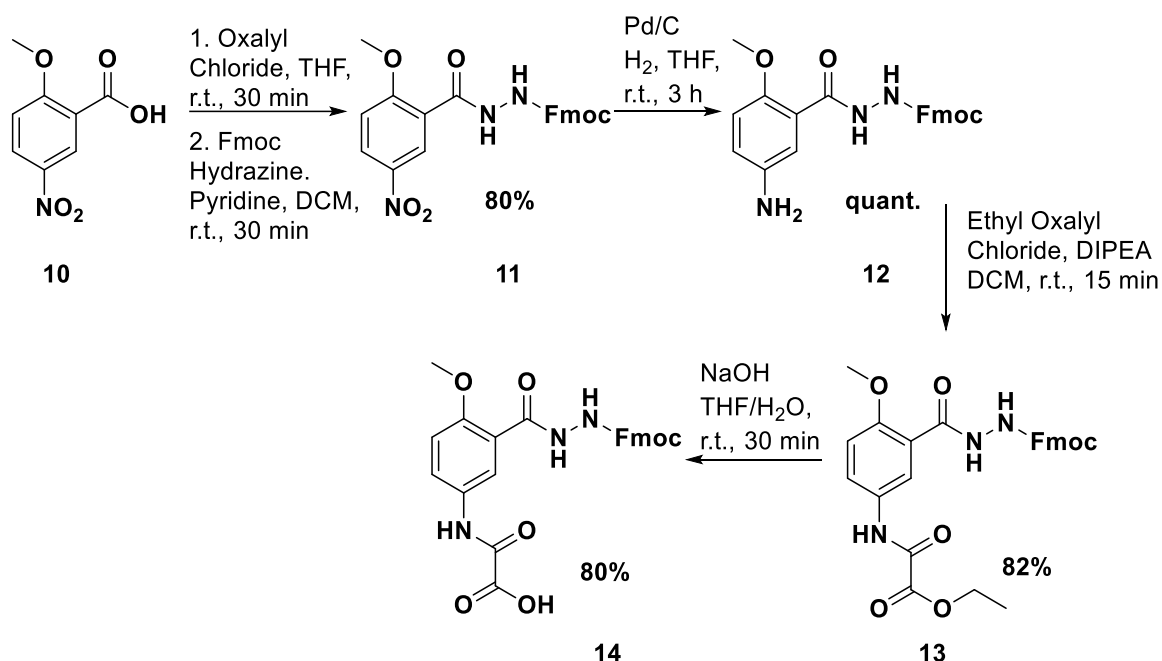


Figure 27. Peptidomimetic design based on the GKAP peptide sequence and incorporating the Hao molecular building block.

Two peptide-based mimetics were designed. The effect of the Hao molecule on peptide structure was to be analysed in both an acyclic and a cyclic format. Ideally the incorporation of the Hao molecule in the acyclic version would provide an opportunity for the peptide to form intermolecular hydrogen bonds and thus pre-organise the peptide in a strand conformation. However, in an aqueous environment the strength of these hydrogen bonds formed between peptide and Hao could be difficult to maintain. As such, a macrocyclic 2nd generation mimetic was designed. The cyclisation of the peptide could potentially force the necessary interactions to occur in order to pre-organise the GKAP peptide.

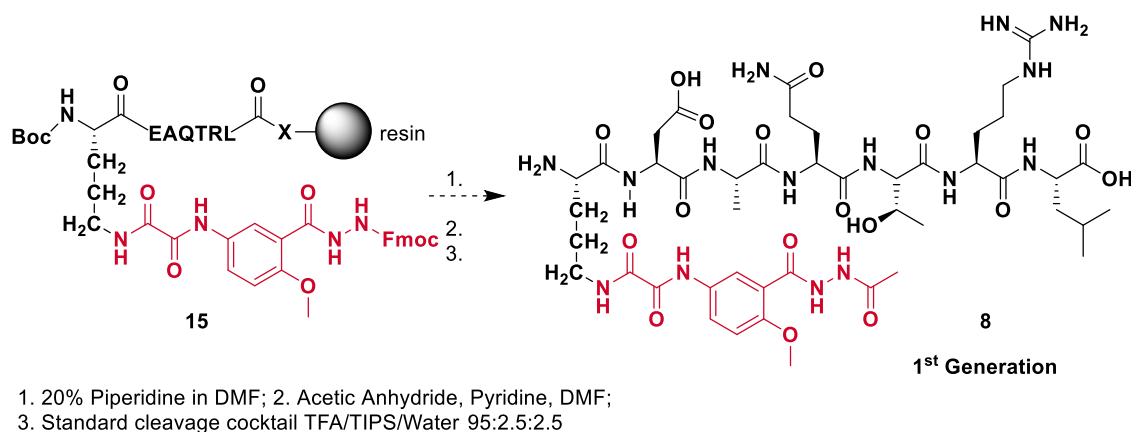
2.6.1.1 Synthesis of First Generation Peptidomimetic

The 1st generation peptidomimetic required an acetylated N-terminus. The acid **10** (2g scale) was treated with oxalyl chloride in THF, and the resulting acid chloride was concentrated *in vacuo*, dissolved in DCM and subsequently reacted with Fmoc-hydrazine in the presence of pyridine to afford intermediate **11**. The Fmoc protected intermediate **11** was subjected to hydrogenation with palladium on carbon to afford **12** in quantitative yield. The resulting amine **12** was reacted with ethyl oxalyl chloride in DCM with DIPEA for 10 minute, after which it was concentrated *in vacuo* and taken further without purification. Saponification of **13** was done in 3:1 ratio of THF/H₂O and a 2M solution of NaOH was added slowly until a pale yellow colour was observed. The crude mixture was then eluted through and acidic ion exchange resin to yield **14** in 80% yield following a previously described synthetic method¹⁸⁰ (Scheme 2).

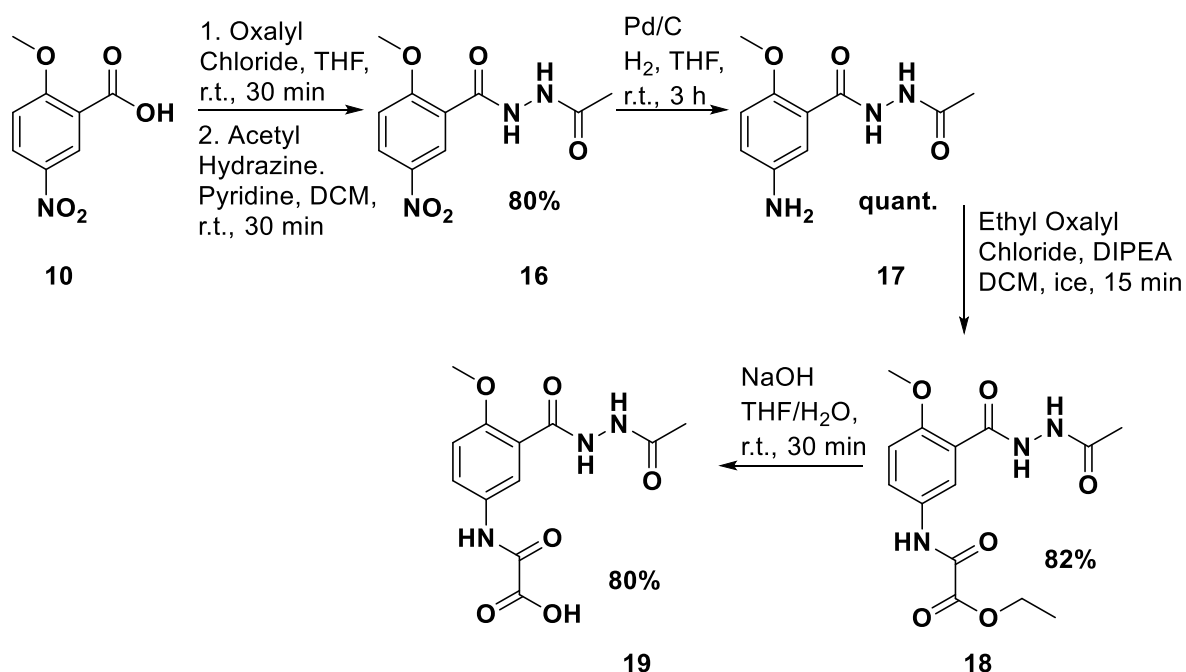


Scheme 2. Synthetic route towards obtaining Fmoc protected Hao.

The GKAP peptide sequence was prepared using the SPPS elongation protocol described previously, and building block **14** was added on resin to the peptide to afford **15** (Scheme 3). Although Hao was coupled successfully to the GKAP peptide chain, the manual Fmoc deprotection and subsequent acetylation at the hydrazine of **15** would not work, as indicated by LCMS analysis (Scheme 3). Several small trials of the Fmoc deprotection and acetylation with varying times seemed to make no difference leading to the conclusion that a new approach was necessary to obtain the 1st generation peptidomimetic. As such, the above synthetic route was modified to include acetyl-Hao to form an acetylated Hao building block (Scheme 4).

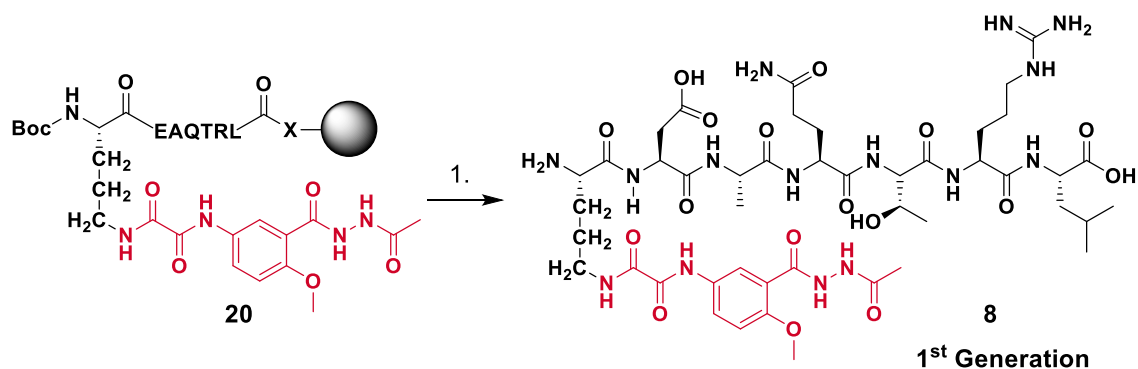


Scheme 3. Failed Fmoc deprotection and acetylation of Hao building block after its manual coupling to the GKAP peptide.



Scheme 4. Synthetic route to obtain the acetylated Hao building block.

The synthetic route followed the original steps and similar results were obtained when starting with 2 g of starting material **10**. Molecule **19** was obtained in high yield and was manually coupled to the GKAP peptide which was still attached to the resin following SPPS elongation to obtain **20** (Scheme 5). In future the Hao building block could be added to the peptide synthesiser and added to the GKPA peptide through SPPS elongation methods. The successful coupling was verified by LCMS and following cleavage and HPLC purification, the 1st Generation peptidomimetic was acquired (Scheme 5).



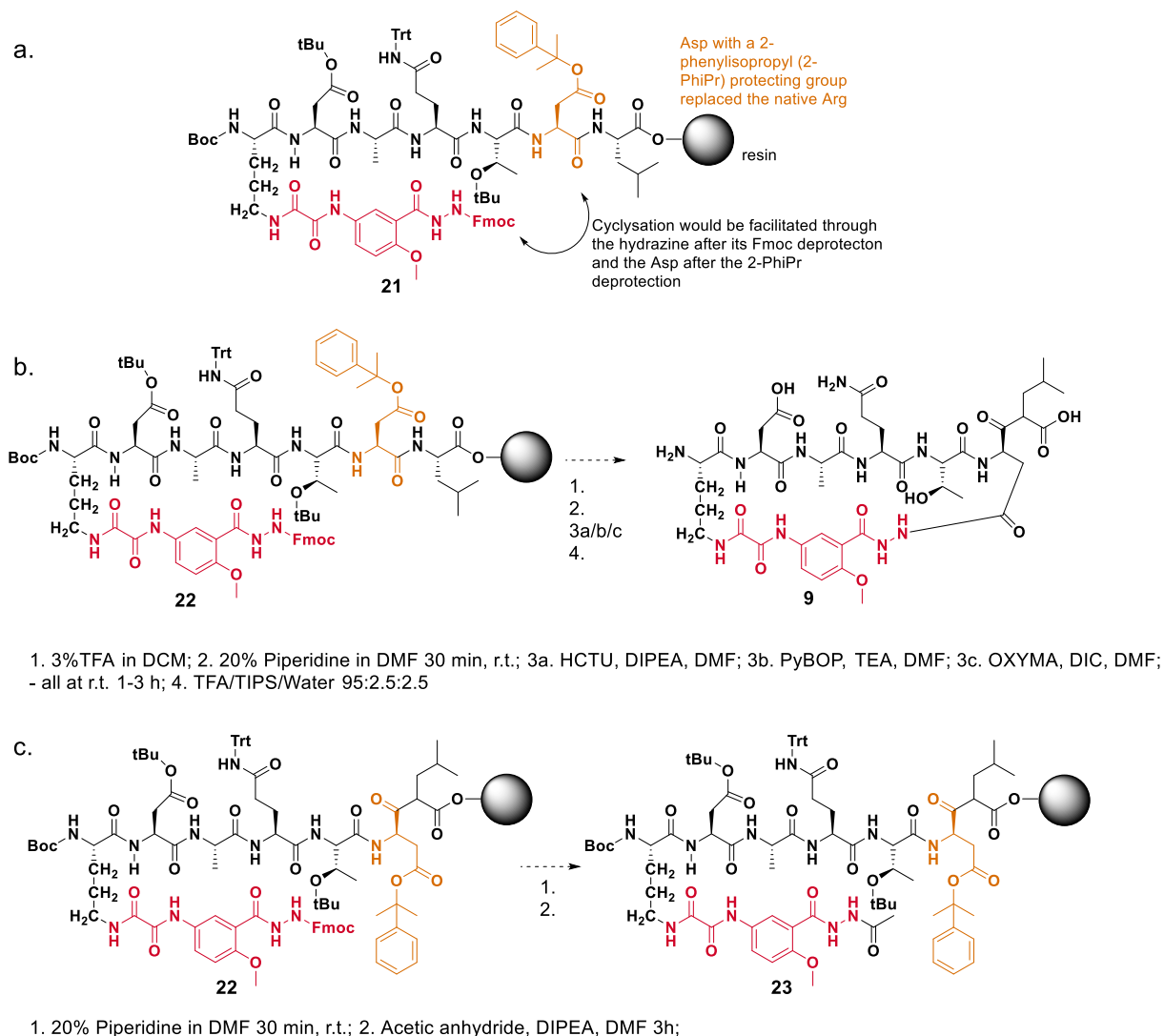
Scheme 5. Final step, standard peptide cleavage off resin, to obtain the 1st Generation Peptidomimetic.

2.6.1.2 Synthesis of Second Generation Peptidomimetic

In order to obtain a macrocyclic version of the peptidomimetic, it was necessary to use the Fmoc protected Hao building block, **14**, again (Scheme 2). The target molecule had been previously synthesised and in hand. The GKAP peptide sequence was synthesised using SPPS however with the arginine position R5 substituted for aspartic acid with a 2-phenylisopropyl (2-PhⁱPr) protecting group, compound **21** (Scheme 6a). This R5 variation was possible since it was evident from the alanine and selected peptide variants that this position was not involved in binding interactions with SHANK1. The use of Asp with the 2-PhⁱPr protecting group at the R5 position in the sequence, ensured selective deprotection for subsequent cyclisation with the hydrazide of the Hao building block, specifically through amide coupling of the Asp residue to the hydrazine. The cyclisation reaction would take place on resin before cleavage.

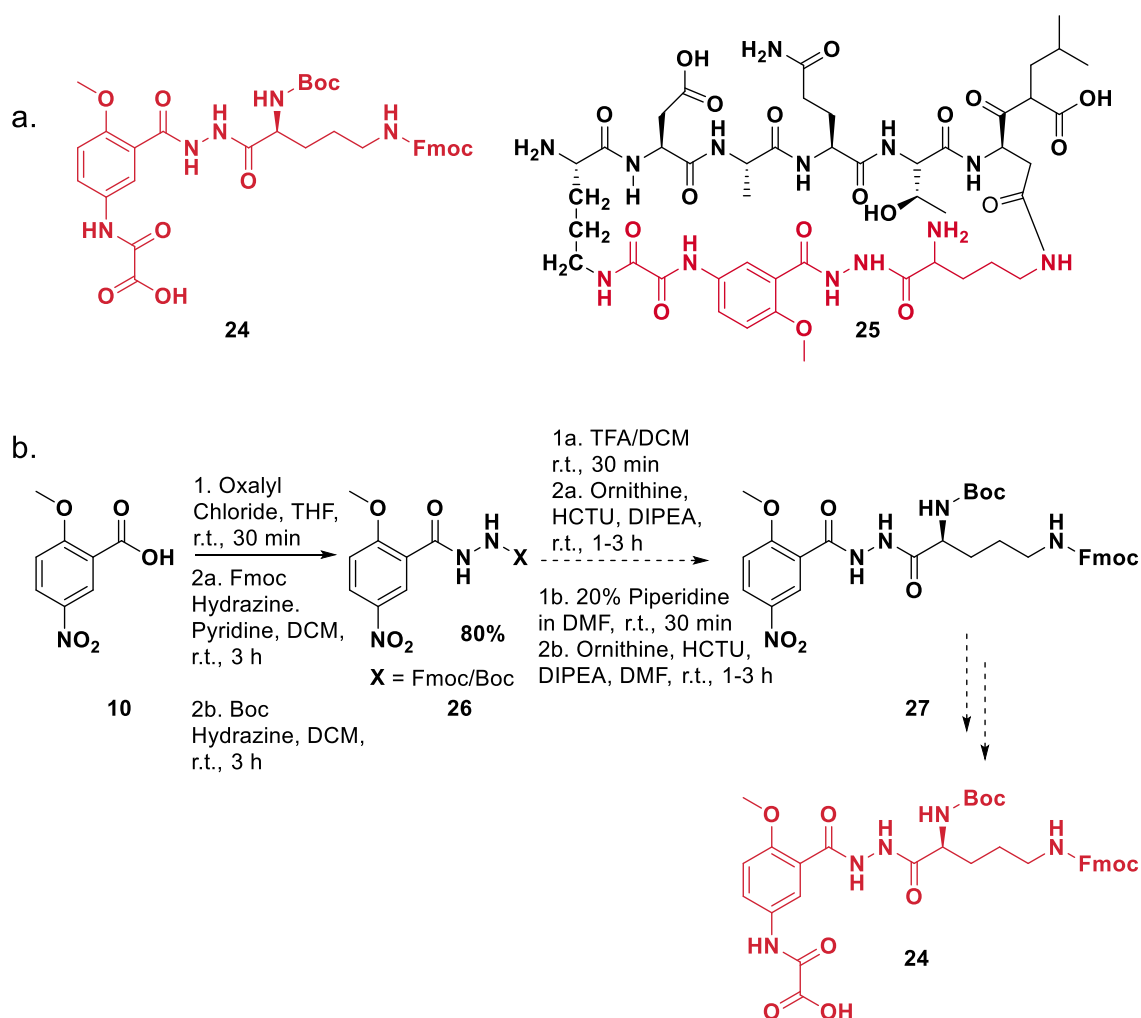
The subsequent coupling of **14** on resin to the GKAP peptide sequence was successful, as indicated by LCMS. The next step was to deprotect at the Asp at the 4 position where the 2-PhⁱPr would be removed using a 3%TFA in DCM solution. This was followed by Fmoc deprotection of the N-terminus. However, the deprotection and cyclization reactions did not afford the expected macrocycle **9** as indicated by LCMS following peptide cleavage off resin (Scheme 6b). To address the synthetic issues, compound **22** was subjected to Fmoc deprotection treated with 20% piperidine in DMF and subsequent acetylation (Scheme 6c). This standard acetyl capping method was also unsuccessful, which hinted

towards the possibility that cyclisation through the hydrazine was the problem. The result was not completely unexpected since the synthesis of the 1st Generation peptidomimetic gave similar results when the Hao Fmoc deprotection and subsequent acetylation failed.



Scheme 6. Synthetic efforts to deprotect and cyclize 15 to obtain the 2nd Generation peptidomimetic 2: a) Linear peptide design with the Asp replacing Arg at the 5 position in the GKAP peptide sequence; b) Initial deprotection and cyclisation of 15 which was unsuccessful; c) Attempted Fmoc deprotection and acetylation to test if this was the location of the synthetic issues encountered.

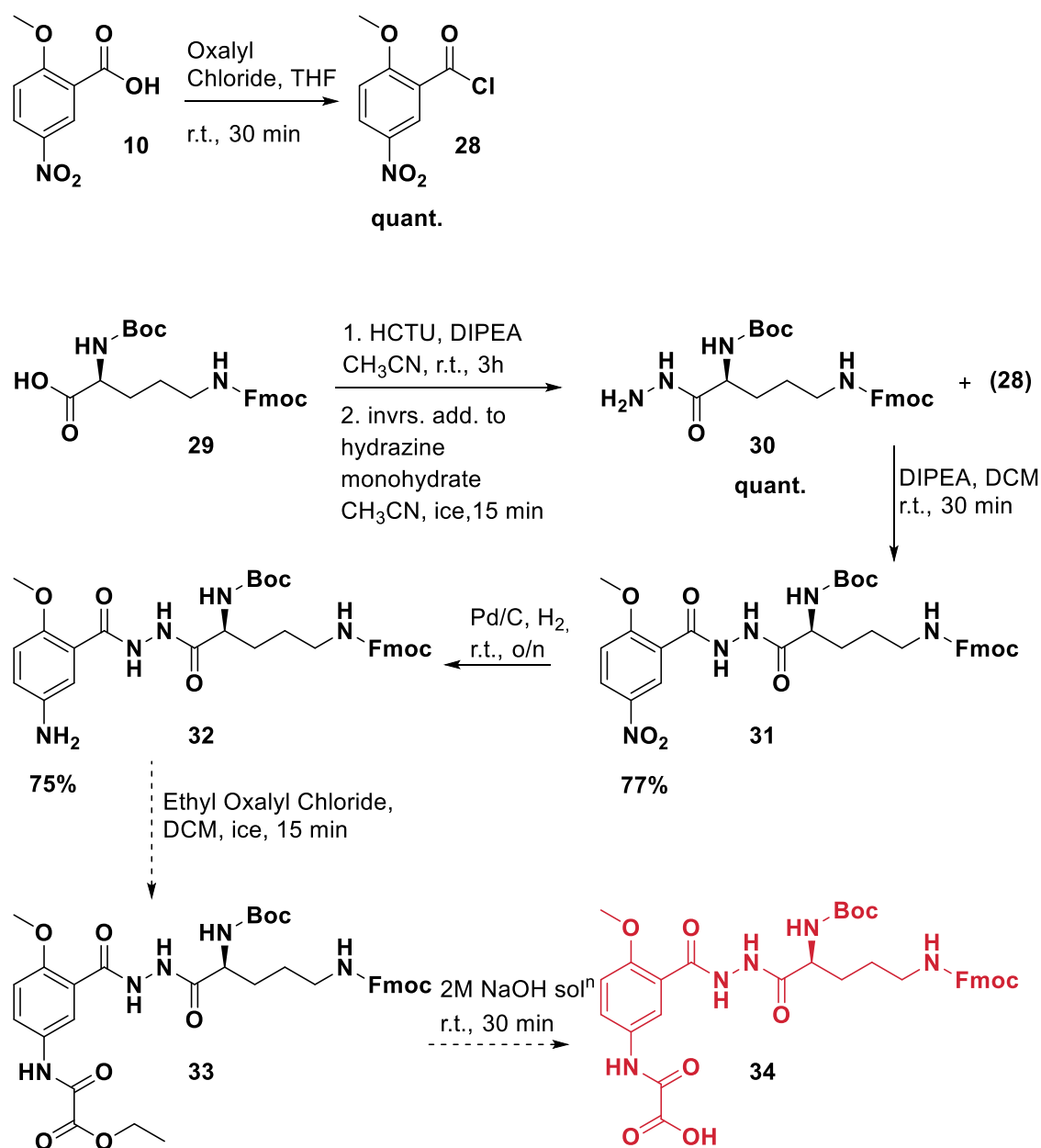
Due to the difficulty in obtaining a macrocyclic peptidomimetic, the Hao building block **14** was redesigned to contain an ornithine linker as highlighted in red in compound **24**. This way the cyclization reaction would be done through the amine of the linker and would hopefully avoid the issues faced earlier in the synthesis (Scheme 7a). A new synthetic route was developed based on newly designed **27** (Scheme 7b).



Scheme 7. New design for the 2nd Generation peptidomimetic: a) the GKAP peptide with the insertion of a Hao building block with a ornithine linker attached; b) Synthetic route for ornithine containing Hao building block.

Similar to previous routes, starting material **10** was converted into an acid chloride and this was subsequently coupled to either Fmoc hydrazine (reagents 2a Scheme 7) or Boc hydrazine (reagents 2b Scheme 7) to give two different protected versions of compound **26**. The subsequent Boc and Fmoc deprotections and immediate coupling with ornithine gave none of the desired product **27** and a mixture of starting material and impurities due to potential decomposition was recovered. This was consistent with the previous failed results to couple at this position when acetylation at the hydrazine did not afford the 1st generation peptide-based inhibitor and synthesis of the building block was redesigned. As such, the synthesis was once again re-designed to contain the

ornithine liker within the Hao molecule and attempt cyclisation through the linker as opposed to the Hao hydrazine.



Scheme 8. Synthetic route for building block 26.

The ornithine linker, **29**, was coupled to hydrazine monohydrate to give compound **30** (Scheme 8). Interestingly, the reaction would not proceed unless the carboxylic acid was first activated with HCTU, DIPEA in acetonitrile and then added to a pot of hydrazine monohydrate stirring on ice in acetonitrile. Compound **28** was then reacted with **30** to obtain compound **31** which was further subjected

to a reduction with palladium on carbon to afford amine **32**. The subsequent coupling of **32** with ethyl oxalyl chloride failed. This reaction had previously worked in both Scheme 2 and 4, albeit with a slightly different material to **32**, within 10 minutes of stirring. The order of adding materials was trialled both with the addition of **32** into a stirring solution of ethyl oxalyl chloride on ice and reversed. This seemed to have no effect on the reaction going forward with compound **32** recovered. Unfortunately, at this point in the work the decision to abandon the macrocyclic design was made. As such, testing of the acyclic peptidomimetic **8**, and comparison to wt-GKAP₁₋₆ was undertaken before anymore synthetic efforts were to be made.

2.6.1.3 Testing and Comparison of the First Generation Peptidomimetic to wt-GKAP₁₋₆

Molecule **8** was tested in a competition FA assay and ITC, and compared to the native peptide, wt-GKAP₁₋₆. No significant difference was observed for the peptidomimetic structure **8**, IC₅₀ of 11 ± 2 μM compared to the 10 ± 1 μM IC₅₀ for wt-GKAP₁₋₆ in the FA competition assay. Similarly, ITC analysis gave a K_d = 3.5 ± 1.2 μM for **8**, comparable with that of the wt-GKAP₁₋₆ K_d = 8.8 ± 1.0 μM. The FA and ITC data showed no improved binding. It was hypothesised that the aqueous media in which the assay was performed in was potentially interrupting the key hydrogen bonding between the Hao building block and the peptide backbone.

NMR analysis was subsequently undertaken to investigate if compound **8** showed signs of pre-organisation through expected NOESY interactions (Figure 29). It was believed that NMR analysis could help elucidate the molecular behaviour of the peptide-based inhibitor in water, and as such the experiment was run in 10% D₂O in H₂O.

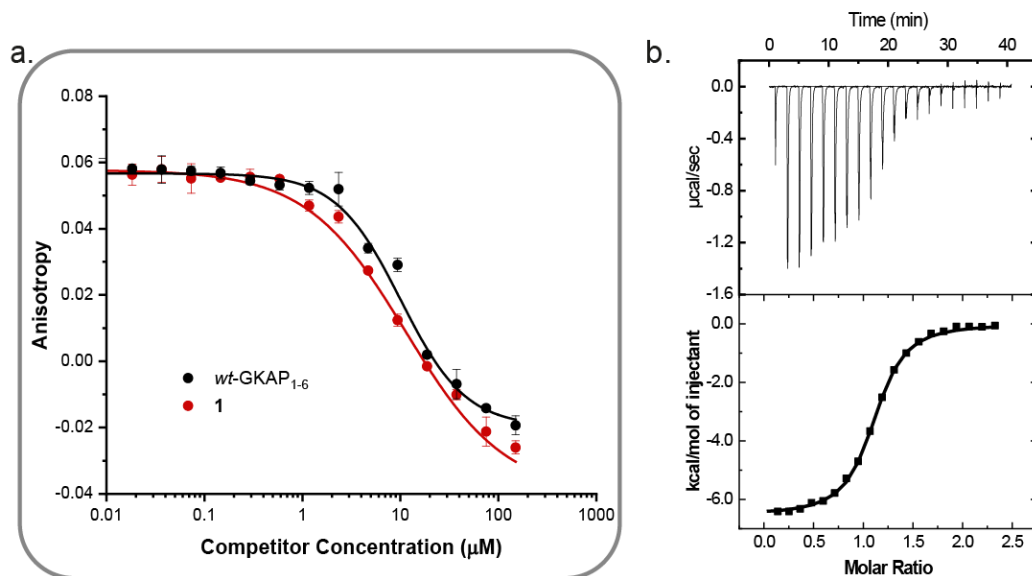


Figure 28. Analysis of compound 1 against *wt*-GKAP1-6; a) Competition assay in 50 mM Tris, 150 mM NaCl using 50 nM tracer and 1 μM SHANK1; b) ITC data in 50 mM Tris, 150 mM NaCl, pH 7.5, at 25 $^{\circ}\text{C}$.

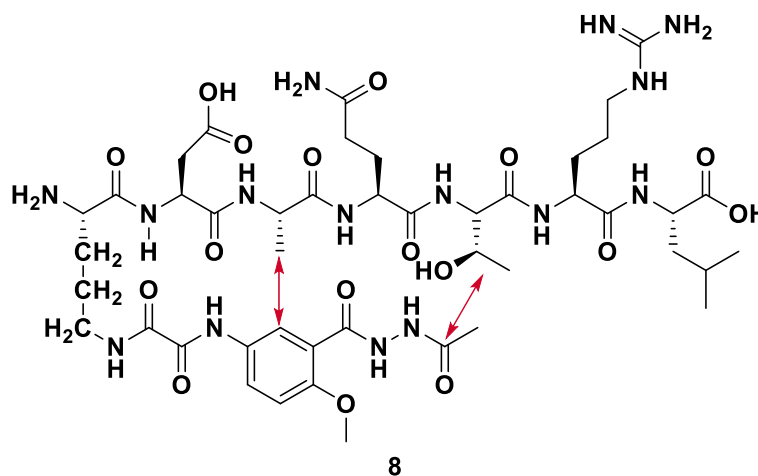


Figure 29. Structure of **8** with red arrows displaying the potential key interactions to look for in NMR as signs of pre-organisation.

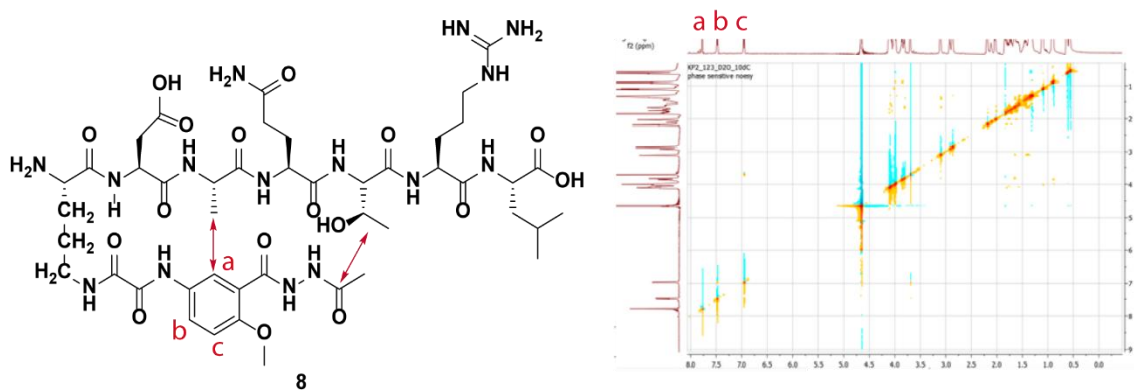


Figure 30. NOESY NMR analysis of **8**.

The NMR experiments were run in 10% D₂O in water. Following the assignment of the aromatic peaks of the Hao building block, it was easy to see there were no correlations in this region in the phase sensitive NOESY. No correlation between the alanine CH₃ and the aromatic **a** peak suggested no pre-organisation in an extended strand conformation. Most likely the hydrogen bonding between the GKAP peptide strand and the Hao building block was not strong enough to hold the peptidomimetic in the desired conformation. This was most likely due to the aqueous environment hindering the ability of the molecule to form hydrogen bonding interactions. Analysis of **8** in other, organic, solvents was sought, simply to verify that the molecule was indeed capable of forming the desired hydrogen bonds with the Hao building block. Unfortunately, **8** was not soluble in chloroform or acetonitrile. Dimethyl sulfoxide was also trialled however a gel-like substance formed. Naturally, in this design case, to force the hydrogen bonds to take place despite the aqueous media would require the cyclised version of **8**. However, synthesis of a macrocyclic structure was unsuccessful even after multiple synthetic routes for Hao were attempted were attempted.

2.7 Conclusions

The residues of the extended β -strand project on alternating sides of the backbone, with maximum distances from each-other minimizing steric clashes exposing the backbone atoms to hydrogen bond with the receptor. As well, hot-spot residues are residues which contribute more towards binding with the receptor. The experimental alanine scan was a good starting point in order to design and run the FA assay and identify hotspot residues. Computational tools were used by researchers at the University of Bristol to see if *in silico* predicted hotspots mirrored experimental results using BUDE. When BUDE failed to predict one of the three GKAP ligand hotspots the team endeavoured to use multiple computational tools, showing that the use of multiple tools . Ultimately, the need for experimental validation of hot spot residues remains pertinent.

The short GKAP peptide sequence allowed for the experimental testing of peptide variants quite easily. Subsequent design, synthesis and testing of GKAP peptide variants established residues A2 and R5 as completely uninvolved in the binding interaction while the T4S variant seemed to suggest this residue could be involved is specificity. Lastly, the L6F variation showed similar binding to SHANK1 as the native GKAP sequence with the crystal structure showing Phe well accommodated within the hydrophobic pocket of SHANK1.

The requirement list for the design of a β -strand mimetic to target the GKAP/SHANK1 PPI seemed to be agreeable. Two peptidomimetic structures, molecules **1** and **2**, were rationally designed to incorporate the Hao building block capable, in theory, to pre-organise of the GKAP β -strand. However, significant difficulty was encountered with the synthesis of the macrocycle which was not obtained. However, successful synthesis of peptidomimetic **1** was achieved. Although no improvement in binding was observed in FA or ITC against the SHANK1 protein target no detrimental effect on binding was observed either. Further efforts into obtaining a Hao-containing molecule amenable to cyclisation could provide a mimetic to target this PPI. However, time constraints diminished the possibility of success.

Chapter 3: Design of Novel Peptide Interacting Motifs Using Predictive Saturation Variation Scanning (PreSaVS)

3.1 Introduction

Decoding, characterizing and understanding PPIs is pivotal for the development and application of selective and successful therapeutics. Progress has been achieved in targeting PPIs^{8,181} yet these are still fairly difficult to target due to the large, flat surface area over which they occur^{1,6,182,183}. However, not all PPI interfaces are flat. PPIs have been characterised based on the secondary or tertiary structure at the interfaces,^{43,184}. Using these secondary structured motifs as templates offers a strategy for the development of peptide-based therapeutics^{185,186}. Peptides generally provide a good starting point for probing, understanding and development of PPI inhibitors. Peptides can target the large surface area of a PPI and can be endowed with selectivity, specificity and low toxicity^{185,187}. However, peptides suffer from low bioavailability and low cellular permeability^{188–190}. Screening and identifying novel peptide sequences, that can bind to the target ligand, can also be used as a launching point for discovery of PPI inhibitors. Through peptide phage displays^{106,136,139,185} or non-antibody affirmer technology^{140,141,191} and split-intein circular ligation of peptides and proteins (SICLOPPS)¹⁴³ such sequences have been discovered. These technologies enable vast sequence space to be explored and functional peptides to be identified. Specifically a wide range of peptide sequences, are screened against protein targets to help with the analysis of the binding interaction and identification of bioactive target-specific peptides^{136,192}. Although these techniques provide excellent experimental sequence optimization, there is an opportunity for positive change by enhancing methods to explore sequence space using computation.

Alongside combinatorial biology efforts, the use of computational methods can facilitate design of peptide- and protein-based ligands. For example computational tools paired with experimental characterization have been successfully implemented to accelerate the rational design of functional biomolecules such as helix bundles¹⁹³ and coiled-coils¹⁹⁴. The helix bundle was grafted with important side chains from known binders of different BCL2 proteins

yielding protein inhibitors for all 6 pro-survival BCL2 homologs¹⁹³. Similarly, a designed coiled-coil with hot-spot residues of NOXA-B was shown to interrupt the BID/MCL-1 interaction¹⁹⁵. However, these examples rely on the initial design of structurally functional and stable mini protein-like assemblies. Much less encountered are integrated computational and experimental strategies to design peptide interacting motifs (PIMs). Performing a Computational Alanine Scan (CAS) has become a standard procedure in identifying potential 'hot-spots'.^{19,20,196,197} As mentioned in the introduction, a 'hot-spot' is a residue which contributes significantly to binding within a PPI evidenced after its computational variation to alanine shows a subsequent loss in binding^{19,20}. Furthermore, Gellman and co-workers discussed a complementary hydrophile scan, where systematic variation to charged residues of the BIM peptide sequence, offered insights into the driving forces and specificity of binding to both its targets; BCL-xL and MCL-1²⁹. Combining the results of their hydrophile scan with the information gathered from CAS, the group was able to design variant BIM peptides with minimal sequence variations that achieved selectivity towards either BCL-xL or MCL-1. Thus, integrated computational methods, such as CAS or hydrophile scanning, and experimental characterization can be successfully used in the design of upgraded PIMs. Lastly, Arora and co-workers developed and applied AlphaSpace¹⁹⁸, a strategy to identify unoccupied space around key amino acid residues. This enabled the design of peptides with non-natural residues that can better explore this unoccupied space. They applied this tool to the of the KIX/MLL PPI. Following truncation of a KIX binding MLL sequence insertion of natural and non-natural amino acid variants, allowed identification of optimized binding sequences that maximized binding pocket coverage¹⁹⁹.

3.2 Objectives

The objective of this work was to develop a methodology where PIMs could be rapidly and efficiently generated. The design and development of successful peptidomimetics comes from the marriage of experimental and computational tools. Gathering knowledge of the PPI is imperative; the secondary structure topology at the interface, the 'hot-spots' and the chemistries applied previously to target similar PPIs. This work looked to expand upon the essential data required to target PPIs – develop a relatively simple computational and experimental workflow to alter the peptide sequence offering more insight into the flexibility of the binding interface.

Drawing inspiration from the work done in Chapter 2, where selected peptide variants of the GKAP ligand were shown to still bind SHANK1 with comparable affinities to the wild-type sequence, a more methodical approach was proposed. Thus, the design of the *in silico* PREdictive Saturation Variation Scanning (PreSaVS), that would computationally substitute each amino acid in a peptide sequence of interest to all 20 other available residues in an expedient manner, was proposed (Figure 31). The PreSaVS approach would generate PIMs for two PPI target systems, driven by the previously described BUDE tool, ensuring smooth progression of computational labors.

To establish the PreSaVS workflow, the project aimed to achieve two goals:

- Apply the computational PreSaVS on at least two diverse PPIs
- Experimentally validate the PreSaVS approach

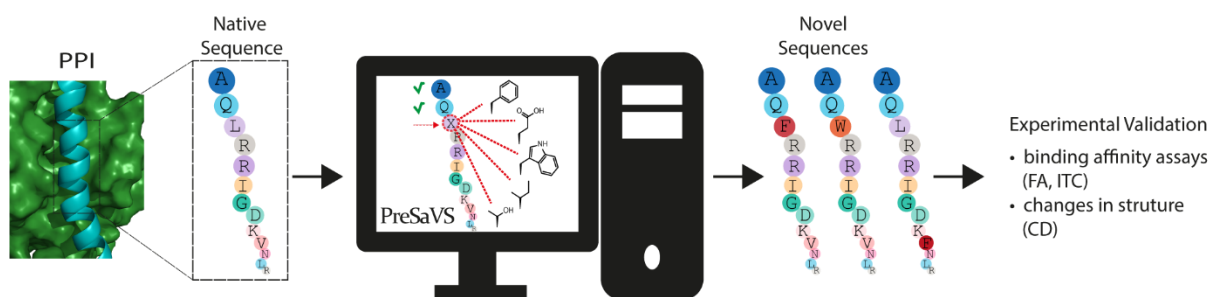


Figure 31. PreSaVS workflow: following target selection the native ligand peptide sequence is computationally placed through the PreSaVS software and new peptide sequences with point variations suggested that require experimental validation.

3.3 Targets of Interest

To establish the PreSaVS workflow, the method was applied to both an α -helix and a β -strand mediated PPI. The NOXA/MCL-1 PPI was selected as a model α -helix mediated PPI, and the SIM/SUMO PPI chosen as the β -strand mediated PPI – both had been investigated previously by other Wilson and Woolfson group members in their development of an experimentally validated approach to CAS¹⁹. NOXA and MCL-1 are proteins of the B-cell Lymphoma 2 (BCL-2) family of apoptotic proteins that are mis-regulated in cancer and have been the focus of significant drug-discovery efforts.^{69,200} Similarly, Small ubiquitin-like modifiers (SUMO) regulate many cellular processes, however the mechanisms of regulation need further investigation^{201,202}. The driving forces for these PPIs were readily available in the literature because they had been previously studied^{195,203} and hot-spot residues identified¹⁹

3.3.1 α -Helix Mediated PPI: NOXA/MCL-1

NOXA and MCL-1 are proteins from the B-cell Lymphoma 2 (BCL-2) family of proteins which control apoptosis (programmed cell death) a healthy process that is mis-regulated in cancer. NOXA is an apoptosis inducing BH3-only protein^{45,204} which binds selectively to MCL-1 over other BCL-2 family members²⁰⁵. NOXA adopts a helical conformation upon binding to MCL-1 and tips the balance of pro- and anti-apoptotic signals toward apoptosis²⁰⁶. Considerable efforts have been placed in targeting MCL-1, as this anti-apoptotic protein is overexpressed in cancers and provides drug resistance during cancer therapy^{69,200,207}. The NOXA B sequence (AAQLRRIGDKVNLRQKLLN) was measured to bind with an affinity of $K_d = 80$ nm in fluorescence anisotropy assay¹⁹. The hydrophobic and electrostatic driving forces that determine this PPI were previously studied¹⁹⁵ and the hot-spot residues identified (L78, R79, I81, D83 and V85), and recently compared through different software tools¹⁹ (Figure 32).

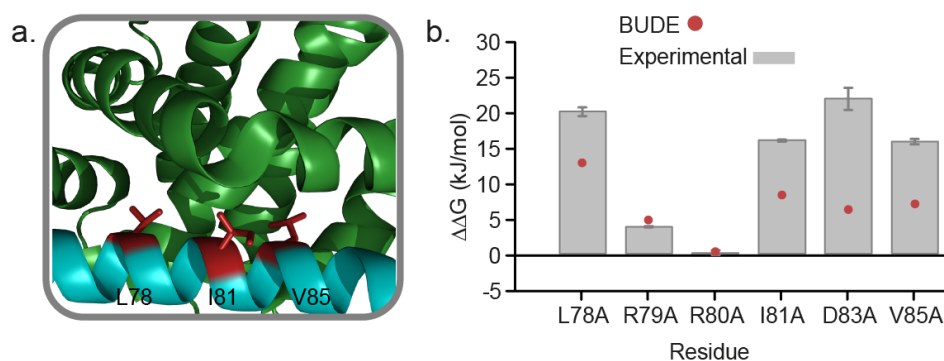


Figure 32. Structural, computational and experimental data for NOXA/MCL-1; a) NMR structure of NOXA/MCL-1 (PDB 2JM6¹⁰) showing NOXA in its α -helical conformation upon binding, with the hot-spot residues highlighted in red; b) The computational BUDE AlaScan results compared with experimental data¹⁹.

Although a number of residues were identified as hot-spots using BUDE Alanine Scanning in the NOXA sequence¹⁹⁵ focus was placed on studying: L78, I81, D83 and V85. Among the hot-spots, L78 is a highly conserved residue among the proapoptosis BH3-only members^{44,45}. The CAS results were confirmed though experimental testing previously by Katherine Horner.

3.3.2 β -Strand Mediated PPI: SIM/SUMO

Small ubiquitin-like modifier (SUMO) regulates many cellular processes including: cell cycle development, DNA repair and transcription^{208–210}. Malfunction within the SUMO PPI network results in cancer progression²¹¹. However the mechanisms of regulation are complex and need further investigation²⁰². SUMO-interacting motifs (SIMs) bind to SUMO through an extended β -strand. Recently a systematic computational and experimental analysis of the SIM peptide sequence and the driving force for binding to SUMO was described¹⁹.

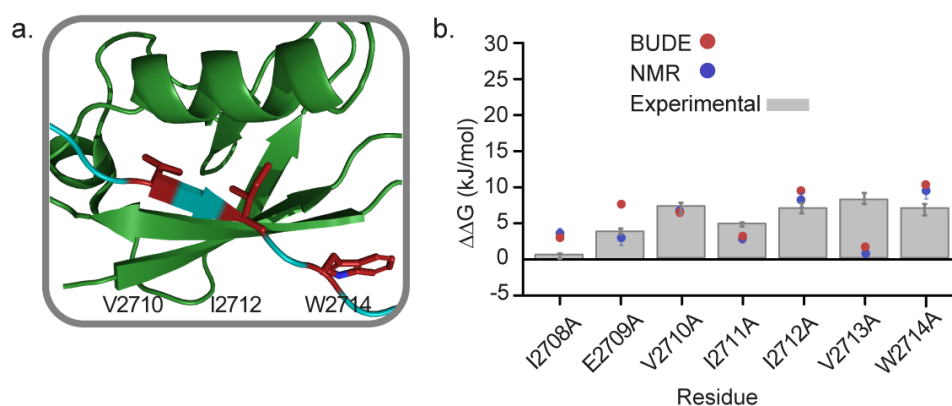


Figure 33. SIM/SUMO; a) NMR structure of SIM/SUMO (PDB 2LAS²⁰³) showing SIM in an extended β -strand conformation upon binding, with the hot-spot residues highlighted in red; b) The computational BUDE AlaScan results compared with the experimental data.

Three hot-spot residues were identified within the SIM sequence using BudeAlaScan: V2710, I2712 and W2714. Experimental testing showed that BudeAlaScan, tended to overestimate the importance of charged residues and thus E2709A, which was on the border, was dismissed as a hot-spot particularly after the NMR ensemble was used (blue dots Figure 33). These were all confirmed following synthesis of the alanine variant peptides through testing in biophysical techniques.

3.4 Results and Discussion

The structural and binding information established the NOXA/MCL-1 and SIM/SUMO PPIs as ideal model systems to test and develop our PreSaVS workflow which substitutes each amino acid in a peptide sequence of interest, to all 20 other available residues computationally, in an expedient manner. The computational PreSaVS was conducted by Richard Sessions and Amaury A. Ibarra at Bristol University for both PPIs. Like CAS, PreSaVS is based on the difference in binding free energy ($\Delta\Delta G$) values where the approach calculates the potential energy based on force fields. The $\Delta\Delta G$ values can be used to determine whether an amino acid substitution would be tolerated or detrimental for binding or enhance binding to its target ligand. In PreSaVS the $\Delta\Delta G$ values are obtained by subtracting the binding energy for the variant peptide from that of the wild-type peptide; $\Delta\Delta G = \Delta G_{wt} - \Delta G_{variant}$. Thus, the $\Delta\Delta G$ values of residues are predicted to be tolerated or have enhanced affinity to the target protein are assumed to be those with $\Delta\Delta G$ values higher than 0 kJ/mol. The scan also assumes that the

change in binding energy arises from the side chain only. The innate challenge in improving rather than knocking-out affinity coupled with the absence of a defined $\Delta\Delta G$ threshold value rendered benchmarking of the method challenging, hence experimental validation was essential. The complete results of the scan can be observed in the Experimental section (5.1).

3.4.1 Computational PreSaVS Scan and Design of Novel Peptide Binders

The two target PPIs were subjected to the PreSaVS computational scan and the outcome was visually inspected (Data for key residues Figure 34a,c with full scan data in Experimental). A $\Delta\Delta G$ value > 0 would be predicted to be a tolerated and, potentially beneficial, variation to make in the peptide sequence, while $\Delta\Delta G$ value < 0 would represent a variation predicted to have a disruptive effect on the PPI. By analogy to the criterion for hotspot residues ($\Delta\Delta G \geq -4.5$ kJ/mol), we reasoned that a threshold of $\Delta\Delta G \geq 4.5$ kJ/mol would represent a reasonable value for a favourable amino acid side chain. Based on the predictions it was reasoned that the NOXA sequence could be modified at the L78 and V85 positions with predicted favourable variations (both hot-residues, Figure 34a). The positive $\Delta\Delta G$ values for L78 varied to F and W were encouraging and selected for experimental assessment. Furthermore, the L78 to Y variation predicted a very large and negative $\Delta\Delta G$ value and was therefore also selected as a negative control. Finally, the V85 position also gave encouraging predictions and the F variation was chosen for synthesis to test the limitations of the predicted $\Delta\Delta G$ values; the value was slightly below 0 and we were interested to use this to further define the window of accuracy. A total of four NOXA peptide variants was thus selected for synthesis and evaluation (Figure 34b).

a.

AA	D70	L78	V85	L91
D	0.00	-8.19	-5.44	-1.44
F	-4.24	4.92	-1.71	0.41
W	-3.98	4.79	1.83	0.18
Y	-3.28	-31.55	-12.83	0.41

$\Delta\Delta G = \Delta G_{wt} - \Delta G_{variant}$

b.

Peptide	Sequence
NOXA ₇₅₋₉₃	AAQLRRIGDKVNLKQLLN
L78F	AAQFRRIKDKVNLKQLLN
L78W	AAQWRRIKDKVNLKQLLN
L78Y	AAQYRRIKDKVNLKQLLN
V85F	AAQLRRIGDKFNLRKQLLN

c.

AA	I2708	V2710	I2711	V2713
D	8.71	6.23	3.17	0.80
E	6.71	3.21	7.36	4.94
F	1.52	-24.33	-0.18	0.69
I	0.00	4.12	0.00	0.31

$\Delta\Delta G = \Delta G_{wt} - \Delta G_{variant}$

d.

Peptide	Sequence
SIM ₂₇₀₅₋₂₇₁₇	DNEIEVIIVWEKK
I2708D	DNEDEVIIVWEKK
V2710I	DNEIEIIVWEKK
I2711E	DNEIEVEIVWEKK
V2713E	DNEIEVIIEWEKK

Figure 34. Design of variant peptides from the computational PreSaVS; a) Excerpt of the computational PreSaVS results of the NOXA peptide; b) NOXA variant peptide sequences selected for synthesis and further testing; c) Excerpt of the computational PreSaVS results of the SIM peptide; d) SIM variant peptide sequences selected for synthesis and further testing.

The SIM peptide sequence gave interesting computational results with favorable variations identified for both hotspots and non-hotspots as worthy for investigation (Figure 34c). Whereas the NOXA results were consistent in the type of amino acid variation, ie from L to F and W all hydrophobic residues, the PreSaVS outcome for the SIM peptide suggested changes from hydrophobic to charged amino acids. As such the predictions fed into the design of I2708D, V2710I (hot-spot residue), I2711E and V2713E peptide variants for experimental investigation (Figure 34d).

Using this approach could have potential to optimize known hotspot residues or to theoretically increase binding affinity through the introduction of favourable interactions for non-hotspot residues. Albeit, the variation of hotspot residues in such a way to improve binding was foreseen to be more difficult because of potential of steric clashes due to bulkier side chain variations within the binding pockets of the target proteins. Thus potential tolerance in variation was perceived as a positive outcome with respect to hotspot residues. However, variation in non-hotspot residues could potentially pick-up untapped interactions with residues form neighboring side chains in the binding interface.

3.4.2 Experimental Validation of the Novel Peptide Sequences

The peptide variants were prepared using microwave assisted Fmoc solid phase synthesis on Rink Amide MBHA resin. Acetyl capped C-terminus NOXA-B and SIM variant peptide sequences were prepared. Fluorescently labelled BIM and SIM peptides were used as tracers which were previously prepared by Katherine Horner and Som Dutt respectively. The direct titration Fluorescence Anisotropy (FA) assay for each target was performed, affording a $K_d = 204 \pm 16$ nM for the fluorescently labelled wt-BID control peptide against MCL-1 and $K_d = 1.5 \pm 0.2$ μ M for the fluorescently labelled wt-SIM against SUMO (Figure 35). Following the direct titration assays, competition FA assays for both targets were set-up and validated to test the effect of peptide sequence variation.

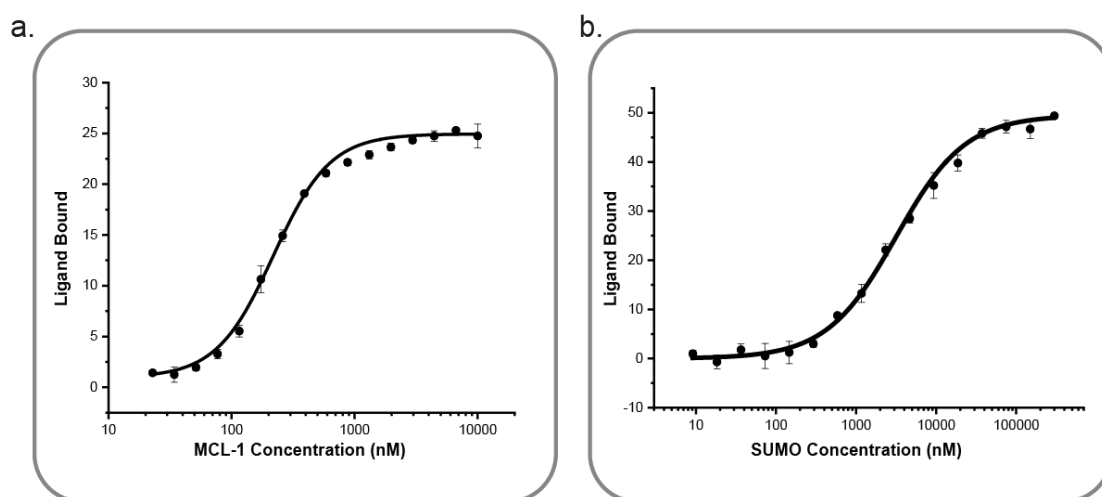


Figure 35. FA direct titration curves; a) Direct titration assay for NOXA/MCL-1 in 50 mM Tris, 150 mM NaCl buffer using 25 nM tracer concentration; b) Direct titration assay for SIM/SUMO in 50 mM Tris, 150 mM NaCl buffer using 50 nM tracer concentration.

3.4.3 NOXA/MCL-1

The peptide variants were prepared using SPPS, described earlier (2.3.1). These were titrated and a constant concentration of protein, followed by a constant concentration of tracer were added. The variant peptides would thus be tested for their ability to displace the tracer molecule from MCL-1. The competition FA studies revealed that NOXA₇₅₋₉₃L78F and NOXA₇₅₋₉₃L78W variations were tolerated well with IC₅₀ values of $2.69 \pm 0.05 \mu\text{M}$ and $2.17 \pm 0.03 \mu\text{M}$ respectively, comparable to $1.17 \pm 0.04 \mu\text{M}$ for wt-NOXA₇₅₋₉₃. In contrast the NOXA₇₅₋₉₃V85F variant displayed weaker inhibitory activity against FAM-BID tracer to MCL-1 (

Figure 36a) with an IC₅₀ >20 μM . To validate these results the peptide variants were also tested in Isothermal Calorimetry (ITC), and showed comparable K_d values to wt-NOXA₇₅₋₉₃ which supported the computational predictions for these variations to bind to MCL-1 and not be disruptive to binding (Table 1). All binding peptides exhibited lower entropic values when compared to the wild-type peptide upon binding based on ITC (

Figure 36 and Table 4). The less unfavourable entropy of binding was possibly due to the more rigid structure of the F and W side chains compared to the original L and V residues. The ΔG values for NOXA₇₅₋₉₃L78F and NOXA₇₅₋₉₃L78W were slightly lower than that of the wild-type peptide, however enthalpy/entropy compensation was also observed. Lastly, as expected, the NOXA₇₅₋₉₃L78Y exhibited loss in binding to MCL-1; an IC₅₀ could not be determined from the starting concentration of 15 μM , providing further confidence in the computational workflow. The NOXA peptide binds to MCL-1 selectively over all other BCL-2 family members. As such, the peptide variants were all tested in FA against another BCL-2 family member protein; BCL-xL. It was reasoned that a change in sequence could also result in a change in specificity. The FA results showed no binding to BCL-xL (

Figure 36c). Given the objective to identify tolerated changes in the NOXA peptide, the observation that binding specificity remained unaltered was an encouraging outcome.

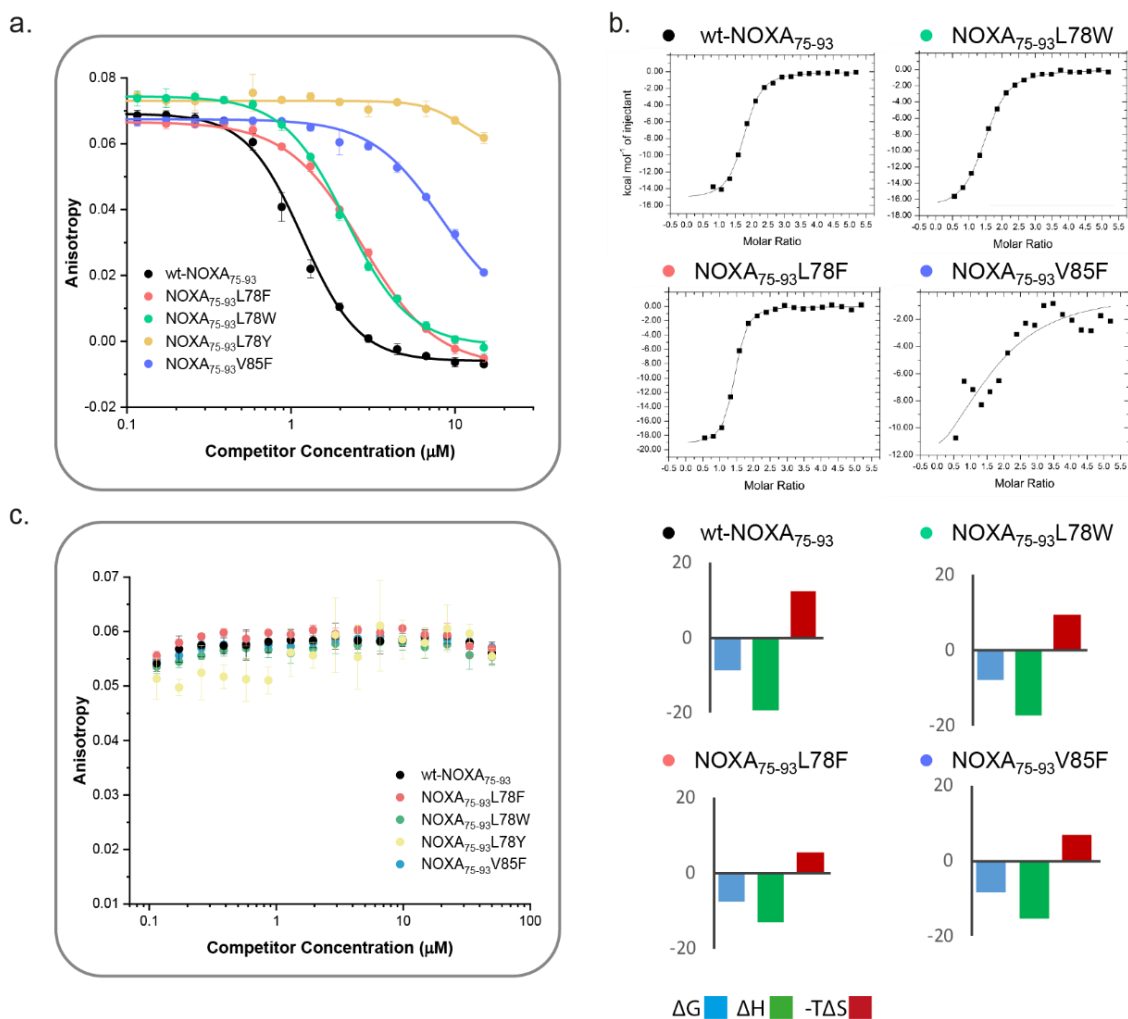


Figure 36. Data set for wt-NOXA₇₅₋₉₃ and the variant sequences presented in **Table 4**; a) Competition FA data for all NOXA peptides in 50 mM Tris, 150 mM NaCl buffer using 25 nM tracer concentration and 150 nM MCL-1; b) ITC data for the peptide variants in 50 mM Tris, 150 mM NaCl buffer using 500 μM peptide concentration and 20 μM MCL-1 followed by plotted thermodynamic profiles from the ITC data for the NOXA peptide variants and wild type NOXA; c) Competition assays of the NOXA variant peptides and wt-NOXA peptide against the BCL-x_L protein in 50 mM Tris, 150 mM NaCl buffer using 25 nM tracer concentration and 150 nM BCL-x_L protein

Table 4. Summary of data of wt-NOXA and variants.

NOXA	wt	L78F	L78W	L78Y	V85F
K _d (μM)	0.4 ± 0.1	1.7 ± 0.1	2.4 ± 0.2	N/A	9.5 ± 0.8
IC ₅₀ (μM)	1.2 ± 0.1	2.7 ± 0.1	2.2 ± 0.1	N/A	> 20
ΔG (kJ/mol)	-8.7	-7.9	-7.7	N/A	-8.4
ΔH (kJ/mol)	-19 ± 1	-17 ± 3	-13 ± 4	N/A	-15 ± 3
-TΔS (kJ/mol)	12	9.5	5.5	N/A	6.9
%helicity	15	8	12	-	9
%helicity (30%TFE)	60	27	37	-	16

Lastly, structural effects of these variations, were investigated using Circular Dichroism (CD). The peptides exhibited CD spectra consistent with a random coil in 20 mM potassium phosphate buffer at 150 μM peptide concentration (Figure 37a). CD spectra in 30% TFE in 20 mM potassium phosphate buffer (Figure 37b) were also acquired to provide an impression of the relative maximum helicity for each variant peptide (Table 4). All the variant peptides exhibited increased helicity (27% NOXA₇₅₋₉₃L78F, 37% NOXA₇₅₋₉₃L78W and 16% NOXA₇₅₋₉₃V85F) in 30% TFE from the helicity observed for the peptides solely in buffer (8% NOXA₇₅₋₉₃L78F, 12% NOXA₇₅₋₉₃L78W and 9% NOXA₇₅₋₉₃V85F). The helicity for NOXA₇₅₋₉₃L78F and NOXA₇₅₋₉₃L78W was surprising since the helix enforcing L residue²¹² was substituted for aromatic residues, whereas the V85 residue exhibited a less significant increase in helicity when changed to F. The variation in sequence clearly had a significant effect on the propensity of the peptides to adopt a helical conformation. When compared to wt-NOXA₇₅₋₉₃, which showed 15% helicity in buffer but a significantly higher propensity of 60% in the 30%TFE environment, the variant peptides did not experience such a significant helical propensity. Since, the effect of the side chain variation did not impact binding affinity of the peptides, this suggested that the change in helicity between variants does not significantly contribute towards binding affinity measured.

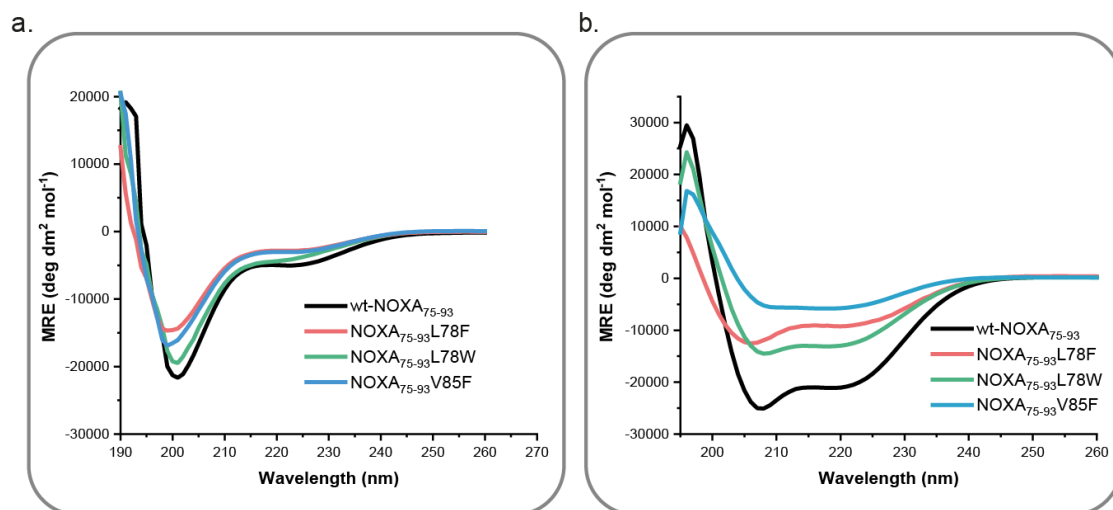


Figure 37. CD data for the wt-NOXA and binding peptide variants; a) CD in 20 mM potassium phosphate buffer, 150 μ M peptide concentration; b) CD in 20 mM potassium phosphate buffer with 30% TFE, 150 μ M peptide concentration.

3.4.4 SIM/SUMO

The SIM peptide variants were also evaluated using FA competition where a constant concentration of protein and tracer were added to each titrated peptide variant (Figure 38a). The ability of each variant to displace the tracer was thus acquired. Most striking was the FA result for SIM₂₇₀₅₋₂₇₁₇I2708D, which showed slightly better inhibitory activity than the wt-SIM, $IC_{50} = 14 \pm 0.7 \mu$ M and $IC_{50} = 22 \pm 0.3 \mu$ M respectively (

Table 5). Usually, hydrophobic residues pack in the surface of their ligand away from the aqueous environment, while hydrophilic residues, charged at neutral pH, can form salt bridges. Interestingly, the I2708 hydrophobic residue was swapped for a charged residue and binding was not affected. In the SIM/SUMO PPI, the SIM peptide binds in an extended β -strand conformation to the SUMO protein. This conformation allows for the side chains to alternate on either side of the backbone. The I2708 residue side chain points towards the SUMO protein (Figure 38b), however upon further analysis of the SIM/SUMO NMR derived structural ensemble, it is evident that the side chain of I2708 is proximal to an LKKL sequence of SUMO (Figure 38c). As such, variation at this position for D, with a negatively charged sidechain at neutral pH, could potentially introduce a salt-bridge with K46, a positively charged residue at neutral pH. Perhaps less surprising was the SIM₂₇₀₅₋₂₇₁₇V2710I variant peptide, which also maintained binding to SUMO with an $IC_{50} = 16 \pm 2 \mu$ M. (Figure 38a). Finally, the SIM₂₇₀₅₋

I2717I2711E and SIM₂₇₀₅₋₂₇₁₇V2713E variant peptides experienced some loss in binding affinity based on the competition FA results. In the NMR ensemble I2711 appears to be solvent exposed, however, flanking the SUMO protein (Figure 38c). I2711 residue sits between the hydrophobic Y21 and the charged K37 SUMO residues (Figure 38c). This offers insight into both the computational and experimental results. The proximity of I2711 to K37 would suggest the I2711E variation could allow for a potential salt bridge at this position. However, the hydrophobic aromatic ring Y21 residue could potentially repel this variation. Lastly, V2713E is also solvent exposed. In this case the residue does not flank the SUMO protein at all, (Figure 38c). The weaker binding affinity cannot thus be readily reconciled by considering potential interactions it may make with the SUMO interface. Earlier work established that the variation V2713A exhibited loss in binding affinity with a $K_d = 100 \pm 20$ nM, a significant difference from the wild-type K_d of 3.7 ± 0.3 nM as measured under¹⁹. As well, it is well known that certain amino acids have the propensity to favour β -strand conformations, with valine and isoleucine among such amino acids. The effect of changing structurally imposing residues in the SIM sequence could affect the overall conformation of the peptide and thus the binding affinity.

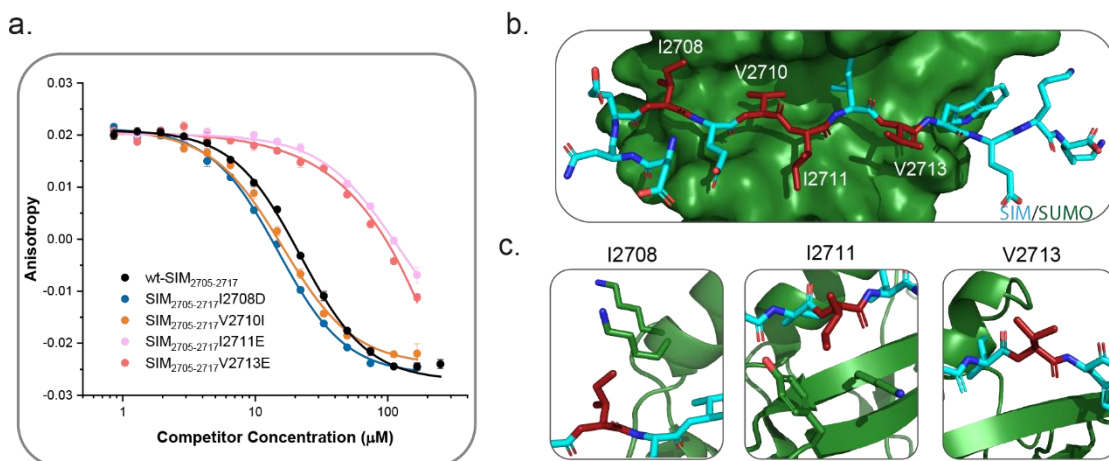


Figure 38. FA data for the SIM peptide variants; a) Competition assay data for the wt-SIM₂₇₀₅₋₂₇₁₇ and SIM peptide variants, 50 mM Tris, 150 mM NaCl buffer, 150 μM competitor concentration, 25 nM tracer, 100 nM SUMO; b) SIM/SUMO (PDB: 2LAS²⁰³) with I2708, V2710, I2711, V2713 residues in red c) I2708 proximity to SUMO K residues, I2711 proximity to SUMO residues Y21 and K37, V2713 pointing towards solvent.

Table 5. IC₅₀ values for the SIM peptides.

SIM	Wt	I2708D	V2710I	I2711E	V2713E
IC ₅₀ (μM)	21.9 ± 0.29	14.8 ± 0.73	16.1 ± 1.60	> 200	> 200

3.5 Conclusions and Future Work

The PreSaVS workflow was applied to both an α -helix and β -strand mediated PPI. Targeting α -helix mediated PPIs has become more accessible, however, PIMs were generated and tested for binding against the model system NOXA/MCL-1 in a quick and efficient manner. Although β -strand mediated PPIs are more difficult to target, we have also generated novel peptide sequences that still bind the SUMO target protein. The PreSaVS scan was successful in identifying sequence variations that were tolerated, that maintain binding potencies to target proteins, rather than favourable ones that result in an increase binding affinity. However, this tool is very useful in generating novel peptide sequences in a fast and efficient manner from which further properties could be tuned later on, such as proteolytic stability and cellular permeation. The FA data, and ITC data where applicable, demonstrate this is a reliable tool in the generation of novel peptide sequences that can still bind desired protein targets. Understanding the impact of sequence variation is not only important in explaining why the new variant peptides still bind but also in exploring the physiochemical changes experienced by each. Through the combination of different computational and experimental tools, we can strive to design and synthesise PIM sequences that could be used as starting points in understanding PPIs but also in the design of PPI inhibitors. The PreSaVS tool could be an easily accessible method for developing PIMs, incorporated alongside other computational tools such as CAS in the quest for more insight into the driving elements of binding.

Chapter 4: Design, Synthesis and Evaluation of Stapled peptide inhibitors of the HIF-1 α /p300 PPI

4.1 HIF1- α /p300 Introduction

Hypoxia Inducible Factor-1 α (HIF-1 α) is an important member of the hypoxic sensing pathway. Continually expressed in cells, HIF-1 α is degraded as fast as it is produced under normoxia via an oxygen-dependent mechanism^{56,213–215}. Under hypoxic conditions, however, HIF-1 α accumulates enabling it to form a dimer with HIF-1 β that is termed HIF-1²¹⁶. The HIF-1 translocates to the nucleus and recruits p300, a transcriptional co-factor, together binding to the hypoxic response elements (HRE) on DNA resulting in transcription of genes to relieve hypoxia^{49,51,217–219} (Figure 39).

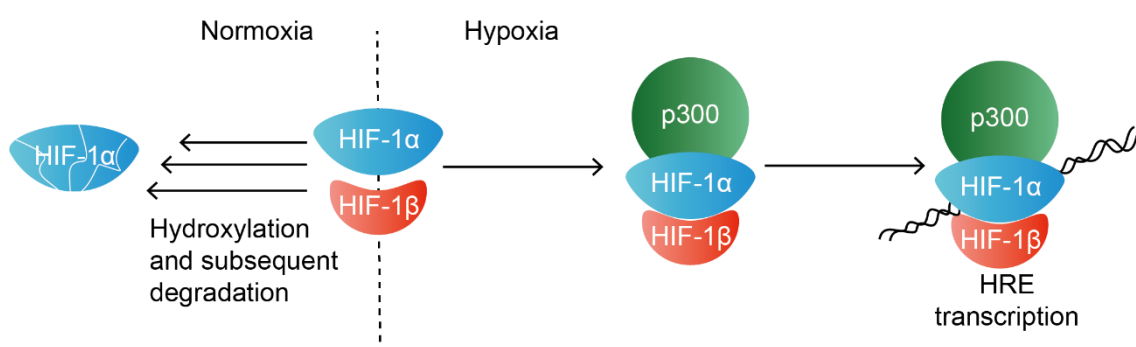


Figure 39. Condensed summary cartoon of the Hif-1 α regulation pathway: with HIF-1 α degradation taking place under normal oxygen levels (left) and HIF-1 α activity under hypoxia (right).

The downstream effects of this PPI have consequences in the promotion of tumor growth. Rapidly dividing cancer cells deplete local oxygen supply resulting in hypoxia²²⁰. Consequently, overexpression of HIF-1 α takes place and the downstream events associated with HIF-1 α activity alleviate the hypoxic state by generating new vasculature (angiogenesis) to supply the tumour with oxygen^{220,221}. The role of HIF-1 α in cancer progression indicates it may be a viable target for therapeutic intervention^{51,52,222}. Moreover, clinical data have linked overexpression of HIF-1 α with increased resistance to cancer therapy^{220,223}. Thus understanding and developing inhibitors of this PPI could bring high reward in cancer therapeutics.

4.1.1 HIF-1 α /p300 Binding Interaction

There are no crystal structures of the HIF-1 α /p300 complex to date. However, structures of the HIF-1 α C-terminal transactivation domain (CTAD) bound to the cysteine/histidine domain (CH1) of p300 were determined by two different groups using solution NMR studies (PDBs: 1L8C and 1L3E)^{34,224}. In the NMR structures, p300 CH1 is observed to consist of 5 helices, quite compactly bound to 3 zinc atoms (Figure 40a). The interaction is observed when HIF-1 α CTAD binds to p300 CH1 and forms three helical regions connected by unstructured segments²²⁴ (Figure 40b). Targeting this extensive interaction interface with a small molecule would be anticipated to be difficult (Figure 40c).

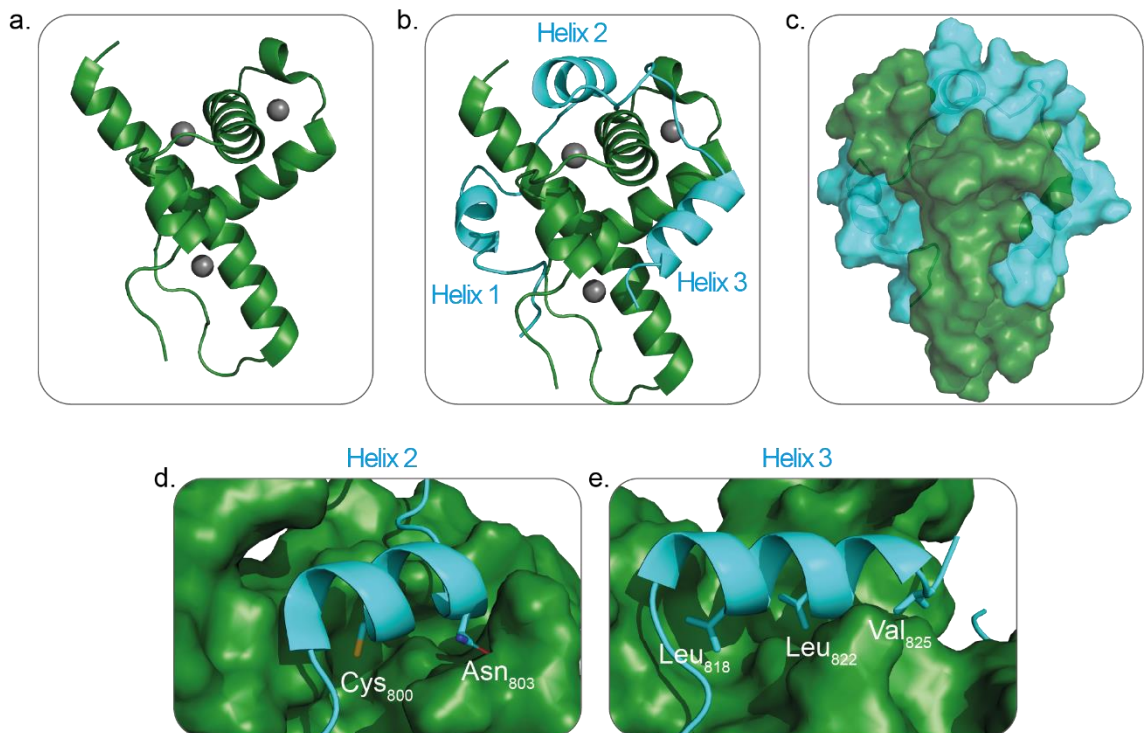


Figure 40. Structural features of the HIF-1 α /p300 interaction (HIF-1 in cyan and p300 in forest green); PPI determined through solution NMR experiments (PDB: 1L8C): a) p300 in a very compact 4 helix bundle structurally supported through 3 zinc atoms; b) HIF-1 α wrapped around p300; c) both proteins shown as surfaces to emphasize the large surface area which a small molecule would not be able to cover; d) Helix 2 with Cys₈₀₀ and Asn₈₀₃ and e) Helix 3 with Leu₈₁₈, Leu₈₂₂ and Val₈₂₅ residues which are believed to be important for binding p300.

Cellular studies where HIF-1 α was expressed under normoxic and hypoxic conditions pointed towards key residues for binding p300 as Cys₈₀₀ and Asn₈₀₃^{56,57,119} on helix 2 while experimental alanine scanning mutagenesis

studies pointed towards Leu₈₁₈, Leu₈₂₂ and Val₈₂₅ on helix 3, as potentially important for binding to p300⁵³ (Figure 40d,e). In addition, Asp₈₂₃ and Gln₈₃₄ were two other residues on helix 3 flagged through Computational Alanine Scan (CAS) as potentially important for binding²²⁵. However, adding to the difficulty of deciphering the driving forces of binding for this PPI, some of the studies mentioned, and others, showed contradictory results regarding the importance of these residues on the HIF-1 α sequence^{53,55,225,226}. More recently in the Wilson group, a thorough investigation using computational alanine scanning (CAS) followed by experimental synthesis and testing of single, double and triple alanine peptide variants of HIF-1 α showed no significant loss in binding due to these variations (Fruzina Hobor, 2019, personal communication). As such none of the aforementioned hotspot residues were validated and no new ones identified either. Although no hotspot residues have been identified, HIF-1 α truncation studies have identified regions of the protein, which are necessary for interaction with p300. Truncation of the shortest helical segment, helix 1, resulted in a moderate effect on activity⁵⁸. However, truncation of the largest helical segment, helix 3, resulted in significant of HIF-1 α affinity⁵⁹.

4.1.2 Inhibitors of HIF-1 α /p300

Active compounds that indirectly modulate the downstream effects of the HIF-1 α /p300 PPI work through: targeting HIF-1 α protein synthesis^{227–229}, inhibiting the HIF-1 α /HIF-1 β dimerization interface^{230,231} or promoting HIF-1 α degradation²³². Efforts to directly target the HIF-1 α /p300 interaction have identified limited examples. Chetomin, a natural product, initially identified as a modulator of this PPI was, after further studies, shown to disrupt the p300 CH1 domain through zinc ejection^{233,234} (Figure 41). Compounds which work via this mechanism, are problematic as they typically lack selectivity resulting in high toxicity. As well, p300 has multiple binding partners and is necessary in mediating a number of PPI pathways all of which would be affected if p300 was structurally compromised by the inhibitor. The other small molecule identified through HTS labelled KCN-1 was also reported to inhibit the HIF-1 α /p300 interaction²³⁵ (Figure 41). However, there is no direct evidence that this molecule binds either p300 or

HIF-1 α ²³⁶. Furthermore, the KCN-1 was synthesised and tested in a competition FA assay by George Burslem (University of Leeds) and although solubility issues were encountered, the compound failed to inhibit the HIF-1 α /p300 PPI at a starting concentration of 25 μ M²³⁷.

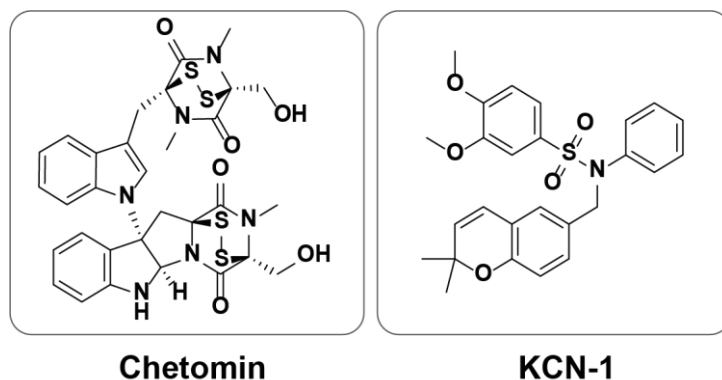


Figure 41. Small molecules thought to target the HIF-1 α /p300 PPI.

Peptide-based approaches and peptidomimetics have also been employed to target the HIF-1 α /p300 interaction. Hydrogen bond surrogate (HBS) helices were designed by the Arora group based on helix 2¹¹⁹ and helix 3²²⁵ segments of HIF-1 α (Figure 4). Both helix 2 and helix 3 based HBS structures showed improved helicity through CD analysis and modest p300 binding potencies as demonstrated by ITC and tryptophan fluorescence spectroscopy experiments respectively (Figure 42). Furthermore, the ability of the HBS helices to downregulate HIF-1 α promoter activity in cellular assays was successful. The group further measured the effects of the HBS helix based on helix 3, on a mouse xenograft model to assess *in vivo* efficacy. The results showed that mice treated with the HBS exhibited 53% tumour reductions when compared with mice from the control group.

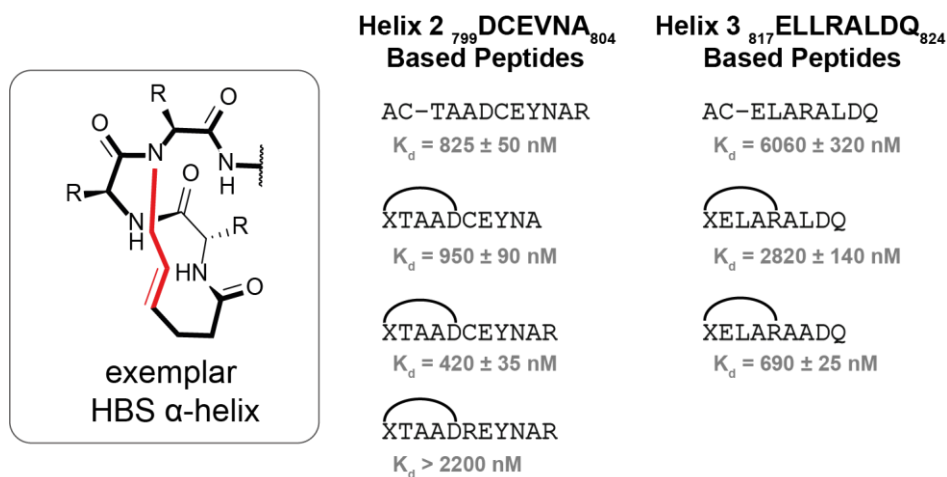


Figure 42. HBS structures derived from helix 2 and helix 3 of HIF-1 α to target the HIF-1/p300 for inhibition with K_d values determined using ITC. The best inhibitors from each helix were used to treat mice xenografts models and showed tumour reduction when compared to the control mice.

In the Wilson group, the HIF-1 α /p300 interaction was previously targeted using a proteomimetic approach with designed trimeric 3-O-alkylated aromatic oligoamides mimicking helix 3 of HIF-1 α ²³⁸. The scaffold replicated the *i*, *i*+4 and *i*+7 side chain projections of an α -helical peptide conformation, and had been decorated with groups mimicking presumed hotspot residues. Fluorescence anisotropy competition showed the best decorated compound to inhibit the HIF-1 α /p300 interaction with an $IC_{50} = 9.2$ μ M however with a similar potency towards *hDM2*/p53 ($IC_{50} = 16$ μ M) (Figure 43a). To circumvent encountered selectivity issues, the group inserted the proteomimetic compound in place of helix 3 into the HIF-1 α peptide sequence creating a hybrid structure labelled a “bionic protein”²³⁹ which showed a significantly improved selectivity towards HIF-1 α /p300 over *hDM2*/p53 (Figure 43b). The use of non peptidic, topographical helix mimics was also employed by the Arora group using oxopiperazine helix mimetic (OHM) scaffolds targeting the helix 3 binding site on p300^{97,226} (Figure 43c). Through the use of molecular modelling the group managed to design and test several decorated scaffolds that validated the importance of the key residues Leu₈₁₈, Leu₈₂₂ and Gln₈₂₄, through alanine OHM variants (Figure 43c). NMR data confirmed the binding of the mimetics to the helix 3 region on p300 and luciferase-based reporter assay confirmed reduced HIF-1 α promoter activity. Lastly, the

ability of the best binding OHM to reduce tumour growth in mouse xenograft models was assessed and deemed successful.

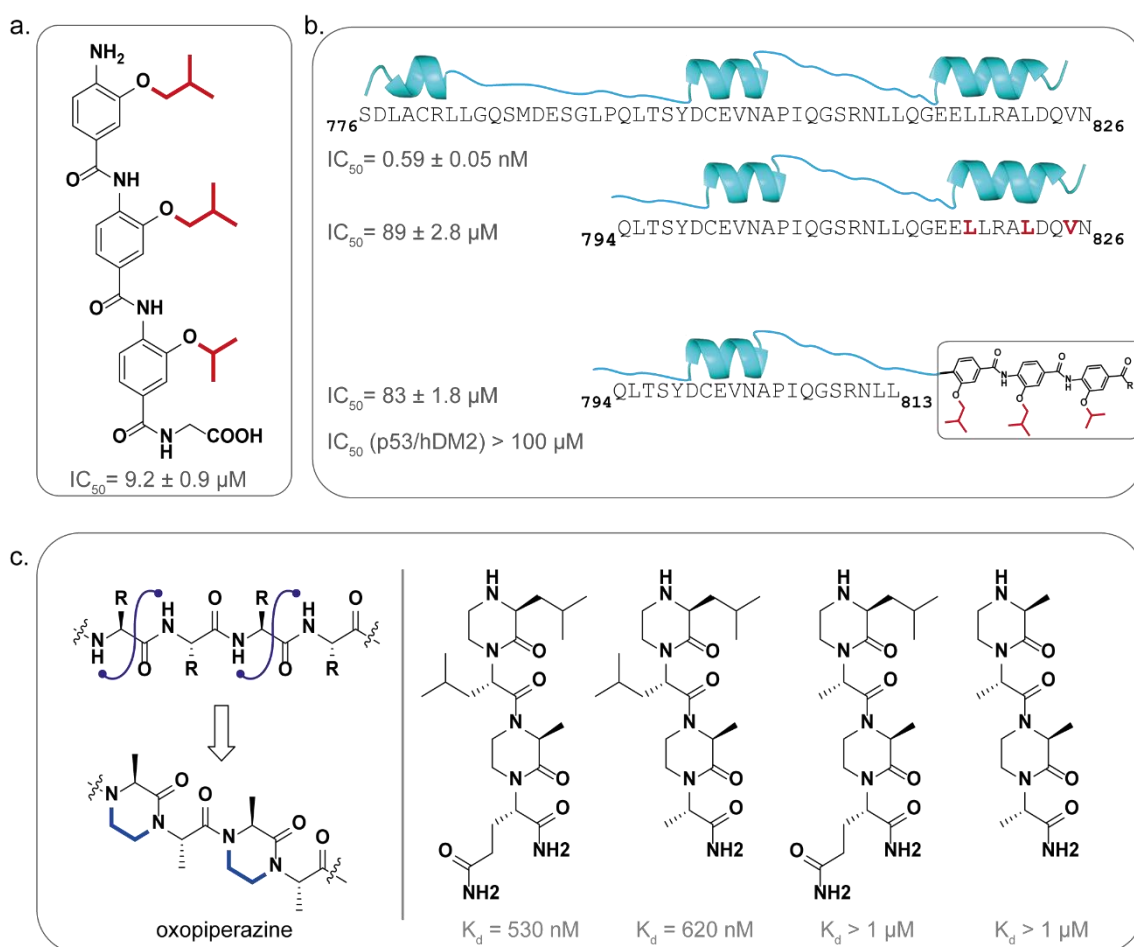


Figure 43. Peptidomimetic approaches targeting the HIF-1/p300 PPI: a) The best oligoamide structure targeting the helix 3 interface on p300; b) Comparison of binding affinities for the full length HIF-1 α ₇₇₆₋₈₂₆, HIF-1 α ₇₉₄₋₈₂₆ and the bionic protein; c) The general structure of oxopiperazines designed from a linear peptide are displayed on the left panel, with blue squiggly lines denoting where the linkage is present; A list of the OHM structures designed and tested by the Arora group on the right hand side of the figure.

4.2 Objectives

Previously, the inhibitory potencies of different lengths of HIF-1 α were tested to verify the whereabouts of the most productive region to target using small molecules⁵⁸. The study confirmed that the HIF-1 α helix 3 region had the highest affinity for p300.

The objective of this work was to expand the application of the dibromomaleimide stapling technique to the HIF-1 α /p300 PPI. The aim was to be completed through the design, synthesis and testing of stapled peptide variants of the HIF-1 α helix 3 sequence (Figure 44).

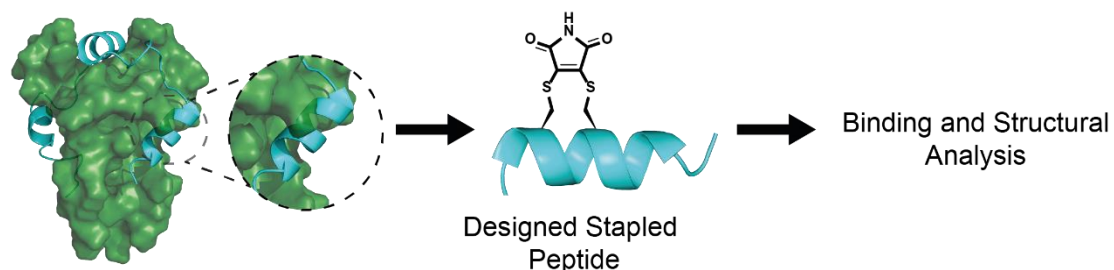


Figure 44. Project aim: design and test stapled peptides based on helix 3 of HIF-1 α for binding to p300.

Furthermore, targeting the helix 3 binding groove on p300 using constrained peptides derived from HIF-1 α could additionally establish this location as amenable to inhibitor modulation. Stapled peptides, (see: introduction, section 1.5.1), are advantageous for targeting PPIs. Through conformationally constraining the peptide, higher binding affinities are generally observed due to the lowered entropic cost needed for binding for a preorganized peptide in a supposedly bioactive conformation^{99,100,104}. Dibromomaleimide stapling was developed as an effective technique for peptide stapling, having this technique in house allowed for further testing of its robustness and efficiency.

4.3 Results and Discussion

4.3.1 Design and Synthesis of HIF-1 α Peptides

As mentioned, truncation of helix 1 did not affect HIF-1 α binding to p300 significantly thus FITC-Ahx-HIF-1 α ₇₈₆₋₈₂₆ was used as the tracer molecule; a molecule with fluorescein as the fluorophore attached to it (Figure 45a). Although the group previously worked with a segment of helix 3 (HIF-1 α ₈₁₆₋₈₂₆) a slightly longer sequence HIF-1 α ₈₁₂₋₈₂₆ was chosen for this work to accommodate the incorporation of staple features at two different locations (Figure 45a).

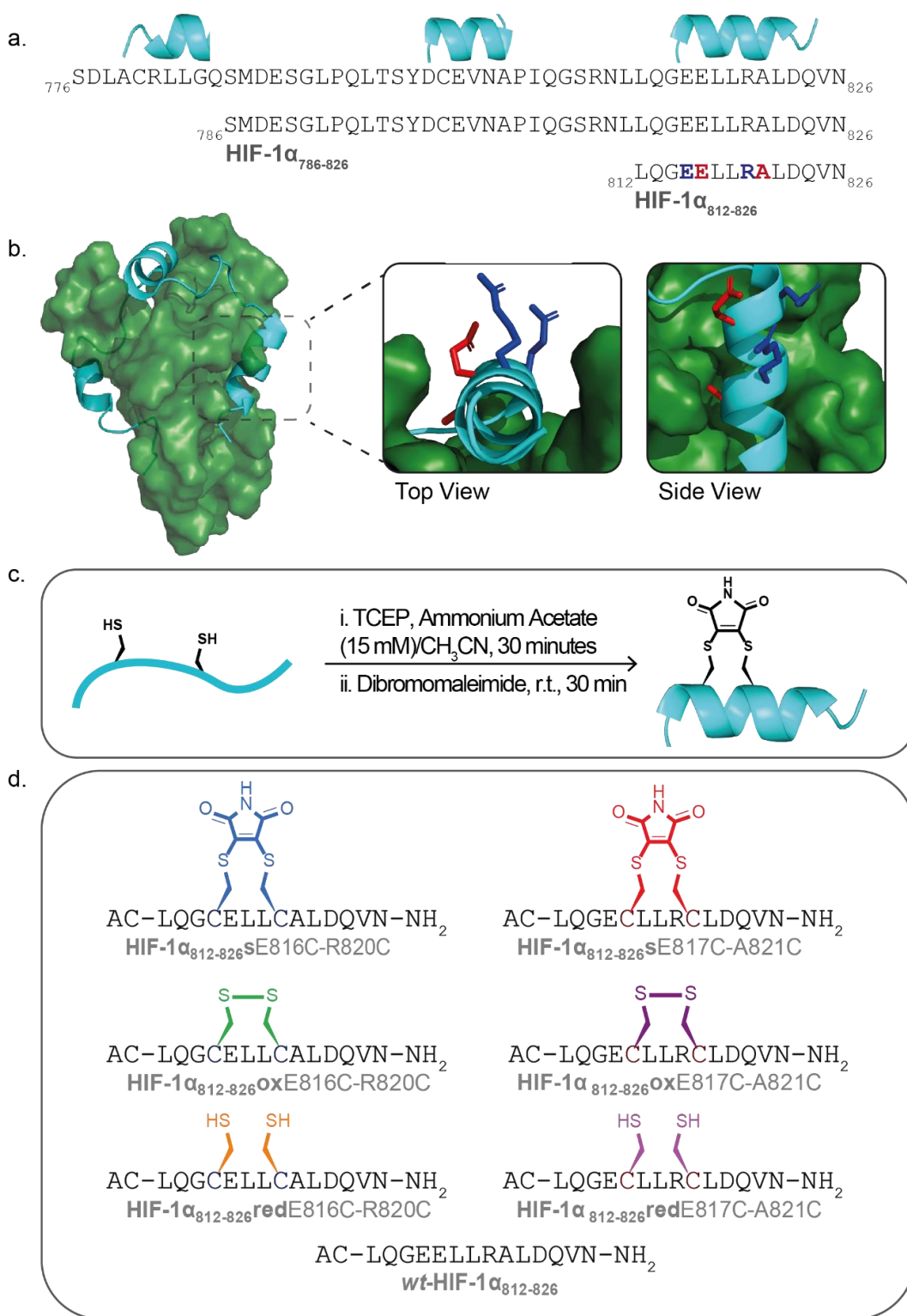


Figure 45. Peptide design and synthesis based on helix 3 of HIF-1α: a) The two main lengths of HIF-1α that were used in this work; b) Residues chosen to be replaced by cysteines and stapled following a methodology previously reported; c) Stapling reaction which was improved upon by lowering the overall reaction time from 16 hours to 30 minutes; d) All variant peptides synthesised and tested for binding to p300.

The two stapling locations, at i , $i+4$ locations, were chosen to probe for best staple placement. Upon visual inspection of the NMR structure, E816 and R820 (blue residues) project out into the solvent while E817 and A821 (red residues) project nearer the binding cleft of the peptide (Figure 45b). The highlighted residues were replaced by cysteine during solid phase peptide synthesis (SPPS). The peptides were reduced, oxidised and stapled (Figure 45c). Interestingly, this stapling methodology showed a faster stapling time compared with the 16 hours previously reported⁶. Following the stapling reaction the peptides were purified for a final time. The other reservoirs were used to obtain fully oxidised and reduced variants for comparison in binding assays. A total of 6 cysteine containing peptides was synthesised: HIF-1 $\alpha_{812-826}$ **s**E817C-A821C (stapled), HIF-1 $\alpha_{812-826}$ **ox**E816C-R820C (oxidised – disulphide bridge formed) and HIF-1 $\alpha_{812-826}$ **red**E816C-R820C (reduced – thiols not engaged in disulphide bond) as well as the other series HIF-1 $\alpha_{812-826}$ **s**E816C-R820C, HIF-1 $\alpha_{812-826}$ **ox**E816C-R820C, HIF-1 $\alpha_{812-826}$ **red**E816C-R820C (Figure 45d). Lastly, the wild-type HIF-1 $\alpha_{812-826}$ acetylated and fluorescently labelled peptides were also synthesised. With all 8 peptides in hand, binding assays were undertaken.

4.3.2 Binding Assays of HIF-1 α Peptides

The effect of stapling of HIF-1 $\alpha_{812-826}$ and the ability to inhibit the HIF-1 α /p300 interaction was to be assessed using the previously described FA technique. A constant concentration of tracer (25 nM) was added to titrated p300. Following a direct titration assay a $K_d = 10 \pm 4$ nM was measured for FITC-Ahx-HIF-1 $\alpha_{786-826}$ (Figure 46a). Next, the competition mode of the assay was established where non-labelled HIF-1 $\alpha_{812-826}$ was titrated and used to displace the FITC-Ahx-HIF-1 $\alpha_{786-826}$ from p300. An inhibitory potency $IC_{50} = 58 \pm 3$ nM was obtained for the non-labelled HIF-1 $\alpha_{812-826}$ (Figure 46b).

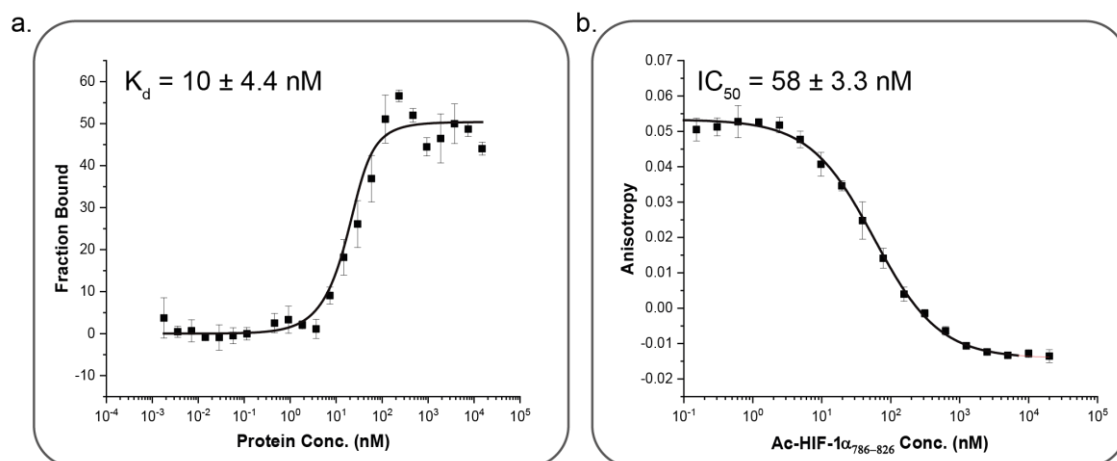


Figure 46. FA data for HIF-1 $\alpha_{786-826}$: a) Direct titration assay of 25 nM FITC-Ahx-HIF-1 $\alpha_{786-826}$ (tracer) in 20 mM Tris, 100 mM NaCl, 0.1 mM DTT, pH 7.46; b) Competition assay using the acetylated HIF-1 $\alpha_{786-826}$ as competitor using 50 nM tracer, 100 nM p300 in 20 mM Tris, 100 mM NaCl, 0.1 mM DTT, pH 7.46.

4.3.2.1 Competition Assays

The HIF-1 $\alpha_{812-826}$ E817C-A821C peptide series was tested in competition mode. With a starting peptide concentration of 250 μ M, the HIF-1 $\alpha_{812-826}$ S817C-A821C, HIF-1 $\alpha_{812-826}$ O816C-R820C peptides were not observed to displace the HIF-1 $\alpha_{786-826}$ tracer in the competition assays (Figure 47a). Based on the NMR structure of the HIF-1 α /p300, the stapling position has the potential to introduce a steric clash with p300. As such, it was unsurprising to see lack of inhibition from this series of peptides. Strangely, HIF-1 $\alpha_{812-826}$ R817C-A821C seemed to show the start of a steady displacement of the tracer at high competitor concentrations

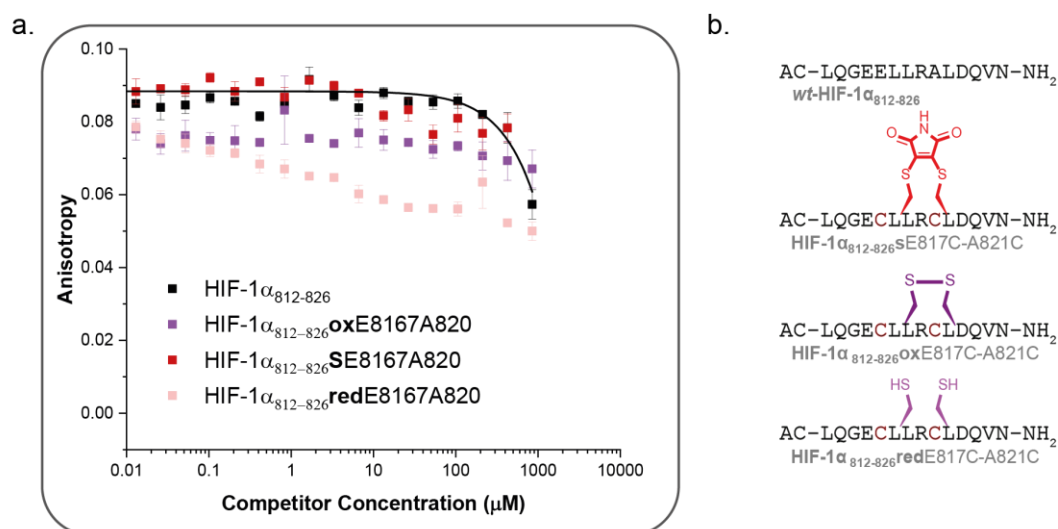


Figure 47. FA competition data for the HIF-1 $\alpha_{812-826}$ E817C-A821C peptide series: a) Competition assay for the peptide variants (25 nM) FITC-Ahx-HIF-1 α 776-826, (100 nM) p300 in 20 mM Tris, 100 mM NaCl, 0.1 mM DTT, pH 7.46; b) Peptide sequences.

The HIF-1 $\alpha_{812-826}$ E816C-R820C peptide series generated more interesting results in the competition assays. Based on the NMR structure, this stapling location was oriented towards the solvent and thus no steric clashes between the staple and the binding interface were anticipated. At the maximal inhibitor concentration of 250 μM , the wt-HIF-1 $\alpha_{812-826}$ (black curve) peptide started to display inhibitory activity (Figure 48a). The HIF-1 $\alpha_{812-826}$ oxE816C-R820C peptide showed similar inhibitory behaviour to the wt peptide with a weak inhibitory potency $\text{IC}_{50} \gg 500 \mu\text{M}$. It became clear that in order to reach the bottom plateau of the binding curves for both peptides the maximal concentration would have to be significantly increased. In contrast, HIF-1 $\alpha_{812-826}$ sE816C-R820C peptide showed significant improvement in inhibitory potency with $\text{IC}_{50} = 30.3 \pm 4.9 \mu\text{M}$. Unexpectedly HIF-1 $\alpha_{812-826}$ redE816C-R820C also displayed an inhibitory potency similar to the stapled peptide with an $\text{IC}_{50} = 9.9 \pm 0.6 \mu\text{M}$ (Figure 48 orange trace). The shallow inhibitory curve for HIF-1 $\alpha_{812-826}$ redE816C-R820C may be due to a more complex or non-specific mode of interaction, one that will be discussed below. After the assays were read, aliquots were removed and analysed by HRMS. The analysis proved that both HIF-1 $\alpha_{812-826}$ redE816C-R820C and HIF-1 $\alpha_{812-826}$ oxE816C-R820C were still in their reduced and oxidised states respectively, after the 45 minute assay window (Figure 49). As the redox state of the peptides was not an issue, structural information for the peptides in

solution was pursued in order to explain the behaviour of HIF-1 $\alpha_{812-826}$ redE816C-R820C.

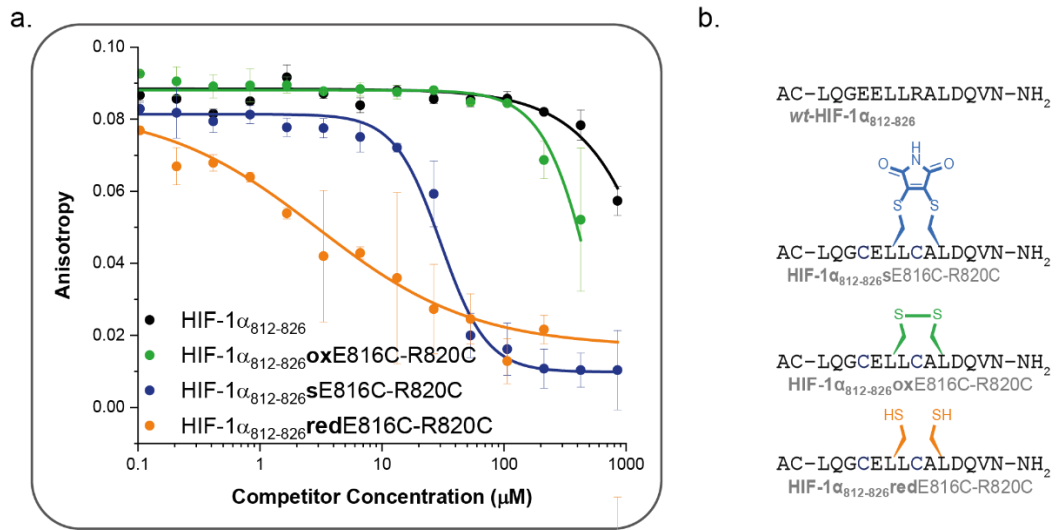


Figure 48. FA data for the HIF-1 $\alpha_{812-826}$ E816C-R820C peptide series: a) Competition assay for the peptide variants using (25 nM) FITC-Ahx-HIF-1 $\alpha_{776-826}$, (100 nM) p300, in 20 mM Tris, 100 mM NaCl, 0.1 mM DTT, pH 7.46; b) Peptide sequences.

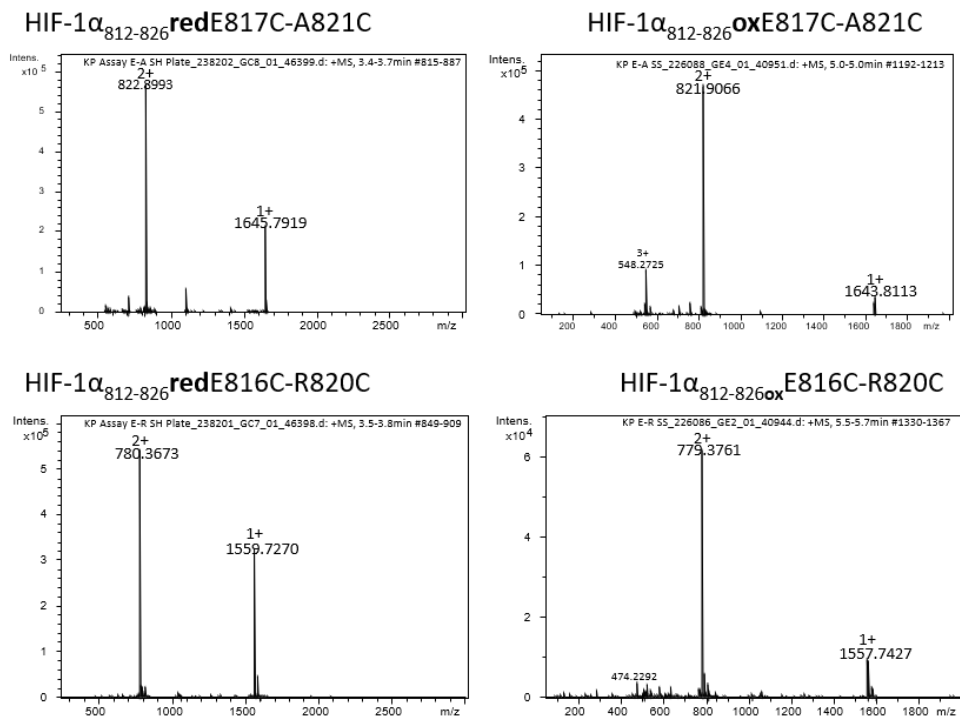


Figure 49. HRMS data of the assay wells following the 45 minute incubation period to ensure peptides are still in their oxidised and reduces states: top two panels show HIF-1 $\alpha_{812-826}$ E817C-A821C series with the reduced peptide mass showing the two extra hydrogen atoms and oxidised peptide; similarly the bottom two panels show the HIF-1 $\alpha_{812-826}$ E816C-R820C reduced and oxidised peptides maintaining their respective states.

4.3.3 Structural Analysis of HIF-1 α Peptides: Investigation into Structure to Explain the Observed Binding Behaviours

4.3.3.1 Circular Dichroism of HIF-1 α Peptides

Often the insertion of a judiciously placed constraint has been shown to enhance target binding affinity through pre-organisation, due to a reduced entropic cost of binding¹⁰⁰. The increased inhibitory potency of HIF-1 α ₈₁₂₋₈₂₆**sE816C-R820C** could be due to a higher degree of helicity arising due to pre-organisation. Thus, all the peptides were analysed by Circular Dichroism (CD) spectroscopy at room temperature to assess the extent to which the inhibitory potency correlated with helicity content of each peptide. Further insight into HIF-1 α ₈₁₂₋₈₂₆**redE816C-R820C** structural behaviour was also expected from this experiment.

Surprisingly, the CD spectra showed very little structural difference between peptide variants in solution with all adopting a predominantly random coil conformation and only HIF-1 α ₈₁₂₋₈₂₆**sE816C-R820C** ($[\theta]_{\text{MRE-222}}$ -6658, 20% helicity) showing slightly higher helicity than the other variants (wt-HIF-1 α ₈₁₂₋₈₂₆; $[\theta]_{\text{MRE-222}}$ -3954, 12 % helicity, HIF-1 α ₈₁₂₋₈₂₆**oxE816C-R820C**; $[\theta]_{\text{MRE-222}}$ -3766, 11% helicity, HIF-1 α ₈₁₂₋₈₂₆**redE816C-R820C**; $[\theta]_{\text{MRE-222}}$ -3375, 10 % helicity) (Figure 50, Table 6).). The difference in helicity between HIF-1 α ₈₁₂₋₈₂₆**sE816C-R820C** and the other variants was not a satisfactory explanation for the significant inhibitory potency observed. Furthermore, no correlation between structure and binding for HIF-1 α **redE816C-R820C** was observed. Thus, thermal unfolding behaviour was observed next.

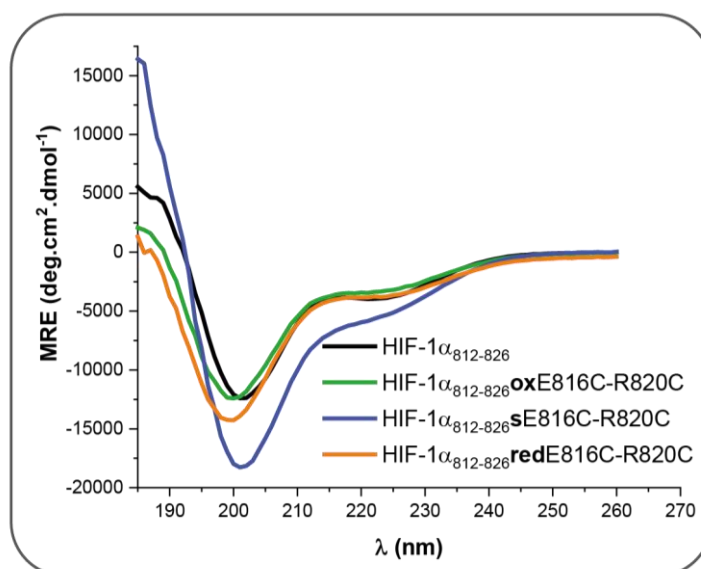


Figure 50. CD spectra for wt-HIF-1 α ₈₁₂₋₈₂₆, HIF-1 α ₈₁₂₋₈₂₆oxE816C-R820C, HIF-1 α ₈₁₂₋₈₂₆sE816C-R820C and HIF-1 α ₈₁₂₋₈₂₆redE816C-R820C at 250 μ M in 20 mM sodium phosphate, pH 7.55.

Table 6. Summary of inhibitory potencies and helicity of peptide variants.

Peptide	IC ₅₀	Helicity %
HIF-1 α ₈₁₂₋₈₂₆	> 500 μ M	12
HIF-1 α ₈₁₂₋₈₂₆ oxE816C-R820C	> 500 μ M	11
HIF-1 α ₈₁₂₋₈₂₆ sE816C-R820C	30 \pm 4.9 μ M	20
HIF-1 α ₈₁₂₋₈₂₆ redE816C-R820C	9.9 \pm 0.6 μ M	10

4.3.3.2 Thermal CD Analysis of HIF-1 α E816R820 Series of Peptides with and without p300

Thermal unfolding CD experiments of the peptides at 250 μ M concentrations in sodium phosphate buffer were run from 20°C to 90°C. The structural response in relation to the increasing temperature of the peptides in isolation and in complex with p300 were obtained.

In isolation the wt-HIF-1 α ₈₁₂₋₈₂₆ exhibited little variation in structure with varying temperature implying that the peptide is mainly unstructured to begin with (Figure 51a). In contrast, a more significant response observed from the other peptides

with HIF-1 $\alpha_{812-826}$ sE816C-R820C displaying the greatest variation in structure with increasing temperature (Figure 51b-d).

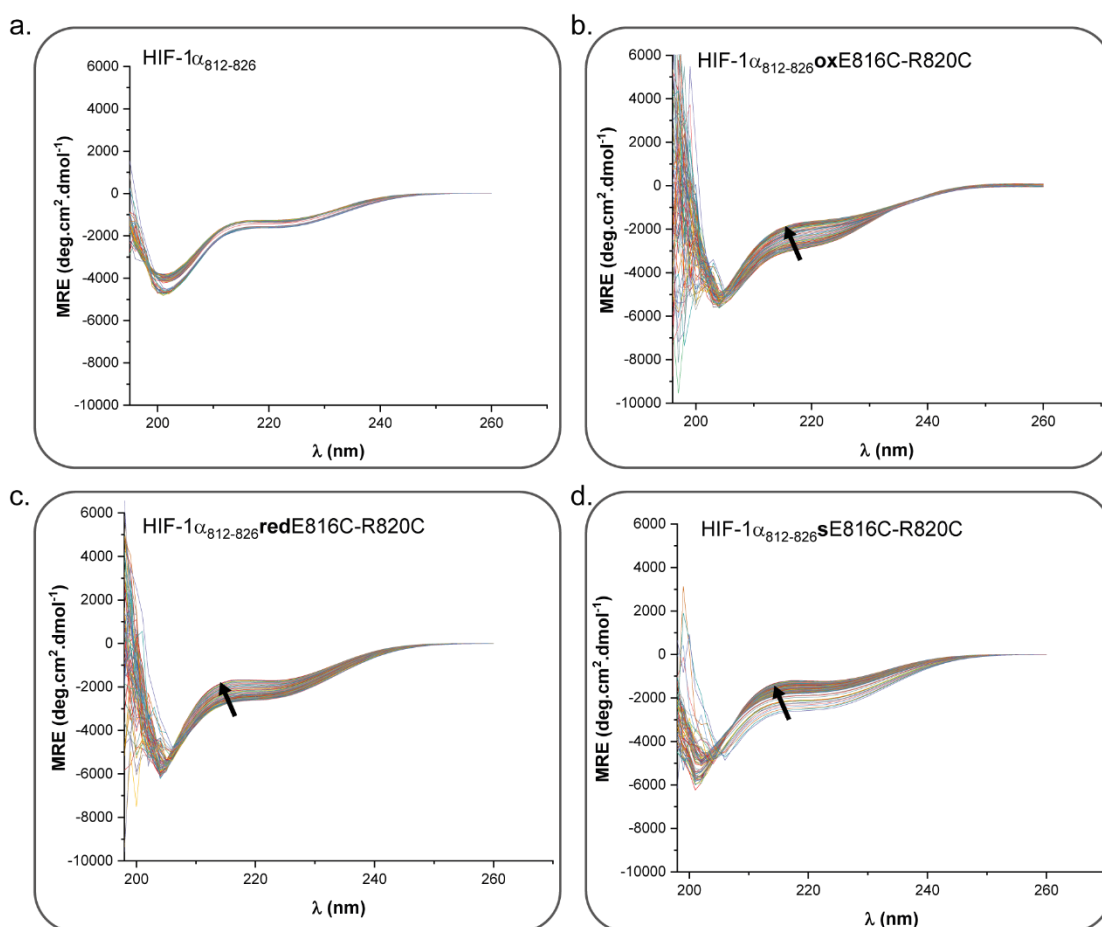


Figure 51. CD thermal melting experiments from 25 to 90°C data, labelled with a black arrow in the direction of unfolding with increasing temperature, of peptides at 250 M, in 20 MM sodium phosphate, pH 7.55: a) Thermal unfolding for wt-HIF-1 $\alpha_{812-826}$; b) Thermal unfolding for HIF-1 $\alpha_{812-826}$ oxE816C-R820C; c) Thermal unfolding for HIF-1 $\alpha_{812-826}$ redE816C-R820C; d) Thermal unfolding HIF-1 $\alpha_{812-826}$ sE816C-R820C.

In an attempt to explain the binding behaviour observed previously for HIF-1 $\alpha_{812-826}$ sE816C-R820C and HIF-1 $\alpha_{812-826}$ redE816C-R820C, thermal analysis of HIF-1 $\alpha_{812-826}$ sE816C-R820C and HIF-1 $\alpha_{812-826}$ redE816C-R820C was conducted in the presence of p300 recorded at a 1:1 ratio of peptide to protein from 20-90°C. For comparison, the p300 CD spectra by itself was also recorded. The thermal denaturation of p300 in isolation, shows gradual unfolding with increasing temperature as expected (Figure 52a). Interestingly, this gradual unfolding pattern exhibits a more significant signal change when p300 and HIF-1 $\alpha_{812-826}$ sE816C-R820C are present in the sample (Figure 52b). In stark contrast, loss

in signal was observed for HIF-1 $\alpha_{812-826}$ redE816C-R820C with p300, indicating there was less structural signature to lose in the first place during denaturation (Figure 52c).

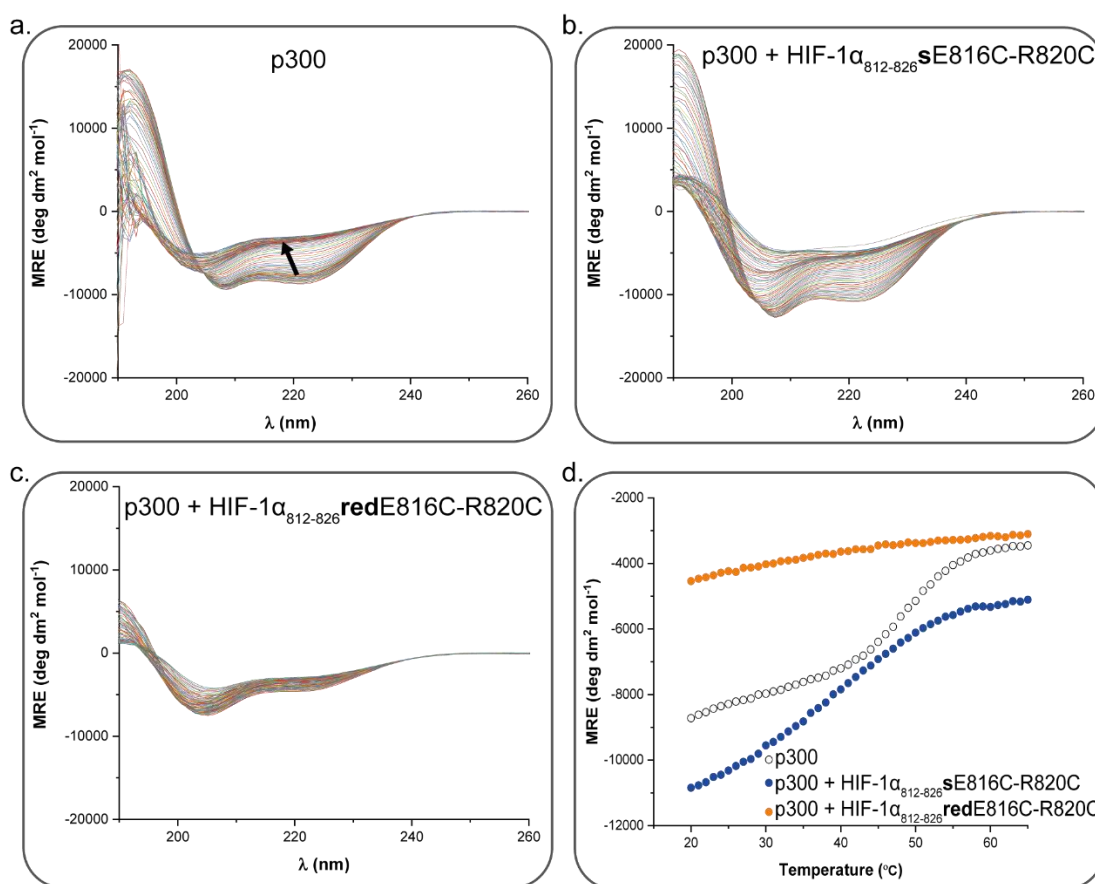


Figure 52. Thermal denaturation experiments in 20 mM sodium phosphate, pH 7.55: a) Thermal CD of p300 at 20 μ M concentration showing a smooth and consistent transition between folding and unfolding CD traces from 20-90°C; b) Thermal CD of p300/HIF-1 $\alpha_{812-826}$ sE816C-R820C, 1:1 ratio at 20 μ M concentration exhibiting a moderate increase in thermal stability; c) Thermal CD of p300/HIF-1 $\alpha_{812-826}$ redE816C-R820C, 1:1 ratio at 20 μ M where a lack of structural folding is observed over the temperature range; d) CD thermal denaturation curve showing structural effects of HIF-1 $\alpha_{812-826}$ sE816C-R820C and HIF-1 $\alpha_{812-826}$ redE816C-R820C on p300: the hollow points show p300 denaturation curve, with defined folded and unfolded states; the orange trace of p300 in a 1:1 ratio with HIF-1 $\alpha_{812-826}$ redE816C-R820C, showing severe loss of structure indicated by the lack of transition; lastly, the blue trace of p300 in a 1:1 ratio with HIF-1 $\alpha_{812-826}$ sE816C-R820C with a reduced T_m but with transition between folding and unfolding still evident.

The spectral change with temperature observed above was quantified by plotting the spectral change for one wavelength against temperature. Thus thermal ramping data at 222 nm for p300, p300/HIF-1 $\alpha_{812-826}$ sE816CR820C and p300/HIF-1 $\alpha_{812-826}$ redE816C-R820C was plotted against the temperature (Figure

52d). The data showed a sigmoidal curve for p300 with a $T_m = 47.3^\circ$. The HIF-1 $\alpha_{812-826}$ sE816C-R820C/p300 complex had a larger mean residue ellipticity at 222 nm, a broader unfolding profile, however a $T_m = 37.3^\circ\text{C}$. The reduced T_m could be due to a change in p300 due to stapled peptide binding, or the effects of the stapled peptide unfolding. The effect of HIF-1 $\alpha_{812-826}$ redE816C-R820C was observed next. The markedly reduced MRE at 222 nm and lack of definition between folded and unfolded states was interpreted as a loss in structural integrity of p300 after its interaction with HIF-1 $\alpha_{812-826}$ redE816C-R820C. Based on the thermal denaturation data, it was hypothesised that the free thiols of HIF-1 $\alpha_{812-826}$ redE816C-R820C were destabilising p300 in some way, possibly zinc ejection. The shallow inhibitory FA curve for this peptide, together with the thermal denaturation data was consistent with non-specific binding behaviour. Chetomin, a previously discovered inhibitor of the HIF-1 α /p300 PPI was shown to act as a non-specific inhibitor though zinc ejection²³³. Since then, other reagents were shown to elicit similar zinc ejection behaviour²³⁴. As such the HIF-1 $\alpha_{812-826}$ redE816C-R820C peptide was eliminated from our enquiries and not pursued in further studies.

4.3.3.3 Molecular Dynamics Analysis of HIF-1 α Peptides: Investigation of HIF-1 α stapE816CR820C Enhanced Binding Affinity

Molecular Dynamics (MD) simulations have been developed as a relatively fast method for obtaining structural information on proteins and their interaction with various ligands²⁴⁰⁻²⁴². MD acquisition of protein folding patterns responsible for the three-dimensional (3D) structure aids the understanding of protein function/structure relationship, while MD simulations of protein/ligand complexes offers more than a static picture of the interaction for deeper understanding of binding mechanism^{240,243,244}. Unsurprisingly, MD simulations have also been applied successfully to shorter constructs such as peptide-based chemical tools^{241,245,246}. Typically, peptide conformation correlates with its function and accurate prediction of structural and conformational flexibility would help progress in peptide-based drug design²⁴⁰. For example, anticancer peptides that were entirely designed computationally, were tested and found to bind PPI targets²⁴⁵.

The computational design improved the chemical properties of the peptides, such as their computationally generated lipophilicity^{243,247,248}. However, it is worth noting that the accuracy and reliability of MD simulations depend on several factors such as the treatment of water molecules, timescale and conformational sampling^{240,241}.

It was decided to complement the CD experiments with further analysis to understand the enhanced potency of HIF-1 $\alpha_{812-826}$ sE816C-R820C to p300. Thus, CD spectra as well as CD thermal denaturation spectra were acquired for both isolated peptides and their p300 complexes. The CD experiment showed the conformation of the peptides in isolation as largely unstructured. The thermal analysis of the HIF-1 $\alpha_{812-826}$ sE816CR820C/p300 complex confirmed a binding event was occurring because a change in T_m and overall shape of the unfolding curve was observed. Corroboration of a binding event, and thus insight on the consequences of introducing the constraint, was sought through MD simulations of the peptides in solution and in complex with p300 over a 100 ns timeframe.

The MD simulations of the peptides were performed by our collaborator in Bristol, Dr Richard Sessions and percentage helicity of each peptide over the last 50 ns timeframe were generated (Table 7). These simulations indicated that the wt-HIF-1 $\alpha_{812-826}$ had a significant helical conformation of 43% with a moderate increase to 61% when in complex with p300 (Figure 53a). The MD simulation of wt-HIF-1 $\alpha_{812-826}$ in isolation was in disagreement with the largely unstructured behaviour observed in the CD. However, it should be noted that MD simulations provide qualitative data of peptide conformational changes. The HIF-1 $\alpha_{812-826}$ oxE816C-R820C peptide MD simulations showed a helicity of 21% in solution and a moderate increase to 37% when in complex with p300 (Figure 53b). Finally, MD simulations of HIF-1 $\alpha_{812-826}$ sE816CR820C generated two different conformers of the peptide when in complex with p300. Over the course of the simulation the maleimide bridge was oriented in opposing directions with respect to the helix (Figure 53d). Both conformers showed low helicity of 24% for conformer 1 and 29% for conformer 2 with the helical region mainly around the staple location (Figure 53c). This data was in agreement with the CD data for HIF-1 $\alpha_{812-826}$ sE816C-R820C in solution. In complex with p300 a more dramatic shift in helicity was observed. The helical character for HIF-1 $\alpha_{812-826}$ sE816C-R820C was

propagated over a greater number of residues over a greater timeframe of the simulation when compared to the unbound state. As well, the simulations suggested that the binding affinity of the stapled peptide was not due to direct interactions of the staple with p300 – no additional non-covalent interactions between peptide staple and protein were observed during the simulation – which has been the case in other designed stapled peptide systems^{109,249,250}. The MD simulations seemed to suggest that higher binding affinity of HIF-1 $\alpha_{812-826}$ sE816C-R820C was achieved due to the ability of the staple to accommodate a conformation when bound to p300 as evidenced by the increased helicity of the peptide when bound. Validation of the MD observations through experimental investigation was sought.

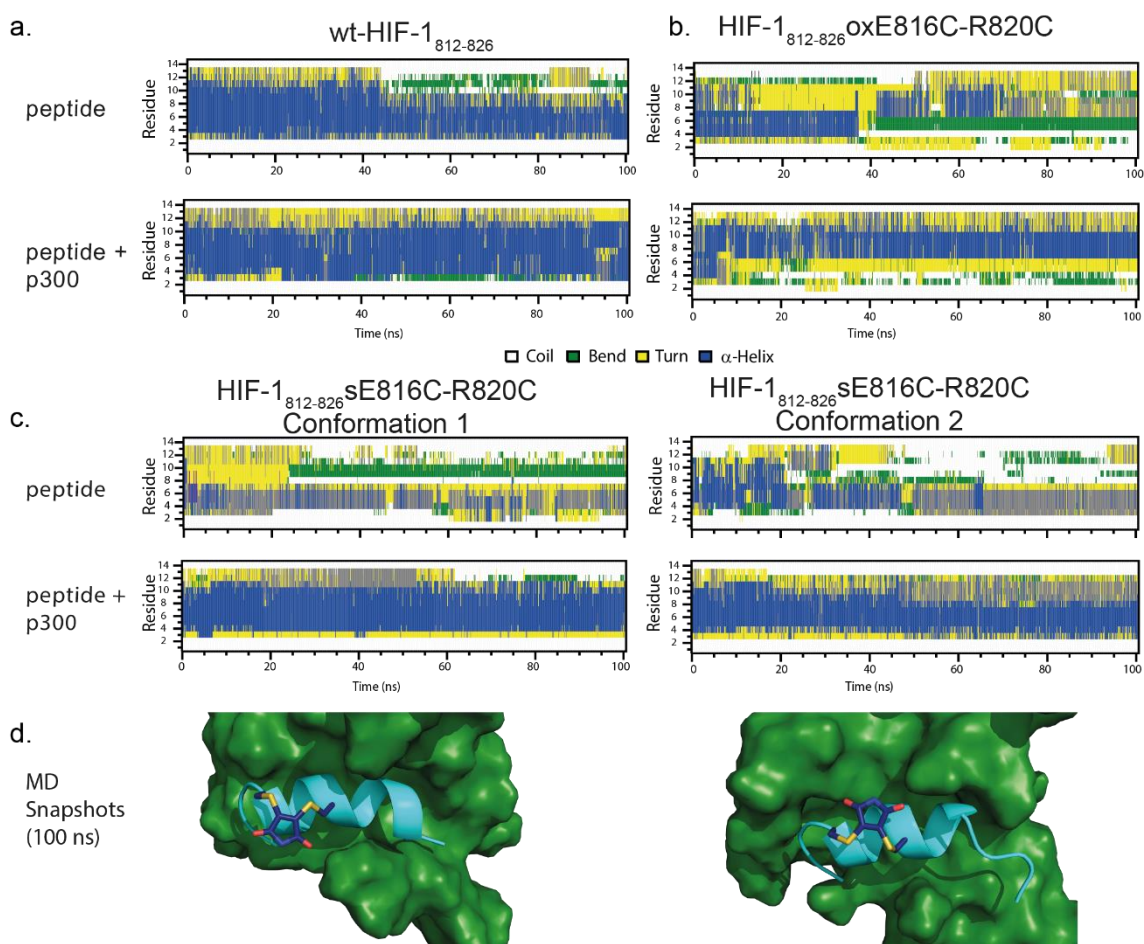


Figure 53. Conformational analyses of HIF-1 $\alpha_{812-826}$ E816C-R820C: a) 100 ns MD simulation of wt- HIF-1 $\alpha_{812-826}$ peptide in the absence (top left panel) and presence (bottom left panel) of p300; b) 100 ns MD simulation of HIF-1 $\alpha_{812-826}$ oxE816C-R820C peptide in the absence (top right panel) and presence (bottom right panel) of p300; c) MD simulations of HIF-1 $\alpha_{812-826}$ sE816C-R820C peptide in isolation and in complex with p300 in the two dominant conformations adopted; d) MD snapshots of two stable p300 bound conformations of HIF-1 $\alpha_{812-826}$ oxE816C-R820C.

Table 7. Percentage Helicity based on MD simulations over the first 50 ns timeframe for each peptide.

Peptide	Helicity % Isolation	Helicity % + p300
HIF-1 $\alpha_{812-826}$	43	61
HIF-1 $\alpha_{812-826}$ oxE816C-R820C	21	37
HIF-1 $\alpha_{812-826}$ sE816C-R820C conf.1	24	56
HIF-1 $\alpha_{812-826}$ sE816C-R820C conf. 2	29	54

4.3.3.4 Difference CD of Constrained HIF-1 α Peptides

To support the MD simulations, design of an alternative CD experiment was envisioned to obtain CD difference spectra (Figure 54a). It was envisioned that the CD difference data could help uncover the overall conformational effects of p300 on the peptides; by subtracting the additive CD traces from the CD of the mixture where the peptide and protein would experience physical interaction.

This difference was plotted as mean residue ellipticity versus wavelength (Figure 54b). The experiment was performed on wt-HIF-1 $\alpha_{812-826}$, HIF-1 $\alpha_{812-826}$ oxE816C-R820C and HIF-1 $\alpha_{812-826}$ sE816C-R820C at a concentration of 40 μ M, close to the IC₅₀ value for the stapled HIF-1 $\alpha_{812-826}$ sE816C-R820C peptide. Firstly, the outcome of the experiment confirms direct interaction between stapled peptide and protein target. Secondly, for HIF-1 $\alpha_{812-826}$ sE816C-R820C, the difference spectrum is consistent with an idealized α -helical conformation with minima at 208 and 222 nm²⁵¹. The MD simulations do not suggest the constraint makes significant non-covalent contacts with p300. As well, conformational changes in p300 that increase helicity in p300 (relative to the largely helical apo p300) are unlikely to lead to such a significant difference in mean residue ellipticity. Therefore the significant increase in MRE in the difference spectra are consistent with a significant increase in helical stability of the peptide ligand in the bound state supporting the conclusions of the MD simulation. The difference spectrum for the wt-HIF-1 $\alpha_{812-826}$ is weaker, consistent with the MD simulations that imply minimal change in helicity upon binding. Lastly, the difference spectrum for HIF-1 $\alpha_{812-826}$ oxE816C-R820C, also indicated an increase in helicity, however the minima at ~ 204 nm present for the unbound peptide persisted, suggesting that

this peptide is unable to adopt an idealized and fully helical conformation in the bound state. Again, these data are consistent with the conclusions of the MD analyses.

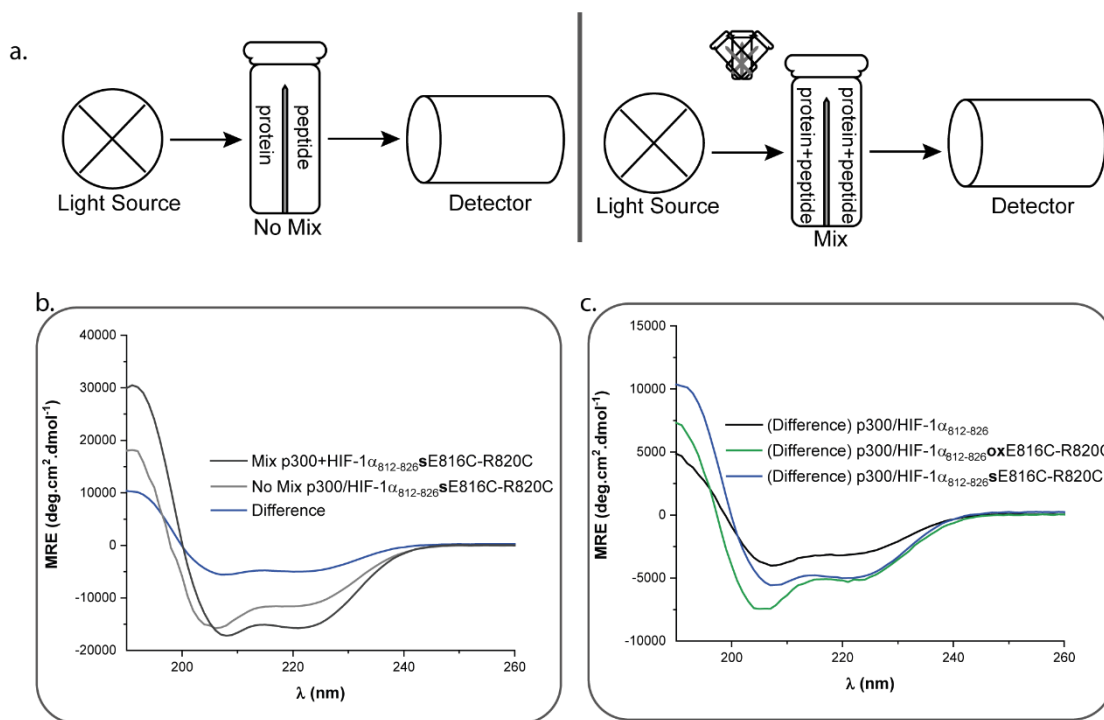


Figure 54. CD analysis: a) Cartoon representation of the method for acquiring difference CD data; b) CD data for the additive (grey), combined (black) and difference spectra (blue) for the HIF-1 α ₈₁₂₋₈₂₆sE816C-R820C variant peptide in the presence of p300; c) Difference CD: spectra for binding of wt-HIF-1 α ₈₁₂₋₈₂₆, (black) stapled peptide HIF-1 α ₈₁₂₋₈₂₆sE816C-R820C (blue) and HIF-1 α ₈₁₂₋₈₂₆oxE816C-R820C (green) to p300.

4.4 Conclusions

The HIF-1 α /p300 PPI remains an attractive target in oncology with little progress of inhibitor design. Through peptides derived from HIF-1 α , a maleimide stapled peptide inhibitor of this PPI was developed. The improved inhibitory potency of the peptide led to an investigation into the relationship between binding affinity and structural conformation. Although the stapled peptide showed little structural integrity in solution, MD simulations pointed towards a more stable helical conformation adopted when bound to p300. This behaviour was corroborated through, CD difference spectra where the impact of p300 on the peptide conformation supported the increase in helicity upon binding. This study may have implications for stapled peptide design in other systems as the effect of

stapling on peptide conformation both in solution and in the presence of the protein target may need to be assessed. Conformational pre-organisation should be considered a balancing act to maintain a degree of peptide plasticity; to avoid constraining the peptide in an unfavourable state or preventing key residues from forming important binding interactions. The results of this line of enquiry suggest that binding affinity could be enhanced by simply stabilizing the bound conformation of peptides as opposed to their unbound one. Further research should place more onus on understanding the relationship between thermal stability of the peptide/protein complex, binding affinity as well as more corroborating evidence for the importance of the bound conformational stability of stapled peptides. Ultimately, this approach and application should be refined and re-applied for the introduction of constraints in peptides. Finally, further efforts should be placed into targeting the HIF-1 α /p300 PPI using peptidomimetic approaches.

Thesis Summary and Future Work

This thesis began with the introduction of PPIs and the strategies used to target these for therapeutic intervention. Although inhibitor campaigns have resulted in successful PPI modulation, PPIs remain challenging targets due to the relatively large and flat surface area over which they occur. However, not all PPIs are equally as difficult to target. The motivation for this thesis was to target secondary structure mediated PPIs, specifically α -helix and β -strand mediated PPIs, using SBD peptide-based inhibitors. A total of four PPI systems were chosen and studied - GKAP/SHANK1 linked to neurodegenerative diseases; NOXA B/MCL-1 relevant in oncology; SIM/SUMO relevant in a number of response pathways including inflammation and oncology; HIF-1/p300 relevant in oncology.

Chapter 2 focused on the design and synthesis of a peptide-based inhibitor targeting the GKAP/SHANK1 β -strand mediated PPI. In order to design the inhibitor further study of the PPI was necessary. In summary, the successful experimental alanine scan of the GKAP ligand identified the hotspot residues (T4, L6 and carboxylate). Comparison of the experimental results with the BUDE computational predictions highlighted issues with the computational tool as this failed to predict T4 as a hotspot. However, the tool successfully predicted the remaining peptides in the GKAP sequence as either hotspots or non-hotspots. Ultimately, these types of comparisons should continue in order to aid the improvement of computational tools. Lastly, the design, synthesis and testing of select GKAP peptide sequences validated the virtually non-existent involvement of the A2 and R5 positions in binding. Additionally T4 involvement in target specificity was implied when loss of binding was observed for the T4S variant peptide. Based on the knowledge acquired following the alanine scan and testing of the select peptide sequences, two peptide-based inhibitor versions were designed using the Nowick Hao building block, acyclic and cyclic. The successfully synthesised acyclic inhibitor showed binding comparable to the wild-type GKAP sequence while the cyclic inhibitor posed significant synthetic challenges and was not obtained. Future efforts should be placed in completing the synthesis of a peptide-based macrocycle utilising the Hao building block. The following experiments would be a natural progression for this work:

- Finish Hao synthesis (Scheme 8), and if obtained attach this building block **24** to the peptide chain still on resin to afford the peptide-based inhibitor **25** and attempt deprotection and cyclisation before cleaving the inhibitor chain off the resin.
- If obtained, **25** should be tested in FA and ITC to study the macrocycle against the wild-type-GKAP peptide in the PPI system and obtain biophysical data and assess the binding affinity.
- NMR studies to confirm the conformations of **25** should also be undertaken regardless if binding is improved to the SHABK1 PDZ domain target.
- If successful this strategy should be applied to another β -strand mediated PPI
- However, in the unfortunate event that the macrocycle is not obtained through the synthetic steps described in Scheme 8, the re-design of the Hao building block devoid of the hydrazine should be considered.

Chapter 3, partly inspired by work done in Chapter 2, explored the successful development of a new methodology applied to both α -helix and β -strand mediated PPIs, NOXA B/MCL-1 and SIM/SUMO respectively. The generated PIM sequences, although did not bind with stronger affinity to their respective target proteins, were all successfully tolerated. In particular, hotspot residues of both NOXA B and SIM ligands were varied and binding was successfully preserved. Future work should look at the application of the PreSaVS workflow to other diverse PPIs to establish this methodology as a routine and potentially a key protocol in the design of peptide-based inhibitors. For this particular work the following experiments would push the project forward:

- Test the PIMs of Noxa B and SIM peptides for their ability to permeate through a cell line potentially using a Caco-2 model test
- Further advance the PIMs of NOXA B and SIM though N-methylation to endow the peptides with proteolytic stability and tested against enzymes such as Proteinase K and Trypsin. This would also show a natural progression towards peptidomimetic design from generated PIMs.

- The PreSaVS protocol should be launched, by the Bristol team, as an online tool for others to explore and make use.

Chapter 4 focused on expanding the application of the dibromomaleimide stapling technique to generate constrained peptides targeting the HIF-1 α /p300 PPI. In summary, a dibromomaleimide stapled peptide of HIF-1/p300 was designed, synthesised and tested based on the largest helical region of HIF-1, helix3. The stapled peptide HIF-1 $\alpha_{812-826}$ S~~E~~E816R820, showed improved inhibitory potency ((HIF-1 $\alpha_{812-826}$ S~~E~~E816R820 at IC₅₀~ 30 μ M), greater than one order of magnitude when compared to the unconstrained or disulfide variants. However, very little increase in helicity was observed by CD for the unbound form following constraint. Since the helicity of the peptide did not increase upon stapling the theory of constraining the peptide into a bioactive conformation was inadequate. Further studies, MD simulations, were able to show that p300 bound form of the peptide adopts a more stable helix as a consequence of introducing the staple. These MD simulations revealed that whilst stabilization of an unbound peptide in a helical conformation can readily be achieved, this could adversely affect binding affinity by favouring metastable conformations that incur a reorganizational penalty on target engagement, or preventing key side-chains from adopting the orientation required for binding. They also demonstrated that the combination of a constraint and judicious sequence modification promoted solution conformations that matched the ideal bound conformation. Thus, although HIF-1 $\alpha_{812-826}$ S~~E~~E816R820 showed improved inhibitory potency, future optimisation will be required to develop chemical probes. As well, future studies should look to incorporate the importance of the bound conformation of inhibitors to their target protein in future design strategies. To sum up the following experiments would help complete this work:

- NMR studies on the bound complex HIF-1 $\alpha_{812-826}$ S~~E~~E816R820/HIF-1 α should be performed in an attempt to better understand the conformation of the stapled peptide when bound. This should also be compared and contrasted with the unconstrained, native peptide. A computational overlap of the two peptides for a visual comparison could help aid the understanding of the binding potency observed. As well, NMR studies

could help strengthen the argument for consideration of the the bound conformation of inhibitors to their target.

- Lastly, a comparison between different staples could be applied to the HIF-1 α /p300 system. This could provide an in depth assessment of different staples and their effects on the bound vs. unbound states of the inhibitors.

In summary, experimental alanine scanning remains necessary for the accurate study of PPI interfaces. The results of such studies should continue to be compared with computational tools in order to improve the accuracy of *in silico* predictions. Computational tools offer the advantage of rapid results when studying PPI interfaces, new methodologies such as PreSaVS should continue to flourish and test binding interface flexibility while generating PIMs which could be further used as starting points for inhibitor design. Lastly, the binding behaviour of the dibromomaleimide stapling peptide inhibitor of HIF-1/p300 was elucidated using MD simulations – re-enforcing the common theme observed throughout the thesis of using computational and experimental approaches in combination.

Chapter 5: Experimental

5.1 General Experimental Remarks

All commercial solvents and reagents were purchased from Sigma–Aldrich or Alfa Aesar and used without further purification unless stated otherwise. All non-aqueous reactions were performed under an atmosphere of nitrogen. Water-sensitive reactions were performed in oven-dried glassware, cooled under nitrogen before use, or flame dried and cooled, under vacuum if stated. Solvents were removed under reduced pressure using a Büchi rotary evaporator. Ether refers to diethyl ether and petrol refers to petroleum spirit (b.p. 40-60 °C). Flash column chromatography was carried out using silica (35-70 µm particles), with crude reaction mixtures loaded in dichloromethane or the initial solvent system.

Chromatography. Thin Layer Chromatography (TLC) was performed on Merck Kieselgel 60 F24 0.25 mm precoated aluminium plates. Product spots were visualised under UV light ($\lambda_{\max} = 254$ nm) and/or staining with anisaldehyde. Purifications were performed with either silica gel 60 (0.043-0.063 mm VWR) using head bellows or by flash chromatography using an Isolera Four Biotage®. Ion-exchange purifications were performed using Supleco Discovery SPE DSC-SAX columns. HPLC experiments were run on an Agilent 1290 Infinity Analytical Preparative system spectrometer.

NMR Spectroscopy. ^1H NMR spectra were recorded on Bruker Advance 500 at a proton frequency of 400, 500 or 600 MHz at 25 °C or at a temperature and frequency stated in each experiment. ^{13}C NMR spectra were carried out at 101 MHz or 126 MHz. Chemical shifts (δ) are expressed as parts per million (ppm) with residual solvent signal used as a reference ((CD_3) $_2$ SO at 2.50 ppm for ^1H NMR and 39.52 ppm for ^{13}C NMR, CD_2Cl_2 at 5.32 ppm for ^1H NMR and 53.84 ppm for ^{13}C NMR, CDCl_3 at 7.26 ppm for ^1H NMR and 77.16 ppm for ^{13}C NMR), and coupling constants are expressed in Hz. The following abbreviations are used: br = broad, s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet.

Preparative HPLC. Preparative HPLC was performed by Waters Fraction Lynx with ZQ MS detector on either a Waters Xbridge C18 OBD 5 μ m column (19 \times 150 mm, flow rate 30 mL/min or 30 \times 150 mm, flow rate 60 mL/min) using a gradient of 5–95% MeCN with 0.2% NH₃ at pH 10 or a Waters SunFire C18 OBD 5 μ m column (19 \times 150 mm, flow rate 30 mL/min or 30 \times 150 mm, flow rate 60 mL/min) using a gradient of 5–95% MeCN with 0.1 M formic acid or on a Gilson Preparative HPLC with a UV/VIS detector 155 on a Kromasil C8 10 μ m column (20 \times 250 mm, flow rate 19 mL/min, or 50 \times 250 mm, flow rate 100 mL/min) using a varying gradient of MeCN with 0.1% formic acid (FA) in water or 0.2% trifluoroacetic acid (TFA) in water or 0.2% acetic acid (AcOH) in water or 0.2% ammonia (NH₃) in water. Mass spectrometry (HR-ESI-MS) was conducted on a Shimadzu LCMS-2020 instrument (ESI+).

5.2 PreSaVS Information and Full Scan Results

Input for PreSaVS is simply a PDB file of the complex in question. We used the first models of the NMR structures of the NOXA/MCL-1 complex (2JM6) and the SIM/SUMO complex (2LAS). A program, `saturation_mutagenesis.py`, was written in Python3 by Dr. A. Avila Ibarra to perform PreSaVS, leveraging functions and methods from BudeAlaScan^{174,175}. The data presented here were generated with the two commands:

```
saturation_mutagenesis.py full -v -p 2jm6.pdb -l A > 2jm6_full.log
```

```
saturation_mutagenesis.py full -v -p 2las.pdb -l B > 2las_full.log
```

By default, the method requires a PDB file with two protein chains and identifies the interfacial residues between them. Each interfacial residue position is replaced by all naturally occurring 20 amino acids with the exception of Gly, Cys, Pro and scored with the BUDE forcefield with the rotamer correction previously described. Each substitution gives an interaction energy and subtracting this from the interaction energy of the native residue gives a value in kJ/mol such that a residue stabilising the interface with respect to the native residue will give a positive value

The software is available on demand from Dr. Avila Ibarra, and will be made available on github in the future.

Computational output tables from the PreSaVS tool for the NOXA-B sequence followed by the table for the SIM peptide sequence. In Column 1 the native sequence of each peptide and in Row 1 all possible amino acid variations with the computationally generated $\Delta\Delta G$ values for each amino acid variation populating the table.

6	A	D	E	F	H	I	K	L	M	N	Q	R	S	T	V	W	Y
P68		2.97	3.61	1.87	0.88	1.96	0.20	0.83	0.80	1.26	0.53	0.22	0.07	1.10	1.31	0.50	3.56
A69		1.87	6.18	1.10	0.10	0.38	3.72	0.59	0.17	0.41	0.58	1.93	0.12	0.23	0.28	0.27	1.40
D70		0.00	0.96	4.24	4.13	2.61	5.89	4.89	5.40	6.36	3.52	8.23	4.94	3.22	3.18	3.89	3.28
L71		5.07	7.70	4.57	3.64	0.85	8.37	0.00	1.64	5.43	1.26	4.89	4.13	0.63	2.23	9.98	4.62
K72		6.23	5.68	2.60	1.35	2.33	0.00	2.29	2.00	2.35	1.82	1.73	1.64	1.88	1.92	3.72	2.80
D73		0.00	0.81	4.35	5.35	4.02	8.12	4.49	5.59	6.32	4.11	7.77	5.99	4.92	5.15	4.12	4.34
E74		7.77	0.00	35.6 2	25.8 8	16.0 1	25.5 1	14.1 8	16.6 6	13.9 8	14.8 4	20.6 7	19.0 0	21.7 4	19.9 5	60.6 0	55.8 2
C75		6.35	1.88	22.5 0	5.85	4.22	16.5 0	18.2 0	8.78	2.60	3.12	2.64	3.57	5.74	4.90	1.31	11.5 2
A76		0.91	0.04	2.40	1.10	0.41	0.79	2.02	0.09	2.43	0.44	0.13	1.21	0.46	0.52	0.54	2.30
Q77		0.01	0.73	0.80	0.18	1.00	0.37	0.59	0.23	1.68	0.00	0.88	1.48	0.04	0.13	2.25	0.81
L78		8.19	3.10	4.92	5.34	1.45	3.06	0.00	4.25	7.06	2.97	6.83	11.6 0	11.9 2	7.21	4.79	31.5 5
R79		22.0 0	23.5 7	20.7 1	14.0 8	- 8.68	6.25	10.7 6	12.1 2	12.0 8	- 8.42	0.00	9.35 0	12.0 0	11.5 2	19.1 0	25.3 8
R80		0.12	0.40	0.26	0.87	0.55	0.37	0.47	0.37	0.40	0.41	0.00	0.44	0.28	0.53	0.30	1.33
I81		0.75	3.75	8.85	5.66	0.00	10.1 2	1.45	6.70	2.37	2.75	10.4 5	6.78	3.88	2.87	33.1 8	12.5 2
G82		20.2 0	- 2.76	93.6 4	41.3 0	35.1 7	21.8 8	41.1 8	27.7 2	26.1 0	- 0.48	25.0 1	3.89	- 3.67	13.6 8	47.5 0	30.7 4
D83		0.00	4.89	8.52	13.0 6	9.69	9.15	8.75	9.92	8.68	9.32	13.7 4	8.69	8.46	10.6 8	9.06	8.71
K84		1.20	1.24	1.35	0.27	6.92	0.00	5.64	4.99	1.38	2.00	1.62	5.90	4.21	4.57	2.50	0.90
V85		5.44	2.07	1.71	2.87	0.74	1.84	2.01	3.52	4.05	1.96	0.81	7.29	4.25	0.00	1.83	12.8 3
N86		9.12	8.97	5.79	0.63	1.97	4.13	1.25	2.11	0.00	1.70	3.25	1.84	2.05	1.83	2.64	7.54
L87		0.05	0.36	0.16	0.02	0.63	0.02	0.00	0.00	0.05	0.26	0.32	0.03	0.03	0.62	0.45	0.15
R88		12.4 4	12.6 1	- 4.45	- 2.90	- 5.31	2.53	- 6.83	- 6.24	- 6.69	- 5.24	0.00	- 7.57	- 5.70	- 5.51	0.43	3.53
Q89		8.57	6.87	3.78	7.85	2.18	3.56	4.96	1.51	4.87	0.00	1.00	7.84	6.17	2.70	1.11	9.74
K90		0.16	0.03	0.47	0.06	0.83	0.00	0.28	0.09	0.12	0.02	1.09	0.59	0.46	0.05	1.35	0.47
L91		1.44	2.71	0.41	0.16	0.47	3.58	0.00	0.09	0.09	0.09	0.66	0.12	0.02	0.18	0.18	0.41
L92		12.6 5	- 6.90	- 1.80	- 3.38	- 1.76	13.9 1	0.00	3.93	1.47	3.68	8.37	1.15	2.56	2.94	5.85	3.26
N93		6.92	4.80	1.72	5.16	1.36	7.39	0.41	1.86	0.00	0.49	5.79	2.23	0.59	0.13	2.71	11.1 8
M94		1.13	0.11	1.21	1.86	0.20	2.08	0.65	0.00	0.19	0.81	0.29	1.29	0.36	0.04	1.58	1.55

AA	D	E	F	H	I	K	L	M	N	Q	R	S	T	V	W	Y
D27 05	0.00	-0.11	-0.83	-1.14	-1.04	-0.14	-0.91	-0.91	-1.73	-0.80	-0.71	-1.41	-1.48	-1.66	-0.77	-0.63
N27 06	1.18	1.42	1.05	0.88	0.59	1.07	0.70	0.63	0.00	0.85	0.61	0.45	0.53	0.57	0.92	1.12
E27 07	0.98	0.00	0.04	0.48	0.14	0.41	0.31	0.29	0.38	0.14	0.77	0.50	0.39	0.24	0.87	0.04
I270 8	3.65	2.87	0.65	0.03	0.00	1.97	0.05	1.36	1.21	0.98	0.91	1.08	0.30	0.12	2.60	1.72
E27 09	-0.03	0.00	-1.88	-2.38	-1.42	-1.53	-1.73	-1.92	-1.60	-0.42	-1.49	-2.07	-1.56	-1.63	-1.38	-1.62
V27 10	-0.28	1.25	-20.47	-1.18	3.56	-0.84	-1.05	-2.12	1.70	2.17	4.84	2.41	-0.57	0.00	-32.56	-32.52
I271 1	-0.64	0.85	0.90	1.42	0.00	2.59	0.21	0.20	0.35	0.68	1.92	1.75	1.10	0.36	0.42	1.07
I271 2	-6.49	-5.05	-3.93	-5.06	0.00	-4.48	-6.37	-7.20	-2.38	-0.46	-5.77	-5.83	-3.16	-1.25	-45.78	-39.68
V27 13	0.80	1.38	0.76	0.37	0.53	0.92	0.40	0.17	0.10	0.65	0.78	0.31	0.16	0.00	0.90	0.75
W2 714	-5.19	-4.05	-0.49	-5.50	-2.53	-6.00	-10.73	-5.22	-7.02	-4.27	-4.98	-6.43	-9.82	-4.34	0.00	0.06
E27 15	-0.62	0.00	-0.64	-0.33	-0.65	-0.08	-0.86	-0.85	-0.91	-0.79	-0.57	-1.06	-1.04	-0.83	-0.69	-0.63
K27 16	-0.22	-0.23	-0.80	-0.27	0.28	0.00	-0.55	-0.17	-0.39	0.05	0.03	0.51	0.04	-0.34	1.77	1.07
K27 17	-1.74	-1.63	-1.06	-0.41	-1.20	0.00	-1.06	-1.51	-1.82	-1.22	0.08	-1.93	-1.32	-1.36	0.07	0.81

6.1 Peptide Synthesis Procedure

All amino acids and resins were purchased from either Novabiochem (Merck) or Sigma-Aldrich. All amino acids were *N*-Fmoc protected and side chains protected with Boc (Lys); *O*^tBu (Asp, Ser, Thr); Trt (Asn, Gln); Pbf (Arg). Synthesis of all peptides was performed using a microwave assisted automated peptide synthesiser (CEM, Liberty or Liberty Blue) on either Rink Amide MBHA resin or pre-loaded Wang resin. Coupling of 6-aminohexanoic acid, γ -aminobutyric acid and N-terminal labelling were performed manually. DMF used in peptide synthesis was of ACS grade and from Sigma Aldrich. Peptides were synthesised on an 0.1 mmol scale and split before acetylation and coupling fluorescent molecules. Lyophilisation was performed using a BenchTop Pro with Omnitronics™ from VirTis SP Scientific

Cycles for Automated Peptide Synthesis

Peptides that were built on the microwave assisted Liberty CEM peptide synthesiser followed this cycle:

Resin Loading

Clean reaction vessel; wash with DMF, wash with CH₂Cl₂; transfer resin to reaction vessel; wash with DMF, wash with CH₂Cl₂; vessel draining.

Deprotection and Coupling

Clean resin dip tube, wash with DMF (15 mL) add 20% piperidine in DMF (6 mL), microwave method (30 sec), wash with DMF (15 mL), clean resin dip tube, wash with DMF (15 mL), add amino acid (2.5 mL), add coupling reagent (1 mL), add activator base (0.5 mL), microwave method (5 min), wash with DMF (15 mL), drain.

For methods that did not use microwave assistance, the reaction cycle was the same, expect the microwave method for deprotection and coupling was replaced by agitation of the resin at rt for 10 min and 90 min respectively.

After the final residue, the resin was ejected from the reaction vessel and linker coupling, capping, cleavage and deprotection was performed manually.

For the microwave methods used, the temperature and total time is shown below:

Deprotection Microwave Methods

Method	Ramp Time	Total Time	Max Temp
Standard	20-30 sec	1:05	90 °C
75 °C deprotection*	30 sec	0:30	~50 °C
	30-75 sec	3:00	75 °C
Standard	N/A	5:00	rt
	N/A	10:00	rt

Coupling Methods

Method	Ramp Time	Total Time	Max Temp
Standard	20-30 sec	1:05	90 °C
50 °C MW	N/A	2:00	rt
	30-75 sec	4:00	50 °C
Arg* coupling	N/A	25:00	rt
	30-75 sec	2:00	75 °C
75 °C coupling	30 sec	0:30	~50 °C
	30-75 sec	5:00	75 °C

Methods for Manual Solid Phase N-terminal Chain Elongation and Capping

Method A: Coupling of Ahx

Following ejection from the automated synthesiser, the resin was placed in a fritted empty SPE tube and the desired unnatural amino acid (5 equiv.), DIPEA (5 equiv.) and HCTU (5 equiv.) were dissolved in DMF (1 mL) and added to the resin, followed by agitation for 1 h at room temperature. For double couplings, this step was repeated. After removal of the reagents by filtration, the resin was washed with DMF (3 × 2 mL × 2 min) and the success of the coupling determined by a negative colour test (Method C). Deprotection of the Fmoc-protected *N*-terminus then followed (Method B).

Method B: Deprotection of N-Fmoc protecting groups

N-terminal Fmoc-protecting groups were removed by the addition of 20% piperidine: DMF (v/v) (5 × 2 mL × 2 min) at room temperature, followed by rinsing the resin with DMF (5 × 2 mL × 2 min). Successful deprotection was determined by a positive colour test (Method C).

Method C: Kaiser Test²⁵²

The Kaiser Test was used for the determination of the successful coupling or deprotection for any residue coupled manually. A small number of resin beads were rinsed with CH₂Cl₂ and placed in a vial, followed by the addition of two drops of each of the three solutions below:

- 1) Ninhydrin (5% w/v) in ethanol
- 2) Phenol (80% w/v) in ethanol
- 3) 1 mM KCN (aq.) in pyridine (2% v/v)

The solution was then heated to ca. 100 °C for five minutes. A successful coupling gave no change in the colour of the beads, whereas bright blue beads demonstrated a successful deprotection.

Method D: N-terminal acetylation

Acetic anhydride (10 equiv.) and DIPEA (10 equiv.) were dissolved in DMF (1 mL) and the solution was transferred to the resin. After 2 h at room temperature, the resin was drained, washed with DMF (3 × 2 mL × 2 min) and successful capping determined by a negative colour test (Method C).

Method E: N-terminal FITC labelling⁴⁶

Fluorescein isothiocyanate (6 equiv.) was dissolved in 12:7:5 pyridine:DMF:CH₂Cl₂ and the solution transferred to the resin in the dark. After 16 h on the spinner at room temperature, the resin was washed with DMF (3 × 2 mL × 2 min) ahead of cleavage and deprotection. The solvents were of anhydrous grade and distilled before use.

Method F: Cleavage and deprotection of Rink amide MBHA resin

After elongation and N-terminal capping was complete, the resin was washed with CH₂Cl₂ (5 × 2 mL × 2 min), Et₂O (5 × 2 mL × 2 min) and dried under vacuum for ca. 2 h. Peptides were simultaneously cleaved and side-chain deprotected using 'Reagent K' TFA:EDT:Thioanisole:Phenol:H₂O 82:3:5:5:5 (3 × 2 mL × 2 h) at room temperature. The solution was precipitated in ice-cold Et₂O (25 mL) and placed in a centrifuge (3000 rpm × 10 min), the supernatant removed, and the precipitate resuspended in ice-cold Et₂O and placed in a centrifuge again. This process was repeated 3-4 times and the precipitate was dried under a stream of nitrogen, before being dissolved in H₂O and lyophilised.

Peptide Purification

The resulting crude product was dissolved in either, H₂O, DMSO or methanol at an approximate concentration of 20 mg mL⁻¹ and purified using reversed phase mass directed HPLC [Kinetex EVO C18 (250 × 21.2 mm) preparative column (reversed phase) preparative column; variable gradient of MeCN to water (plus 0.1% formic acid v/v in both solvents) and flow rate of 20 mL min⁻¹ during 8 min]. The resulting fractions were concentrated and lyophilized. Purity of peptides was assessed by analytical HPLC and HRMS.

6.2 Peptide Biophysical Binding Assays

6.2.1 General Remarks

Fluorescence Anisotropy

Fluorescence anisotropy assays were performed in 384-well plates (Greiner Bio-one). Each experiment was run in triplicate and the fluorescence anisotropy measured using a Perkin Elmer EnVision™ 2103 MultiLabel plate reader, with excitation at 480 nm (30 nm bandwidth), polarised dichroic mirror at 505 nm and emission at 535 nm (40 nm bandwidth, S and P polarised). All assays were performed in triplicate unless otherwise stated and data analysed following previously published methods.

The data from both the P (perpendicular intensity) and S (parallel intensity) channels, resulting from this measurement and corrected by subtracting the corresponding control wells, were used to calculate the intensity and anisotropy for each well following Equations 1 and 2:

$$I=(2PG)+S \quad \text{Equation 1}$$

$$r=(S-PG) \quad \text{Equation 2}$$

$$L_b=(r-r_{\min})/\lambda(r_{\max}-r)+r-r_{\min} \quad \text{Equation 3}$$

$$y=\{(k+x+[FL])-\sqrt{(k+x+[FL])^2-4x[FL]}\}/2 \quad \text{Equation 4}$$

Fluorescence anisotropy data was processed as described previously.⁴⁶

r = anisotropy, I = total intensity, P = perpendicular intensity, S = parallel intensity, G = instrument factor which was set to 1 for all experiments, L_b = fraction ligand bound, λ = $I_{\text{bound}}/I_{\text{unbound}} = 1$, $[FL]$ = concentration of fluorescent ligand, $k = K_d$, $y = L_b \cdot \text{Flu-trimer}$ and $x = [\text{added titrant}]$, G is an instrument gain factor.

The average anisotropy (across three experimental replicates) and the standard deviation of these values were then calculated and fit to a sigmoidal logistic model (Equation 3) using OriginPro 9.0 which provided the IC_{50} and error values. $y = r_{\max} + (r_{\min}-r_{\max})/(1+(x/x_0)^p)$

Direct binding assays:

Fluorescence anisotropy direct titration assays were performed with protein concentration diluted over 16-24 points using ½ (half) dilutions. 20 µL of buffer was first added to each well. 20 µL of a solution of protein was added to the first column. The solution was well mixed and 20 µL was taken out and added to the next column and so on. This operation consists on serial dilution of the protein

across the plate. Finally, 20 μ L of tracer was added to the wells. For control wells, the tracer peptide was replaced with an identical volume of assay buffer and plates were read after 45 minutes.

Competition binding assays:

FA competition assays were performed in 384 well plates with the concentration of variant peptide competitor, diluted over 16 points in $\frac{1}{2}$ (half) dilutions with fixed protein and tracer concentrations. Tracer was added to each well to give a final concentration of 50 nM. For control wells, the tracer peptide was replaced with an identical volume of assay buffer. The total volume in each well was 60 μ L. Plates were read after 45 minutes of incubation at room temperature.

Assay buffers:

All assays related to GKAP/SHANK1 were carried out in Tris buffer (50 mM Tris, 150 mM NaCl, pH 7.4)

All assays related to NOXA B/Mcl-1 were carried in Tris buffer (50 mM Tris, 150 mM NaCl, 0.01% Triton X-100, pH 7.4).

All assays related to SIM/SUMO1 were carried in Tris buffer (20 mM Tris, 150 mM NaCl, 0.01% Triton X-100, pH 7.4).

All assays related to HIF-1 α /p300 were carried in Tris buffer (20 mM Tris, 100 mM NaCl, 0.1 mM DTT, pH 7.4).

Circular Dichroism

Circular Dichroism was performed on an Applied Photophysics ChiraScan Apparatus and Software. For each scan, the following parameters were used: 180-260 nm range; point time 1 s; 1 nm per point; step = 1; bandwidth 2.5 nm; path length 1 mm; temperature 20 °C. Scans were done in duplicate. Samples were dissolved in 10 mM sodium phosphate buffer pH 7.45 at concentrations of 100 μ M. DMSO was not used in the samples for CD measurements due to its high absorbance below 230 nm. The solvent signal was subtracted to the raw circular dichroism data obtained for the peptides before conversion to the mean residue ellipticity:

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

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

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

Isothermal Calorimetry

ITC experiments were carried out using a Microcal ITC200i instrument (Malvern) at 25°C in 20 mM Tris, 150 mM NaCl, pH 7.5 buffer. Protein of interest was dialysed against the buffer prior to experiment, lyophilized peptides were dissolved in the same buffer. Protein was present in the cell and titrated with peptide solutions loaded into the syringe using 20, 2 μL injections with 120 s spacing between the injections for 20 injections. Heats of peptide dilution were subtracted from each measurement raw data. Data was analysed using Microcal Origin 8 and fitted to a one-binding site model

6.2.2 Chapter 2 Assays: GKAP/SHANK1 PPI System

6.2.2.1 Fluorescence Anisotropy Assays

Direct titration assays were performed in 384 well plates. A total of 20 μL buffer was added followed by titration of 20 μL SHANK1 protein (total concentration in the first well of 15 μM after all components added) and a constant concentration of FITC-Ahx-GKAP₁₋₆ tracer of 100 nM. The total volume in each well was 40 μL . Plates were read after 1 h of incubation at room temperature.

FA competition assays were performed in 384 well plates with the concentration of variant peptide competitor starting from 150 μM , diluted over 16-24 points in 1/2 regime with fixed protein and tracer concentrations. For the GKAP/SHANK1 PDZ FA competition assay, 20 μL of variant GKAP and 20 μL of Shank1PDZ were added to each well to give a final concentration of 50 nM and 1 μM , respectively. For control wells, the tracer peptide was replaced with an identical

volume of assay buffer. The total volume in each well was 60 μ L. Plates were read after 1 h of incubation at room temperature.

6.2.2.2 Isothermal Calorimetry

ITC experiments were carried out using Microcal ITC200i instrument (Malvern) at 25°C in 20 mM Tris, 150 mM NaCl, pH 7.5 buffer. ShankPDZ was dialysed against the buffer prior to experiment, lyophilized peptides were dissolved in the same buffer. 150 μ M shankPDZ was present in the cell and titrated with 1.4-2 mM peptide solutions loaded into the syringe using 20, 2 μ L injections with 120 s spacing between the injections for 20 injection. Heats of peptide dilution was subtracted from each measurement raw data. Data was analysed using Microcal Origin 8 and fitted to a one-binding site model.

6.2.3 Chapter 3 Assays: NOXA B/MCL-1 and SIM/SUMO PPI Systems

6.2.3.1 Circular Dichroism

CD experiments were run as per the general procedure described in 5.4.1 using 150 μ M peptide concentrations. CD experiments of the NOXA B variant peptides were also run in standard buffer with 30% TFE at 150 μ M peptide concentrations.

6.2.3.2 Fluorescence Anisotropy Assays

Direct titration assays were performed in 384 well plates. A total of 20 μ L buffer was added followed by titration of 20 μ L MCL-1 or SUMO protein (total concentration in the first well of 10 μ M and 100 μ M respectively, after all components added) and finally 20 μ L of tracer at 50 nM concentrations of either FAM-Ahx-BID (FAM-Ahx-EDIIRNIARHLA QVGDSMDRSIW) for the MCL-1 system or FAM-Ahx-SIM (FAM-Ahx-DNEIEVIVWEKK) for the SUMO system were added. The total volume in each well was 40 μ L. Plates were read after 1 h of incubation at room temperature.

FA competition assays were performed in 384 well plates with the concentration of variant peptide competitor typically starting from 10-1500 μ M, diluted over 16-24 points in 1/2 regime with fixed protein and tracer concentrations. For the NOXA B/MCL-1 FA competition assay, FAM-BID was added to each well to give a final concentration of 25 nM using 150 nM MCL-1. For the SIM/SUMO1 FA competition assay, 20 μ L FAM-SIM and 20 μ L of SUMO1 were added to each

well to give a final concentration of 25 nM and 3 μ M respectively. For control wells, the tracer peptide was replaced with an identical volume of assay buffer. The total volume in each well was 60 μ L. Plates were read after 1 h of incubation at room temperature.

6.2.4 Chapter 4 Assays and Molecular Dynamics

6.2.4.1 Circular Dichroism

Circular Dichroism: protein/peptide complexes

Spectra were recorded on a chirascan circular dichroism spectropolarimeter (Applied Photophysics), from 20-90 °C, using 1 mm cells and a scan speed of 5 nm/min. The spectra were averaged over 3 repeats with a buffer baseline subtracted. Protein concentrations of approximately 0.2 mg/mL were used. Peptide concentrations were also of approximately 0.2 mg/mL, in a 1:1 ratio with the protein. The helical content of protein/peptide complex was determined from the mean residue ellipticity at 222 nm, $[\theta]$ (deg cm² dmol⁻¹) and compared to that of the protein on its own.

Circular Dichroism: Difference CD

Spectra were recorded on a chirascan circular dichroism spectropolarimeter (Applied Photophysics), at 20 °C, using 10 mm tandem cells and a scan speed of 5 nm/min. The protein and peptide were each dissolved in buffer and added to a sample cell, separated by a partition. This prevented the two solutions from mixing while the CD would acquire a data set which showed an additive signal of the protein and peptide without these two interacting. Following the CD acquisition of the two separated solutions, the cuvette was given a shake and the two solutions mixed. CD spectra was acquired for the mixture. The spectra were averaged over 3 repeats with a buffer baseline subtracted. Each peptide would have three CD traces, one for the signal in the cuvette where the peptide and protein solution were separated (no mix), one of the mixed solutions (mix) and the CD trace of the difference between these two which was calculated by subtracting the raw data. This CD signals was plotted as mean residue ellipticity $[\theta]$ (deg cm² dmol⁻¹) versus wavelength.

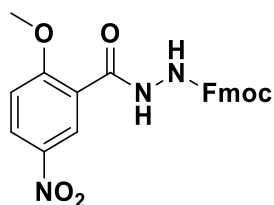
6.2.4.2 Fluorescence Anisotropy Assays

Direct titration assays were performed in 384 well plates. A total of 20 μL buffer was added followed by titration of 20 μL p300 protein (total concentration in the first well of 10 μM after all components added) and a constant concentration of FITC-Ahx-HIF-1 $\alpha_{786-826}$ tracer of 25 nM. The total volume in each well was 40 μL . Plates were read after 1 h of incubation at room temperature.

FA competition assays were performed in 384 well plates with the concentration of variant peptide competitor starting from 850 μM , diluted over 16-24 points in $\frac{1}{2}$ (half) series dilution with fixed protein and tracer concentrations. Total of 20 μL p300 and 20 μL FITC-Ahx-HIF-1 $\alpha_{786-826}$ tracer were added to each well to give a final concentration of 100 nM and 25 nM respectively. For control wells, the tracer peptide was replaced with an identical volume of assay buffer. The total volume in each well was 60 μL . Plates were read after 1 h of incubation at room temperature.

6.3 Synthesis of Hao Building Block

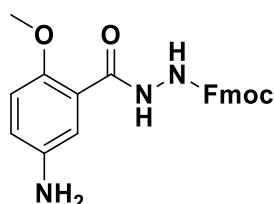
N'-[(9H-fluoren-9-ylmethoxy)carbonyl]-2-methoxy-5-nitrobenzohydrazide
(Hydrazide **11**)



Compound **10** (2.0 g, 10 mmol) was added to a solution of dimethylformamide (2 μL) in anhydrous THF (30 mL), with oxalyl chloride (1.9 g, 16 mmol) and stirred for 1 h. The reaction mixture was then concentrated and re-dissolved in DCM (30 mL). The resulting solution of acid chloride was then added over ca. 2 min to an ice-cold, stirred solution of Fmoc-hydrazine (1.6 g, 6.3 mmol) and pyridine (4.12 mg, 5.22 mmol) in DCM (50 mL). The ice bath was removed and the reaction solution was allowed to stir for 30 min. The reaction mixture was filtered due to precipitation of product, the filtrate washed with H_2O (125 mL), sat. solution of NaHCO_3 (125 mL), and sat. solution of NaCl (125 mL), dried over MgSO_4 , filtered and concentrated *in vacuo* to yield crude product. Precipitate and concentrated

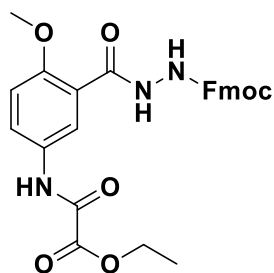
filtrate was combined and purified by column chromatography, eluting with 1:19 - 1:9 DCM:MeOH, to yield compound **11** (1.6 g, 3.8 mmol, 80% yield). ¹H NMR (501 MHz, CDCl₃) δ 9.38 (s, 2H, Dihydrazide-H), 9.10 (d, *J* = 2.7 Hz, 1H, Ar 6-H), 8.40 (dd, *J* = 9.2, 2.7 Hz, 1H, Ar 4-H), 7.75 (app.br.s, 2H, 4- and 5-fluorenyl-H), 7.61 (d, *J* = 6.7 Hz, 2H, 1- and 8-fluorenyl-H), 7.39 (app. br. s, 2H, 3- and 6-fluorenyl-H), 7.31 (app. br. s, 2H, 2- and 7-fluorenyl-H), 7.13 (d, *J* = 9.2 Hz, 1H, Ar 3-H), 4.52 (d, *J* = 6.2 Hz, 2H, Methylene-H), 4.28 (t, *J* = 6.7 Hz, 1H, Fluorenyl 9-H), 4.12 (s, 3H, Methoxy-H); HRMS (ES) *m/z* for C₂₃H₁₉N₃O₆ [M+H]⁺ calcd 434.1352, found 434.1346

N'-[(9H-fluoren-9-ylmethoxy)carbonyl]-2-methoxy-5-aminobenzohydrazide
(Amine **12**)



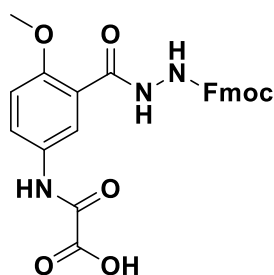
A 500 mL, three-necked round bottom flask equipped with a septum, a stopper, and a stopcock, and was charged with 150 mg Pd/C followed by evacuation and re-filling with N₂ three times. A solution of compound **11** (1.6 g, 3.7 mmol) in THF (175 mL) and MeOH (90 mL) was added carefully via the septum, followed by a balloon of H₂ attached to the stopcock. Flask was evacuated and refilled with H₂ three times, followed by vigorous stirring for 2 h. The reaction mixture was then opened to the air and carefully filtered through Celite, with the Celite bed being thoroughly rinsed with EtOAc (150 mL). Reaction mixture was concentrated *in vacuo*, dry loaded for purification by column chromatography in 9:1 EtOAc:Hexane, to yield compound **12** (1.4 g, 3.6 mmol, 95% yield) as a white, fluffy solid. ¹H NMR (501 MHz, CDCl₃) δ 9.70 (s, 1H, Ar-local hydrazide-H), 7.75 (d, *J* = 7.3 Hz, 2H, 4- and 5-fluorenyl-H), 7.61 (app. br. s, 2H, 1- and 8-fluorenyl-H), 7.55 (d, *J* = 9.0 Hz, 1H, Ar 6-H), 7.39 (d, *J* = 7.3 Hz, 2H, 3- and 6-fluorenyl-H), 7.29 (app. br.s, 2H, 2- and 7-fluorenyl-H), 7.05 (br. s, 1H, Fmoc-local hydrazideH), 6.86 (d, *J* = 9.0 Hz, 2H, Ar-3 and -4-H), 4.47 (d, *J* = 6.6 Hz, 2H, Methylene-H), 4.28 (t, *J* = 6.6 Hz, 1H, Fluorenyl-9-H), 3.93 (s, 3H, Methoxy-H), 3.58 (s, 2H, 5-Amino-H); HRMS (ES) *m/z* for C₂₃H₂₁N₃O₄ [M+H]⁺ calcd 404.1610, found 404.1602.

Ethyl[(3-{N'-[(9H-fluoren-9-ylmethoxy)carbonyl]hydrazinecarbonyl}-4-methylphenyl)carbamoyl]formate (**13**)



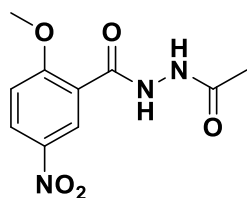
Compound **12** (1.4 g, 3.6 mmol) was added to a solution of ice-cooled DCM (150 mL) and pyridine (0.3 g, 4.3 mmol), and ethyl oxalyl chloride (0.5 mg, 3.9 mmol) added dropwise over 2 min. Following 10 minutes of stirring, the solution was transferred to a separating funnel along with an additional 100 mL DCM and washed with H₂O (100 mL × 2) and sat. NaCl solution (100 mL × 2), followed by drying with MgSO₄, filtered, and concentration *in vacuo*. Purification was performed by dry loading of crude product for column chromatography, eluting in DCM:MeOH (1:0 - 25:1), and requiring a second column eluting with EtOAc, to yield compound **13** (1.5 g, 2.9 mmol, 82% yield). ¹H NMR (400 MHz, CDCl₃) δ 9.63 (br. s, 1H), 9.06 (s, 1H), 8.31 (d, *J* = 9.0 Hz, 1H, Ar 6-H), 8.11 (d, *J* = 2.8 Hz, 1H, Ar 2-H), 7.75 (d, *J* = 7.3 Hz, 2H, 4- and 5-fluorenyl-H), 7.60 (d, *J* = 2.8 Hz, 2H, 1- and 8-fluorenyl-H), 7.39 (t, *J* = 7.3 Hz, 2H, 3- and 6-fluorenyl-H), 7.29 (app. br. s, 2H, 2- and 7-fluorenyl-H), 7.06 (d, *J* = 9.0 Hz, 1H, Ar 3-H), 4.49 (d, *J* = 7.1 Hz, 2H, Methylene-H), 4.46 – 4.38 (m, 6H, Ethyl-H and Methoxy-H), 4.27 (s, 1H, Fluorenyl-9-H), 1.43 (app br. s, 2H, Ethyl-methylene-H); HRMS (ES) *m/z* for C₂₇H₂₅N₃O₇ [M+Na]⁺ calcd 526.1590, found 526.1579.

[(3-{N'-[(9H-fluoren-9-ylmethoxy)carbonyl]hydrazinecarbonyl}-4-methylphenyl)carbamoyl]formic acid (Fmoc-Hao **14**)



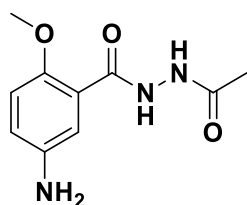
Compound **13** (1.4 g, 2.8 mmol) was added to THF:H₂O (160 mL, 4:1) with 2 M NaOH solution (1.4 mL, 2.8 mmol) added in a single portion. Following 30 min of stirring, the solution was passed through a column of Amberlyst 15 ion exchange resin (4 cm d × 2 cm h, 25mL, 1.7 mmol/mL) and concentrated *in vacuo* to yield compound **14** (1.2 g, 2.6 mmol, 90% yield) as a light brown solid. ¹H NMR (501 MHz, CDCl₃) δ 9.65 (s, 2H), 8.43 (d, *J* = 2.2 Hz, 1H, Ar 6-H), 8.21 (d, *J* = 8.7 Hz, 1H), 7.74 (br. s, 2H), 7.61 (br s, 2H), 7.38 (br s, 2H), 7.27 (br s, 1H), 7.06 (d, *J* = 8.7 Hz, 1H), 4.47 (d, *J* = 7.1 Hz, 2H, Methylene-H), 4.26 (s, 1H, Fluorenyl-9-H), 4.03 (s, 3H, Methoxy), 1.57 (v. br.s, 3H, Carboxy-H); HRMS (ES) *m/z* for C₂₅H₂₁N₃O₇ [M+H]⁺ calcd 476.1458, found 476.1472.

N'-acetyl-2-methoxy-5-nitrobenzohydrazide (**16**)



Compound **10** (2.4 g, 12 mmol) was added to a solution of oxalyl chloride (4.5 mL, 3.9 mmol) in THF (50 mL) and DMF (cat.). The reaction was left stirring for 30 minutes, concentrated *in vacuo* and added to a solution of acethydrazide (3.2 g, 12 mmol), pyridine (0.9 mL) in DCM (50 mL). The reaction was stirred for 2 hours, and a heavy white precipitate formed. The precipitate was filtered, washed with DCM and purified by column chromatography eluting with 100% ethyl acetate. ¹H NMR (400 MHz, MeOD) δ 8.68 (d, *J* = 9.0 Hz, 1H), 8.33 (d, *J* = 9.0 Hz, 1H), 8.34 – 8.31 (m, 1H), 4.00 (s, 3H), 1.97 (s, 3H). LCMS found 254.08, C₁₀H₁₁N₃O₅ requires [M+H]⁺254.07.

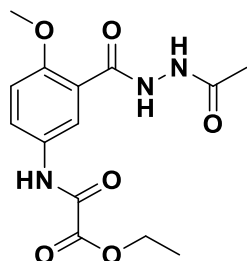
N'-acetyl-5-amino-2-methoxybenzohydrazide (**17**)



Compound **16** (2.6 g, 12 mmol) was added to a solution of palladium on carbon (4.5 mL) in THF (50 mL) under a constant flow of nitrogen. The reaction was left stirring for 2 hours under hydrogen conditions. The mixture was poured over wet

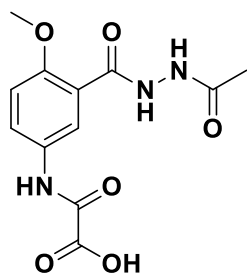
Celite, the filtrate was concentrated *in vacuo* and the product was used in the next step without further purification. ^1H NMR (400 MHz, MeOD) δ 6.79 (m, 1H), 6.30 – 6.39 (m, 2H), 2.71 (s, 3H), 1.46 (s, 3H). LCMS found 224.10, $\text{C}_{10}\text{H}_{13}\text{N}_3\text{O}_3$ requires $[\text{M}+\text{H}]^+224.09$.

Ethyl [3-(N'-acetylhydrazinecarbonyl)-4-methoxyphenyl] formate (**18**)



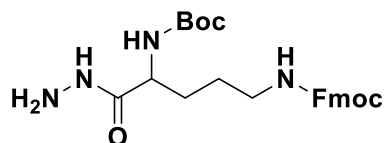
Compound **17** (1.4 g, 4.3 mmol) was added to a solution of ethyl oxalyl chloride 1.0 g, 8.6 mmol) and DIPEA (4.5 mL) in DCM (50 mL). The reaction was left stirring for 30 minutes concentrated *in vacuo* and the product was used in the next step without further purification. ^1H NMR (400 MHz, CDCl_3) δ 8.09 (dd, $J = 9.0, 2.5$ Hz, 1H), 8.00 (dt, $J = 9.0, 2.5$ Hz, 1H), 7.03 (d, $J = 9.0$ Hz, 1H), 4.34 (q, $J = 7.1$ Hz, 2H), 3.96 (s, 3H), 2.03 (s, 3H), 1.35 (t, $J = 7.1$ Hz, 3H). LCMS found 324.10, $\text{C}_{14}\text{H}_{17}\text{N}_3\text{O}_6$ requires $[\text{M}+\text{H}]^+324.11$.

{[3-(N'-acetylhydrazinecarbonyl)-4-methoxyphenyl]carbamoyl}formic acid
(Acetyl Hao **19**)



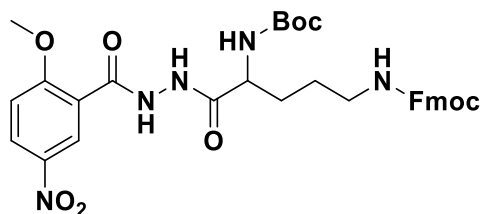
Compound **18** (1.2 g, 3.7 mmol) was stirred in a 1:1 mixture of THF/ H_2O and 2M NaOH was added dropwise until a colour change was observed. The reaction was left stirring for 30 minutes after which 1M HCl was added to pH 3. The product (1.0 g) crashed out of solution, was placed on the freeze drier and used without further purification. ^1H NMR (300 MHz, MeOD) δ 8.45 (s, 1H), 8.14 (dd, $J = 8.9, 2.8$ Hz, 1H), 7.81 (dd, $J = 8.9, 2.8$ Hz, 1H), 3.89 (s, 3H), 1.98 (s, 3H). LCMS found 296.10, $\text{C}_{12}\text{H}_{13}\text{N}_3\text{O}_6$ requires $[\text{M}+\text{H}]^+296.08$.

Tertbutyl-*N*-(4({[(9Fluorenyl)methoxy]carbonyl}amino)1(hydrazinecarbonyl)butyl]carbamate (**29**)



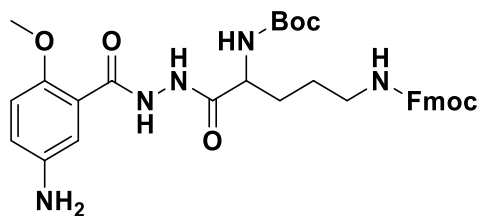
The ornithine linker (2 g, 4.4 mmol) was first activated with HCTU (3.6 g, 8.8 mmol), DIPEA (1.1 g, 8.8 mmol) in acetonitrile and then added to a pot of hydrazine monohydrate (0.2 g, 4.4 mmol) stirring on ice in acetonitrile. The product (1.9 g) crashed out of solution, was concentrated *in vacuo*, further placed on the freeze drier and used in the next synthetic step without purification. ¹H NMR (501 MHz, MeOD) δ 7.69 (d, *J* = 7.5 Hz, Ar-2H), 7.54 (d, *J* = 7.5 Hz, Ar-2H), 7.33 – 7.25 (t, *J* = 7.4 Hz, Ar-2H), 7.21 (t, *J* = 7.4 Hz, Ar-2H), 4.25 (d, *J* = 6.8 Hz, 2H), 4.10 (t, *J* = 6.8 Hz, 1H), 3.87 (dd, *J* = 25.0, 18.7 Hz, 1H), 3.11 – 2.92 (m, 2H), 1.70 – 1.53 (m, 1H), 1.45 (ddd, *J* = 25.0, 18.7, 5.1 Hz, 2H), 1.33 (s, 9H). LCMS found 469.25, C₂₅H₃₂N₄O₅ requires [M+H]⁺469.24.

Tertbutyl-*N*-[5(Fluorenyl)methoxy]carbonyl]amino-1-[2methoxy-5nitrophenyl formohydrazido-1-oxopentan-2-carbamate (**30**)



Compounds **27** (1.7 g, 8.2 mmol) and **29** (1.9 g, 4.1 mmol) were added and stirred in DCM at room temperature followed by slow addition of DIPEA (0.2 mL). The reaction mixture was allowed to stir for 1 hour before being concentrated *in vacuo*. The product was purified by column chromatography eluting with 100% ethyl acetate (2.1 g, 77% yield). ¹H NMR (501 MHz, MeOD) δ 8.72 (s, Ar-1H), 8.39 (dd, *J* = 9.0, 2.2 Hz, Ar-1H), 7.75 (d, *J* = 9.0, 2.2 Hz, Ar-1H), δ 7.63 (d, *J* = 7.5 Hz, Ar-2H), 7.32 (m, Ar-5H), 4.35 (m, 1H), 4.21 (m, 2H), 4.07 (s, 3H), 3.11 – 2.92 (m, 2H), 1.70 – 1.53 (m, 1H), 1.45 (m, 3H), 1.33 (s, 9H). LCMS found 675.26, C₃₃H₃₇N₅O₉ requires [M+H]⁺675.25.

Tertbutyl-*N*-[(5-amino-2-methoxyphenyl)formohydrazido]-5-(fluorenyl)methoxycarbonyl-amino-1-oxopental-2-carbamate (**31**)



A 500 mL, three-necked round bottom flask equipped with a septum, a stopper, and a stopcock, and was charged with 150 mg Pd/C followed by evacuation and re-filling with N₂ three times. A solution of compound **30** (1.8 g, 2.6 mmol) in THF (100 mL) was added carefully via the septum, followed by a balloon of H₂ attached to the stopcock. Flask was evacuated and refilled with H₂ three times, followed by vigorous stirring for 2 h. The reaction mixture was then opened to the air and carefully filtered through Celite, with the Celite bed being previously rinsed with EtOAc (100 mL). Reaction mixture was concentrated *in vacuo*, dry loaded for purification by column chromatography in 9:1 EtOAc:Hexane, to yield compound **30** (1.2 g, 75% yield). ¹H NMR (501 MHz, MeOD) δ 8.54 (s, Ar-1H), 8.40 (dd, J = 9.0, 2.2 Hz, Ar-1H), 7.75 (d, J = 9.0, 2.2 Hz, Ar-2H), δ 7.63 (d, J = 7.5 Hz, Ar-2H), 7.32 (m, Ar-5H), 4.35 (m, 2H), 4.21 (m, 2H), 4.07 (s, 3H), 3.11 – 2.92 (m, 2H), 1.70 – 1.53 (m, 1H), 1.45 (m, 3H), 1.33 (s, 9H).). LCMS found 618.75, C₃₃H₃₉N₅O₇ requires [M+H]⁺618.70.

References

- 1 A. E. Modell, S. L. Blosser and P. S. Arora, *Trends Pharmacol. Sci.*, 2016, **37**, 702–713.
- 2 C. Chung and M. M. Hann, in *Structural Biology in Drug Discovery*, John Wiley & Sons, Ltd, 2020, pp. 479–502.
- 3 B. Gemovic, N. Sumonja, R. Davidovic, V. Perovic and N. Veljkovic, *Curr. Med. Chem.*, 2019, **26**, 3890–3910.
- 4 A. A. Ivanov, F. R. Khuri and H. Fu, *Trends Pharmacol. Sci.*, 2013, **34**, 393–400.
- 5 M. R. Arkin and A. Whitty, *Curr. Opin. Chem. Biol.*, 2009, **13**, 284–290.
- 6 M. R. Arkin and J. A. Wells, *Nat. Rev. Drug Discov.*, 2004, **3**, 301–317.
- 7 J. C. Fuller, N. J. Burgoyne and R. M. Jackson, *Drug Discov. Today*, 2009, **14**, 155–161.
- 8 L. Mabonga and A. P. Kappo, *Biophys. Rev.*, 2019, **11**, 559–581.
- 9 D. E. Scott, A. R. Bayly, C. Abell and J. Skidmore, *Nat. Rev. Drug Discov.*, 2016, **15**, 533–550.
- 10 P. E. Czabotar, E. F. Lee, M. F. van Delft, C. L. Day, B. J. Smith, D. C. S. Huang, W. D. Fairlie, M. G. Hinds and P. M. Colman, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 6217–6222.
- 11 C. Corbi-Verge, M. Garton, S. Nim and P. M. Kim, *Annu. Rev. Pharmacol. Toxicol.*, 2017, **57**, 39–60.
- 12 C. Corbi-Verge and P. M. Kim, *Cell Commun. Signal.*, 2016, **14**, 1–12.
- 13 S. Eyrisch and V. Helms, *J. Med. Chem.*, 2007, **50**, 3457–3464.
- 14 M. Pelay-Gimeno, A. Glas, O. Koch and T. N. Grossmann, *Angew. Chemie Int. Ed.*, 2015, **54**, 8896–8927.
- 15 E. D. Levy and J. B. Pereira-Leal, *Curr. Opin. Struct. Biol.*, 2008, **18**, 349–357.
- 16 J.-Y. Kim and H. S. Chung, *Science (80-.)*, 2020, **368**, 1253 LP – 1257.
- 17 T. Clackson and J. a Wells, *Science (80-.)*, 1995, **267**, 383–386.
- 18 B. C. Cunningham and J. A. Wells, *Science (80-.)*, 1989, **244**, 1081 LP – 1085.
- 19 A. A. Ibarra, G. J. Bartlett, Z. Hegedüs, S. Dutt, F. Hobor, K. A. Horner, K. Hetherington, K. Spence, A. Nelson, T. A. Edwards, D. N. Woolfson, R. B. Sessions and A. J. Wilson, *ACS Chem. Biol.*, 2019, **14**, 2252–2263.
- 20 T. Kortemme, D. E. Kim and D. Baker, *Sci. STKE*, 2004, **2004**, pl2 LP-pl2.
- 21 O. Keskin, B. Ma and R. Nussinov, *J. Mol. Biol.*, 2005, **345**, 1281–1294.
- 22 N. E. Davey, M.-H. Seo, V. K. Yadav, J. Jeon, S. Nim, I. Krystkowiak, C. Blikstad, D. Dong, N. Markova, P. M. Kim and Y. Ivarsson, *FEBS J.*, 2017, **284**, 485–498.

- 23 W. Yang, X. Sun, C. Zhang and L. Lai, *Comput. Struct. Biotechnol. J.*, 2019, **17**, 1396–1403.
- 24 R. Biswas and A. Bagchi, *Comput. Biol. Chem.*, 2017, **70**, 116–124.
- 25 L. T. Vassilev, B. T. Vu, B. Graves, D. Carvajal, F. Podlaski, Z. Filipovic, N. Kong, U. Kammlott, C. Lukacs, C. Klein, N. Fotouhi and E. A. Liu, *Science (80-.)*, 2004, **303**, 844 LP – 848.
- 26 Y. S. Chang, B. Graves, V. Guerlavais, C. Tovar, K. Packman, K.-H. To, K. A. Olson, K. Kesavan, P. Gangurde, A. Mukherjee, T. Baker, K. Darlak, C. Elkin, Z. Filipovic, F. Z. Qureshi, H. Cai, P. Berry, E. Feyfant, X. E. Shi, J. Horstick, D. A. Annis, A. M. Manning, N. Fotouhi, H. Nash, L. T. Vassilev and T. K. Sawyer, *Proc. Natl. Acad. Sci.*, 2013, **110**, E3445--E3454.
- 27 L. Mabonga and A. P. Kappo, *Int. J. Pept. Res. Ther.*, 2020, **26**, 225–241.
- 28 H. Yin, G. I. Lee, S. P. Hyung, G. A. Payne, J. M. Rodriguez, S. M. Sebti and A. D. Hamilton, *Angew. Chemie - Int. Ed.*, 2005, **44**, 2704–2707.
- 29 M. D. Boersma, J. D. Sadowsky, Y. A. Tomita and S. H. Gellman, *Protein Sci.*, 2008, **17**, 1232–1240.
- 30 M. R. Arkin, Y. Tang and J. A. Wells, *Chem. Biol.*, 2014, **21**, 1102–1114.
- 31 J. Gavenonis, B. A. Sheneman, T. R. Siegert, M. R. Eshelman and J. A. Kritzer, *Nat. Chem. Biol.*, 2014, **10**, 716–722.
- 32 N. E. Davey, K. Van Roey, R. J. Weatheritt, G. Toedt, B. Uyar, B. Altenberg, A. Budd, F. Diella, H. Dinkel and T. J. Gibson, *Mol. BioSyst.*, 2012, **8**, 268–281.
- 33 P. Hraber, P. E. O'Maille, A. Silberfarb, K. Davis-Anderson, N. Generous, B. H. McMahon and J. M. Fair, *Trends Biotechnol.*, 2020, **38**, 113–127.
- 34 S. J. Freedman, Z.-Y. Y. J. Sun, F. Poy, A. L. Kung, D. M. Livingston, G. Wagner and M. J. Eck, *Proc Natl Acad Sci U S A*, 2002, **99**, 5367–5372.
- 35 Y. J. Im, J. H. Lee, S. H. Park, S. J. Park, S. H. Rho, G. B. Kang, E. Kim and S. H. Eom, *J. Biol. Chem.*, 2003, **278**, 48099–48104.
- 36 O. Kristensen, S. Guenat, I. Dar, N. Allaman-Pillet, A. Abderrahmani, M. Ferdaoussi, R. Roduit, F. Maurer, J. S. Beckmann, J. S. Kastrup, M. Gajhede and C. Bonny, *EMBO J.*, 2006, **25**, 785–797.
- 37 J. A. Wells and C. L. McClendon, *Nature*, 2007, **450**, 1001–1009.
- 38 V. Azzarito, K. Long, N. S. Murphy and A. J. Wilson, *Nat. Chem.*, 2013, **5**, 161–173.
- 39 W. A. Loughlin, J. D. A. Tyndall, M. P. Glenn and D. P. Fairlie, *Chem. Rev.*, 2004, **104**, 6085–6118.
- 40 A. Ganeshpurkar, R. Swetha, D. Kumar, G. P. Gangaram, R. Singh, G. Gutti, S. Jana, D. Kumar and A. K. and S. K. Singh*, *Curr. Top. Med. Chem.*, 2019, **19**, 501–533.
- 41 J. Banerjee, E. Radvar and H. S. Azevedo, in *Peptides and Proteins as Biomaterials for Tissue Regeneration and Repair*, eds. M. A. Barbosa and M. C. L. Martins, Woodhead Publishing, 2018, pp. 245–281.

- 42 A. L. Boyle, in *Peptide Applications in Biomedicine, Biotechnology and Bioengineering*, ed. S. Koutsopoulos, Woodhead Publishing, 2018, pp. 51–86.
- 43 B. N. Bullock, A. L. Jochim and P. S. Arora, *J. Am. Chem. Soc.*, 2011, **133**, 14220–14223.
- 44 J. Fogha, B. Marekha, M. De Giorgi, A. S. Voisin-Chiret, S. Rault, R. Bureau and J. Sopkova-de Oliveira Santos, *J. Chem. Inf. Model.*, 2017, **57**, 2885–2895.
- 45 C. L. Day, C. Smits, F. C. Fan, E. F. Lee, W. D. Fairlie and M. G. Hinds, *J. Mol. Biol.*, 2008, **380**, 958–971.
- 46 J. A. Miles, D. J. Yeo, P. Rowell, S. Rodriguez-Marin, C. M. Pask, S. L. Warriner, T. A. Edwards and A. J. Wilson, , DOI:10.1039/c5sc04048e.
- 47 A. de Groot, E. de Rosny, C. Juillan-Binard, J.-L. Ferrer, V. Laudet, R. J. Pierce, E. Pebay-Peyroula, J. C. Fontecilla-Camps and F. Borel, *J. Mol. Biol.*, 2005, **354**, 841–853.
- 48 E. A. Campbell, J. L. Tupy, T. M. Gruber, S. Wang, M. M. Sharp, C. A. Gross and S. A. Darst, *Mol. Cell*, 2003, **11**, 1067–1078.
- 49 G. M. Burslem, H. F. Kyle, A. Nelson, T. A. Edwards and A. J. Wilson, *Chem. Sci.*, 2017, **8**, 4188–4202.
- 50 A. L. G. and K. D. Janda, *Curr. Top. Med. Chem.*, 2011, **11**, 258–280.
- 51 R. Chowdhury, A. Hardy and C. J. Schofield, *Chem. Soc. Rev.*, 2008, **37**, 1308–1319.
- 52 I. K. Nordgren and A. Tavassoli, *Chem. Soc. Rev.*, 2011, **40**, 4307.
- 53 J. L. Ruas, *J. Biol. Chem.*, 2002, **277**, 38723–38730.
- 54 I. M. Yasinska and V. V Sumbayev, *FEBS Lett.*, 2003, **549**, 105–109.
- 55 H. Cho, D.-R. Ahn, H. Park and E. G. Yang, *FEBS Lett.*, 2007, **581**, 1542–1548.
- 56 D. Lando, *Science (80-.)*, 2002, **295**, 858–861.
- 57 J. Gu, J. Milligan and L. E. Huang, *J. Biol. Chem.*, 2001, **276**, 3550–3554.
- 58 H. F. Kyle, K. F. Wickson, J. Stott, G. M. Burslem, A. L. Breeze, C. Tiede, D. C. Tomlinson, S. L. Warriner, A. Nelson, A. J. Wilson and T. A. Edwards, *Mol. Biosyst.*, 2015, **11**, 2738–2749.
- 59 P. Carrero, K. Okamoto, P. Coumailleau, S. O'Brien, H. Tanaka and L. Poellinger, *Mol. Cell. Biol.*, 2000, **20**, 402–415.
- 60 O. Khakshoor and J. S. Nowick, *Curr. Opin. Chem. Biol.*, 2008, **12**, 722–729.
- 61 D. P. Fairlie, J. D. A. Tyndall, R. C. Reid, A. K. Wong, G. Abbenante, M. J. Scanlon, D. R. March, D. A. Bergman, C. L. L. Chai and B. A. Burkett, *J. Med. Chem.*, 2000, **43**, 1271–1281.
- 62 A. M. Watkins and P. S. Arora, *ACS Chem. Biol.*, 2014, **9**, 1747–1754.
- 63 Y.-L. Zhao and Y.-D. Wu, *J. Am. Chem. Soc.*, 2002, **124**, 1570–1571.

- 64 S. Malek, D.-B. Huang, T. Huxford, S. Ghosh and G. Ghosh, *J. Biol. Chem.*, 2003, **278**, 23094–23100.
- 65 C. Sala, V. Piëch, N. R. Wilson, M. Passafaro, G. Liu and M. Sheng, *Neuron*, 2001, **31**, 115–130.
- 66 S. M. Shin, N. Zhang, J. Hansen, N. Z. Gerges, D. T. S. Pak, M. Sheng and S. H. Lee, *Nat. Neurosci.*, 2012, **15**, 1655–1666.
- 67 J. H. Lee, H. Park, S. J. Park, H. J. Kim and S. H. Eom, *Biochem. Biophys. Res. Commun.*, 2011, **407**, 207–212.
- 68 R. Singh, A. Letai and K. Sarosiek, *Nat. Rev. Mol. Cell Biol.*, 2019, **20**, 175–193.
- 69 A. Ashkenazi, W. J. Fairbrother, J. D. Levenson and A. J. Souers, *Nat. Rev. Drug Discov.*, 2017, **16**, 273–284.
- 70 J.-C. Carry and C. Garcia-Echeverria, *Bioorg. Med. Chem. Lett.*, 2013, **23**, 2480–2485.
- 71 C. C. Harris, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 1659–1660.
- 72 J. D. Harper and P. T. Lansbury, *Annu. Rev. Biochem.*, 1997, **66**, 385–407.
- 73 J. L. Silverman, S. M. Turner, C. L. Barkan, S. S. Tolu, R. Saxena, A. Y. Hung, M. Sheng and J. N. Crawley, *Brain Res.*, 2011, **1380**, 120–137.
- 74 A. Y. Hung, K. Futai, C. Sala, J. G. Valtchanoff, J. Ryu, M. A. Woodworth, F. L. Kidd, C. C. Sung, T. Miyakawa, M. F. Bear, R. J. Weinberg and M. Sheng, *J. Neurosci.*, 2008, **28**, 1697–1708.
- 75 C. Sala, C. Vicidomini, I. Bigi, A. Mossa and C. Verpelli, *J. Neurochem.*, 2015, **135**, 849–858.
- 76 P. Monteiro and G. Feng, *Nat. Rev. Neurosci.*, 2017, **18**, 147–157.
- 77 J. G. Moffat, F. Vincent, J. A. Lee, J. Eder and M. Prunotto, *Nat. Rev. Drug Discov.*, 2017, **16**, 531–543.
- 78 J. Shortt, A. K. Hsu and R. W. Johnstone, *Oncogene*, 2013, **32**, 4191–4202.
- 79 A. Lopez-Girona, D. Mendy, T. Ito, K. Miller, A. K. Gandhi, J. Kang, S. Karasawa, G. Carmel, P. Jackson, M. Abbasian, A. Mahmoudi, B. Cathers, E. Rychak, S. Gaidarova, R. Chen, P. H. Schafer, H. Handa, T. O. Daniel, J. F. Evans and R. Chopra, *Leukemia*, 2012, **26**, 2326–2335.
- 80 M. Schenone, V. Dančík, B. K. Wagner and P. A. Clemons, *Nat. Chem. Biol.*, 2013, **9**, 232–240.
- 81 P. A. Wadsworth, O. Folorunso, N. Nguyen, A. K. Singh, D. D’Amico, R. T. Powell, D. Brunell, J. Allen, C. Stephan and F. Laezza, *Sci. Rep.*, 2019, **9**, 16890.
- 82 C. Tse, A. R. Shoemaker, J. Adickes, M. G. Anderson, J. Chen, S. Jin, E. F. Johnson, K. C. Marsh, M. J. Mitten, P. Nimmer, L. Roberts, S. K. Tahir, Y. Xiao, X. Yang, H. Zhang, S. Fesik, S. H. Rosenberg and S. W. Elmore, *Cancer Res.*, 2008, **68**, 3421–3428.

- 83 H. Chen, X. Zhou, A. Wang, Y. Zheng, Y. Gao and J. Zhou, *Drug Discov. Today*, 2015, **20**, 105–113.
- 84 M. Cancellieri, M. Bassetto, I. Widjaja, F. van Kuppeveld, C. A. M. de Haan and A. Brancale, *Antiviral Res.*, 2015, **122**, 46–50.
- 85 J. Lyu, S. Wang, T. E. Balius, I. Singh, A. Levit, Y. S. Moroz, M. J. O'Meara, T. Che, E. Alga, K. Tolmacheva, A. A. Tolmachev, B. K. Shoichet, B. L. Roth and J. J. Irwin, *Nature*, 2019, **566**, 224–229.
- 86 D. Seebach, A. K. Beck and D. J. Bierbaum, *Chem. Biodivers.*, 2004, **1**, 1111–1239.
- 87 N. P. Chongsiriwatana, J. A. Patch, A. M. Czyzewski, M. T. Dohm, A. Ivankin, D. Gidalevitz, R. N. Zuckermann and A. E. Barron, *Proc. Natl. Acad. Sci.*, 2008, **105**, 2794–2799.
- 88 A. M. Webster and S. L. Cobb, *Chem. – A Eur. J.*, 2018, **24**, 7560–7573.
- 89 P. N. Cheng, C. Liu, M. Zhao, D. Eisenberg and J. S. Nowick, *Nat. Chem.*, 2012, **4**, 927–933.
- 90 W. A. Loughlin, J. D. A. Tyndall, M. P. Glenn, T. A. Hill and D. P. Fairlie, *Chem. Rev.*, 2010, **110**, PR32–PR69.
- 91 V. Azzarito, K. Long, N. S. Murphy and A. J. Wilson, *Nat. Chem.*, 2013, **5**, 161–173.
- 92 M. K. P. Jayatunga, S. Thompson and A. D. Hamilton, *Bioorg. Med. Chem. Lett.*, 2014, **24**, 717–724.
- 93 I. Arrata, C. M. Grison, H. M. Coubrough, P. Prabhakaran, M. A. Little, D. C. Tomlinson, M. E. Webb and A. J. Wilson, *Org. Biomol. Chem.*, 2019, **17**, 3861–3867.
- 94 O. V Kulikov, C. Incarvito and A. D. Hamilton, *Tetrahedron Lett.*, 2011, **52**, 3705–3709.
- 95 A. L. Rodriguez, A. Tamrazi, M. L. Collins and J. A. Katzenellenbogen, *J. Med. Chem.*, 2004, **47**, 600–611.
- 96 B. P. Orner, J. T. Ernst and A. D. Hamilton, *J. Am. Chem. Soc.*, 2001, **123**, 5382–5383.
- 97 B. B. Lao, K. Drew, D. A. Guarracino, T. F. Brewer, D. W. Heindel, R. Bonneau and P. S. Arora, *J. Am. Chem. Soc.*, 2014, **136**, 7877–7888.
- 98 A. J. Wilson, *Prog. Biophys. Mol. Biol.*, 2015, **119**, 33–40.
- 99 G. L. Verdine and G. J. Hilinski, in *Protein Engineering for Therapeutics, Part B*, eds. K. D. Wittrup and G. L. Verdine, Academic Press, 2012, vol. 503, pp. 3–33.
- 100 N. S. Robertson and D. R. Spring, *Molecules*, , DOI:10.3390/molecules23040959.
- 101 L. D. Walensky and G. H. Bird, *J. Med. Chem.*, 2014, **57**, 6275–6288.
- 102 P. M. Cromm, J. Spiegel, P. Kuchler, L. Dietrich, J. Kriegesmann, M. Wendt, R. S. Goody, H. Waldmann and T. N. Grossmann, *ACS Chem. Biol.*, 2016, **11**, 2375–2382.

- 103 Q. Chu, R. E. Moellering, G. J. Hilinski, Y.-W. Kim, T. N. Grossmann, J. T.-H. Yeh and G. L. Verdine, *Med. Chem. Commun.*, 2015, **6**, 111–119.
- 104 L. D. Walensky, A. L. Kung, I. Escher, T. J. Malia, S. Barbuto, R. D. Wright, G. Wagner, G. L. Verdine and S. J. Korsmeyer, *Science (80-.)*, 2004, **305**, 1466–1470.
- 105 C. J. Brown, S. T. Quah, J. Jong, A. M. Goh, P. C. Chiam, K. H. Khoo, M. L. Choong, M. A. Lee, L. Yurlova, K. Zolghadr, T. L. Joseph, C. S. Verma and D. P. Lane, *ACS Chem. Biol.*, 2013, **8**, 506–512.
- 106 J. [Laxio Arenas], J. Kaffy and S. Onger, *Curr. Opin. Chem. Biol.*, 2019, **52**, 157–167.
- 107 L. K. Henchey, A. L. Jochim and P. S. Arora, *Curr. Opin. Chem. Biol.*, 2008, **12**, 692–697.
- 108 S. Baek, P. S. Kutchukian, G. L. Verdine, R. Huber, T. A. Holak, K. W. Lee and G. M. Popowicz, *J. Am. Chem. Soc.*, 2012, **134**, 103–106.
- 109 C. Phillips, L. R. Roberts, M. Schade, R. Bazin, A. Bent, N. L. Davies, R. Moore, A. D. Pannifer, A. R. Pickford, S. H. Prior, C. M. Read, A. Scott, D. G. Brown, B. Xu and S. L. Irving, *J. Am. Chem. Soc.*, 2011, **133**, 9696–9699.
- 110 M. Moiola, M. G. Memeo and P. Quadrelli, *Molecules*, 2019, **24**, 3654.
- 111 T. L. Joseph, D. P. Lane and C. S. Verma, *PLoS One*, 2012, **7**, e43985–e43985.
- 112 Z. Guo, U. Mohanty, J. Noehre, T. K. Sawyer, W. Sherman and G. Krilov, *Chem. Biol. Drug Des.*, 2010, **75**, 348–359.
- 113 C. H. Douse, S. J. Maas, J. C. Thomas, J. A. Garnett, Y. Sun, E. Cota and E. W. Tate, *ACS Chem. Biol.*, 2014, **9**, 2204–2209.
- 114 H. N. Hoang, R. W. Driver, R. L. Beyer, T. A. Hill, A. D. de Araujo, F. Plisson, R. S. Harrison, L. Goedecke, N. E. Shepherd and D. P. Fairlie, *Angew. Chemie Int. Ed.*, 2016, **55**, 8275–8279.
- 115 A. D. de Araujo, J. Lim, K.-C. Wu, Y. Xiang, A. C. Good, R. Skerlj and D. P. Fairlie, *J. Med. Chem.*, 2018, **61**, 2962–2972.
- 116 D. P. Fairlie and A. de Araujo, *Pept. Sci.*, 2016, **106**, 843–852.
- 117 C. M. Grison, G. M. Burslem, J. A. Miles, L. K. A. Pilsl, D. J. Yeo, Z. Imani, S. L. Warriner, M. E. Webb and A. J. Wilson, .
- 118 Y. H. Lau, P. de Andrade, N. Sköld, G. J. McKenzie, A. R. Venkitaraman, C. Verma, D. P. Lane and D. R. Spring, *Org. Biomol. Chem.*, 2014, **12**, 4074–4077.
- 119 L. K. Henchey, S. Kushal, R. Dubey, R. N. Chapman, B. Z. Olenyuk and P. S. Arora, *J. Am. Chem. Soc.*, 2010, **132**, 941–943.
- 120 A.-M. Leduc, J. O. Trent, J. L. Wittliff, K. S. Bramlett, S. L. Briggs, N. Y. Chirgadze, Y. Wang, T. P. Burris and A. F. Spatola, *Proc. Natl. Acad. Sci.*, 2003, **100**, 11273–11278.
- 121 H. Jo, N. Meinhardt, Y. Wu, S. Kulkarni, X. Hu, K. E. Low, P. L. Davies, W. F. DeGrado and D. C. Greenbaum, *J. Am. Chem. Soc.*, 2012, **134**,

17704–17713.

- 122 X. Shi, R. Zhao, Y. Jiang, H. Zhao, Y. Tian, Y. Jiang, J. Li, W. Qin, F. Yin and Z. Li, *Chem. Sci.*, 2018, **9**, 3227–3232.
- 123 A. Lindsey-crosthwait, D. Rodriguez-lema and M. Walko, 11–27.
- 124 E. Katz and A. L. Demain, *Bacteriol. Rev.*, 1977, **41**, 449–474.
- 125 J. L. Martin, J. Begun, A. Schindeler, W. A. Wickramasinghe, D. Alewood, P. F. Alewood, D. A. Bergman, R. I. Brinkworth, G. Abbenante, D. R. March, R. C. Reid and D. P. Fairlie, *Biochemistry*, 1999, **38**, 7978–7988.
- 126 J. D. A. Tyndall, R. C. Reid, D. P. Tyssen, D. K. Jardine, B. Todd, M. Passmore, D. R. March, L. K. Pattenden, D. A. Bergman, D. Alewood, S.-H. Hu, P. F. Alewood, C. J. Birch, J. L. Martin and D. P. Fairlie, *J. Med. Chem.*, 2000, **43**, 3495–3504.
- 127 M. Prabu-Jeyabalan, E. A. Nalivaika, N. M. King and C. A. Schiffer, *J. Virol.*, 2003, **77**, 1306–1315.
- 128 E. M. Driggers, S. P. Hale, J. Lee and N. K. Terrett, *Nat. Rev. Drug Discov.*, 2008, **7**, 608–624.
- 129 K. Zerbe, K. Moehle and J. A. Robinson, *Acc. Chem. Res.*, 2017, **50**, 1323–1331.
- 130 S. C. Shankaramma, Z. Athanassiou, O. Zerbe, K. Moehle, C. Mouton, F. Bernardini, J. W. Vrijbloed, D. Obrecht and J. A. Robinson, *Chembiochem*, 2002, **3**, 1126–1133.
- 131 W. L. DeLano, M. H. Ultsch, de Abraham M., Vos and J. A. Wells, *Science (80-.)*, 2000, **287**, 1279–1283.
- 132 R. L. A. Dias, R. Fasan, K. Moehle, A. Renard, D. Obrecht and J. A. Robinson, *J. Am. Chem. Soc.*, 2006, **128**, 2726–2732.
- 133 R. Fasan, R. L. A. Dias, K. Moehle, O. Zerbe, J. W. Vrijbloed, D. Obrecht and J. A. Robinson, *Angew. Chemie Int. Ed.*, 2004, **43**, 2109–2112.
- 134 C. A. Lipinski, F. Lombardo, B. W. Dominy and P. J. Feeney, *Adv. Drug Deliv. Rev.*, 1997, **23**, 3–25.
- 135 S. D. Appavoo, T. Kaji, J. R. Frost, C. C. G. Scully and A. K. Yudin, *J. Am. Chem. Soc.*, 2018, **140**, 8763–8770.
- 136 R. C. Ladner, A. K. Sato, J. Gorzelany and M. de Souza, *Drug Discov. Today*, 2004, **9**, 525–529.
- 137 C.-H. Wu, I.-J. Liu, R.-M. Lu and H.-C. Wu, *J. Biomed. Sci.*, 2016, **23**, 8.
- 138 L. Simonetti and Y. Ivarsson, *ACS Cent. Sci.*, 2020, **6**, 336–338.
- 139 G. N. Sundell, R. Arnold, M. Ali, P. Naksukpaiboon, J. Orts, P. Güntert, C. N. Chi and Y. Ivarsson, *Mol. Syst. Biol.*, 2018, **14**, 1–22.
- 140 J. A. Miles, F. Hobor, J. Taylor, C. Tiede, P. R. Rowell, C. H. Trinh, B. Jackson, F. Nadat, H. F. Kyle, B. I. M. Wicky, J. Clarke, D. C. Tomlinson, A. J. Wilson and T. A. Edwards, *bioRxiv*, 2019, 651364.
- 141 H. F. Kyle, K. F. Wickson, J. Stott, G. M. Burslem, A. L. Breeze, C. Tiede, D. C. Tomlinson, S. L. Warriner, A. Nelson, A. J. Wilson and T. A.

- Edwards, *Mol. BioSyst.*, 2015, **11**, 2738–2749.
- 142 K. Deyle, X.-D. Kong and C. Heinis, *Acc. Chem. Res.*, 2017, **50**, 1866–1874.
- 143 A. Tavassoli, *Curr. Opin. Chem. Biol.*, 2017, **38**, 30–35.
- 144 B. Vu, P. Wovkulich, G. Pizzolato, A. Lovey, Q. Ding, N. Jiang, J.-J. Liu, C. Zhao, K. Glenn, Y. Wen, C. Tovar, K. Packman, L. Vassilev and B. Graves, *ACS Med. Chem. Lett.*, 2013, **4**, 466–469.
- 145 Q. Ding, Z. Zhang, J.-J. Liu, N. Jiang, J. Zhang, T. M. Ross, X.-J. Chu, D. Bartkovitz, F. Podlaski, C. Janson, C. Tovar, Z. M. Filipovic, B. Higgins, K. Glenn, K. Packman, L. T. Vassilev and B. Graves, *J. Med. Chem.*, 2013, **56**, 5979–5983.
- 146 J. Kocik, M. Machula, A. Wisniewska, E. Surmiak, T. A. Holak and L. Skalniak, *Cancers (Basel)*, 2019, **11**, 1–39.
- 147 J. Chen, S. Jin, V. Abraham, X. Huang, B. Liu, M. J. Mitten, P. Nimmer, X. Lin, M. Smith, Y. Shen, A. R. Shoemaker, S. K. Tahir, H. Zhang, S. L. Ackler, S. H. Rosenberg, H. Maecker, D. Sampath, J. D. Levenson, C. Tse and S. W. Elmore, *Mol. Cancer Ther.*, 2011, **10**, 2340–2349.
- 148 L. M. Juárez-Salcedo, V. Desai and S. Dalia, *Drugs Context*, 2019, **8**, 212574.
- 149 A. Green, F. Hobor, C. Tinworth, S. Warriner, A. Wilson and A. Nelson, *Chem. – A Eur. J.*, , DOI:10.1002/chem.202002153.
- 150 M. Arkin, *Curr. Opin. Chem. Biol.*, 2005, **9**, 317–324.
- 151 N. R. Christensen, J. Čalyševa, E. F. A. Fernandes, S. Lüchow, L. S. Clemmensen, L. M. Haugaard-Kedström and K. Strømgaard, *Adv. Ther.*, 2019, **2**, 1800143.
- 152 H.-J. Lee and J. J. Zheng, *Cell Commun. Signal.*, 2010, **8**, 8.
- 153 Y. Ivarsson, *FEBS Lett.*, 2012, **586**, 2638–2647.
- 154 S. Lim, S. Naisbitt, J. Yoon and J. Hwang, *J. Biol. Chem.*, 1999, **274**, 29510–29518.
- 155 J. Saupe, Y. Roske, C. Schillinger, N. Kamdem, S. Radetzki, A. Diehl, H. Oschkinat, G. Krause, U. Heinemann and J. Rademann, *ChemMedChem*, 2011, **6**, 1411–1422.
- 156 E. Kim and M. Sheng, *Nat. Rev. Neurosci.*, 2004, **5**, 771–781.
- 157 S. K. Ponna, S. Ruskamo, M. Myllykoski, C. Keller, T. M. Boeckers and P. Kursula, *J. Neurochem.*, 2018, **145**, 449–463.
- 158 M. Sheng and E. Kim, *J. Cell Sci.*, 2000, **113 (Pt 1)**, 1851–1856.
- 159 J. Li, A. Mahajan and M.-D. Tsai, *Biochemistry*, 2006, **45**, 15168–15178.
- 160 F. Ye and M. Zhang, *Biochem. J.*, 2013, **455**, 1–14.
- 161 A. S. Shaw and E. L. Filbert, *Nat Rev Immunol*, 2009, **9**, 47–56.
- 162 S. Naisbitt, K. Eunjoon, J. C. Tu, B. Xiao, C. Sala, J. Valtschanoff, R. J. Weinberg, P. F. Worley and M. Sheng, *Neuron*, 1999, **23**, 569–582.

- 163 C. M. Durand, C. Betancur, T. M. Boeckers, J. Bockmann, P. Chaste, F. Fauchereau, G. Nygren, M. Rastam, I. C. Gillberg, H. Anckarsäter, E. Sponheim, H. Goubran-Botros, R. Delorme, N. Chabane, M. C. Mouren-Simeoni, P. De Mas, E. Bieth, B. Rogé, D. Héron, L. Burglen, C. Gillberg, M. Leboyer and T. Bourgeron, *Nat. Genet.*, 2007, **39**, 25–27.
- 164 E. Kim, S. Naisbitt, Y. Hsueh, A. Rao, A. Rothschild, A. M. Craig and M. Sheng, *J. Cell Biol.*, 1997, **136**, 669–678.
- 165 J. C. Tu, B. Xiao, S. Naisbitt, J. P. Yuan, R. S. Petralia, P. Brakeman, A. Doan, V. K. Aakalu, A. A. Lanahan, M. Sheng and P. F. Worley, *Neuron*, 1999, **23**, 583–592.
- 166 D. Grillo-Bosch, D. Choquet and M. Sainlos, *Drug Discov. Today Technol.*, 2013, **10**, e531–e540.
- 167 C. Nourry, S. G. N. Grant and J.-P. Borg, *Sci. STKE*, 2003, **2003**, re7 LP-re7.
- 168 I. Yao, Y. Hata, K. Hirao, M. Deguchi, N. Ide, M. Takeuchi and Y. Takai, *J. Biol. Chem.*, 1999, **274**, 27463–27466.
- 169 D. Grillo-Bosch, D. Choquet and M. Sainlos, *Drug Discov. Today Technol.*, 2013, **10**, e531–e540.
- 170 X. Liu and E. J. Fuentes, in *International Review of Cell and Molecular Biology*, ed. L. Galluzzi, Academic Press, 2019, vol. 343, pp. 129–218.
- 171 E. Park, M. Na, J. Choi, S. Kim, J. R. Lee, J. Yoon, D. Park, M. Sheng and E. Kim, *J. Biol. Chem.*, 2003, **278**, 19220–19229.
- 172 J. Liu, M. Liu, B. Zheng, Z. Yao and J. Xia, *PLoS One*, 2016, **11**, 1–12.
- 173 M. Zeng, Y. Shang, T. Guo, Q. He, W.-H. Yung, K. Liu and M. Zhang, *Proc. Natl. Acad. Sci.*, 2016, **113**, E3081–E3090.
- 174 C. W. Wood, A. A. Ibarra, G. J. Bartlett, A. J. Wilson, D. N. Woolfson and R. B. Sessions, *Bioinformatics*, DOI:10.1093/bioinformatics/btaa026.
- 175 S. McIntosh-Smith, J. Price, R. B. Sessions and A. A. Ibarra, *Int. J. High Perform. Comput. Appl.*, 2015, **29**, 119–134.
- 176 T. Fröhlking, M. Bernetti, N. Calonaci and G. Bussi, *J. Chem. Phys.*, 2020, **152**, 230902.
- 177 Y. J. Im, G. B. Kang, J. H. Lee, K. R. Park, H. E. Song, E. Kim, W. K. Song, D. Park and S. H. Eom, *J. Mol. Biol.*, 2010, **397**, 457–466.
- 178 B. Z. Harris and W. A. Lim, *J. Cell Sci.*, 2001, **114**, 3219–3231.
- 179 R. Tonikian, Y. Zhang, S. L. Sazinsky, B. Currell, J.-H. Yeh, B. Reva, H. A. Held, B. A. Appleton, M. Evangelista, Y. Wu, X. Xin, A. C. Chan, S. Seshagiri, L. A. Lasky, C. Sander, C. Boone, G. D. Bader and S. S. Sidhu, *PLoS Biol.*, 2008, **6**, e239–e239.
- 180 J. S. Nowick, D. M. Chung, K. Maitra, S. Maitra, K. D. Stigers and Y. Sun, *J. Am. Chem. Soc.*, 2000, **122**, 7654–7661.
- 181 D. Bojadzic and P. Buchwald, *Curr. Top. Med. Chem.*, 2018, **18**, 674–699.

- 182 G. Zinzalla and D. E. Thurston, *Future Med. Chem.*, 2009, **1**, 65–93.
- 183 D. E. Scott, A. R. Bayly, C. Abell and J. Skidmore, *Nat. Rev. Drug Discov.*, 2016, **15**, 533–550.
- 184 A. M. Watkins and P. S. Arora, *ACS Chem. Biol.*, 2014, **9**, 1747–1754.
- 185 J. L. Lau and M. K. Dunn, *Bioorg. Med. Chem.*, 2018, **26**, 2700–2707.
- 186 H. Bruzzoni-Giovanelli, V. Alezra, N. Wolff, C. Z. Dong, P. Tuffery and A. Rebollo, *Drug Discov. Today*, 2018, **23**, 272–285.
- 187 M. Erak, K. Bellmann-Sickert, S. Els-Heindl and A. G. Beck-Sickinger, *Bioorganic Med. Chem.*, 2018, **26**, 2759–2765.
- 188 L. Di, *AAPS J.*, 2015, **17**, 134–143.
- 189 J. L. Lau and M. K. Dunn, *Bioorganic Med. Chem.*, 2018, **26**, 2700–2707.
- 190 E. Biron, J. Chatterjee, O. Ovadia, D. Langenegger, J. Brueggen, D. Hoyer, H. A. Schmid, R. Jelinek, C. Gilon, A. Hoffman and H. Kessler, *Angew. Chemie Int. Ed.*, 2008, **47**, 2595–2599.
- 191 D. J. Hughes, C. Tiede, N. Penswick, A. A.-S. Tang, C. H. Trinh, U. Mandal, K. Z. Zajac, T. Gaule, G. Howell, T. A. Edwards, J. Duan, E. Feyfant, M. J. McPherson, D. C. Tomlinson and A. Whitehouse, *Sci. Signal.*, , DOI:10.1126/scisignal.aaj2005.
- 192 M. S. Zambrano-Mila, K. E. S. Blacio and N. S. Vispo, *Ther. Innov. Regul. Sci.*, 2020, **54**, 308–317.
- 193 S. Berger, E. Procko, D. Margineantu, E. F. Lee, B. W. Shen, A. Zelter, D. A. Silva, K. Chawla, M. J. Herold, J. M. Garnier, R. Johnson, M. J. Maccoss, G. Lessene, T. N. Davis, P. S. Stayton, B. L. Stoddard, W. D. Fairlie, D. M. Hockenbery and D. Baker, *Elife*, 2016, **5**, 1–31.
- 194 C. W. Wood, J. W. Heal, A. R. Thomson, G. J. Bartlett, A. Á. Ibarra, R. L. Brady, R. B. Sessions and D. N. Woolfson, *Bioinformatics*, 2017, **33**, 3043–3050.
- 195 J. M. Fletcher, K. A. Horner, G. J. Bartlett, G. G. Rhys, A. J. Wilson and D. N. Woolfson, *Chem. Sci.*, 2018, **9**, 7656–7665.
- 196 I. Massova and P. A. Kollman, *J. Am. Chem. Soc.*, 1999, **121**, 8133–8143.
- 197 K. A. Barlow, S. Ó Conchúir, S. Thompson, P. Suresh, J. E. Lucas, M. Heinonen and T. Kortemme, *J. Phys. Chem. B*, 2018, **122**, 5389–5399.
- 198 D. Rooklin, C. Wang, J. Katigbak, P. S. Arora and Y. Zhang, *J. Chem. Inf. Model.*, 2015, **55**, 1585–1599.
- 199 D. Rooklin, A. E. Modell, H. Li, V. Berdan, P. S. Arora and Y. Zhang, *J. Am. Chem. Soc.*, 2017, **139**, 15560–15563.
- 200 A. Kotschy, Z. Szlavik, J. Murray, J. Davidson, A. L. Maragno, G. Le Toumelin-Braizat, M. Chanrion, G. L. Kelly, J.-N. Gong, D. M. Moujalled, A. Bruno, M. Csekei, A. Paczal, Z. B. Szabo, S. Sipos, G. Radics, A. Prosenyak, B. Balint, L. Ondi, G. Blasko, A. Robertson, A. Surgenor, P. Dokurno, I. Chen, N. Matassova, J. Smith, C. Pedder, C. Graham, A. Studeny, G. Lysiak-Auvity, A.-M. Girard, F. Gravé, D. Segal, C. D. Riffkin,

- G. Pomilio, L. C. A. Galbraith, B. J. Aubrey, M. S. Brennan, M. J. Herold, C. Chang, G. Guasconi, N. Cauquil, F. Melchiorre, N. Guigal-Stephan, B. Lockhart, F. Colland, J. A. Hickman, A. W. Roberts, D. C. S. Huang, A. H. Wei, A. Strasser, G. Lessene and O. Geneste, *Nature*, 2016, **538**, 477–482.
- 201 A. Kötter, H. D. Mootz and A. Heuer, *J. Chem. Theory Comput.*, 2019, **15**, 6403–6410.
- 202 J. R. Gareau and C. D. Lima, *Nat. Rev. Mol. Cell Biol.*, 2010, **11**, 861–871.
- 203 A. T. Namanja, Y.-J. Li, Y. Su, S. Wong, J. Lu, L. T. Colson, C. Wu, S. S. C. Li and Y. Chen, *J. Biol. Chem.*, 2012, **287**, 3231–3240.
- 204 E. Oda, R. Ohki, H. Murasawa, J. Nemoto, T. Shibue, T. Yamashita, T. Tokino, T. Taniguchi and † Nobuyuki Tanaka, *Science (80-.)*, 2000, **288**, 1053 LP – 1058.
- 205 L. Zhang, H. Lopez, N. M. George, X. Liu, X. Pang and X. Luo, *Cell Death Differ.*, 2011, **18**, 864–873.
- 206 C. Ploner, R. Kofler and A. Villunger, *Oncogene*, 2008, **27**, S84–S92.
- 207 J. Montero, C. Gstalder, D. J. Kim, D. Sadowicz, W. Miles, M. Manos, J. R. Cidado, J. Paul Secrist, A. E. Tron, K. Flaherty, F. Stephen Hodi, C. H. Yoon, A. Letai, D. E. Fisher and R. Haq, *Nat. Commun.*, 2019, **10**, 5157.
- 208 R. [Jürgen Dohmen], *Biochim. Biophys. Acta - Mol. Cell Res.*, 2004, **1695**, 113–131.
- 209 S. P. Jackson and D. Durocher, *Mol. Cell*, 2013, **49**, 795–807.
- 210 G. J. K. Praefcke, K. Hofmann and R. J. Dohmen, *Trends Biochem. Sci.*, 2012, **37**, 23–31.
- 211 Z. J. Han, Y. H. Feng, B. H. Gu, Y. M. Li and H. Chen, *Int. J. Oncol.*, 2018, **52**, 1081–1094.
- 212 P. C. Lyu, J. C. Sherman, A. Chen and N. R. Kallenbach, *Proc. Natl. Acad. Sci. U. S. A.*, 1991, **88**, 5317–5320.
- 213 M. Ivan, K. Kondo, H. Yang, W. Kim, J. Valiando, M. Ohh, A. Salic, J. M. Asara, W. S. Lane and W. G. Kaelin, *Science (80-.)*, 2001, **292**, 464–468.
- 214 H. J. Choi, B.-J. Song, Y.-D. Gong, W. J. Gwak and Y. Soh, *Br. J. Pharmacol.*, 2008, **154**, 114–125.
- 215 P. C. Mahon, K. Hirota and G. L. Semenza, *Genes Dev.*, 2001, **15**, 2675–2686.
- 216 F. Wang, C. B. Marshall and M. Ikura, *Cell. Mol. Life Sci.*, 2013, **70**, 3989–4008.
- 217 M. S. Wiesener, J. A. N. S. Ju, C. Warnecke, S. Mandriota, I. Bechmann, U. A. Frei, C. W. Pugh and K. Eckardt, *FASEB J.*, 2003, **17**, 271–273.
- 218 G. L. Semenza, *Biochem. Pharmacol.*, 2002, **64**, 993–998.
- 219 A. L. Harris, *Nat. Rev. Cancer*, 2002, **2**, 38–47.

- 220 M. Höckel and P. Vaupel, *Semin. Oncol.*, 2001, **28**, 36–41.
- 221 J. C. Jun, A. Rathore, H. Younas, D. Gilkes and V. Y. Polotsky, *Curr. sleep Med. reports*, 2017, **3**, 1–10.
- 222 J. Frost, C. Galdeano, P. Soares, M. S. Gadd, K. M. Grzes, L. Ellis, O. Epemolu, S. Shimamura, M. Bantscheff, P. Grandi, K. D. Read, D. A. Cantrell, S. Rocha and A. Ciulli, *Nat. Commun.*, 2016, **7**, 13312.
- 223 G. L. Semenza, *Oncogene*, 2010, **29**, 625–634.
- 224 S. A. Dames, M. Martinez-Yamout, R. N. De Guzman, H. J. Dyson and P. E. Wright, *Proc. Natl. Acad. Sci.*, 2002, **99**, 5271–5276.
- 225 S. Kushal, B. B. Lao, L. K. Henchey, R. Dubey, H. Mesallati, N. J. Traaseth, B. Z. Olenyuk and P. S. Arora, *Proc. Natl. Acad. Sci.*, 2013, **110**, 15602–15607.
- 226 B. B. Lao, I. Grishagin, H. Mesallati, T. F. Brewer, B. Z. Olenyuk and P. S. Arora, *Proc. Natl. Acad. Sci.*, 2014, **111**, 7531–7536.
- 227 K. Beppu, K. Nakamura, W. M. Linehan, A. Rapisarda and C. J. Thiele, *Cancer Res.*, 2005, **65**, 4775–4781.
- 228 H. Zhang, D. Z. Qian, Y. S. Tan, K. Lee, P. Gao, Y. R. Ren, S. Rey, H. Hammers, D. Chang, R. Pili, C. V Dang, J. O. Liu and G. L. Semenza, *Proc. Natl. Acad. Sci.*, 2008, **105**, 19579–19586.
- 229 P. Sapra, H. Zhao, M. Mehlig, J. Malaby, P. Kraft, C. Longley, L. M. Greenberger and I. D. Horak, *Clin. Cancer Res.*, 2008, **14**, 1888–1896.
- 230 K. Lee, H. Zhang, D. Z. Qian, S. Rey, J. O. Liu and G. L. Semenza, *Proc. Natl. Acad. Sci.*, 2009, **106**, 17910–17915.
- 231 E. Miranda, I. K. Nordgren, A. L. Male, C. E. Lawrence, F. Hoakwie, F. Cuda, W. Court, K. R. Fox, P. A. Townsend, G. K. Packham, S. A. Eccles and A. Tavassoli, *J. Am. Chem. Soc.*, 2013, **135**, 10418–10425.
- 232 M. C. Chan, J. P. Holt-Martyn, C. J. Schofield and P. J. Ratcliffe, *Mol. Aspects Med.*, 2016, **47–48**, 54–75.
- 233 A. L. Kung, S. D. Zabudoff, D. S. France, S. J. Freedman, E. A. Tanner, A. Vieira, S. Cornell-Kennon, J. Lee, B. Wang, J. Wang, K. Memmert, H.-U. Naegeli, F. Petersen, M. J. Eck, K. W. Bair, A. W. Wood and D. M. Livingston, *Cancer Cell*, 2004, **6**, 33–43.
- 234 K. M. Cook, S. T. Hilton, J. Mecinovic, W. B. Motherwell, W. D. Figg and C. J. Schofield, 2009, **284**, 26831–26838.
- 235 C. Tan, R. G. [de Noronha], N. S. Devi, A. A. Jabbar, S. Kaluz, Y. Liu, S. R. Mooring, K. C. Nicolaou, B. Wang and E. G. [Van Meir], *Bioorg. Med. Chem. Lett.*, 2011, **21**, 5528–5532.
- 236 S. Yin, S. Kaluz, N. S. Devi, A. A. Jabbar, R. G. de Noronha, J. Mun, Z. Zhang, P. R. Boreddy, W. Wang, Z. Wang, T. Abbruscato, Z. Chen, J. J. Olson, R. Zhang, M. M. Goodman, K. C. Nicolaou and E. G. Van Meir, *Clin. Cancer Res.*, 2012, **18**, 6623–6633.
- 237 G. M. Burslem, H. F. Kyle, A. L. Breeze, T. A. Edwards, A. Nelson, S. L. Warriner and A. J. Wilson, *ChemBioChem*, 2014, **15**, 1083–1087.

- 238 G. M. Burslem, H. F. Kyle, A. L. Breeze, T. A. Edwards, A. Nelson, S. L. Warriner and A. J. Wilson, *ChemBioChem*, 2014, **15**, 1083–1087.
- 239 G. M. Burslem, H. F. Kyle, A. L. Breeze, T. A. Edwards, A. Nelson, S. L. Warriner and A. J. Wilson, *Chem. Commun.*, 2016, **52**, 5421–5424.
- 240 H. Zhao and A. Caflisch, *Eur. J. Med. Chem.*, 2015, **91**, 4–14.
- 241 H. Geng, F. Chen, J. Ye and F. Jiang, *Comput. Struct. Biotechnol. J.*, 2019, **17**, 1162–1170.
- 242 J. A. McCammon, B. R. Gelin and M. Karplus, *Nature*, 1977, **267**, 585–590.
- 243 C. S. Neuhaus, G. Gabernet, C. Steuer, K. Root, J. A. Hiss, R. Zenobi and G. Schneider, *Angew. Chem. Int. Ed. Engl.*, 2019, **58**, 1674–1678.
- 244 M. Garofalo, G. Grazioso, A. Cavalli and J. Sgrignani, *Molecules*, 2020, **25**, 1756.
- 245 F. Jiang and H. Geng, *Methods Mol. Biol.*, 2019, **2001**, 61–71.
- 246 Y. S. Tan, D. P. Lane and C. S. Verma, *Drug Discov. Today*, 2016, **21**, 1642–1653.
- 247 Z. Ghaemi, D. Alberga, P. Carloni, A. Laio and G. Lattanzi, *J. Chem. Theory Comput.*, 2016, **12**, 4093–4099.
- 248 G. Gabernet, D. Gautschi, A. T. Müller, C. S. Neuhaus, L. Armbrecht, P. S. Dittrich, J. A. Hiss and G. Schneider, *Sci. Rep.*, 2019, **9**, 11282.
- 249 A. Glas, D. Bier, G. Hahne, C. Rademacher, C. Ottmann and T. N. Grossmann, *Angew. Chem. Int. Ed. Engl.*, 2014, **53**, 2489–2493.
- 250 T. L. Joseph, D. Lane and C. S. Verma, *Cell Cycle*, 2010, **9**, 4560–4568.
- 251 N. J. Greenfield, *Nat. Protoc.*, 2006, **1**, 2876–2890.
- 252 E. Kaiser, R. L. Colescott, C. D. Bossinger and P. I. Cook, *Anal. Biochem.*, 1970, **34**, 595–8.

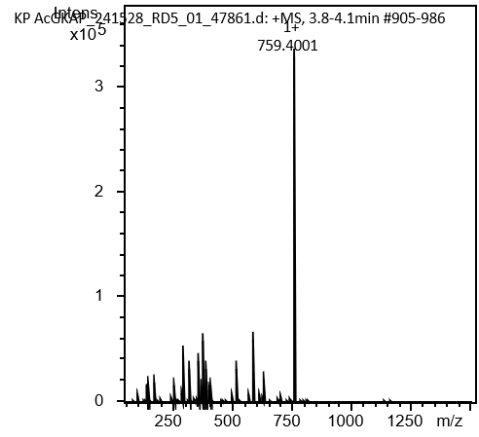
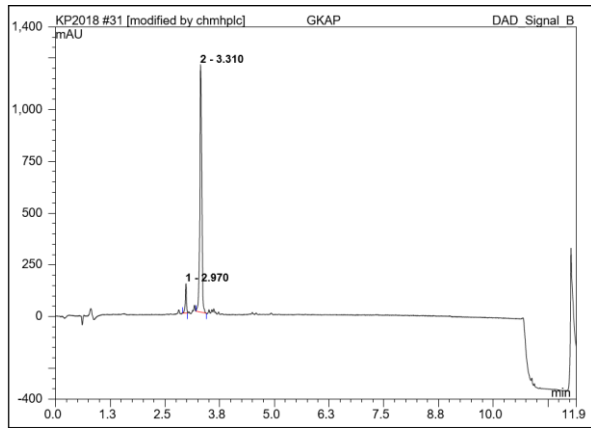
Appendix

A.1 Chapter 2 Peptide Analytical Data

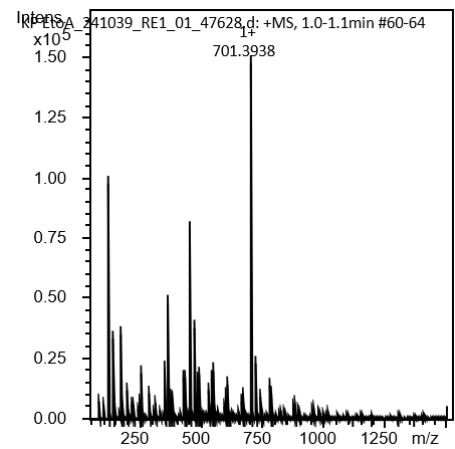
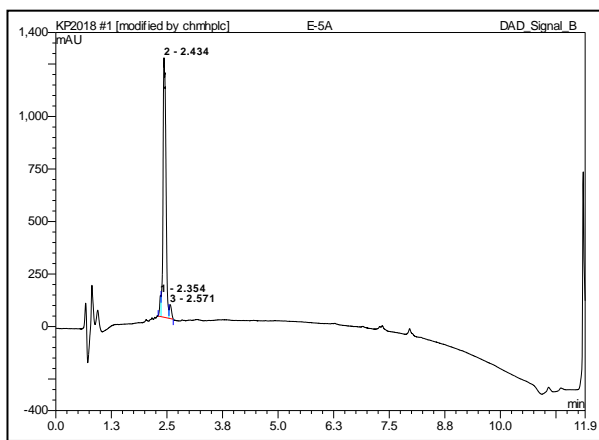
Tabulated HRMS data of synthesised peptides are shown in the table below. Peptide identity was confirmed by the inspection of multiple charge states and are quoted as the monoisotopic peak for the Expected (Exp^d) and Observed (Obs^d) masses.

Peptide	Sequence	[M+1H] ¹⁺	[M+1H] ⁴⁺	Purity (%)
		Exp ^d	Obs ^d	
Wt-GKAP	Ac-EAQTRL-OH	759.39	759.40	96
E1A	Ac-AAQTRL-OH	701.38	701.39	94
A2G	Ac-EGQTRL-OH	745.37	744.37	100
Q3A	Ac-EAATRL-OH	702.37	702.37	95
T4A	Ac-EAQARL-OH	729.38	729.38	95
R5A	Ac-EAQTAL-OH	674.32	674.33	96
L6A	Ac-EAQTRA-OH	717.34	717.35	100
CONH ₂	Ac-EAQTRL-NH ₂	758.40	757.40	94
A2F	Ac-EDQTRL-OH	802.38	803.38	96
A2D	Ac-EFQTRL-OH	834.42	835.42	98
R5W	Ac-EAQTWL-OH	788.37	789.37	100
L6F	Ac-EAQTAFF-OH	792.37	792.35	97
Hao-GKAP	Ac-Hao-EAQTRL-COOH	1108.53	1108.59	98

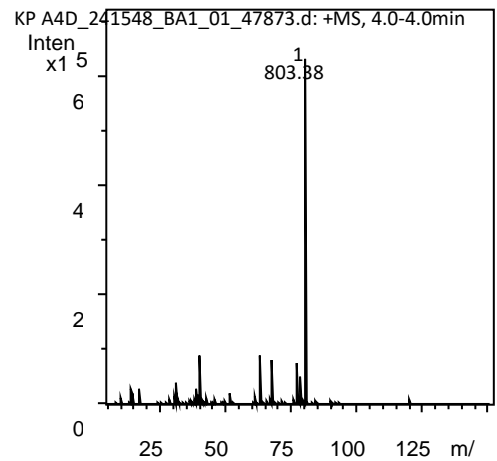
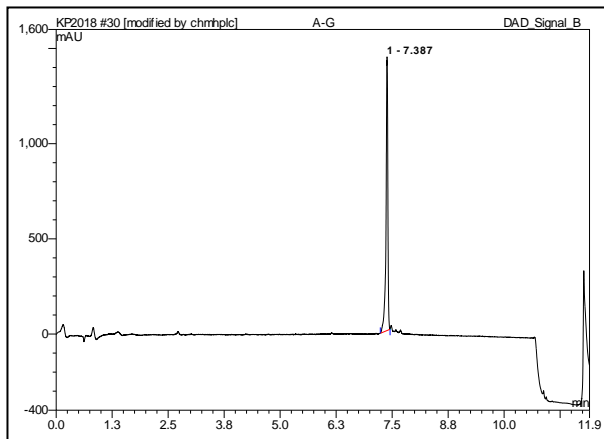
Wt-GKAP



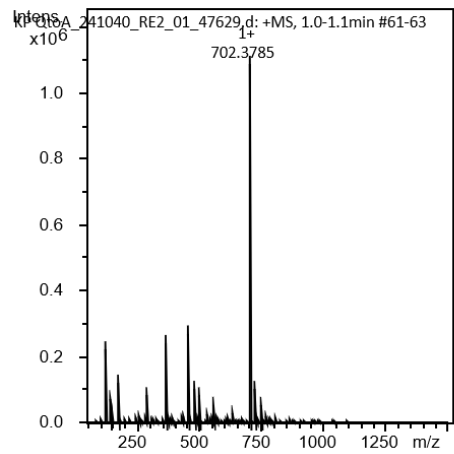
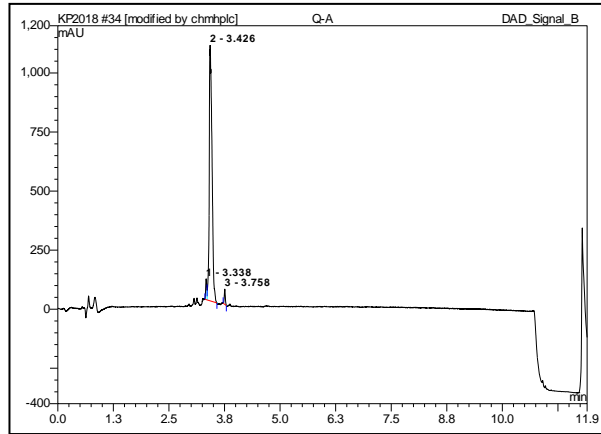
E1A



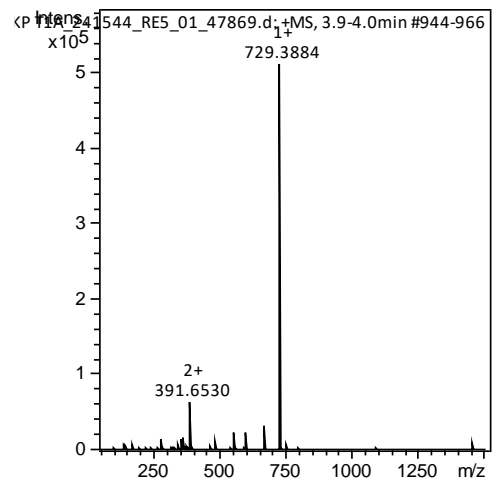
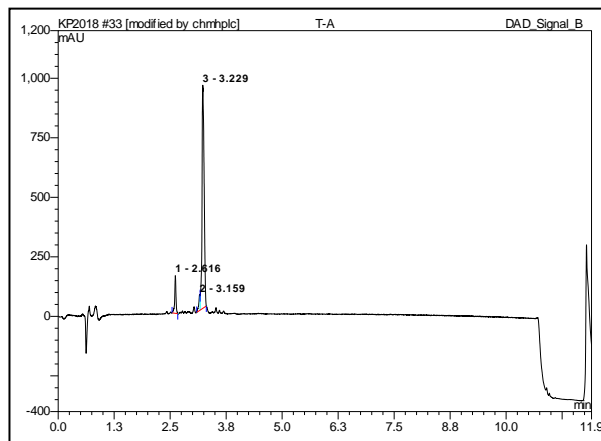
A2G



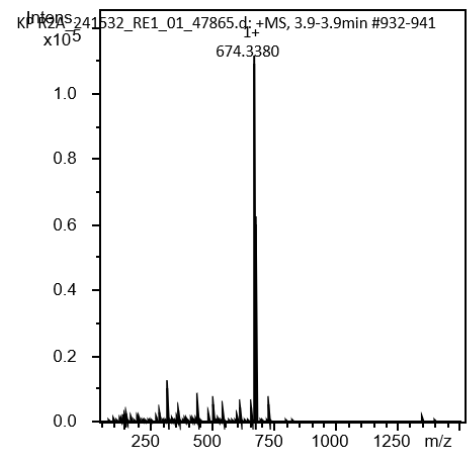
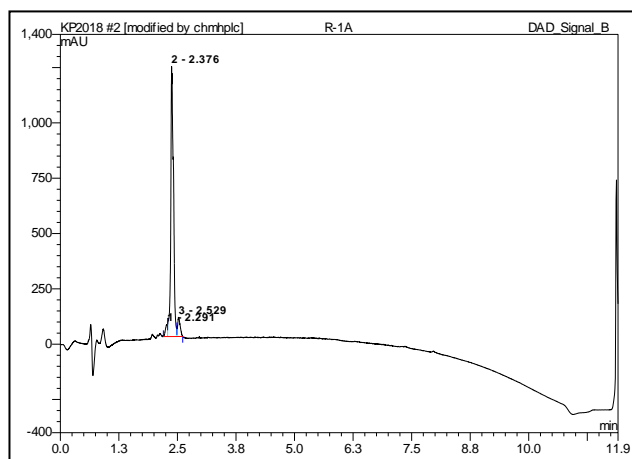
Q3A



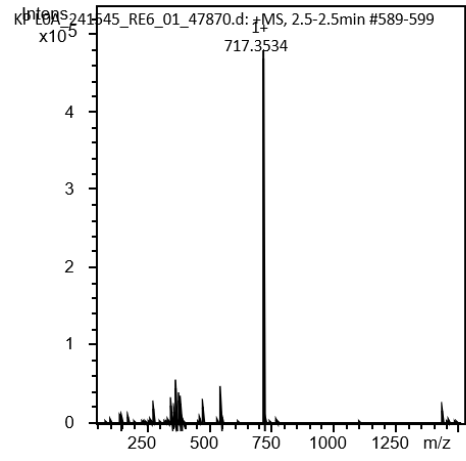
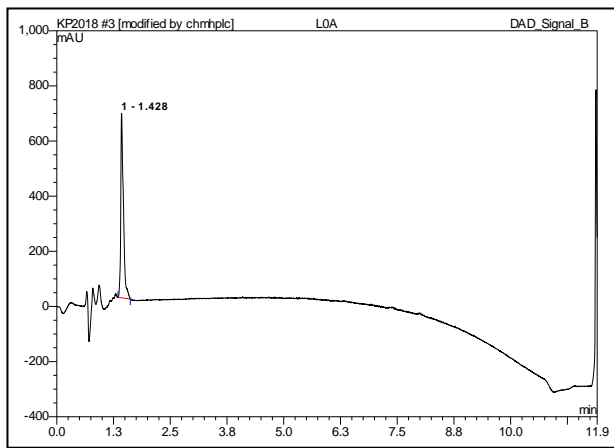
T4A



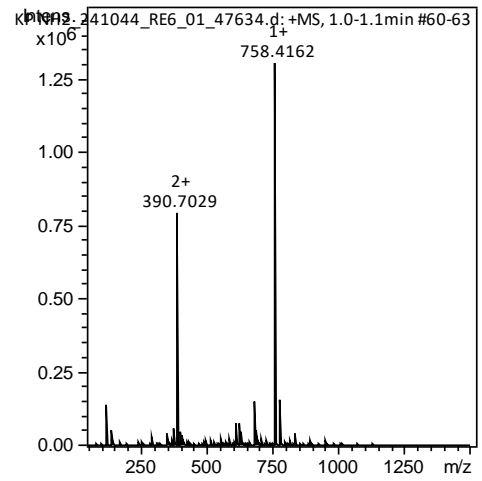
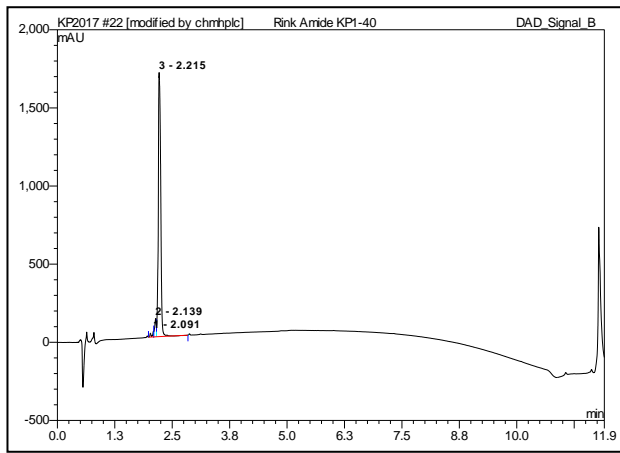
R5A



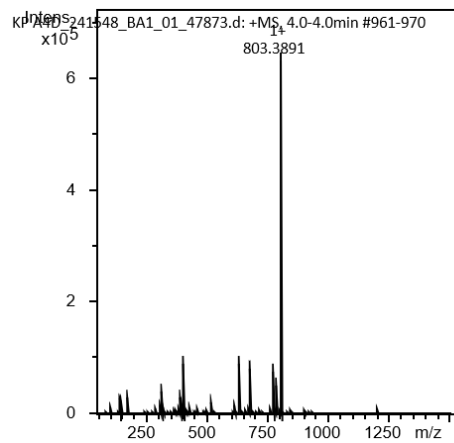
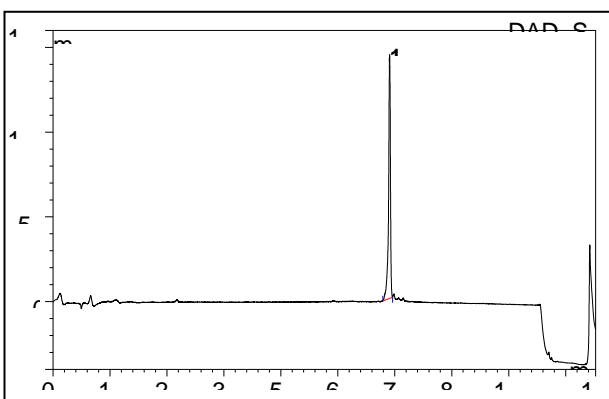
L6A



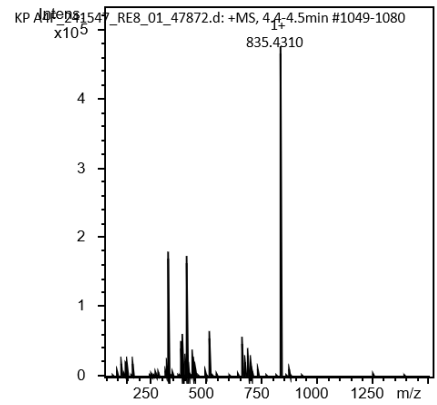
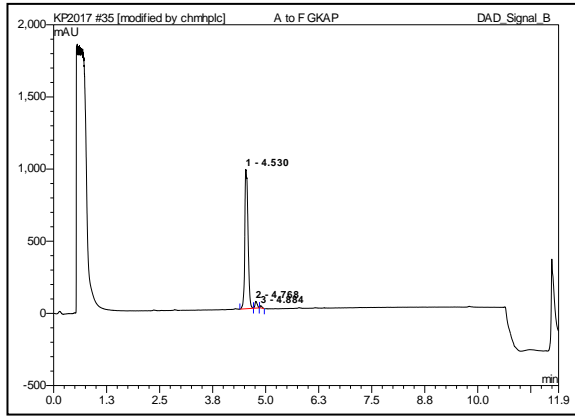
CONH₂



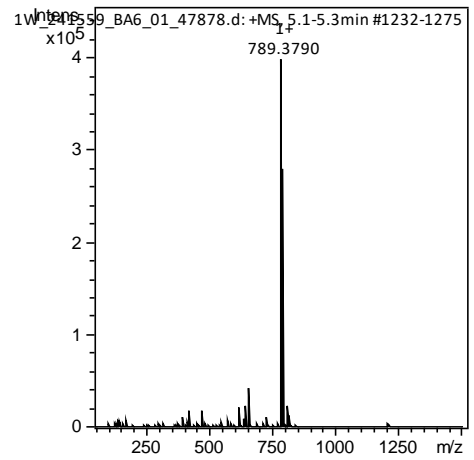
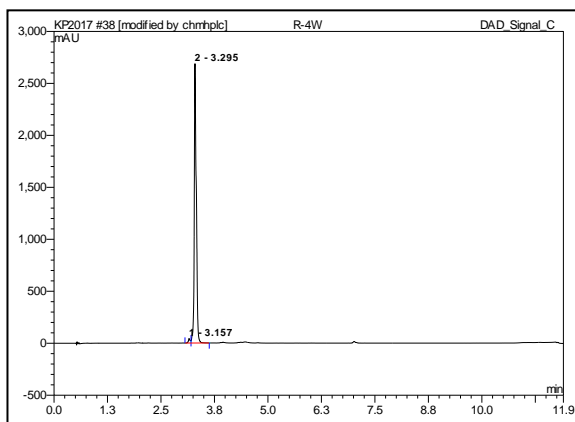
A2D



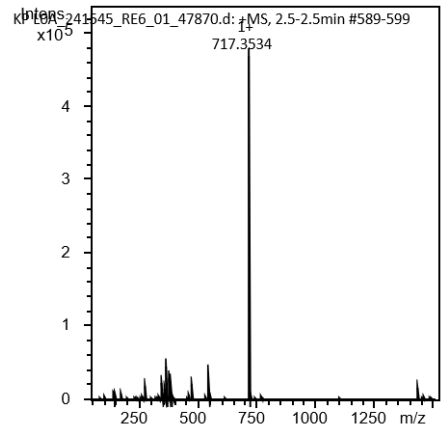
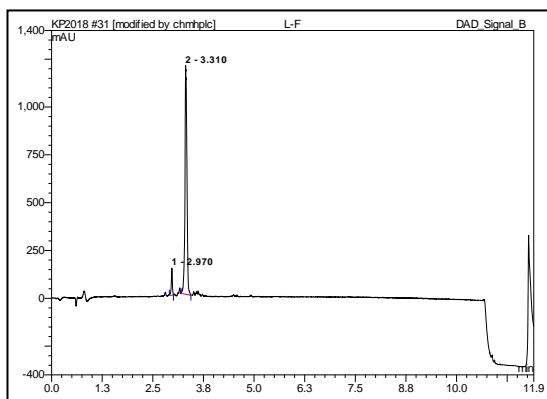
A2F



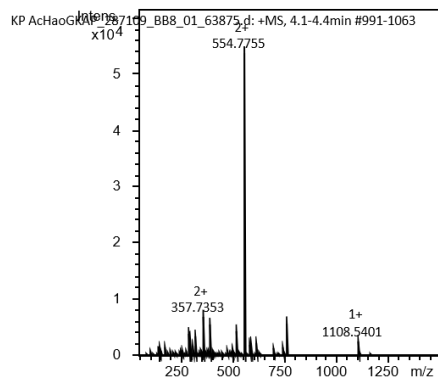
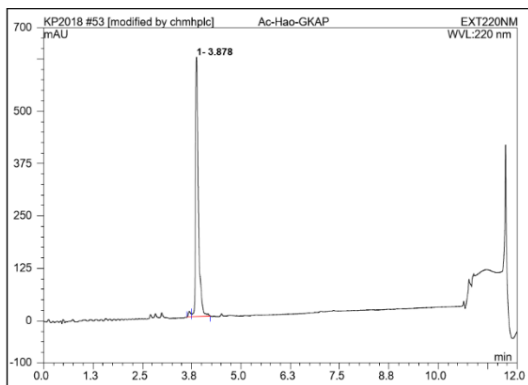
R5W



L6F



Ac-Hao-EAQTRL-COOH



A.2 Chapter 3 Peptide Analytical Data

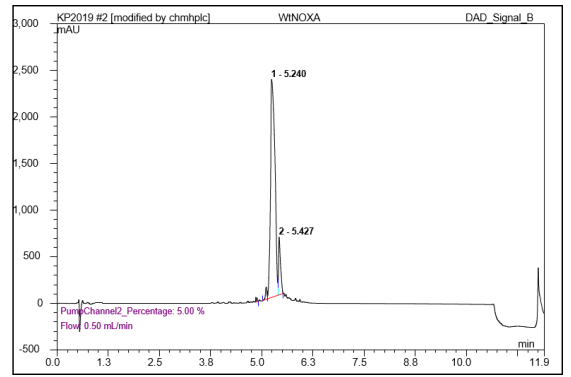
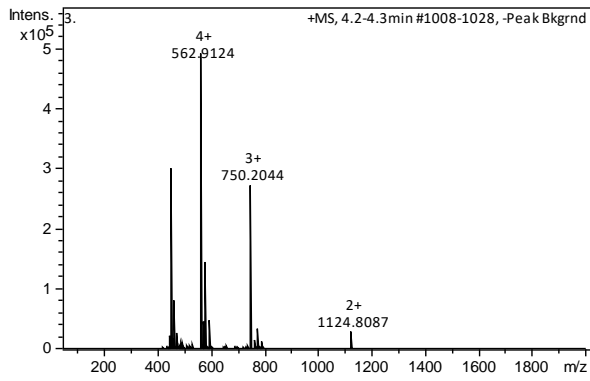
Tabulated HRMS data of synthesised peptides are shown below. Peptide identity was confirmed by the inspection of multiple charge states and are quoted as the monoisotopic peak for the Expected (Exp^d) and Observed (Obs^d) masses.

Peptide	Sequence	[M+4H] ⁴⁺ Exp ^d	[M+4H] ⁴⁺ Obs ^d	Purity (%)
Wt-NOXA ₇₄₋₉₃	AAQLRRIGDKVNLQRKLLN	562.50	562.91	90
NOXA ₇₄₋₉₃ L78F	AAQFRRIGDKVNLQRKLLN	571.60	571.65	94
NOXA ₇₄₋₉₃ L78W	AAQWRRIGDKVNLQRKLLN	580.17	581.09	94
NOXA ₇₄₋₉₃ L78Y	AAQYRRIGDKVNLQRKLLN	575.12	575.34	95
NOXA ₇₄₋₉₃ V85F	AAQLRRIGDKFNLRKLLN	574.58	574.84	95

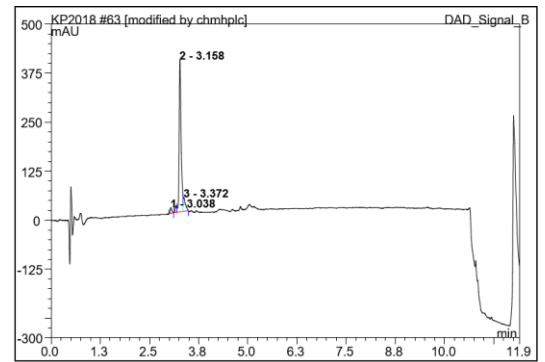
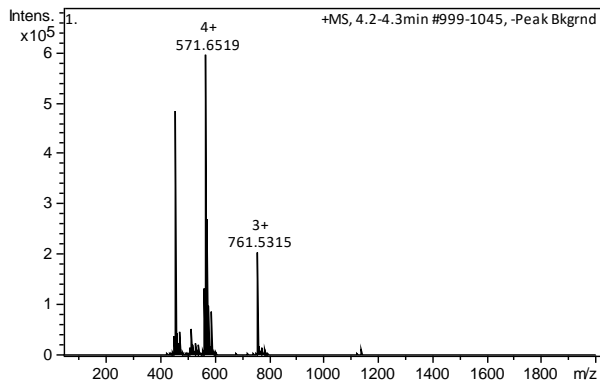
Peptide	Sequence	[M+2H] ²⁺ Exp ^d	[M+2H] ²⁺ Obs ^d	Purity (%)
Wt-SIM ₂₇₀₅₋₂₇₁₇	DNEIEVIIVWEKK	828.45	828.40	100
SIM ₂₇₀₅₋₂₇₁₇ I2708D	DNEDEVIIVWEKK	829.42	829.43	96
SIM ₂₇₀₅₋₂₇₁₇ V2710I	DNEIEIIIVWEKK	835.46	835.52	97
SIM ₂₇₀₅₋₂₇₁₇ I2711E	DNEIEVEIVWEKK	836.43	836.44	98
SIM ₂₇₀₅₋₂₇₁₇ V2713E	DNEIEVIIWEKK	843.44	843.44	91

NOXA peptides

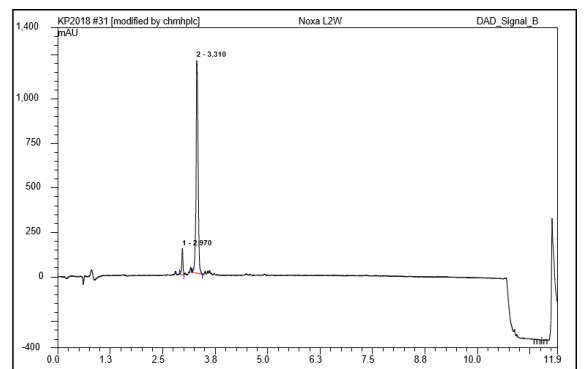
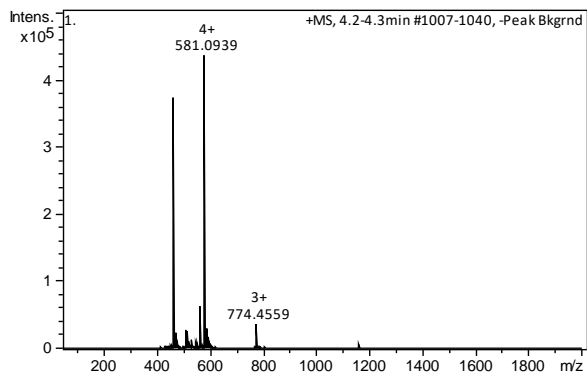
Wt-NOXA



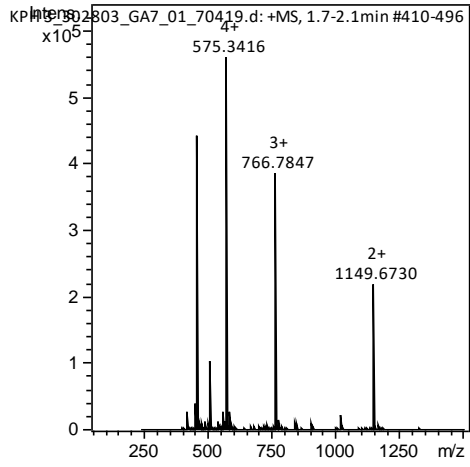
L78F



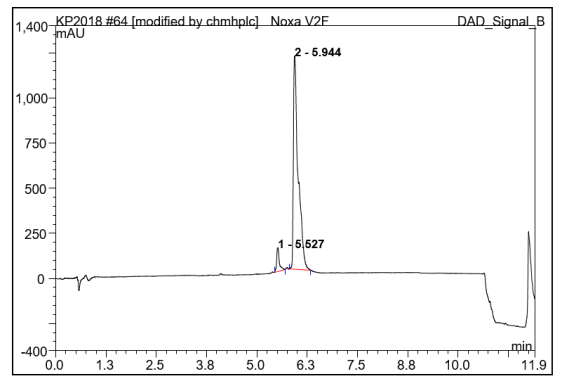
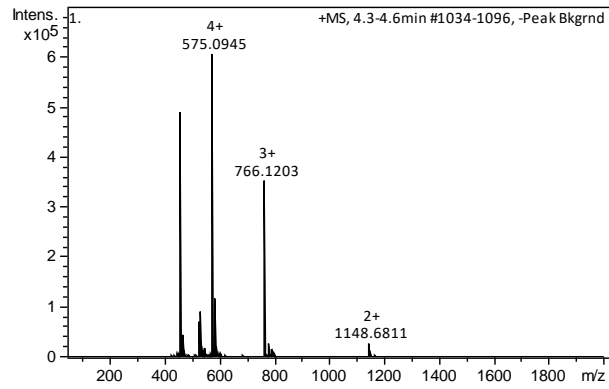
L78W



L78Y

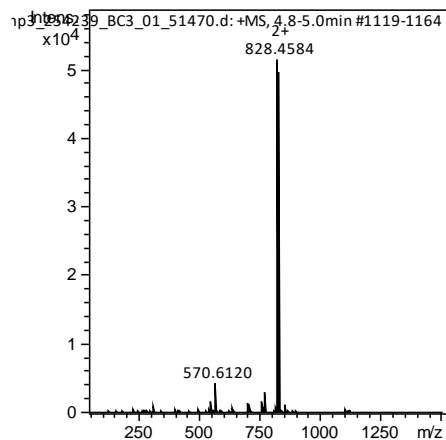
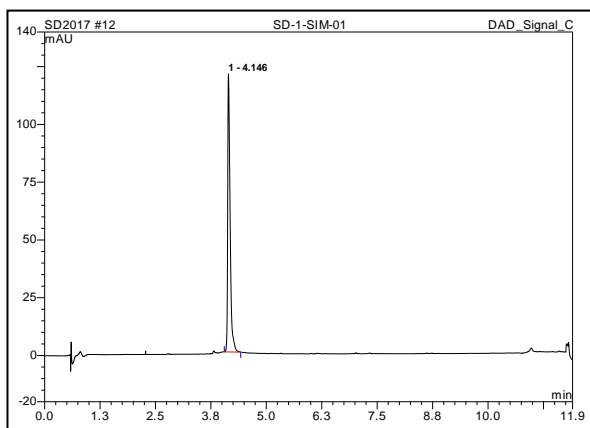


V85F

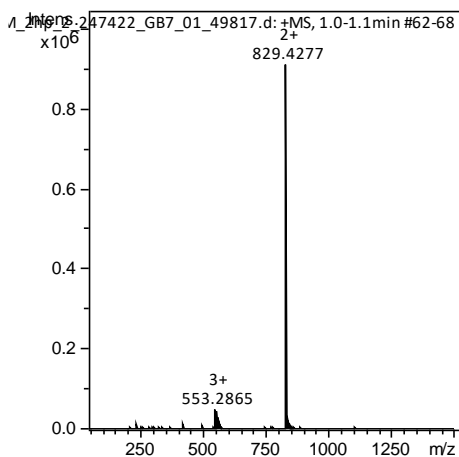
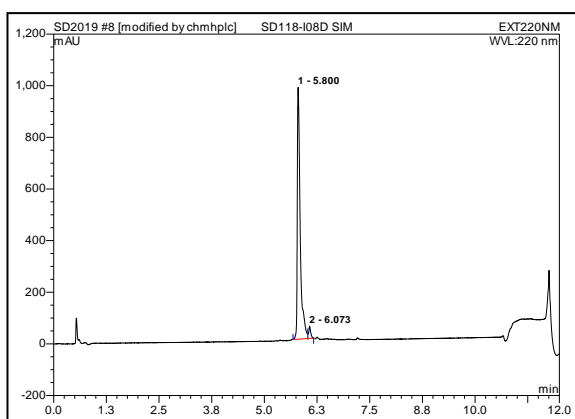


SIM peptides

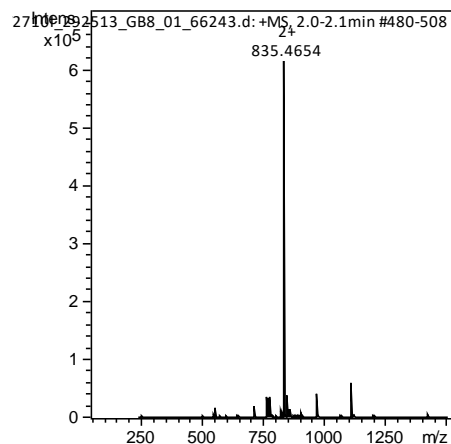
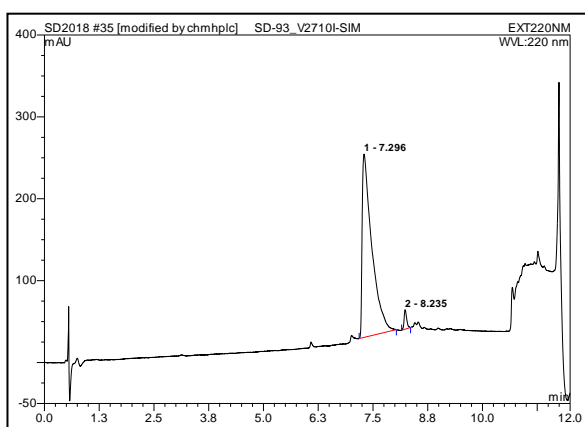
Wt-SIM



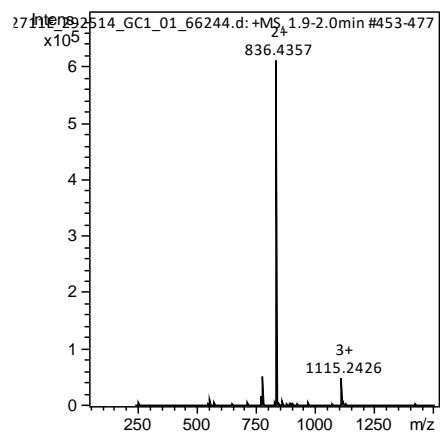
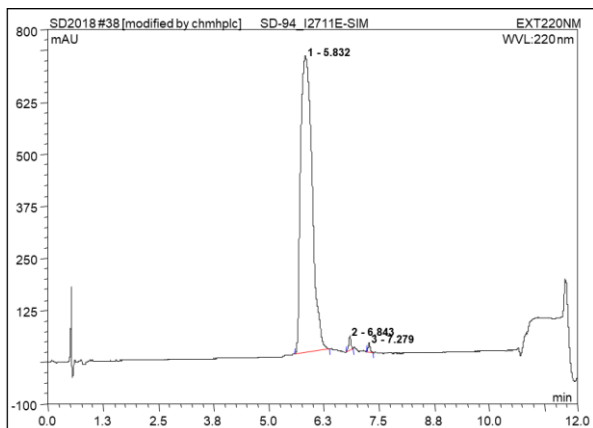
I2708D



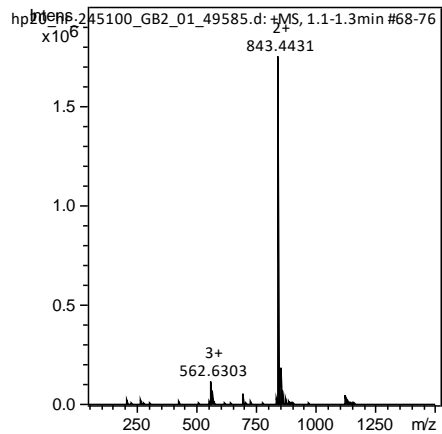
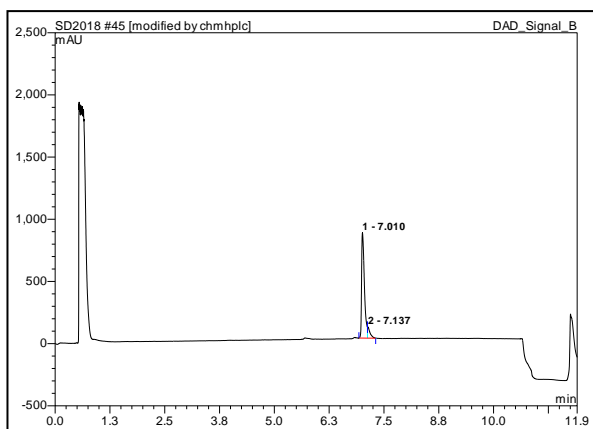
V2710I



I2711E

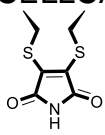

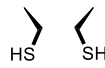
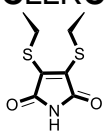

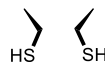


V2713E

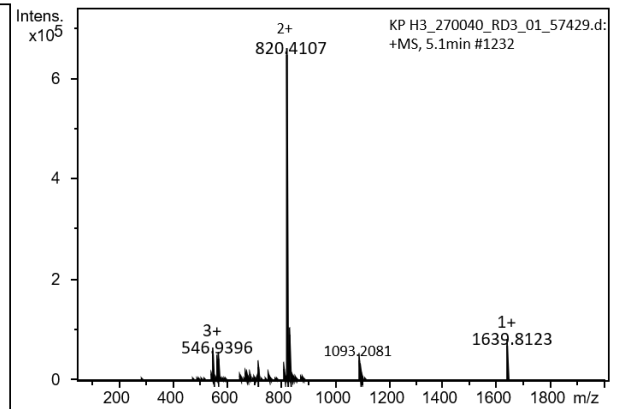
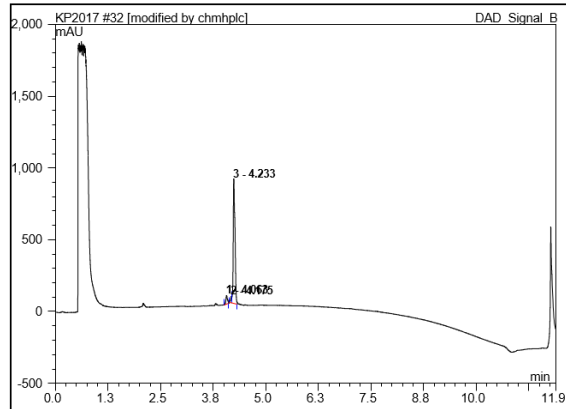


A.3 Chapter 4 Peptide Analytical Data

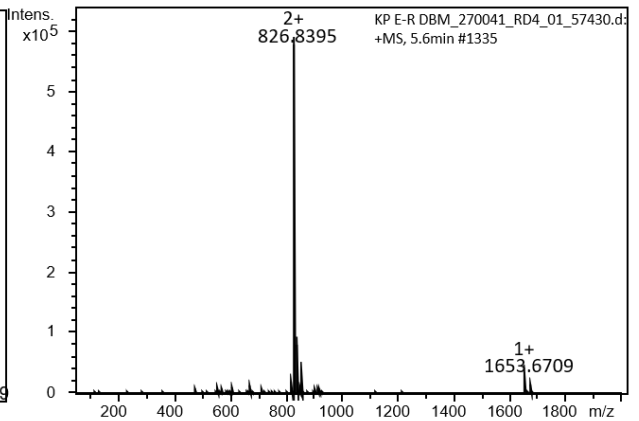
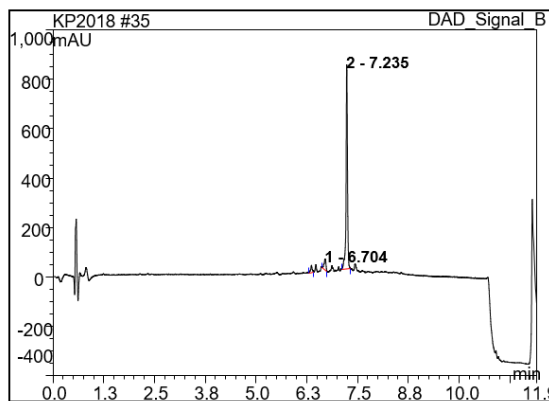
Tabulated HRMS data of synthesised peptides are shown below. Peptide identity was confirmed by the inspection of multiple charge states and are quoted as the monoisotopic peak for the Expected (Exp^d) and Observed (Obs^d) masses. Purity of peptides higher than 90% by analytical HPLC.

Peptide	Sequence	[M+2H] ²⁺ Exp ^d	[M+2H] ²⁺ Obs ^d
wt-HIF-1 α ₈₁₂₋₈₂₆	Ac-LQGEELLRALDQVN-NH ₂	820.39	820.41
HIF-1 α ₈₁₂₋₈₂₆ sE816C-R820C	Ac-LQG C ELL C ALDQVN-NH ₂ 	826.80	826.84
HIF-1 α ₈₁₂₋₈₂₆ oxE816C-R820C	Ac-LQG C ELL C ALDQVN-NH ₂ 	779.40	779.37
HIF-1 α ₈₁₂₋₈₂₆ redE816C-R820C	Ac-LQG C ELL C ALDQVN-NH ₂ 	780.37	780.36
HIF-1 α ₈₁₂₋₈₂₆ sE817C-A821C	Ac-LQGE C LLR C LDQVN-NH ₂ 	869.47	869.40
HIF-1 α ₈₁₂₋₈₂₆ oxE817C-A821C	Ac-LQGE C LLR C LDQVN-NH ₂ 	821.89	821.90
HIF-1 α ₈₁₂₋₈₂₆ redE817C-A821C	Ac-LQGE C LLR C LDQVN-NH ₂ 	822.94	822.89

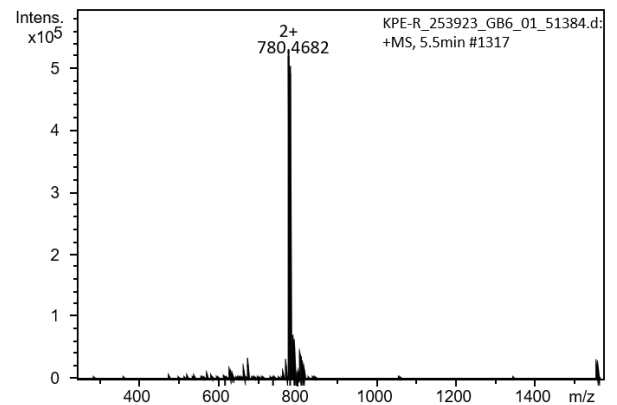
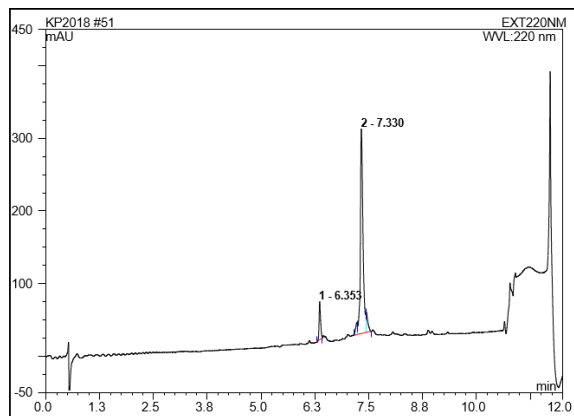
wt- HIF-1 α ₈₁₂₋₈₂₆



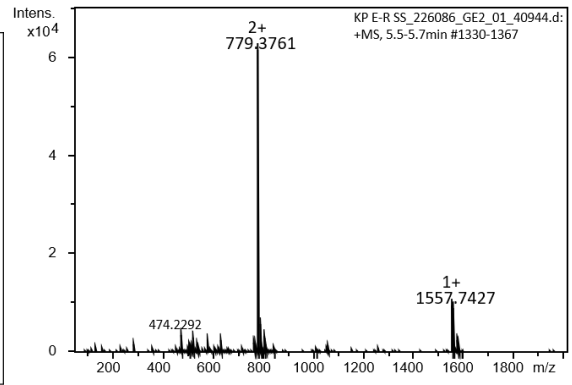
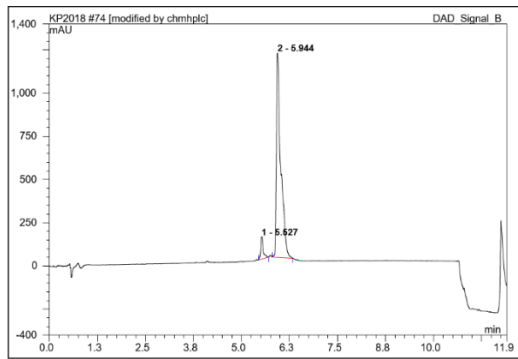
HIF-1 α ₈₁₂₋₈₂₆SE816R820



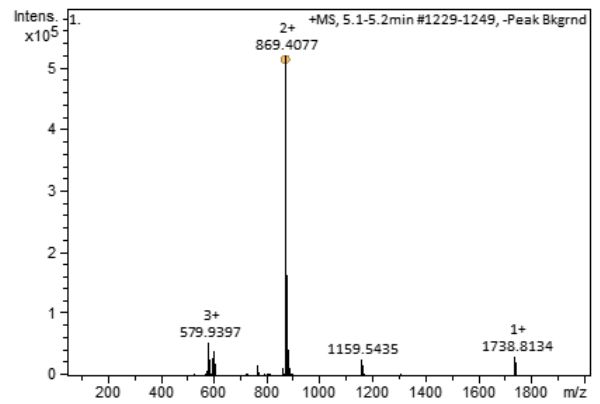
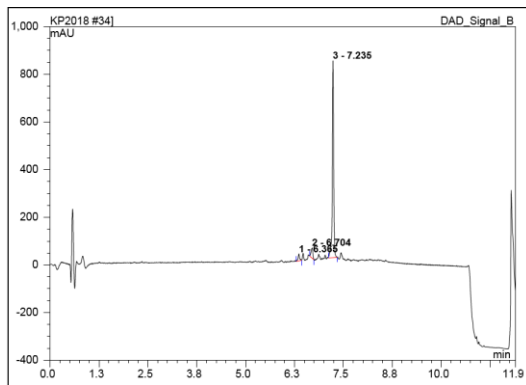
HIF-1 α ₈₁₂₋₈₂₆E816R820



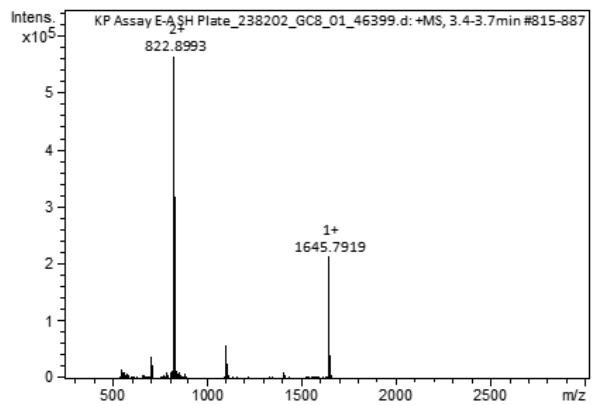
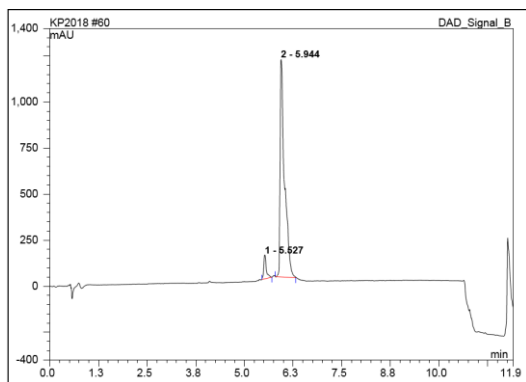
HIF-1 α ₈₁₂₋₈₂₆Ox_{E816R820}



HIF-1 α _{SE817A821}



HIF-1 α _{RedE817A821}



HIF-1 α xE817A821

