

University of Southampton

Faculty of Engineering and Physical Sciences

School of Chemistry

The Development of Cyclic Peptide Inhibitors of Hypoxia-Inducible Factors

By

Reece Michael Gardner

Thesis for the degree of Doctor of Philosophy

November 2020

University of Southampton

Abstract

Faculty of Engineering and Physical Sciences

Chemistry

Thesis for the degree of Doctor of Philosophy

The Development of Cyclic Peptide Inhibitors of Hypoxia-Inducible Factors

Reece Michael Gardner

Hypoxia inducible factors (HIFs) are transcription factors that act as the master regulators of oxygen homeostasis. Diseases such as cancer can disrupt this homeostasis leading to areas with low environmental oxygen (hypoxia) especially in tumour cores. HIFs control the expression of hundreds of genes involved in processes such as metabolism, erythropoiesis, angiogenesis and pH regulation, all of which aid in the adaptation of cells to survive in hypoxic conditions. In the case of cancer this adaptation promotes the progression and metastasis of the disease, leading to HIF becoming a prominent target for therapeutic intervention.

In this work, the structure activity relationship of the cyclic peptide HIF inhibitor cyclo-CLLFVY has been explored. A series of derivatives were synthesised via the incorporation of unnatural amino acids, which led to the development of a new lead peptide that was shown to bind to HIF with a 30-fold greater affinity than cyclo-CLLFVY. The improved peptide leads have been used in the development of several new *in vitro* assays for the quantification of HIF binding and inhibition of HIF activity.

Determination of the structural basis of cyclo-CLLFVY binding was also attempted via an x-ray crystallographic approach, which led to the development of a novel HIF structure.

Table of Contents

Table of Contents	i
Table of Tables	vii
Table of Figures	ix
Research Thesis: Declaration of Authorship	xv
Acknowledgements	xvii
Abbreviations	xix
Chapter 1 Introduction	1
1.1 Hypoxia-Inducible Factor	1
1.1.1 The Role of HIF	1
1.1.2 HIF Structure.....	1
1.1.3 HIF Regulation	5
1.1.3.1 Oxygen-dependent mechanisms of HIF regulation	6
1.1.3.2 Oxygen-Independent mechanisms of HIF regulation	8
1.1.4 HIF Activity	10
1.1.4.1 Angiogenesis	11
1.1.4.2 Metabolism	11
1.1.4.3 Cellular pH.....	12
1.2 Targeting the HIF Pathway.....	12
1.2.1 Inhibitors of DNA binding.....	13
1.2.2 Inhibitors of HIF Transcriptional activity.....	14
1.2.3 Indirect Inhibitors	14
1.2.4 Direct Inhibitors of HIF	16
1.2.5 HIF-2 Inhibitors	17
1.3 Targeting Protein-Protein Interactions.....	20
1.3.1 Peptides as Therapeutics	20
1.3.2 DNA Encoded Methods for the Generation of Cyclic Peptides.....	23
1.3.2.1 Phage Display	23
1.3.2.2 mRNA Display.....	24
1.3.2.3 SICLOPPS.....	26

Table of Contents

1.4	CLLFVY	28
1.5	Project Aims.....	29
Chapter 2 Investigating the Structure Activity Relationship of CLLFVY.....		31
2.1	Synthesis of Cyclo-CLLFVY	31
2.1.1	Solid-Phase Peptide Synthesis	31
2.2	Recombinant Expression of HIF-1 α PAS-B Domain.....	34
2.3	Thermal Shift.....	36
2.4	The Development of an MST Assay	39
2.4.1	Introduction to Microscale Thermophoresis.....	39
2.4.2	Analysis of CLLFVY by MST.....	41
2.5	Alanine Scanning of CLLFVY	43
2.6	Investigating Cysteine Analogues	44
2.7	Investigating the Potential of Disulphide Mediated Binding	46
2.7.1	Dithiothreitol (DTT)	46
2.7.2	Cysteine Mutants	47
2.8	Investigating Phenylalanine Analogues.....	50
2.9	CLLF(CF ₃)VY Alanine Scan	56
2.10	Tyrosine Analogues	57
2.11	Valine Analogues.....	60
2.12	L2 Analogues.....	62
2.13	Investigating Artefacts.....	63
2.14	Introduction of Charged Residues	64
2.15	Investigating HIF Isoform Specificity	67
Chapter 2 Summary.....		68
Chapter 3 Development of Further <i>In Vitro</i> Assays.....		69
3.1	Enzyme-Linked Immunosorbent Assay	69
3.1.1	Expression of HIF-1 β PAS-B Domain.....	71
3.2	Expression of Three-Domain HIF Heterodimer	77
3.3	Fluorescence Polarisation.....	82
3.3.1	Development of a HIF FP Assay.....	83

3.4	Chapter 3 Summary	88
	Chapter 4 Development of CLLFVY Fluorescent Probes.....	90
	Chapter 5 Determining the Structural Basis of CLLFVY Binding.....	95
5.1	Introduction to Protein X-ray Crystallography.....	95
5.2	Crystallising HIF	99
5.2.1	Seeding Experiments	100
5.2.2	Soaking HIF-1 α PAS-B crystals with CLLFVY	105
5.2.3	Co-Crystallisation of HIF-1 α PAS-B and CLLFVY	105
5.2.4	Optimisation of crystallisation conditions	107
5.2.5	Development of a new protein model for crystallisation.....	108
5.3	Chapter 5 Summary	110
	Chapter 6 Conclusions	113
6.1	Exploring the Structure Activity Relationship of CLLFVY	113
6.2	Development of <i>In Vitro</i> HIF Assays	114
6.3	Development of a Fluorescent CLLFVY Derivative	115
6.4	Investigating the Structural Basis of CLLFVY Binding	115
	Chapter 7 Future Directions	117
	Chapter 8 Experimental.....	119
8.1	Equipment and Materials	119
8.2	Preparation of Materials.....	120
8.2.1	Lysogeny Broth (LB)	120
8.2.2	LB Agar plates	120
8.2.3	Super optimal culture (SOC) media	120
8.2.4	Antibiotics.....	120
8.2.5	Preparation of chemically competent cells.....	121
8.2.6	Agarose gels.....	121
8.2.7	Preparation of SDS-PAGE gels	122
8.2.8	SDS-PAGE loading buffer	123
8.2.9	SDS-PAGE running buffer	123
8.2.10	Coomassie protein stain solution	123

Table of Contents

8.2.11	Coomassie de-staining solution	124
8.2.12	Protein buffers	124
8.2.13	Annealing buffer.....	125
8.3	General biology procedures	125
8.3.1	Polymerase Chain Reaction (PCR)	125
8.3.2	Colony PCR	126
8.3.3	Plasmid Purification.....	126
8.3.4	Transformation of chemically competent cells	127
8.3.5	Restriction Digests.....	127
8.3.6	Ligations.....	127
8.3.7	Site-Directed Mutagenesis.....	127
8.3.8	Running SDS-PAGE gels.....	128
8.3.9	Annealing HRE oligonucleotides	128
8.3.10	Construction of the HIF-1 α PAS-B ₍₂₃₈₋₃₄₂₎ vector	128
8.4	Protein procedures	129
8.4.1	Expression and Purification of HIF-1 α PAS-B Domain	129
8.4.2	Expression and Purification of HIF-1 Three Domain	129
8.4.3	Protein labelling	130
8.4.4	Thrombin/ TEV Cleavage	130
8.5	Assay protocols	130
8.5.1	Thermal shift	130
8.5.2	HIF PAS-B MSTs	130
8.5.3	Enzyme-Linked Immunosorbent Assay	131
8.5.4	HIF Three Domain MSTs	131
8.5.5	Fluorescence Polarization Assays.....	132
8.5.6	MST SD-Test	132
8.6	Chemical procedures.....	132
8.6.1	Wang Resin SPPS	132
8.6.2	Cleavage of Wang Resin.....	133
8.6.3	2-Chlorotriyl chloride resin Loading.....	133
8.6.4	Determining Fmoc-AA resin loading	133

8.6.5	2-Chlorotrityl chloride resin SPPS.....	133
8.6.6	2-Chlorotrityl chloride resin cleavage.....	133
8.6.7	Cyclisation via EDC/ Oxyma Pure.....	134
8.6.8	Cyclisation Via HATU.....	134
8.6.9	Cys(StBu) deprotection	134
8.6.10	Deprotection of protected cyclic peptides	134
8.6.11	Kaiser free amine test.....	134
8.6.12	Purification via HPLC.....	134
8.6.13	Analytical HPLC.....	135
8.6.14	Purification via Flash Chromatography.....	135
8.6.15	ESI*MS	136
8.6.16	High Resolution Mass Spectrometry (HRMS).....	136
8.6.17	¹ H and ¹³ C NMR	136
8.7	Hela cell experiments	136
8.7.1	Assessment of peptide permeability.....	136
8.7.2	HIF-1 α immunofluorescence experiments	137
8.8	Crystallography.....	138
8.8.1	Crystallisation of HIF-1 α PAS-B 4x using Dr Andrew Fosters conditions	138
8.8.2	Crystal soaking experiments	138
8.8.3	Co-Crystallisation experiments.....	139
8.8.3.1	Succinic acid conditions.....	139
8.8.3.2	Broad screen conditions	139
8.9	Compound Characterisation	140
8.9.1	CLLFVY	140
8.9.2	CLLFVY Alanine Scan Derivatives	140
8.9.3	CLLFVY Phenylalanine derivatives	142
8.9.4	Trifluoromethylphenylalanine alanine scan derivatives.....	145
8.9.5	CLLF(CF ₃)VY Tyrosine Derivatives	147
8.9.6	CLLF(CF ₃)VF(Br) Valine Derivatives	150
8.9.7	CLLF(CF ₃)IF(Br) Leucine Derivatives	152
8.9.8	Polar Derivatives.....	154

Table of Contents

8.9.9 Coumarin Synthesis	155
Bibliography.....	159

Table of Tables

Table 1. Diffraction statistics obtained for the HIF-1 α PAS-B crystal on the I04 beamline at Diamond Light Source. Highest resolution shell shown in parentheses.	102
Table 2. Refinement statistics for the HIF-1 α PAS-B data set.....	103
Table 3. Stock and working concentrations of antibiotics used.	120
Table 4. TFB I buffer composition	121
Table 5. TFB II buffer composition	121
Table 6. TAE buffer composition.....	122
Table 7. SDS-PAGE running gel (15 %) composition.....	122
Table 8. SDS-PAGE stacking gel composition.....	122
Table 9. SDS-PAGE loading buffer (2x) composition	123
Table 10. SDS-PAGE running buffer composition	123
Table 11. Coomassie protein staining solution composition.....	123
Table 12. Coomassie de-staining solution composition	124
Table 13. HIF buffer.....	124
Table 14. HIS-trap protein loading buffer	124
Table 15. HIS-trap protein elution buffer	124
Table 16. Labelling buffer.....	124
Table 17. HIF MST Assay Buffer	125
Table 18. HIF FP Assay Buffer	125
Table 19. Annealing buffer	125
Table 20. Composition of PCR reaction mixture.....	125
Table 21. PCR programme.....	126
Table 22. Colony PCR master mix	126

Table of Figures

Figure 1. Domain structures of the HIF- α isoforms and their binding partner HIF-1 β	2
Figure 2. Images of the different solved HIF structures.....	4
Figure 3. A) Structure of the C-TAD domain of HIF-1 α (magenta) bound to p300 (green).....	7
Figure 4. Pathways involved in the degradation of HIF- α under normoxic conditions.....	8
Figure 5. Key genes regulated by HIF. ²⁶⁰	10
Figure 6. Inhibitors of HIF DNA binding.....	13
Figure 7. Inhibitors of p300 binding.....	14
Figure 8. Indirect inhibitors of HIF.	15
Figure 9. Inhibitors of HSP90.	15
Figure 10. Direct inhibitors of HIF.	17
Figure 11. Structural basis of PT2385 binding to the HIF-2 α PAS-B domain.	18
Figure 12. Structures of the initial hit compound THS-004 and intermediates that have led to the HIF-2 α PAS-B inhibitor PT2977. ¹¹⁵⁻¹²⁰	19
Figure 12. Comparison of the internal cavities of the HIF- α PAS-B domains.	19
Figure 14. Structures of some of the most common moieties used as amide mimetics.....	22
Figure 15. Overview of phage display.	24
Figure 16. Overview of the FIT-RAPID system for the discovery of cyclic peptide binders.....	26
Figure 17. Mechanism of SICLOPPS intein splicing.	27
Figure 18. RTHS used to discover cyclic peptide inhibitors of HIF-1 α /HIF-1 β dimerisation.....	28
Figure 19. Plasmid map of the pET-28 vector used for expression of HIF-1 α PAS-B domain.....	34
Figure 20. Analysis of expression of the HIF-1 α PAS-B domain.....	35
Figure 21. Mass spectrometry analysis of obtained HIF-1 α PAS-B domain.	36
Figure 22. Outline of the principle of thermal shift.	37

Table of Figures

Figure 23. Thermal shift analysis of the HIF-1 α PAS-B domain in the presence of DMSO.	38
Figure 24. Thermal shift analysis of the HIF-1 α PAS-B domain in the presence of CLLFVY.	38
Figure 25 . Fluorescence over time of a typical MST experiment.	41
Figure 26. Cyclo-CLLFVY binding to HIF-1 α PAS-B domain as determined by MST.	42
Figure 27. MST analysis of CLLFVY specificity.	43
Figure 28. MST binding curves of the six CLLFVY alanine scan analogues against the HIF-1 α PAS-B domain.	44
Figure 29. Structures of the seven amino acids that were substituted into the cysteine position of CLLFVY.	44
Figure 30. MST binding curves of cysteine derivatives and bar chart comparison of the obtained K _d values.	45
Figure 31. Assessing the impact of DTT on the HIF-1 α PAS-B domain and HIF-1 α PAS-B-CLLFVY binding.	46
Figure 32. MST binding curve of DTT and the HIF-1 α PAS-B domain.	47
Figure 33. Location of the three cysteine residues to be mutated in the HIF-1 α PAS-B domain (PDB: 4H6J). ¹¹⁸	48
Figure 34. Thermal shift analysis of the HIF-1 α PAS-B cysteine mutants.	48
Figure 35. MST traces of CLLFVY binding to the mutant proteins C255A, C334A and C337A. ...	49
Figure 36. Structures of the amino acids substituted into the phenylalanine position of CLLFVY.	50
Figure 37. UV traces from LCMS analysis of cyclo-C(StBu)LL-Nal-VY cyclised using different coupling reagents.	51
Figure 38. Proposed mechanism of the neighbouring group effect of HATU couplings.	52
Figure 39. Comparison of the K _d values obtained for the phenylalanine derivatives.	52
Figure 40. MST binding curves of the 16 phenylalanine derivatives of CLLFVY.	53
Figure 41. MST binding curve of CLL-dF(CF ₃)-VY.	55
Figure 42. MST binding curves for the CLLF(CF ₃)VY alanine scan, and comparison of obtained K _d values.	56

Figure 43. Comparison of the K _d values obtained for the tyrosine derivatives.	57
Figure 44. MST binding curves of the 16 tyrosine analogues of CLLF(CF ₃)VY.....	59
Figure 45. Structures of the amino acids substituted for valine.....	60
Figure 46. MST binding curves for the valine derivatives.	61
Figure 47. MST binding curves obtained for L derivatives.	62
Figure 48. MST binding curves for CL-NL-F(CF ₃)IF(Br) with and without DTT.....	63
Figure 49. Dibenzofulvene-piperidine adduct that absorbs at a wavelength of 301 nm.....	64
Figure 50. MST binding curves of the glutamic acid (A) and lysine derivatives (B).	65
Figure 51. MST binding curves of CK-NL-F(CF ₃)IF(Br) and CE-NL-F(CF ₃)IF(Br) in the presence of DTT.	66
Figure 52. MST binding curves of CK-NL-F(CF ₃)IF(Br) and CE-NL-F(CF ₃)IF(Br) against the HIF-2 α PAS-B domain.	67
Figure 53. Overview of an ELISA assay.	70
Figure 54. Analysis of the structure of the HIF-1 α PAS-B mutant.....	71
Figure 55. A) First derivative of thermal shift data obtained for the HIF-1 α PAS-B 4x mutant, with a melting temperature of 54.25 °C. B) Dose-response curve for CLLFVY binding to the HIF-1 α PAS-B 4x mutant.	72
Figure 56. A) Plasmid map of the construct used for expression of HIF-1 β PAS-B domains. B) First derivative of the thermal shift data for the WT and mutant HIF-1 β PAS-B domains.....	73
Figure 57. Assessment of dimer formation of HIF-1 α PAS-B and HIF-1 β PAS-B by ELISA.	74
Figure 58. ELISA binding curve for CLLFVY inhibition of HIF-PAS-B dimerisation.	74
Figure 59. A) ELISA dose-response curve for CE-NL-F(CF ₃)IF(Br). B) ELISA dose-response curve for CK-NL-F(CF ₃)IF(Br).	75
Figure 60. Structure of the HIF-1 three-domain heterodimer.....	77
Figure 61. A) Plasmid map of the pET21-d vector containing the HIF-1 β three-domain gene. B) Plasmid map of the pET28-a vector containing the HIF-1 α three-domain gene.....	77

Table of Figures

Figure 62. Analysis of the expression of the HIF three-domain protein.....	78
Figure 63. MST binding curve of CLLFVY against HIF three-domain.....	79
Figure 64. Capillary scans of initial fluorescence for CK-NL-F(CF ₃)IF(Br) and CEK-NL-F(CF ₃)IF(Br) where 1 is the highest concentration (500 μM) and 16 is the lowest (15 nM).79	
Figure 65. SD-Test of CK-NL-F(CF ₃)IF(Br) and CE-NL-F(CF ₃)IF(Br).	80
Figure 66. A) MST dose-response curve obtained for CE-NL-F(CF ₃)IF(Br) with the HIF-1 three-domain protein. B) MST dose-response curve obtained for CK-NL-F(CF ₃)IF(Br) with the HIF-1 three-domain protein.	81
Figure 67. Outline of the principles of fluorescence polarisation.	82
Figure 68. A) Dose-response curve of HIF three-domain protein with FAM-HRE. B) Dose-response curve from the competition assay between FAM-HRE and HRE.....	84
Figure 69. A) Dose-response curve of CLLFVY in the FAM-HRE FP assay. Dose-response curve of ALLFVY in the FAM-HRE FP assay.	85
Figure 70. A) Dose-response curve for CE-NL-F(CF ₃)IF(Br) in the FAM-HRE FP assay. B) Dose-response curve for CK-NL-F(CF ₃)IF(Br) in the FAM-HRE FP assay.....	85
Figure 71. Effect of NaCl concentration on HIF-1 FAM-HRE binding as determined by fluorescence polarisation.	86
Figure 72. A) Effect of increasing NaCl concentration on CK-NL-F(CF ₃)IF(Br) FAM-HRE binding. B) Effect of increasing NaCl concentration on CE-NL-F(CF ₃)IF(Br) FAM-HRE binding.	87
Figure 73. Dose-response curve for CLLF(CF ₃)IF(Br) in the FAM-HRE FP assay.	88
Figure 74. Structures of the fluorescent peptides CLL-hCou-VY (A) and CLLF(CF ₃)V-hCou (B). ..	91
Figure 75. Analysis of the homo-coumaryl CLLFVY derivatives.....	91
Figure 76. A) HeLa cells dosed with CLL-hCou-VY (50 μM) B) HeLa cells dosed with CLLF(CF ₃)V-hCou (50 μM).	92
Figure 77. Immunofluorescence of HIF-1α in HeLa cells with and without DFX treatment. Cells were also treated with nuclear DAPI stain and an overlay of the two channels is also shown.	93

Figure 78. Immunofluorescence of HIF-1 α in HeLa cells with and without CLL-hCou-VY treatment. An overlay of the two images is also shown.	94
Figure 79. Illustration of a sitting drop and hanging drop crystallisation set up.	96
Figure 80. 2D visualisation of the solubility of protein as a function of precipitant concentration.	97
Figure 81. Image of HIF-1 α PAS-B crystals obtained using previously discovered crystallisation conditions.	100
Figure 82. Images of crystals obtained from seeding experiments.	101
Figure 83. A) Loop containing a crystal of HIF-1 α PAS-B. Loop size 0.4-0.5 mm, beam size 20 x 20 μ m. B) Associated diffraction pattern from 12 keV synchrotron radiation. .	101
Figure 84. Polygon plot of refinement statistics for the HIF-1 α PAS-B data set. Values in red are upper and lower limits of these parameters found in the PDB.	103
Figure 85. Protein structure obtained upon solving the obtained data set.	104
Figure 86. Structural waters present in the internal cavity of HIF-1 α PAS-B domain.	105
Figure 87. Image of a poor quality protein crystal obtained during the attempted co-crystallisation of HIF-1 α PAS-B with CLLFVY.	106
Figure 88. Mass spectrum obtained from thrombin cleaved HIF-1 α PAS-B domain.	107
Figure 89. Amino acid sequences corresponding to the different peaks observed in the mass spectrum of the thrombin cleaved HIF-1 α PAS-B domain.	108
Figure 90. Plasmid map of pETM11 vector containing HIF-1 α PAS-B ₍₂₃₈₋₃₄₂₎ and a TEV cleavage site to enable removal of the N-terminal HIS-tag.	108
Figure 91. Mass spectrum obtained for TEV cleaved HIF-1 α PAS B domain.	109
Figure 92. Sequences of HIF-1 α PAS-B ₍₂₃₈₋₃₄₂₎ and cleaved HIF-1 α PAS-B ₍₂₃₈₋₃₄₂₎ protein.	109
Figure 93. Crystals of TEV-cleaved HIF-1 α PAS-B ₍₂₃₈₋₃₄₂₎ protein.	110

Research Thesis: Declaration of Authorship

Print name: Reece Gardner

Title of thesis: The Development of Cyclic Peptide Inhibitors of Hypoxia-Inducible Factors

I declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. None of this work has been published before submission.

Signature:

Date:

Acknowledgements

Firstly, I would like to thank Prof Ali Tavassoli for the opportunity to carry out this research, as well as for the support and advice provided throughout my PhD. I would also like to thank Prof Peter Roach and Dr Josh Prince for the helpful discussions on crystallising proteins, as well as allowing us to use their beamtime.

Thanks also to members of the Tavassoli group, past and present, in particular Dr Andy Ball for his help with HIF, Dr Cyrielle Doigneaux for running the cell experiments and James Craswell for the helpful discussions.

A special thanks to Dr Andy Foster, not only for teaching a chemistry graduate molecular biology, but also for proofreading this thesis.

Finally, I would like to thank my family and friends, especially Dr Kirsty Maskell for her support, patience and encouragement.

Abbreviations

ACF	Acriflavine
AhR	Aryl hydrocarbon receptor
ACN	Acetonitrile
ARNT	Aryl hydrocarbon receptor nuclear translocator
ATP	Adenosine triphosphate
bHLH	basic helix-loop-helix
Boc	tert-Butoxycarbonyl
CBP	CREB-binding protein
CDI	Carbonyl diimidazole
Cbz	Carboxybenzyl
DCM	Dichloromethane
DIC	Diisopropylcarbodiimide
DIPEA	<i>N,N</i> -Diisopropylethylamine
DMF	Dimethylformamide
DMSO	Dimethyl sulfoxide
DNA	Deoxyribosenucleic acid
DTT	Dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-linked Immunosorbent assay
EPO	Erythropoiten
ESI	Electrospray ionisation
FIH	Factor-inhibiting HIF
Fmoc	Fluorenylmethyloxycarbonyl
FP	Fluorescence polarisation
FPLC	Fast protein liquid chromatography
GLUT1	Glucose transporter 1
GPCR	G protein-coupled receptor
HAF	HIF-associated factor
HATU	Hexafluorophosphate azabenzotriazole tetraethyl uronium
HFIP	1,1,1,3,3,3-Hexafluoro-2-propanol
HIF	Hypoxia-inducible factor

Abbreviations

HOAt	1-Hydroxy-7-azabenzotriazole
HPLC	High pressure liquid chromatography
HRE	Hypoxia responsive element
Hsp70	Heat shock protein 70
Hsp90	Heat shock protein 90
IC ₅₀	Half maximal inhibitory concentration
iNOS	Inducible nitric oxide synthase
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ITC	Isothermal titration calorimetry
IVTT	<i>In vitro</i> transcription-translation
K _d	Dissociation constant
LB	Lysogeny broth
LCMS	Liquid chromatography mass spectrometry
LDHA	Lactose dehydrogenase A
mRNA	Messenger ribonucleic acid
MST	Microscale thermophoresis
NHS	N-hydroxysuccinimide
NMR	Nuclear magnetic resonance
OD	Optical density
ODDD	Oxygen dependent degradation domain
PAS	Per-ARNT-Sim
Pbf	2,2,4,6,7-Pentamethyldihydrobenzofuran-5-sulfonyl chloride
PCR	Polymerase chain reaction
PDB	Protein data bank
PDK1	Pyruvate dehydrogenase kinase-1
PHD	Prolyl hydroxylase domain
PPI	Protein-protein interaction
RACK-1	Receptor of activated protein kinase 1
R _t	Retention time
RTHS	Reverse two-hybrid system
SAR	Structure-activity relationship
SDM	Site-directed mutagenesis
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SICLOPPS	Split-intein cyclisation of peptides and proteins
SOC	Super optimal broth with catabolite repression

SPPS	Solid phase peptide synthesis
StBu	t-Butylthio
TAD	Transactivational domain
TCEP	tris(2-carboxyethyl)phosphine hydrochloride
TFA	Trifluoroacetic acid
TFB	Transformation buffer
TIS	Triisopropylsilane
TIC	Total ion count
T _m	Melting temperature
tRNA	Transfer ribonucleic acid
VEGF	Vascular endothelial growth factor
VHL	Von Hippel Lindau tumour suppressor protein

Abbreviations

Unnatural amino acids

Amino acid	Abbreviation
Homo-Cysteine	hC
D-Cysteine	dC
Thioprolino	TP
Homo-phenylalanine	hF
Phenylglycine	Phg
D-phenylalanine	dF
4-Fluoro-phenylalanine	F(F)
4-Chloro-phenylalanine	F(Cl)
4-Bromo-phenylalanine	F(Br)
4-Iodo-phenylalanine	F(I)
4-Nitro-phenylalanine	F(NO ₂)
4-Trifluoromethyl-phenylalanine	F(CF ₃)
4-Chloro-phenylalanine	F(tBu)
4-tert-Butyl-phenylalanine	F(Ph)
1-Naphthylalanine	Nal
4-pyridylalanine	Pal
Cyclohexylalanine	Cha
O-methyl-tyrosine	Y(Me)
Homo-leucine	hL
Norleucine	NL
Norvaline	NV
Aminoisobutyric acid	Aib
2-amino-3-methoxypropionic acid	Prop

Chapter 1 Introduction

1.1 Hypoxia-Inducible Factor

1.1.1 The Role of HIF

Being able to maintain oxygen homeostasis is essential for mammalian cells in order to carry out aerobic metabolism. As a result, it is a tightly regulated process, with all nucleated cells possessing mechanisms for sensing and responding to changes in oxygen levels.¹ Cellular oxygen levels can vary during normal biological function, such as embryonic development and wound healing, but they can also be impacted by disease.^{2,3} Cancer, heart disease and pulmonary obstructive disorder can all impair the cells ability to maintain balanced oxygen levels, ultimately leading to low environmental oxygen levels within the cell (hypoxia).⁴ Hypoxia is particularly prevalent in cancerous tumours, where the rapid proliferation of cells leads to a demand in oxygen that exceeds the supply provided by the local vasculature. This leads to cancerous tumour cells having an average of 1-2 % oxygenation compared to 4.6-14.5 % in healthy normoxic cells.^{5,6} Whilst cells in extreme hypoxic stress often become necrotic, those in less severe hypoxia often become more aggressive and treatment resistant.^{7,8} This is a consequence of the physiological changes induced by the hypoxic microenvironment which are brought about via several signalling pathways, including PI3K/AKT/mTOR, MAPK and NFκB. Between these pathways cell-proliferation, survival, apoptosis, migration, inflammation and metabolism are regulated.^{9,10}

Another pathway activated in hypoxia is that of the hypoxia-inducible factor (HIF). HIF is a heterodimeric transcription factor that acts as the master regulator of the cellular response to hypoxia. HIF has been shown to regulate the expression of more than 300 genes many of which are involved in tumour growth and progression.¹¹

The importance of HIF in the cellular response to hypoxia, and the prevalence of hypoxia in cancerous tumours, has made HIF an attractive therapeutic target for the development of novel cancer treatments.

1.1.2 HIF Structure

There are three isoforms of HIF; HIF-1, HIF-2 and HIF-3. All three comprise of two subunits, an alpha subunit which is unique in each, and a conserved beta subunit, HIF-1β, also known as the aryl hydrocarbon receptor nuclear translocator (ARNT). Both subunits of HIF are part of the basic

Chapter 1

helix-loop-helix-Per-ARNT-Sim (bHLH-PAS) protein family, and the domain structure of the different HIF subunits are shown in Figure 1.¹²

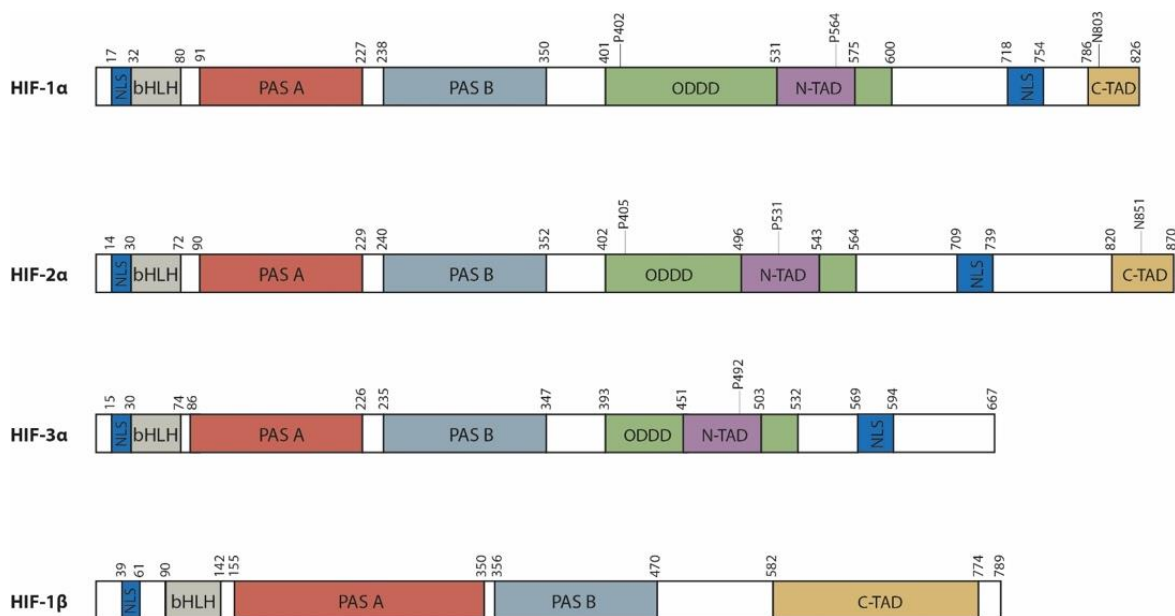


Figure 1. Domain structures of the HIF- α isoforms and their binding partner HIF-1 β .

Domains: NLS: Nuclear Localization Sequence, required for import into the cell nucleus.²⁵⁵ bHLH: basic Helix-Loop-Helix, responsible for engaging DNA via basic residues contained in one of its two characteristic alpha helices.²⁵⁶ PAS A/B: Per ARNT Sim, key in mediating PPIs and have also been shown to act as molecular sensors.²⁵⁷ ODDD: Oxygen dependent degradation domain, region of HIF containing the proline residues 402 and 564 which are hydroxylated in normoxia leading to the degradation of the HIF protein.²⁵⁸ N/C-TAD: N/C-terminal transactivation domain, responsible for recruiting other proteins involved in regulating transcription and formation of an active transcription complex.²⁵⁹

The basic helix-loop-helix (bHLH) and two Per-ARNT-Sim (PAS) domains are found in all α isoforms as well as the binding partner HIF-1 β , whereas the oxygen-dependent degradation domains (ODDD) are present only in the α subunits, and are responsible for the breakdown of the HIF- α subunits in normoxia. HIF-1 α and HIF-2 α also possess two transactivation domains, N-TAD and C-TAD, required for the formation of a functional transcription complex. Notably HIF-3 α lacks the C-TAD domain and so cannot recruit the required coactivators for transcription, which has led to the suggestion that HIF-3 α acts as a competitive inhibitor to the other two isoforms.¹³

The first reported structure of HIF was obtained via NMR and detailed the structure of the isolated PAS-B domains of both HIF-2 α and HIF-1 β . The structure showed that the two domains interfaced via an antiparallel interaction between the beta sheets of the two proteins.¹⁴ Analysis of the NMR protein model revealed an interfacial salt bridge between the HIF-2 α PAS-B R247 and HIF- β PAS-B E362. Mutations of these residues to reverse the salt bridge (R247E and E362R) led to

the development of a stabilised protein model that had higher affinity for heterodimer formation than the wild type PAS-B domains. This new protein model enabled the crystallisation of the HIF-2 α PAS-B and HIF-1 β PAS-B heterodimer (Figure 2A).¹⁵ The structure of HIF-1 has been less well characterised and there are only two published structures at the time of writing. The first of which was published by Cardoso *et al.* in 2012 and shows the structure of a heterodimer of the PAS-B domains of HIF-1 α and HIF-1 β (Figure 2B).¹⁶ The previously mentioned work, the development of a mutant of the HIF-2 α PAS-B domain via a salt bridge mutation, was used as a guide to develop a more stable HIF-1 α protein model. This involved introducing the analogous point mutation (R245E) in the HIF-1 α PAS-B domain which significantly improved the solution characteristics and stability of the domain.¹⁷ Cardoso *et al.* then showed that by creating the complementary mutation of the HIF-1 β PAS-B domain (E362R) the PAS-B domains would form a stable dimer, something that was not found to be true of the wild type domains.¹⁶ This structure was then crystallised to reveal that the dimerisation was mediated via the same anti-parallel contact of the beta sheets of each protein as was observed in the HIF-2 structure. The second HIF-1 structure was published in 2015 by Wu *et al.*, which showed the bHLH, PAS-A and PAS-B domains of HIF-1 α and HIF-1 β as a heterodimer bound to DNA.¹⁸ This structure revealed that the HIF-1 β domains were wrapped around a core consisting of the HIF- α domains (Figure 2C).

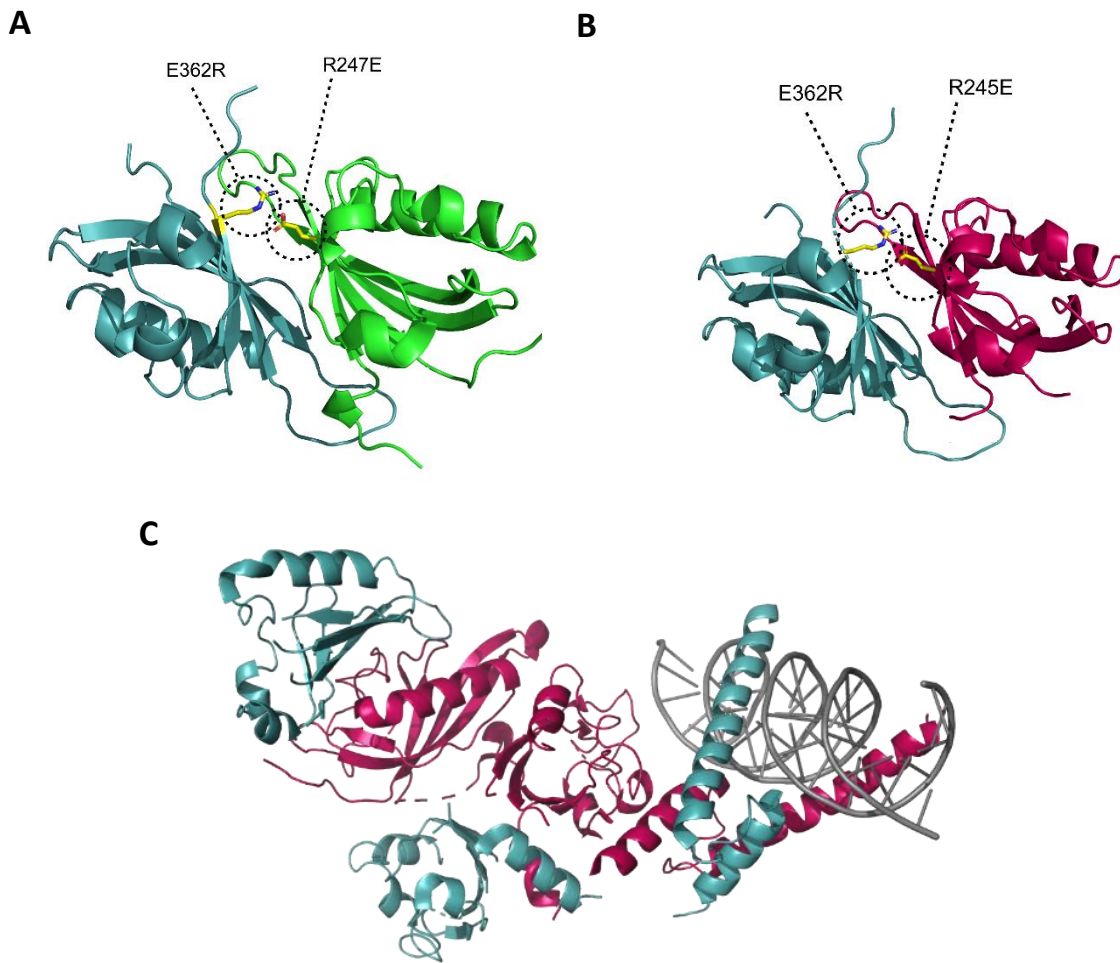


Figure 2. Images of the different solved HIF structures. A) Structure of the HIF-2 α PAS-B (green) and HIF-1 β PAS-B (cyan) heterodimer (PDB: 3F1N).¹⁵ B) Structure of the HIF-1 α PAS-B (magenta) and HIF-1 β PAS-B (cyan) heterodimer (PDB: 4H6J).¹⁶ C) Structure of the HIF-1 α bHLH-PAS A-PAS-B (magenta) and HIF-1 β bHLH-PAS A-PAS-B (cyan) heterodimer bound to the HRE (PDB: 4ZPR).¹⁸

Despite the asymmetry present in the bHLH-PAS-A-PAS-B structure, it is clear that the two bHLH domains converge to form a functional DNA reading head, with the domains inserting into the major groove faces of the DNA. HIF binds to DNA at a specific hypoxia responsive element (HRE), containing the core DNA sequence RCGTG (where R represents either A or G).¹⁹ A number of amino acids interact specifically with the HRE, with the HIF-1 β bHLH utilising the residues R102, E98 and H94 to form base pair specific hydrogen bonds that recognise the GTG half of the HRE. This is identical to how DNA is engaged by HIF-1 β in other transcription factors.²⁰ The HIF-1 α bHLH confers the selectivity for binding only HREs by interacting with the AC base pairs in the first half of the HRE via its A25 and R30 residues. Interestingly, the PAS-A domain of HIF- α was also found to make contact with the DNA, six base pairs downstream of the HRE, significantly increasing the area over which HIF interacts with the DNA.¹⁸

PAS domains had not previously been observed to interact with DNA and are generally thought of as signalling proteins. First described in 1991, PAS domains are defined by a conserved global fold consisting of ~100 amino acids, and have since been found in all kingdoms of life where they detect a wide range of chemical and physical stimuli.^{21,22} Whilst an external signal for the PAS domains of HIF has not yet been found, mutational studies have shown both PAS domains are crucial for mediating the successful dimerisation of the alpha and beta subunits of HIF.²³ The PAS domains of HIF- α have also been demonstrated to regulate HIF activity via interaction with other proteins. Binding of the heat shock protein HSP90 to HIF- α has been shown to be mediated by the PAS domains, thus preventing the degradation of the subunit by receptor of activated kinase 1 (RACK-1) and increasing HIF activity.²⁴ Conversely, the PAS domains can also bind members of the minichromosome maintenance protein complex (MCM) protein family in particular MCM7 which stabilises the interaction between HIF and an E3 ligase complex promoting its degradation.²⁵

Once the active HIF dimer has bound to the HRE within the nucleus, it next needs to recruit transcriptional coactivators in order to form a functional transcription complex. This recruitment is mediated by the two transactivation domains located within the HIF- α subunit. The C-terminal transactivation domain (C-TAD) is oxygen regulated whereas the N-terminal transactivation domain (N-TAD) is more centrally located and not oxygen dependent.²⁶ Both domains are known to recruit the coactivator p300, or the closely related Creb-Binding Protein (CBP), which is essential for linking HIF to other transcription factors and transcriptional machinery.²⁷ This interaction is therefore critical for transcriptional activity, and as such has been well studied. The C-TAD domain engages the CH1 domain of p300/CBP and is the primary activator of transcription. The N-TAD domain by comparison binds to the CH3 domains of p300/CBP and is a less effective activator. Despite being a less potent activator, the N-TAD domain has been implicated in conferring gene specificity between HIF-1 and HIF-2 by binding to different transcription factors.²⁸ The structural basis for the interaction between the HIF C-TAD domain and the p300/CBP CH1 domain was first described by Freedman *et al.* in 2002.²⁹ They found that despite being intrinsically disordered, in the presence of p300/CBP the C-TAD domain wraps itself around p300/CBP to form three short helices that fit in grooves of the p300 surface, as shown in Figure 3A.²⁹

Lastly, there is the oxygen-dependent degradation domain (ODDD) which is only present in HIF- α subunits and is responsible of the degradation of these proteins under normoxic conditions.

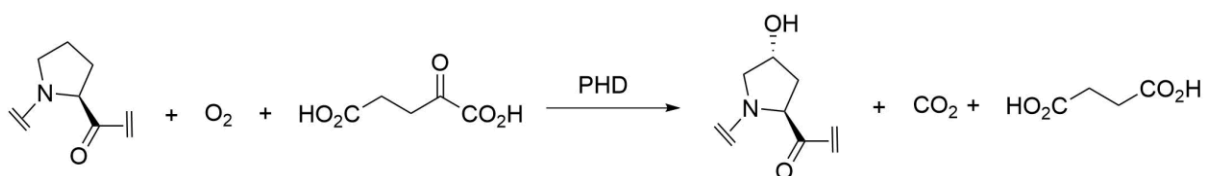
1.1.3 HIF Regulation

To enable cells to adapt rapidly to changes in oxygen concentration both subunits of the HIF protein are constitutively expressed. However, the activity and concentration of the alpha subunit

is tightly regulated by a range of oxygen-dependent and independent mechanisms. This includes the degradation of the alpha subunit under normoxic conditions, with the half-life of the HIF-1 α subunit having been reported to be as little as 5-8 minutes.³⁰ This constant expression and degradation is a resource expensive process and thus demonstrates the importance of the cell being able to react rapidly to a drop in oxygen concentration.

1.1.3.1 Oxygen-dependent mechanisms of HIF regulation

The major regulatory pathway of HIF occurs via the degradation of the alpha subunit under normoxic conditions. The dominant mechanism of degradation is initiated by a class of enzymes called prolyl hydroxylases (PHDs), that utilise environmental oxygen, along with 2-oxoglutarate, in order to hydroxylate two proline residues, yielding CO₂ and succinate as shown in Scheme 1.³¹⁻³³



Scheme 1. Hydroxylation of proline by PHD enzymes utilising environmental oxygen and 2- α -ketoglutarate yielding the hydroxylated proline carbon dioxide and succinic acid.

The proline residues are located in the ODDD of the HIF-1 α subunit (P402 and P564) and appear in the conserved sequence LXXLAP (where X is any amino acid).³² Once hydroxylated, these proline residues act as a recognition motif for the recruitment of the von Hippel Lindau (VHL) tumour suppressor protein. The VHL protein in turn recruits the proteins elongin C, elongin B, cullin-2 and ring-box-1 to form an E3 ubiquitin ligase complex. The complex then poly-ubiquitinates the alpha subunit, targeting it for degradation by the cells proteasome (Figure 4A).³⁴⁻³⁶ In addition to the VHL-mediated regulation of HIF activity, there is a second oxygen-dependent mechanism that prevents HIF activity. This mechanism is mediated by a protein called factor inhibiting HIF (FIH), and also utilises environmental oxygen, this time to hydroxylate an asparagine residue (N803) located in the C-TAD domain of HIF (Figure 4B, N803 is highlighted blue in Figure 3B).^{37,38} The asparagine residue is located in a deep hydrophobic pocket and the hydroxyl group prevents the interaction of this region with p300/CBP. Notably the C-TAD domain adopts a very different conformation when binding FIH with an extended structure presenting a single strand centred upon the asparagine residue (Figure 3B). This is in contrast to the previous structure seen in Figure 3A and demonstrates the plasticity of this domain.³⁹

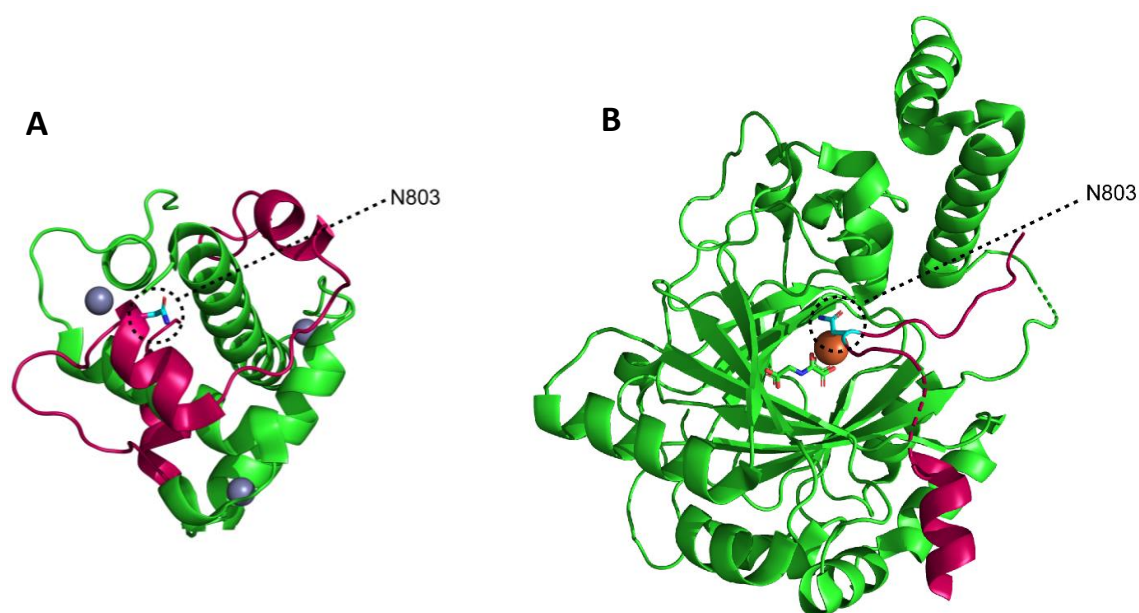


Figure 3. A) Structure of the C-TAD domain of HIF-1 α (magenta) bound to p300 (green) (PDB: 1L8C).³⁰ B) Structure of factor inhibiting HIF bound to HIF-1 α C-TAD domain with succinate substrate. (PDB: 1H2K).⁴⁰ The position of N803, which is hydroxylated by p300, is shown in blue in both structures. Orange = Fe²⁺, Grey = Zn²⁺.

Interestingly, the concentration of oxygen required for the hydroxylation of N803 by FIH is less than that required for the proline hydroxylation carried out by PHDs.⁴⁰ This enables finer transcriptional control of HIF genes, with HIF accumulating at intermediate oxygen concentrations, but not transcriptionally active until oxygen concentrations fall further. At this point FIH is inhibited and the C-TAD domains are free to recruit p300 and form an active transcription complex. It has also been suggested that the higher affinity of FIH for molecular oxygen leads to preferential gene expression via the N-TAD domain. This enables FIH to regulate the transcription of HIF-dependent genes over a range of oxygen concentrations.⁴¹

In the absence of environmental oxygen, the above mentioned enzymatic hydroxylations can no longer occur and so the HIF-1 α protein is free to accumulate within cells. It can then bind to its partner protein HIF-1 β , which is unaffected by changes in oxygen levels due to its role in other cellular pathways.^{42,43} The heterodimeric HIF can then recruit p300 and other transcriptional co-factors and express the genes needed to adapt to the cellular new environment (Figure 4C).

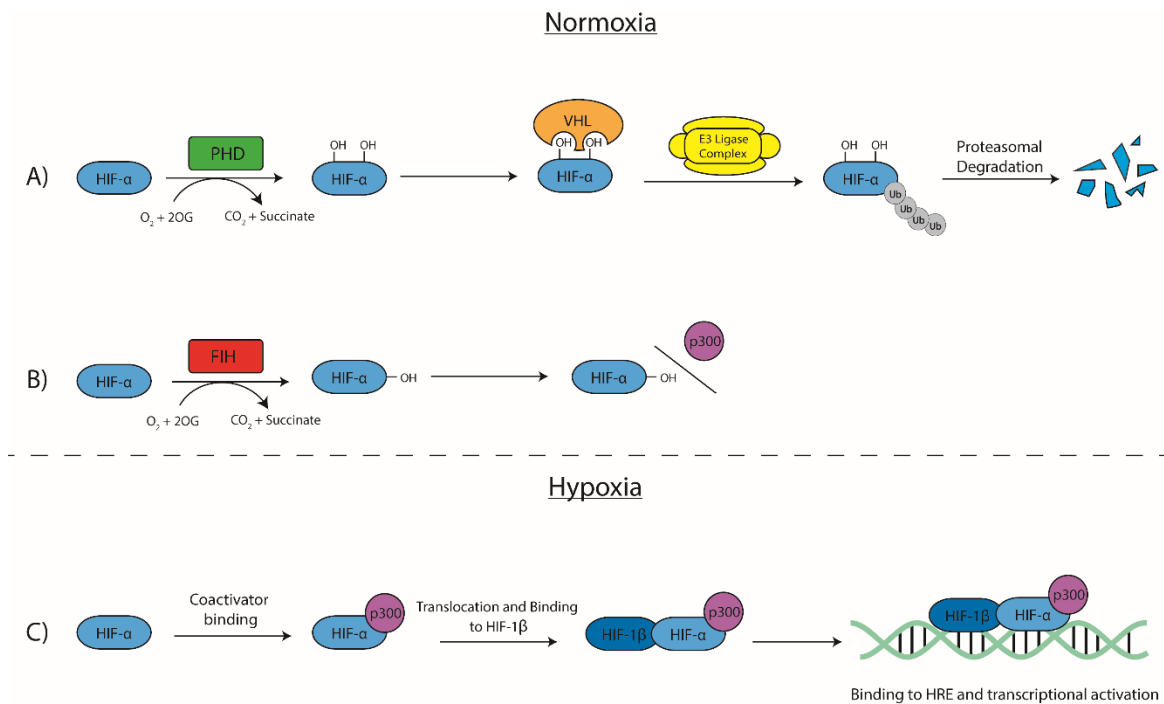


Figure 4. Pathways involved in the degradation of HIF- α under normoxic conditions. A) PHD mediated hydroxylation of proline residues (402 and 564) leading to VHL binding and subsequent proteasomal degradation. B) FIH mediated hydroxylation of asparagine 803 inhibits p300 binding and prevents the formation of an active transcription complex. C) Under hypoxic conditions, HIF- α is stabilised and able to bind co-factors, translocate to the nucleus and bind to HIF-1 β forming an active transcription complex. ^{35–39}

1.1.3.2 Oxygen-Independent mechanisms of HIF regulation

Whilst the VHL-mediated degradation of the alpha subunits is the predominant method of HIF regulation in normoxic environments, there also exists a number of processes that are capable of regulating HIF activity independent of oxygen concentration. These mechanisms regulate HIF in normoxic conditions as well as causing a reduction in HIF protein levels under prolonged hypoxia. ^{44,45}

One such mechanism is mediated by the RACK1, which binds to the PAS-A domain of HIF-1 α . This leads to the recruitment of an E3 ligase complex which subsequently poly-ubiquitinates HIF, promoting its degradation by the proteasome. The protein HSP90 has been shown to compete with RACK-1 for binding the HIF-1 α PAS-A domain, stabilising the HIF-1 α protein by preventing RACK-1-mediated degradation. ⁴⁶

In addition, the alpha subunit of HIF has also been shown to be degraded via the oncogenic E3 ligase mouse double minute 2 homolog (MDM2) in a p53 dependent manner. P53 has also been

implicated in attenuating the transcriptional activity of HIF by acting as competitor for binding the coactivator p300.^{47,48}

More recently, a novel E3 ligase called hypoxia-associated factor (HAF) has been described that binds and ubiquitinates HIF-1 α regardless of oxygen tension. HAF was shown to perform this function in cells that were VHL competent or deficient and is therefore acting via an independent mechanism.⁴⁹ Interestingly, HAF was later found to also bind to HIF-2 α , but at a different site from 1 α . When binding HIF-2 α , HAF causes an increase in HIF-2 α activity, which combined with its ability to degrade HIF-1 α enables HAF to effectively switch the hypoxic response from HIF-1 α to HIF-2 α .^{49,50} The switch between active HIF isoforms via differentiated degradation mechanisms is also shown with the interactions of HIF- α with HSP70 and carboxyl terminus of Hsp70-interacting protein (CHIP). Both HIF-1 α and HIF-2 α were shown to bind to HSP70 with comparable affinity, however only the HIF-1 α -HSP70 complex was capable of recruiting CHIP, which then recruits an E3 ligase complex leading to the degradation of HIF-1 α but not HIF-2 α .⁵¹

Before HIF- α can become transcriptionally active, it must first form a dimer with HIF-1 β , with this process thought to be mediated by the PAS domains. This interaction presents another point for HIF regulation. The protein MgcRacGAP, a regulator of Rho proteins, was shown to bind directly to the HIF-1 α PAS-B domain, thus inhibiting its activity by preventing HIF-1 β binding.⁵² Similarly, the protein COMM domain-containing 1 (COMMD1) was shown to bind to the bHLH-PAS region of HIF-1 α and prevent HIF-1 α binding to HIF-1 β .⁵³ The aryl hydrocarbon receptor (AhR), another protein of the bHLH-PAS family, also binds to HIF-1 β to form a transcription factor, and as such can inhibit HIF activity by competing for HIF-1 β .⁴³ In addition the prevention of HIF dimerisation can be achieved via post-translational modification (PTM) of a serine residue located in the PAS-B domain of HIF-1 α (S247). This PTM is carried out by casein kinase 1 which phosphorylates the serine preventing the formation of a critical salt bridge between 1 α and 1 β .⁵⁴

Once formed, the activity of the α/β dimer is also modulated via interactions with a range of proteins that affect the recruitment of different cofactors. The ability of HIF to recruit p300 has been shown to be modulated by the activity of a number of proteins such as iNOS, MUC1, Raptor, PMK2 and EAF2.⁵⁵⁻⁵⁹

The transcriptional activity of HIF is also modulated by post-translational modifications. Kinases in particular have been heavily implicated in the normoxic regulation of HIF- α , with over 17 different kinases having been described as directly or indirectly regulating HIF.⁶⁰ These kinases phosphorylate HIF at many different positions which can lead to the regulation of HIF nuclear translocation (S641, S643), as well as stabilising or destabilising the HIF- α sub unit (S551, T555, S589). PTMs also regulate HIF's ability to bind DNA, with the methylation of HIF-1 α at K32 by a

SET domain-containing lysine methyltransferase called Set7 preventing HIF from engaging with the DNA.⁶¹⁻⁶⁴

1.1.4 HIF Activity

Under hypoxic conditions HIF- α is stabilised and the two subunits of HIF can come together and form an active dimer. HIF then mediates the transcription of hundreds of genes involved in processes such as metabolism, erythropoiesis, angiogenesis and pH regulation, with some of the key genes shown in Figure 5.

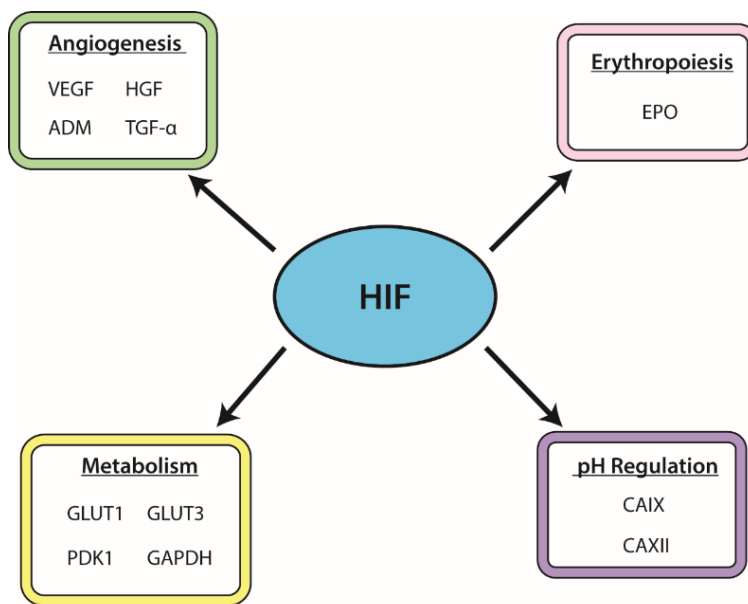


Figure 5. Key genes regulated by HIF.²⁶⁰

HIF-1 α has been shown to be ubiquitously expressed; the HIF-2 α isoform however, is expressed in a more tissue specific manner.¹¹ In addition, both isoforms follow different temporal patterns of induction. These differences amount to the two isoforms performing distinct roles that result in differing phenotypes, as demonstrated in knockout mouse studies.⁶⁵⁻⁶⁷ HIF-1 α was found to be involved in the transcription of genes involved in the glycolytic pathway, whereas HIF-2 α was not. By comparison, HIF-2 α was found to drive the transcription of many growth factors, including dominant control over EPO expression.⁶⁸ The two isoforms have also been found to be antagonistic of one another in some processes such as nitric oxide (NO) production, with HIF-1 α upregulating iNOS thereby increasing nitric NO production, whereas HIF-2 α was shown to induce ABL2 expression which inhibits NO production.⁶⁹

1.1.4.1 Angiogenesis

As the master regulator of hypoxic responses, HIF regulates many pro-angiogenic genes involved in the adaptation to longer term hypoxia. Many of these genes are essential during embryonic development, but a number are also relevant in a range of pathologies. Cancerous tumour cores in particular are commonly hypoxic as the rapid proliferation of cells outstrips the local vasculature's ability to provide enough oxygen. This leads to the so called 'angiogenic switch' as the tumour transitions to a state of active vessel growth.⁷⁰ This switch is mediated by a number of genes including the vascular endothelial growth factor (VEGF), Angiopoietin-1 (Ang-1), Angiopoietin-2 (Ang-2), and Tyrosine-protein kinase receptor (Tie-2). VEGF is considered the primary effector of HIF-mediated angiogenesis with *in vivo* studies of Hepa-1 xenografts showing the direct regulation of VEGF by HIF-1.⁷⁰ Angiogenesis is essential for tumour progression as demonstrated in several inducible tumour models, where inhibition of angiogenesis early in tumorigenesis impairs tumour progression.⁷¹ Despite this, these new tumour vessels are often irregular, leaky and sometimes lacking red blood cells entirely. They are therefore often poorly functioning, leading to hypoxic areas where HIF is stabilised, even in highly vascular tumours.^{72–74}

HIF also plays a role in controlling oxygen delivery via the regulation of the enzyme inducible nitric oxide synthase (iNOS). iNOS is responsible for producing NO, which acts as a regulator within the cardiovascular system causing vasodilation of the blood vessels and thus increasing blood flow to that area.⁷⁵ In addition, HIF also regulates the concentration of the hormone erythropoietin (EPO), which promotes the production of red blood cells thus increasing the oxygen carrying capacity of the blood.⁷⁶

1.1.4.2 Metabolism

A shift away from oxidative carbon metabolism is a key adaptation to hypoxic environments as molecular oxygen is required as an electron acceptor in the final step of the electron transport chain. HIF-1 α acts to reprogram cells to use anaerobic methods of ATP generation such as glycolysis, thus reducing oxygen demand. This is achieved via the upregulation of a number of proteins involved in glucose metabolism, including glucose transporters, glycolytic enzymes and lactate dehydrogenase A (LDHA). In addition, HIF has also been shown to upregulate the PDK1 gene which codes for the protein pyruvate dehydrogenase kinase-1. This acts to divert carbon away from the mitochondria by preventing the conversion of pyruvate into acetyl-CoA.⁷⁷ The end product of glycolysis is lactate, which if not removed will hinder further glycolysis. As an anion, lactate requires transporters to cross cell membranes, a process facilitated by mono-carboxylate-proton symporters (MCT). The expression of these transporters has been demonstrated to be under the control of HIF-1 α , further solidifying its role in metabolic reprogramming.^{78,79}

Chapter 1

As well as increasing glycolysis, HIF has also been shown to prevent oxidative metabolism. One way in which this is achieved is via the expression of the proteins BNIP3 and BNIP3L, which eliminate mitochondria carrying out oxidative metabolism of glucose and fatty acids.^{80,81} In addition, HIF is also capable of upregulating microRNAs (miRNA). One such miRNA, miR-210, inhibits the expression of proteins required for the formation of sulfur clusters critical for the function of the TCA enzyme aconitase, as well as the electron transport chain Complex I.^{82,83}

Cancer cells were first shown to have dramatically increased glucose uptake and lactate production in 1927 by Warburg *et al.*⁸⁴ Interestingly, it was noted that these increases were seen even when oxygen was present at concentrations high enough for respiration, a phenomenon that has since been named the Warburg effect.⁸⁴ Often the Warburg effect arises in cell lines that lack the VHL protein, impeding the oxygen-dependent degradation of the HIF-1 α sub unit. HIF-1 dependent metabolic reprogramming can then occur via the regulation of the aforementioned proteins promoting glycolysis.⁸⁵ In addition, a study in 2002 by Lu *et al.* revealed that the HIF-1 α protein was stabilised by the presence of lactate and pyruvate, all key components of anaerobic glycolysis, providing a potential feedback loop for the malignant transformation and survival of cancerous cells.⁸⁶

1.1.4.3 Cellular pH

A side effect of the shift towards anaerobic metabolism is the build-up of lactic acid which, as a result, lowers the cellular pH. In order to maintain a functional cellular pH, HIF also upregulates the expression of two carbonic anhydrases, CAIX and CAXII.⁸⁷ Carbonic anhydrases are zinc metalloenzymes that catalyse the reversible hydration of carbon dioxide, thereby providing the HCO₃⁻ ion responsible for buffering biological systems.⁸⁷ CAIX is a known diagnostic marker of tumour hypoxia as high levels of CAIX expression are common in a large proportion of carcinomas. A high level of CAIX expression is also indicative of poor survival for a number of cancers.⁸⁸⁻⁹¹

1.2 Targeting the HIF Pathway

Given the critical role HIF plays in the regulation of a cellular response to hypoxia, numerous efforts have been made to identify inhibitors of the HIF pathway. The majority of existing inhibitors have been discovered using cell-based screening techniques, and these act on a number of different targets along the HIF-1 pathway.⁹²⁻⁹⁴ Examples of some of the most prominent targets, and the efforts to develop inhibitors for them, are detailed below.

1.2.1 Inhibitors of DNA binding

One method of inhibiting HIF is to prevent the dimeric HIF from binding to the HRE. Screening of 128 previously identified compounds via an enzyme-linked immunosorbent assay (ELISA) identified the compound echinomycin as capable of inhibiting HIF-HRE binding in a dose-dependent manner (Figure 6). Echinomycin is a natural cyclic peptide that was originally extracted from *Streptomyces echinatus* and is part of the quinoxaline family of antibiotics. It was shown to bind in a sequence specific manner to HREs, and was also shown to inhibit hypoxic induction of luciferase, as well as downregulating VEGF transcription in hypoxic conditions. However, cytotoxicity issues encountered during clinical trials have stalled the development of echinomycin as a therapeutic.⁹⁵ The compound doxorubicin, and the related anthracycline daunorubicin, have been shown to intercalate DNA and induce topoisomerase II facilitated inhibition of DNA replication (Figure 6). This activity has been used to treat cancer in a general manner but was also shown to reduce the expression of HIF-1 target genes in a cell-based reporter assay. Leading to the inhibition of transcription of a HIF-1-dependent reporter gene, as well as the endogenous HIF gene VEGF in the Hep3B-c1 cell line when dosed at 0.2 μM .⁹⁶

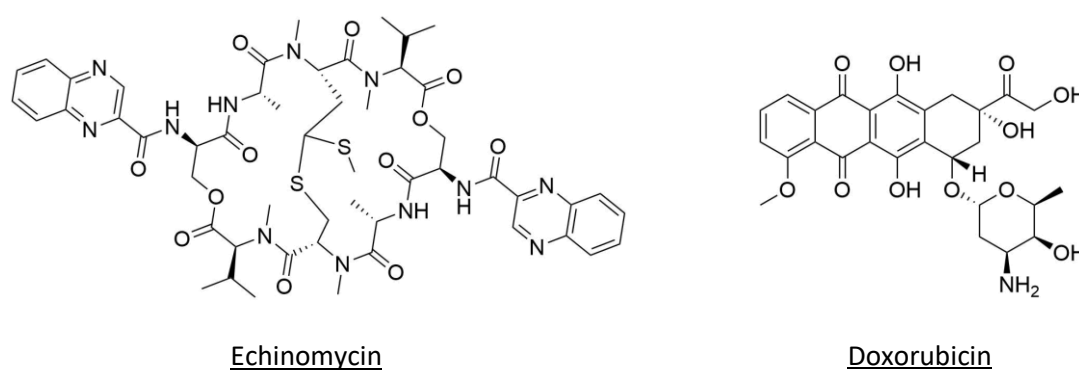
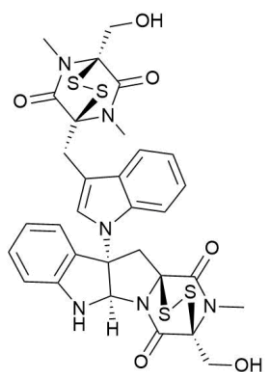


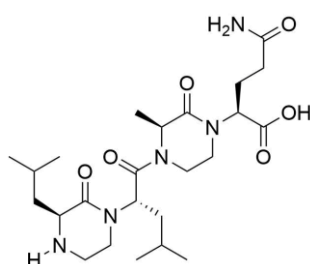
Figure 6. Inhibitors of HIF DNA binding. Echinomycin binds selectively to the HRE preventing HIF binding.⁹⁵ Doxorubicin was shown to bind to the HRE and induce topoisomerase II-mediated strand breaks. It was also shown to reduce HIF activity in both a luciferase reporter assay and qPCR in Hep3B-c1 cells dosed at a concentration of 0.2 μM .

1.2.2 Inhibitors of HIF Transcriptional activity

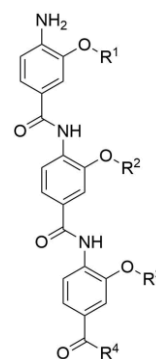
Another strategy of HIF inhibition is to prevent the formation of an active transcription complex. As has been previously discussed, the protein p300 is essential for HIF transcriptional activity, and as such this interaction has attracted attention as a therapeutic target. A study by Kung *et al.* screened a library of 600,000 small molecules in a high-throughput screening assay which led to the discovery of the compound Chetomin (Figure 7).⁹⁷ Chetomin is a dithiodiketopiperazine from the fungus *Chaetomium* that affects the global fold of the CH1 domain of p300, preventing it from binding to HIF-1 α . Chetomin was also shown to decrease HIF-specific reporter activity in mice, and to attenuate the expression of Glut1 and VEGF.⁹⁷ There has also been recent progress in inhibiting p300 binding using peptidomimetics of the HIF-1 α C-TAD (794-826). These include the synthesis of aromatic oligobenzamide foldamers and oxopiperazine helix mimetics that mimic the binding helices of the C-TAD domain (Figure 7). The oxopiperazine OHM 1 (Figure 7) was found to bind to the p300 CH1 domain with an affinity of 530 nM by intrinsic fluorescence spectroscopy, whereas the best oligobenzamide was found to bind with an IC₅₀ of 136 nM. The peptidomimetics then compete with HIF for p300 binding and therefore reduce HIF activity.^{98,99}



Chetomin



Oxopiperazine OHM 1



Oligobenzamide

Figure 7. Inhibitors of p300 binding. Chetomin, a dithiodiketopiperazine, was found to disrupt the interaction between HIF and p300 at concentrations as low as 10 nM. The oxopiperazine OHM1 and oligobenzamide derivatives were designed to mimic a key α -helical domain at the interface of HIF and p300, binding to p300 with a k_d of 530nM and 136 nM respectively.⁹⁷⁻⁹⁹

1.2.3 Indirect Inhibitors

HIF activity can also be indirectly modulated by compounds that affect the enzymatic hydroxylation of the alpha-subunit. One such example is the compound KRH102053 that was found to decrease HIF protein levels in a dose-dependent manner, and consequently mRNA levels of HIF regulated genes, by increasing the activity of PHD2 (Figure 8).¹⁰⁰ The FIH enzyme has also been targeted with the compound YC-1 which upregulates FIH activity, thus preventing the C-TAD

from binding p300 (Figure 8). YC-1 was shown to reduce the activity of a luciferase reporter assay by 50 % at a dose of 2 μ M.¹⁰¹ In addition, Bortezomib has also been shown to upregulate FIH activity, and was shown to enhance FIH-HIF-1 α interactions at concentrations as low as 0.1 nM via immunoprecipitation assays. Although it also functions as an inhibitor of the 26S proteasome, and a suppressor of the PI3K/Akt/mTOR pathway leading to a reduction in HIF protein levels (Figure 8).¹⁰² This means Bortezomib's role in HIF inhibition could be multifaceted, however HIF- α N803A mutant studies have shown that FIH-mediated hydroxylation is the predominant mechanism of HIF inhibition.¹⁰³

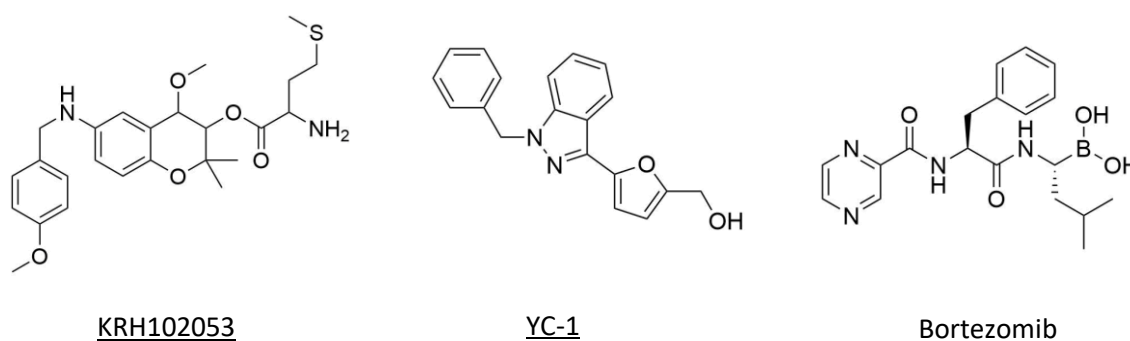


Figure 8. Indirect inhibitors of HIF. KRH102053 was found to be a potent activator of PHD2, that led to a decrease in HIF protein levels as well as a decrease in mRNA levels of HIF-regulated genes. YC-1 was found to prevent recruitment of p330 by the HIF CTAD domain. Bortezomib was also shown to function via preventing p330 recruitment.¹⁰⁰⁻¹⁰³

The interaction between HSP90 and HIF-1 α has been shown to be critical for HIF stabilisation against RACK-1-mediated degradation.⁴⁶ This is therefore another point of potential therapeutic intervention.¹⁰⁴ Inhibitors of HSP90 binding such as Geldanamycin, Ganetespib and Radicolol all function by inhibiting the binding of ATP to HSP90 (Figure 9). This in turn prevents HSP90 from binding to HIF and thus promotes its degradation, which was observed as a decrease in HIF protein levels in hypoxic conditions.¹⁰⁴⁻¹⁰⁶

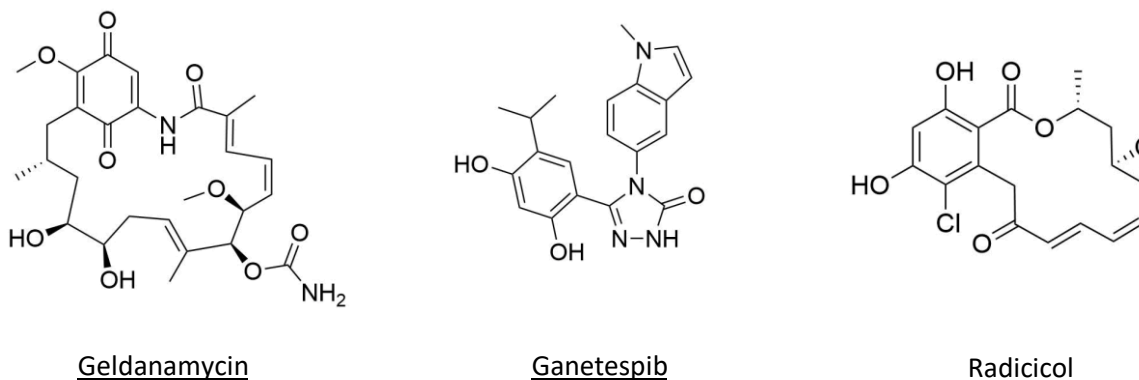


Figure 9. Inhibitors of HSP90. Geldanamycin, Ganetespib and Radicolol all function by inhibiting the binding of ATP to HSP90, preventing it from stabilising the HIF- α protein.

1.2.4 Direct Inhibitors of HIF

HIF activity can also be inhibited by targeting the HIF protein directly. This has the obvious benefit of minimising off target effects, although it has also proven to be the most challenging target for therapeutic intervention. As has been previously discussed, the inhibition of the interaction between the HIF- α C-TAD domain and p300 is an effective method for the suppression of HIF gene expression. However, whilst there are multiple examples of inhibitors that bind to p300, finding inhibitors that bind to the C-TAD domain of HIF has been more challenging. This can be attributed to the conformational flexibility of the C-TAD domain which, when not bound to a partner protein, is intrinsically disordered, making this domain difficult to target therapeutically. In addition, the C-TAD domain has been shown to exist in different conformations depending on which partner it is binding, adopting an alpha-helical structure when bound to FIH, and a beta sheet structure when binding to p300. More success has been had targeting the N-terminal region of HIF. The N-terminal region of the HIF heterodimer is known to be critical for mediating HIF dimerisation, and so inhibitors targeting these domains aim to disrupt the dimerisation of the alpha and beta subunits, thereby preventing HIF mediated gene transcription.¹⁰⁷

One such attempt by Park *et al.* utilised an ELISA assay to screen two small molecule libraries that led to the discovery of the compound NSC 50352 (Rolitetracycline)(Figure 10). Rolitetracycline was shown to inhibit the interaction between the HIF-1 α PAS-A and HIF-1 β PAS-A domains with an IC₅₀ of 1.4 μ M, demonstrating the potential of targeting the PAS domains of HIF.¹⁰⁸

Further evidence of the potential of targeting the PAS domains was provided by Lee *et al.* in 2009, when a library of 200 compounds were screened in a split luciferase assay. The compound Acriflavine (ACF) was found to inhibit the dimerisation of the HIF-1 α /HIF-2 α PAS-B and HIF-1 β PAS-B domains by binding directly to the HIF- α PAS-B domain (Figure 10). The effect of Acriflavine on HIF transcriptional activity was also investigated via a HIF-dependent luciferase assay. This showed that the compound inhibited luciferase expression in a dose dependent manner with an IC₅₀ of approximately 1 μ M. Furthermore, Acriflavine was shown to prevent vascularisation and expression of angiogenic cytokines in mice bearing prostate cancer xenografts, resulting in the prevention of tumour growth in these models.¹⁰⁹ More recently a crystal structure of ACF bound to HIF was obtained that indicated that ACF bound between the HIF-1 β PAS-A and HIF- α PAS-B domains of HIF-2 via a pi-stacking interaction with a tryptophan residue.¹⁸ Whilst Acriflavine has been shown to have potential DNA intercalating properties,¹¹⁰ its use as an anti-cancer therapeutic is still being actively pursued.¹¹¹⁻¹¹⁴

Most recently Cardoso *et al.* performed a reactomics screen of the HIF-1 α PAS-B domain against a library of small molecules containing electrophilic motifs to specifically target the three cysteine

residues present in the HIF-1 α PAS-B domain. The screen identified 'Compound 5' which was shown to be incorporated into the HIF-1 α PAS-B domain at >50 % when dosed at 200 μ M (Figure 10). Tandem mass spectrometry (MS/MS) experiments then determined that the compound was binding to Cys255. Interestingly this cysteine residue appears inaccessible in the x-ray structure 4H6J, however it is located within 8 Å of a water molecule that if displaced could permit access to the thiol side chain. It was shown by surface plasmon resonance (SPR) and an amplified luminescent proximity homogeneous assay (AlphaScreen) that binding of Compound 5 led to a 10-fold loss of affinity for heterodimer formation between the HIF-1 α PAS-B domain and HIF-1 β PAS-B domain. NMR analysis revealed that the binding of Compound 5 to Cys255 led to the perturbation of many residues in the local environment, with most located on the beta-sheet region of the PAS domain which is the interacting interface in the HIF-PAS-B crystal structure.¹⁶

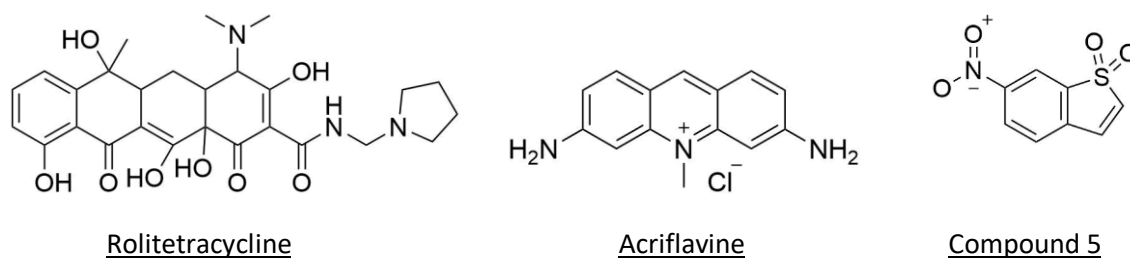


Figure 10. Direct inhibitors of HIF. Rolitetracycline was found to inhibit the interaction between the HIF- α and HIF- β PAS-A domains with an IC₅₀ of 1.4 μ M. Acriflavine is an inhibitor of the interaction between the PAS-B domains of HIF- α and HIF- β with an IC₅₀ of \sim 1 μ M. Compound 5 was discovered via a reactomics screen and was shown to bind covalently to the Cys255 residue of the HIF-1 α PAS-B domain preventing it from interacting with the HIF- β PAS-B domain.

1.2.5 HIF-2 Inhibitors

In addition to these attempts to inhibit HIF-1, it is also worth noting a relatively new breakthrough in the inhibition of HIF-2. Work carried out by Gardner *et al.* in 2009 showed that the PAS-B domain of HIF-2 α contained a large internal pocket of approximately 290 Å³. Having discovered this internal pocket, an NMR-based assay was used to screen for an unnatural ligand capable of binding within it. The study found a number of ligands that bound with K_d values of approximately 10 μ M, the best of which, designated THS-044, bound with a K_d of 2 μ M.¹¹⁵

Having verified that the internal cavity could bind a ligand, an ALPHAscreen was designed to test for compounds that bound the cavity and disrupted the mutant PAS-B-PAS-B heterodimer. A library of over 200,000 compounds was tested using this assay and 70 verified hits were identified, many of which having similar scaffolds to the previously discovered THS-044

compound. The best hit named 'compound 2' was shown to bind to the HIF-2 α PAS-B domain via isothermal titration calorimetry (ITC) with a K_d of 81 nM, a significant increase on THS-044.

Compound 2 was also shown to selectively disrupt HIF-2 activity in 786-O and Hep3B cell lines.¹¹⁶

Through a program of structure-based design, compound 2 was modified to yield the compounds PT2385 and PT2399, which showed promising results in animal clear cell renal cell carcinoma models.^{117,118} The structural basis of PT2385 was investigated with a crystal structure of the HIF bHLH-PAS A-PAS-B heterodimer bound to PT2385 having been obtained. The bound PAS-B domain is shown in Figure 11.¹¹⁹ This highlighted that the compound binds via a number of hydrophobic interactions as well as a hydrogen bond between the hydroxyl group and the side chain of the H293 residue. Compound binding was also found to perturb the Y278 and M252 residues, with the M252 being forced out of the internal cavity and moved towards the junction between the HIF-1 β PAS-B domain and HIF-2 α PAS-B domain.¹¹⁹

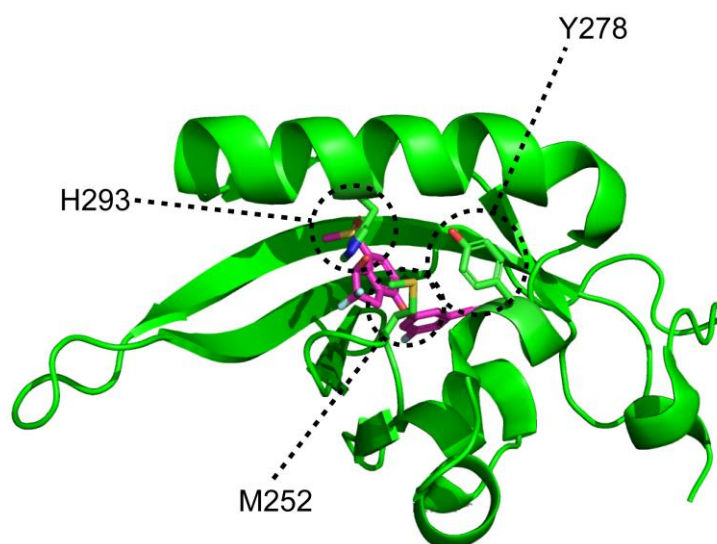


Figure 11. Structural basis of PT2385 binding to the HIF-2 α PAS-B domain. PT2385 was found to form a hydrogen bond with the residue H293. PT2385 was also found to perturb the position of the Y278 and M252 residues. Structure adapted from PDB: 6E3S.¹²³

Most recently, the inhibitor PT2385 was modified to improve the pharmacokinetic profile and prevent the formation of a glucuronide metabolite. This yielded the compound PT2977, which is currently undergoing phase II clinical trials for the treatment of renal cell carcinomas.¹²⁰ The compound series leading to PT2977 is shown in Figure 12.

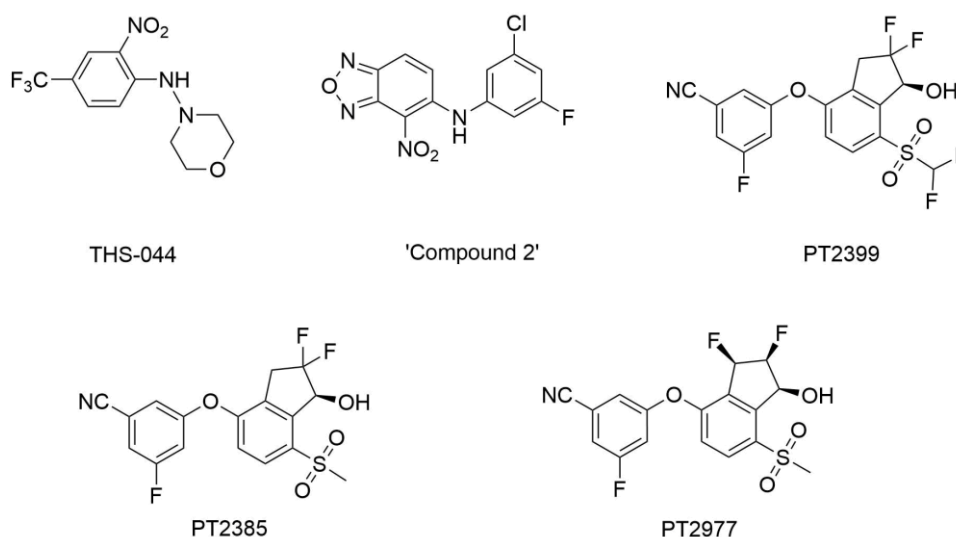


Figure 12. Structures of the initial hit compound THS-004 and intermediates that have led to the HIF-2 α PAS-B inhibitor PT2977.¹¹⁵⁻¹²⁰

Notably, the internal pocket targeted by these compounds in the HIF-2 α PAS-B domain is missing in the HIF-1 α PAS-B domain. The HIF-1 α PAS-B domain has two much smaller internal cavities at 59 and 33 \AA^3 , compared to the single 290 \AA^3 cavity present in the HIF-2 α PAS-B domain. A comparison of the internal pockets of both domains is shown in Figure 13. The difference in cavity size is also evidenced by the presence of eight structural water molecules within the HIF-2 α PAS-B domain (Figure 13B) compared with only 3 in HIF-1 α (Figure 13A).^{16,121}

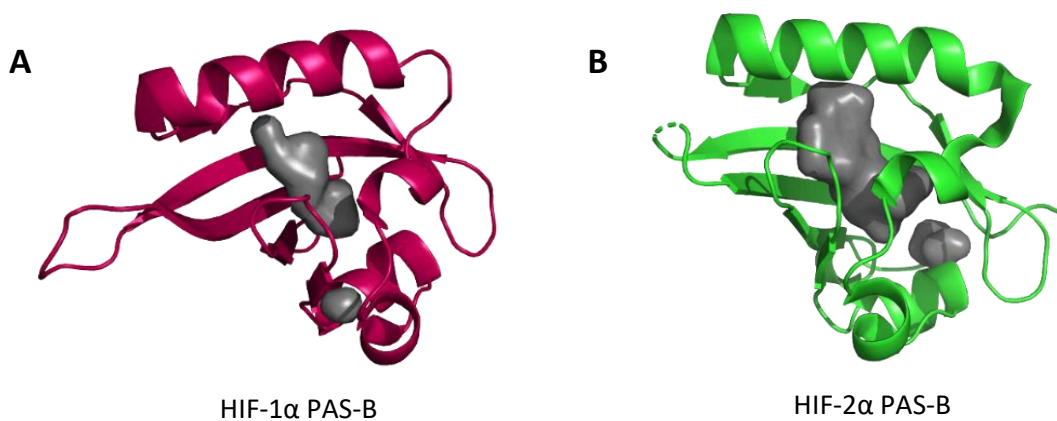


Figure 13. Comparison of the internal cavities of the HIF- α PAS-B domains. A) Internal cavities of the HIF-1 α PAS-B domain (59 and 33 \AA^3). B) Internal cavity for the HIF-2 α PAS-B domain (290 \AA^3). Structures adapted from the PDB files 4H6J and 3F1N.^{15,16}

1.3 Targeting Protein-Protein Interactions

Proteins are integral to all biological processes, and as such so are the interactions between them.

With between 40,000 and 200,000 PPIs predicted to exist within the human interactome, molecules capable of binding to PPIs, for use as probes or therapeutics, are in high demand.¹²²

However, PPIs have proven to be challenging targets for conventional small molecule 'Lipinski like' drugs, as they struggle to bind to the large flat featureless surfaces typical of PPIs.^{123,124}

PPIs are typically composed of surfaces of approximately 1600 Å² that are devoid of appropriate binding pockets for small molecule drugs. However, it has become established that within these large surfaces are 'hot spot' regions that contain a few critical residues that, whilst still not optimal for targeting by small molecules, do present shallow pockets and grooves for the development of ligands.¹²⁵ Identification of these hot spots can be performed via alanine scanning mutational analysis, and has led to the development of molecules capable of binding to these regions, often with better affinities than the corresponding amino acids of the binding partner.^{126–129} It is also worth noting that whilst upon first inspection many PPIs are flat and devoid of binding pockets, many of the protein surfaces are conformationally plastic and can adapt, via the movement of side chains and small perturbations of loops, thus presenting binding pockets that are not present in static protein models.¹³⁰

Despite these challenges, PPIs have become increasingly attractive targets as the understanding of communication between proteins has improved, allowing the molecular basis for a number of diseases to be elucidated. This is particularly true in complex diseases such as cancer, and so new therapeutic scaffolds have started to be investigated.

1.3.1 Peptides as Therapeutics

In recent years, the pharmaceutical industry has been shifting away from classical drug targets such as G-protein coupled receptors, enzymes and ion channels, in order to pursue new target classes. These new classes, such as transcription factors and protein-protein interactions have in the past been labelled as intractable; however, as new drug modalities have become available these targets are being re-examined. One class of molecules showing promise with these difficult targets are peptides. Peptides are a unique class of compounds that are positioned between small molecules and proteins. The ability of peptides to act as signalling molecules within the cell presents the opportunity to develop therapeutics that mimic natural pathways. In addition, the amino acid composition of peptides means that they can mimic the protein surfaces used to mediate many cellular processes, providing another route to therapeutic intervention.¹³¹

Peptides also possess secondary structures that arise from intramolecular hydrogen bonding, as well as the chirality of the amino acid building blocks. This enables polypeptide sequences to form three-dimensional structures that can bind to the previously mentioned hot-spot regions of a PPI in a way small molecules cannot. A further advantage of their amino acid composition is that peptides often have little to no toxicity.¹³²

Despite these advantages, natural polypeptide sequences typically have poor absorption, distribution, metabolism and excretion (ADME) properties. Their susceptibility to proteases causes rapid breakdown and clearance with typical half-lives of less than 30 minutes, and their poor solubility and passive permeability means they typically have a bioavailability of below 2%.¹³³ As the field of peptide therapeutics has evolved, a number of strategies have been developed for the improvement of these ADME properties, in particular increasing the peptides proteolytic stability.¹³⁴ One class of protease responsible for the breakdown of peptide ligands *in vivo* are exopeptidases which act to cleave from either the N or C terminus of a peptide. As such, chemical modification of these termini can hamper exopeptidase activity. N-acetylation, C-amidation and PEGylation (attaching the polymer polyethylene glycol) have been shown to increase proteolytic resistance in a number of cases.^{135,136} For example, the N-acetylated glucagon-like peptide GLP-1-(7-34)-amide analogue was reported to have greatly increased stability when compared to the native peptide.¹³⁷ Similarly, PEGylation of the peptide MART-1 was shown to vastly improve its resistance to proteolysis and plasma stability.¹³⁸

Another strategy for increasing proteolytic resistance is substituting L-amino acids for D-amino acids. This reduces the substrate recognition of proteolytic enzymes thus increasing the peptides stability. One example of this can be seen with the peptide vasopressin which has a half-life of 10 to 35 minutes.¹³⁹ Swapping an arginine residue with D-arginine increased this figure to over 3.7 hours.¹⁴⁰

Chapter 1

In addition to the strategies for protecting hydrolysis of the amide bond, an alternative method for increasing the stability of peptides is to replace the amide bond with a mimetic. There is a diverse set of chemical modifications that can be made and some are shown in Figure 14. All of these mimetics prevent proteolysis by blocking recognition of the peptide by proteases, but can also serve to increase the peptides permeability and potency.

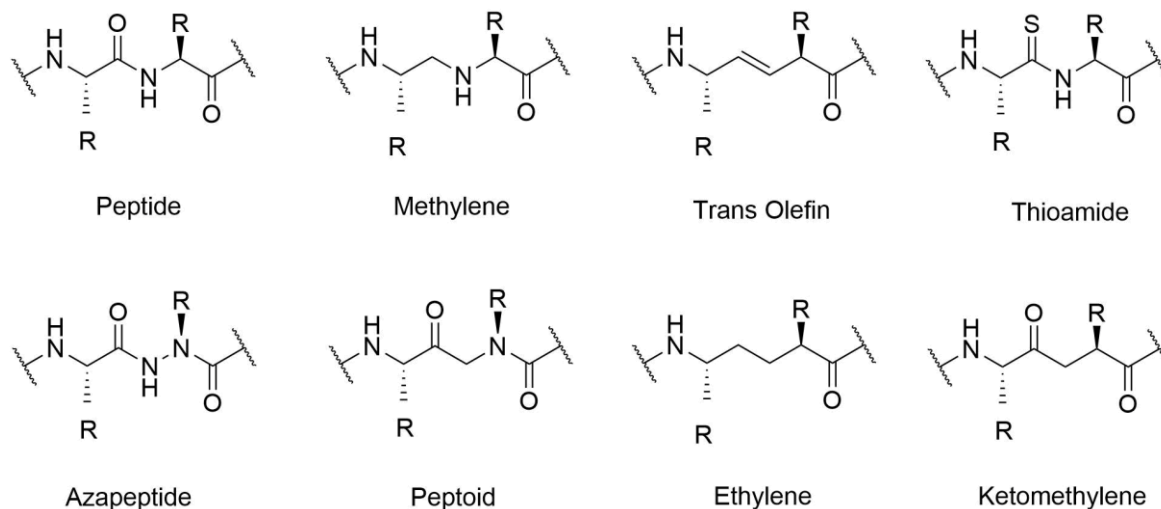


Figure 14. Structures of some of the most common moieties used as amide mimetics.

Lastly, the stability and ADME properties of peptides have been shown to improve upon cyclisation of the peptide sequence. Many cyclic peptides have had success as therapeutics, including the immunosuppressant cyclosporine,¹⁴¹ antibiotics such as vancomycin and daptomycin,^{142,143} as well as the synthetic analogue of somatostatin, octreotide,¹⁴⁴ all of which demonstrate the potential of this scaffold.

Cyclic peptides possess many attributes that make them desirable therapeutics. One advantage is that cyclic peptides have a large surface area when compared to both linear peptides and small molecule drugs. This makes them better suited for tackling the hot-spots that mediate PPIs.¹⁴⁵ Their cyclic form also confers an element of rigidity to the molecule, minimising entropy loss upon binding. The larger size of the cyclic peptide scaffold also conveys the ability to be highly selective, with cyclic peptides having been shown to be capable of distinguishing between target isoforms displaying a high degree of homology.^{146,147} Crucially, unlike their linear counterparts, cyclic peptides are more resistant to degradation by proteases as there is no N or C terminus.

Despite the promise of this scaffold, there is still one significant issue when progressing these hits toward the clinic, and that is their poor bioavailability. Whilst methods have been developed to circumvent this issue, such as cell penetrating peptide sequences, it is still a major disadvantage, with oral administration being a key goal of the drug development process.¹⁴⁸ The poor bioavailability of cyclic peptides is not surprising as they break most, if not all, of the established rules for 'drug-likeness' such as the Lipinski's rule of five.¹⁴⁹ Considerable effort is being devoted

towards the development of new rules, using naturally occurring examples such as cyclosporine to develop a new understanding of the link between structure and bioavailability for these beyond rule of five (bRo5) compounds.^{150–152}

1.3.2 DNA Encoded Methods for the Generation of Cyclic Peptides

Cyclic peptides are relatively easy to synthesise, with automated solid phase methods enabling easy access to peptide scaffolds. If a suitable biological assay is available for the target of interest, then this can be used to screen a chemically synthesised peptide library. However this is a costly approach that requires complex logistics impeding the screening of very large libraries.¹⁵³ In order to synthesise and screen libraries of unprecedented size, genetically encoded screening platforms are required. The biological nature of peptides mean that they are compatible with genetically encoded methods of library generation. This typically involves the generation of a library of DNA with a randomised region incorporated via randomised oligonucleotides which, upon translation, give rise to many different peptide sequences. The encoded nature of the library is then utilised to create a phenotype-genotype link between each member of the library and the DNA that codes for the peptide sequence. This is the defining feature of genetically encoded platforms that enables them to successfully manipulate libraries of unprecedented size, as when coupled with next generation sequencing, it enables the rapid identification of hit sequences. Some of the most popular DNA encoded methods of cyclic peptide library generation are detailed below.

1.3.2.1 Phage Display

First developed by George P Smith in 1985, phage display fuses a molecule of interest to a native phage protein, which is subsequently displayed on the phage surface.¹⁵⁴ This is most commonly achieved with the non-cell-lytic F-specific filamentous phage.¹⁵⁵ The phage possess five different coat proteins all of which have been demonstrated to be compatible with phage display, but it is the minor coat protein pIII that is most commonly used due to its tolerance to a high variety of substituents.¹⁵⁶

The ability of the displayed molecule to bind a particular target can then be assessed by allowing the phage to bind to an immobilised target. A wash step is then used to remove any non-binding sequences, before amplification of successful binders in host cells. This process termed 'biopanning', is typically repeated 3-5 times before sequencing is used to elucidate the highest affinity sequence. This process is outlined in Figure 15A. The display of cyclic peptides was first achieved via the incorporation of two cysteines, typically in the format of CX_nC with X representing any other proteinogenic amino acid, which post-translation would oxidize to form a disulphide linked cyclic peptide.^{157,158} This technique has been successful in the identification of

Chapter 1

numerous cyclic peptides that bind a wide variety of diverse targets, with affinities typically in the micromolar range.^{159–161}

A modification of this technique involves utilising chemical linkers such as 1,3,5-tris(bromomethyl) benzene (TBMB) to enable the screening of bicyclic peptide libraries.¹⁶² Bicyclic peptides enhance many of the advantages conferred to monocyclic peptides, such as the reduction in flexibility which can lead to a reduced entropic barrier to binding, as well as higher resistance to proteases. This technique requires three cysteines to be present in a displayed linear sequence, which upon addition of a linker reagent bearing three thiol-reactive groups will form the desired bicycle scaffold (Figure 15B). These scaffolds retain their rigidity at greater molecular weights than monocycles, enabling larger peptides and therefore greater chemical diversity.¹⁶³

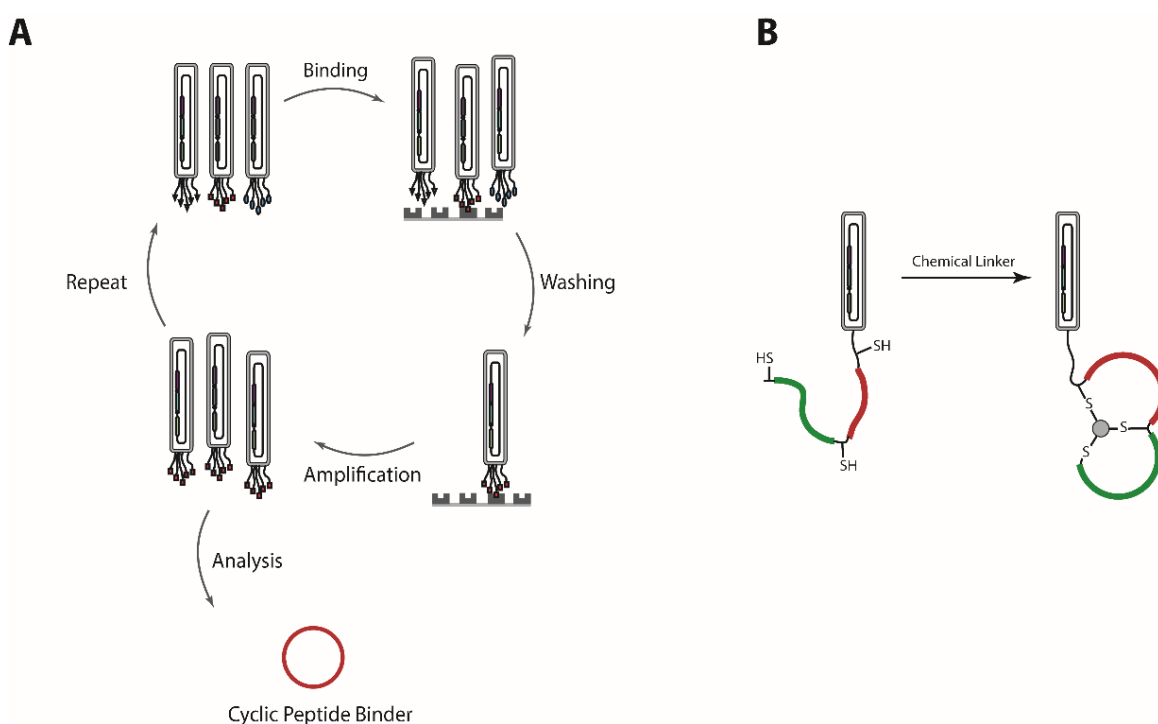


Figure 15. Overview of phage display. A) Summary of the steps involved in phage biopanning to enable identification of peptide binders. B) Formation of bicyclic phage displayed peptide libraries via chemical linkers.

1.3.2.2 mRNA Display

Another genetically encoded method of cyclic peptide library generation is mRNA display, the defining feature of which is that it is an *in vitro* selection technique, which provides the advantage of avoiding the need to transform genetic libraries into a host organism.¹⁶⁴ The transformation step is typically the limiting factor for genetically encoded library sizes, with typical

transformation efficiencies in *E. coli* limited to $\sim 10^9$ transformants.¹⁶⁴ The limiting factors for mRNA display are based on how much material can be generated and handled, enabling mRNA display to create libraries of up to 10^{13} members.¹⁶⁵ mRNA display maintains a phenotype-genotype link by utilising a puromycin linker, which joins the peptide sequence to the 3' end of its own mRNA via a small linker. When the junction between the two reaches the ribosome, translation pauses and the puromycin is taken into the A site of the ribosome. It then covalently bonds the 3' end of the mRNA to the C-terminus of the peptide chain.¹⁶⁶

A library of random mRNA sequences can then be translated to yield a huge variety of peptides, which can then be subjected to rounds of screening and enrichment to isolate those that display a high affinity for the target. Reverse transcription and PCR amplification of the mRNA tag can then be used to reveal the sequence of the isolated peptides.¹⁶⁴ With the C-terminus of the peptide bound to puromycin, head-to-tail cyclisation of peptides was initially incompatible with mRNA display. However, work by Suga *et al.* has demonstrated that, via the incorporation of a number of unnatural amino acids, this issue can be circumvented.¹⁶⁷ This work utilised a ribozyme capable of charging tRNAs with unnatural amino acids termed a flexizyme.¹⁶⁷ Flexizymes have been developed to lack side chain recognition, and so have a high tolerance for different side chain moieties, which enables the synthesis of tRNAs that are loaded with a variety of natural and unnatural amino acids.¹⁶⁸ A system called flexible *in vitro* translation (FIT) is then utilised whereby selected amino-acyl tRNA synthetases and amino acids are omitted from the *in vitro* transcription translation (IVTT) mixture, effectively vacating the corresponding codons. The flexizyme-generated tRNAs bearing non-proteinogenic amino acids can then be substituted in along with an anti-codon matching that of the omitted tRNAs.¹⁶⁹ Whilst this enables a number of synthetic routes to cyclic peptides, the most widely utilised method uses the FIT system to charge tRNAs with unnatural amino acids containing a chloroacetyl moiety. This residue is then susceptible to nucleophilic substitution with any cysteine residues present in the peptide forming a thioether-closed macrocycle. This reaction occurs spontaneously in mild aqueous conditions to yield the cyclic peptide almost quantitatively.¹⁷⁰⁻¹⁷²

Integrating the FIT system into an mRNA display platform allows for the creation of very diverse libraries that are no longer restricted to the 20 canonical amino acids. This vastly increases the utility of the platform as it allows for the creation of macrocyclic peptide libraries containing unnatural residues, including D-stereochemistry and N-methylated peptides.^{170,173} This process has been named as random non-standard peptide integrated discovery (RAPID), and has already demonstrated itself as a powerful tool for the discovery of peptide therapeutics. The RAPID system is outlined in Figure 16.¹⁷⁰ For example, the thioether-closed peptide Pakti-L1 was found to bind to the target protein Akt2 with an IC_{50} value of approximately 100 nM. Furthermore, the

peptide was found to exhibit a specificity of greater than 1000-fold for the Akt2 isoform over the closely related Akt1 and Akt3.¹⁴⁷

1.3.2.3 SICLOPPS

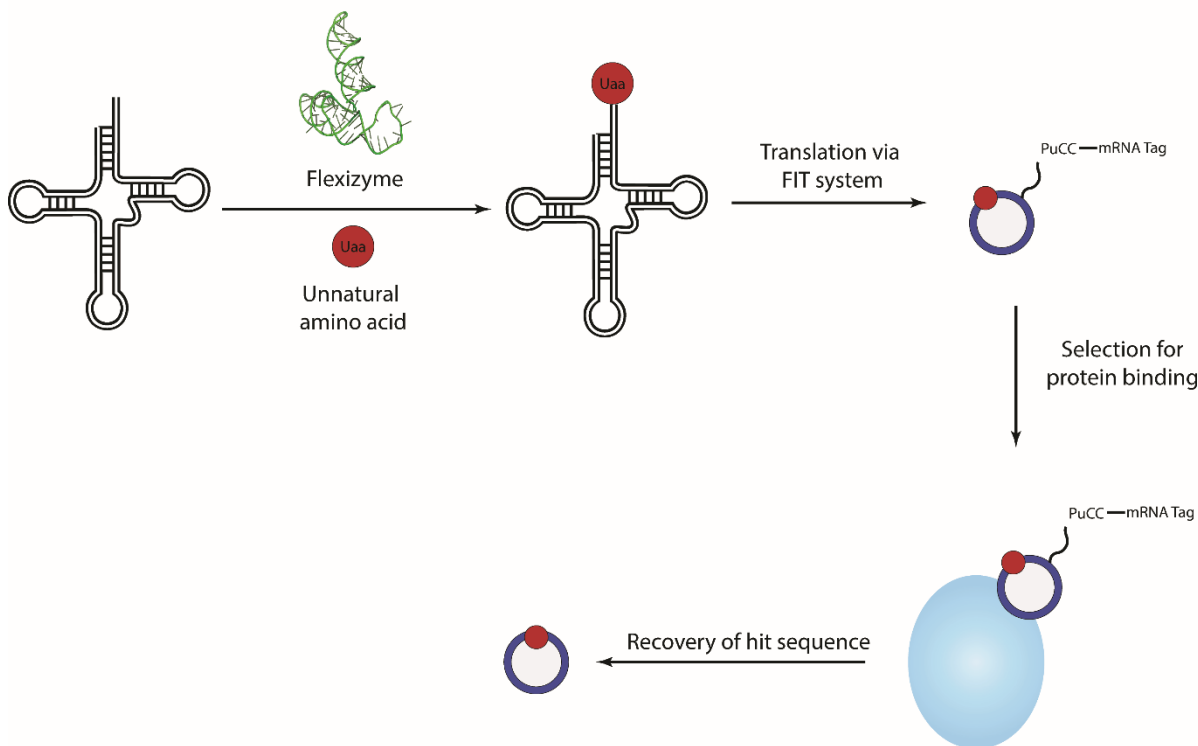


Figure 16. Overview of the FIT-RAPID system for the discovery of cyclic peptide binders. tRNA's are charged with unnatural amino acids that allow for cyclisation of the peptide whilst retaining the mRNA tag. Screening against target proteins then enables the identification of peptide binders.

One final method for the generation of peptide libraries is split-intein circular ligation of peptides and proteins (SICLOPPS), which utilises intein splicing to generate cyclic peptides within cells. The SICLOPPS construct consists of an artificially split intein, resulting in a C-terminal intein and an N-terminal intein which flank a polypeptide sequence as the extein. Post translation, the inteins rearrange to yield an active *cis*-intein, that upon splicing results in the cyclisation of the polypeptide sequence as shown in Figure 17.¹⁷⁴ This cyclic peptide can be varied at the DNA level on the plasmid, whereby a randomised oligonucleotide can be incorporated as typically either NNS or NNB (where N represents A,T,C or G. S represents C or G, and B representing C,G or T). This allows for the creation of libraries of up to 10^8 members containing any of the 20 canonical amino acids.¹⁷⁵ Whilst there is an example of using orthogonal aminoacyl-tRNA synthetase/tRNA_{CUA} pairs to incorporate an unnatural amino acid into the SICLOPPS screening platform,¹⁷⁶ it

could be argued that IVTT techniques are better suited for this purpose, and as such most SICLOPPS screens have used proteinogenic building blocks.^{146,177–181}

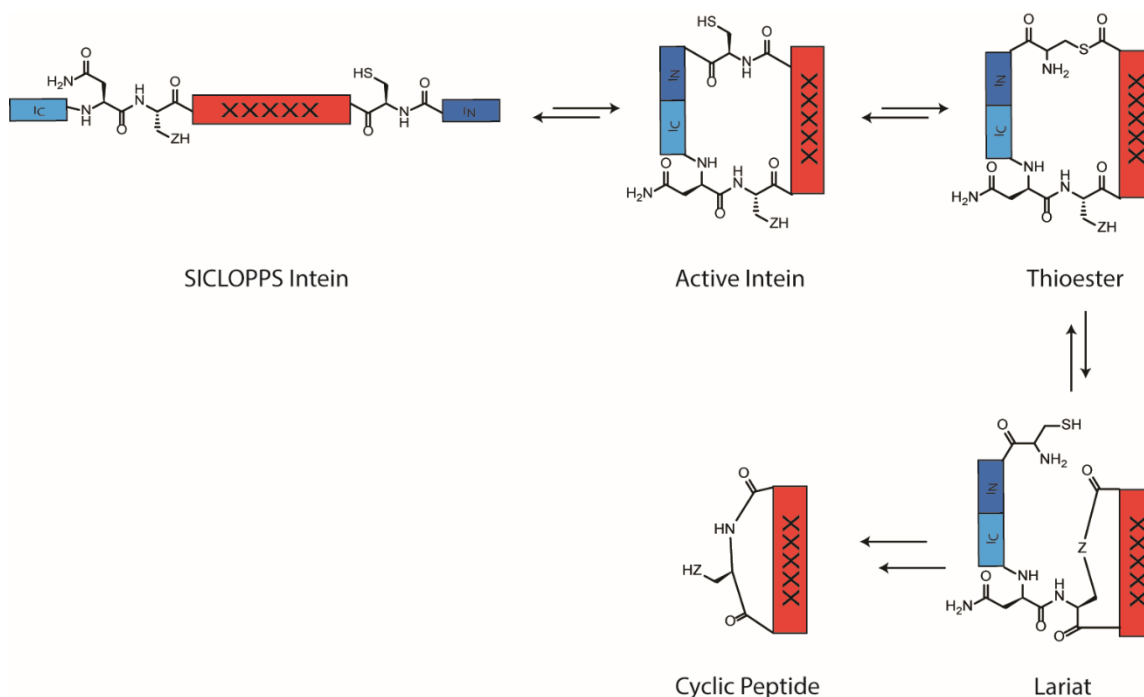


Figure 17. Mechanism of SICLOPPS intein splicing. The two inteins I_C and I_N associate to form an active intein that splices to yield the cyclised product, in this case CXXXXX, where X is any proteinogenic amino acid.

SICLOPPS libraries have most commonly been employed in concert with a bacterial reverse two-hybrid system (RTHS). This bacteriophage regulatory system links the disruption of a protein dimer with the expression of a reporter cassette typically consisting of the genes $HIS3$, Kan^R and $LacZ$. $HIS3$ encodes the protein imidazoleglycerolphosphate dehydratase, without which an auxotrophic *E. coli* strain cannot synthesise its own histidine, and so the use of a growth medium lacking histidine inhibits the growth of cells not expressing this gene.^{177,178} The Kan^R gene encodes the protein aminoglycoside 3'-phosphotransferase which provides resistance to the antibiotic Kanamycin. The use of media containing Kanamycin will therefore inhibit the growth of *E. coli* not expressing this protein.^{177,178} The last gene of the reporter cassette is $LacZ$, which enables quantification of the disruption via a β -galactosidase assay.^{177,178} The coupling of SICLOPPS and RTHSs has been used in the identification of many inhibitors of protein-protein interactions (PPIs).^{146,179–182}

One advantage of using SICLOPPS in *E. coli* is that there is a greatly reduced risk of identifying off-target inhibitors when screening for eukaryotic proteins. This is due to the upstream and downstream protein partners being absent in *E. coli*, thus circumventing the main disadvantage of cell-based functional screening. Another advantage of this system is that it is a functional screen

as only cyclic peptides that disrupt the dimer of interest are selected for. This is in contrast to other techniques which identify hit compounds based solely on their affinity for binding the target, which does not necessarily equate to having the desired function.¹⁸³ Further, the use of a positive selection system enables the growth and enrichment of the library post-screen, a workflow further benefited by the fast growth rates of *E. coli*.

1.4 CLLFVY

Previous work by Miranda *et al.* within the Tavassoli group discovered the cyclic peptide cyclo-CLLFVY as being capable of disrupting dimer formation between HIF-1 α and HIF-1 β .¹⁴⁶ This was achieved by screening a SICLOPPS generated cyclic peptide library of 3.2 million members against a bacterial RTHS. The RTHS consisted of the HIF-1 α bHLH-PAS-A-PAS-B and HIF-1 β bHLH-PAS-A-PAS-B conjugated to the proteins P22 and 434 respectively. When the two proteins dimerise, the P22 and 434 proteins come together to form a functional repressor preventing expression of the downstream reporter cassette. If a cyclic peptide is capable of inhibiting the dimerisation of HIF-1 α and HIF-1 β then the repressor complex cannot form and the reporter cassette is expressed. This process is outlined in Figure 18.

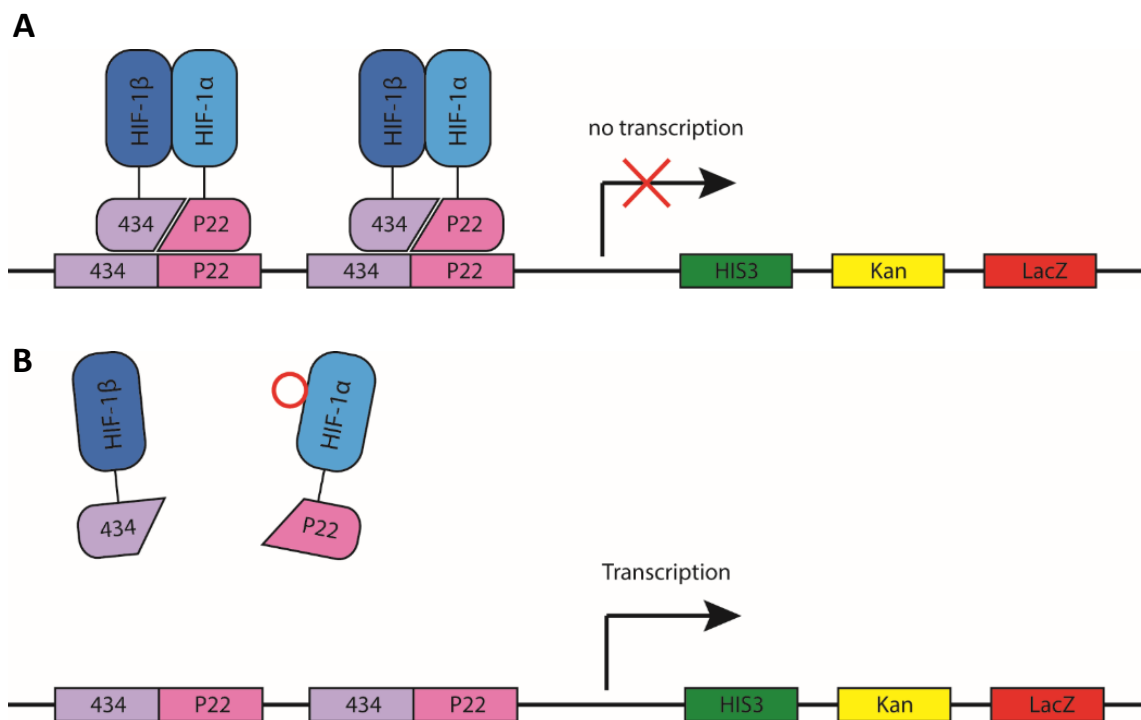


Figure 18. RTHS used to discover cyclic peptide inhibitors of HIF-1 α /HIF-1 β dimerisation. A) In the absence of an inhibitor, a functional repressor complex is formed preventing expression of the downstream reporter cassette. **B)** If a cyclic peptide is capable of disrupting the dimerisation of HIF-1 α and HIF-1 β , expression of the downstream reporter cassette occurs, enabling survival of that *E. coli* colony.¹⁴⁶

Sequencing of the SICLOPPS plasmid of the surviving *E. coli* revealed the sequence cyclo-CLLFVY. The peptide was then obtained synthetically, and a cell-penetrating peptide tag (TAT-tag) was attached to enable a series of *in vitro* and *in vivo* assays to be carried out. A luciferase-reporter assay with the TAT-tagged cyclo-CLLFVY (P1) found a dose-dependent reduction in luciferase signal when compared to untreated cells in normoxia. P1 was also shown to reduce transcription of VEGF and CAIX in a dose-dependent manner via qPCR analysis. In order to determine to which domain of HIF-1 α P1 was binding to, a fluorescently tagged derivative of P1 was created. The compound was then added to wells containing plate bound HIF domains and the fluorescence measured to assess binding to each domain. This assay showed that P1 preferentially binds to the PAS-B domain of HIF-1 α over the PAS-A and bHLH domains.

Further work on cyclo-CLLFVY was published in 2016 by Mistry *et al.*, where it was demonstrated that a modified line of HEK-293 cells could express chromosomally encoded cyclo-CLLFVY.¹⁸⁴ This expression was placed under the control of both a tetracycline-sensitive promoter and hypoxia itself. This dual control enabled both selective expression in hypoxic cells as well as temporal control, as expression would require treatment with doxycycline. This system was capable of inhibiting HIF-1 dimerisation in a proximity ligation assay, as well as reducing the transcription of the HIF-1 regulated genes VEGF and CAIX as determined by qPCR.¹⁸⁴

Perhaps most notable is P1's ability to selectively inhibit HIF-1 α hypoxia signalling whilst having no observed effect on HIF-2 α regulated VEGF, LOX or CITED2 mRNA transcription by qPCR. It was also demonstrated that cyclo-CLLFVY had no effect in a HIF-2 α luciferase reporter assay. Specifically targeting one isoform of HIF is extremely challenging due to their high sequence homology, and is something that has thus far eluded other inhibitors of HIF.¹⁴⁶

1.5 Project Aims

HIF represents a critical point for therapeutic intervention in cancerous tumours, and therefore molecules capable of binding to and inhibiting HIF are of particular interest. The distinct and sometimes antagonistic roles of the HIF-1 α and HIF-2 α isoforms makes isoform specific inhibitors especially valuable. Despite this, there are few compounds that are capable of inhibiting HIF by directly interacting with the HIF-1 α protein. Of these, only CLLFVY is capable of binding to the HIF-1 α protein in an isoform specific manner at the time of writing.

This work aims to improve the CLLFVY hit peptide via an *in vitro* approach, whereby a series of derivatives of the parent peptide will be made and their affinity and activity will be assessed with a recombinant HIF protein. Through this process, a lead compound will be selected that can be used in the development of further *in vitro* assays that will enable a more detailed assessment of

Chapter 1

CLLFVY activity. The incorporation of a fluorophore to enable the use of CLLFVY as a tool for understanding the mechanism of HIF inhibition will also be investigated.

The structural basis for CLLFVY will also be probed in an attempt to gain structural data that will enable a rational approach to the development of future HIF targeting peptides.

Chapter 2 Investigating the Structure Activity

Relationship of CLLFVY

2.1 Synthesis of Cyclo-CLLFVY

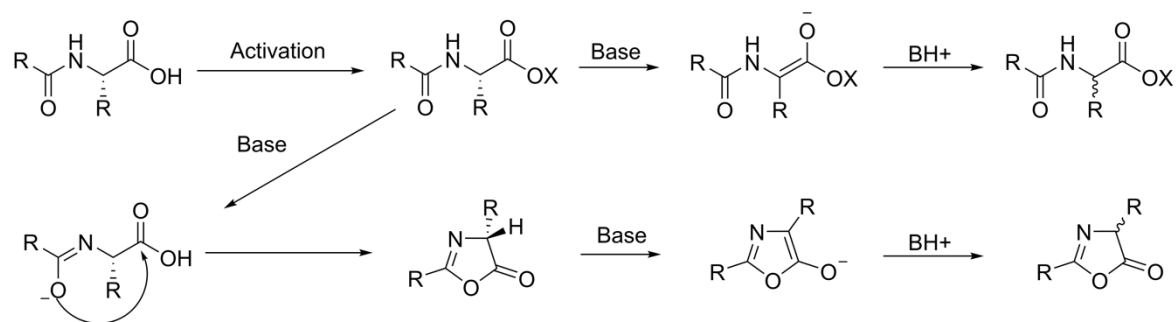
2.1.1 Solid-Phase Peptide Synthesis

Solid-Phase Peptide Synthesis (SPPS) has revolutionised peptide synthesis since the first Merrifield resin method was published in 1963.¹⁸⁵ SPPS has several advantages over solution phase techniques, most prominent is the ease of which the reagents can be removed with filtration. This enables the use of excess reagents that can be filtered off and the resin then washed to remove any unreacted reagents. Merrifield resin utilises a Boc strategy with the N α of the amino acids protected with a Boc group and side chains requiring acid stable protecting groups. The N-terminus is then deprotected with TFA before being reacted with an activated carboxyl of a second amino acid. Repetition of these two steps is then carried out to build out the desired linear sequence. The major drawback of this method is the need to use HF to cleave sequences from the solid support which requires special handling and safety precautions.

Merrifield's SPPS has since been adapted with an array of different resins, protecting group strategies and coupling reagents,¹⁸⁶ with one of the most common methods being the Fmoc strategy. Fmoc SPPS utilises the fluorenylmethoxycarbonyl group to protect the N α of amino acid building blocks, which is deprotected with a solution of 20 % piperidine in DMF. The base lability of the Fmoc protecting group enables cleavage of the linear sequence under milder acidic conditions, typically a solution of TFA, avoiding the need to have the laboratory set up for handling HF. In contrast to Boc synthesis, this strategy requires base stable, acid labile protecting groups.

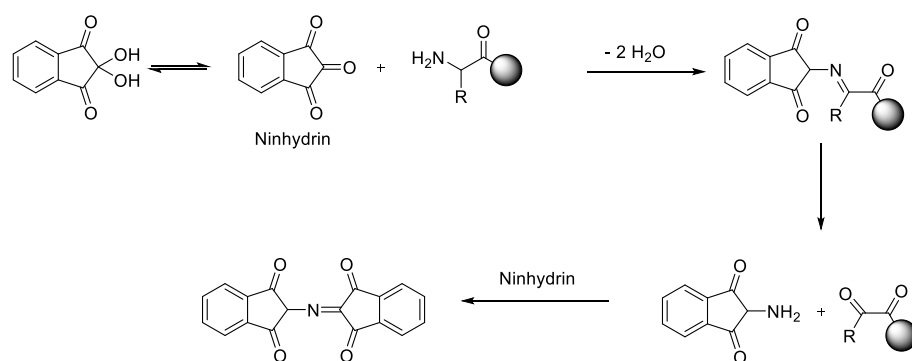
After deprotection of the N-terminus of the peptide, the next amino acid in the sequence requires its C-terminus to be activated for it to undergo aminolysis. There are a plethora of different activating agents available, however the most traditional approach is the use of carbodiimides.¹⁸⁶ Reaction of carboxyl groups with carbodiimides forms a highly reactive O-acyl urea intermediate. Whilst this activates the C-terminus to nucleophilic attack from the free amine, it can however have the unintended consequence of a loss of stereochemistry of the amino acid. There are two routes via which activated amino acids are known to racemise, both of which are base mediated,

and these are oxazolone formation and direct enolisation. These mechanisms are outlined in Scheme 5 below.



In order to suppress racemisation an additional reagent is commonly added to form an activated species with modulated reactivity. Finding reagents that will sufficiently activate the C-terminus whilst maintaining the stereochemistry of the amino acid has become a key aspect of peptide synthesis, with this area being extensively studied.^{187–189} Most prominent among these reagents has been the benzotriazole derivatives HOBt and HOAt, which react with the O-acyl urea intermediate to form an activated ester, suppressing racemisation during the coupling reaction. More recently, a new reagent called Oxyma Pure has been reported that has equal or better racemisation suppression than the benzotriazoles.¹⁹⁰ Oxyma Pure is also cheaper and less hazardous, due to it not being classified as explosive which the benzotriazole derivatives are.¹⁹¹ For these reasons Oxyma Pure was selected for use in the synthesis of CLLFVY and its derivatives.

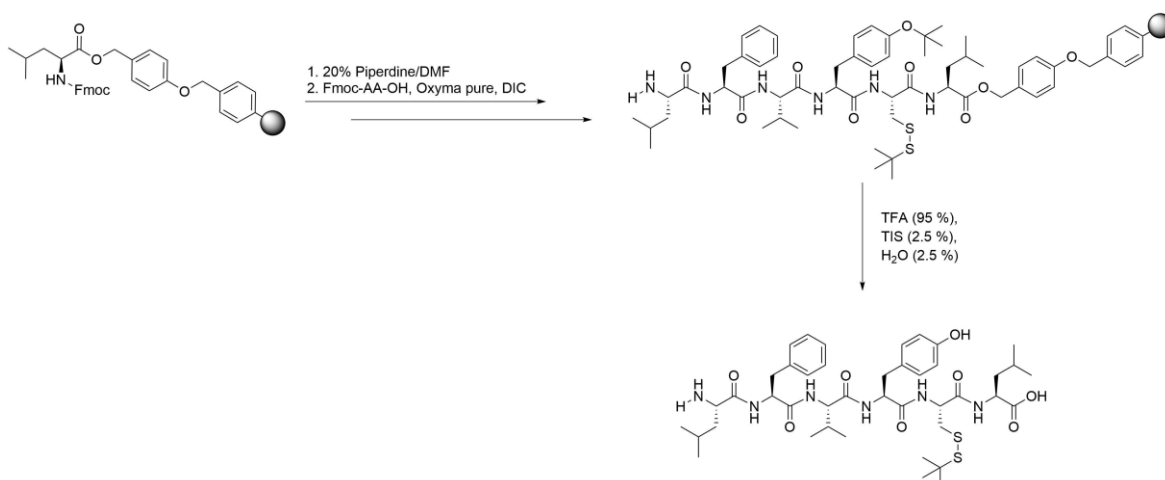
In order to monitor SPPS without having to cleave the peptide after each coupling or deprotection step, a Kaiser test is instead carried out. This involves the use of the dye compound Ninhydrin which in the presence of a free amine is converted into the blue compound Ruhemann's purple as shown in Scheme 6. This test can therefore be used to determine if coupling/ deprotection steps are complete whilst the peptide is still bound to the resin.



Scheme 6. Mechanism of formation of Ruhemann's purple which allows the Kaiser test to detect the presence of free amines.

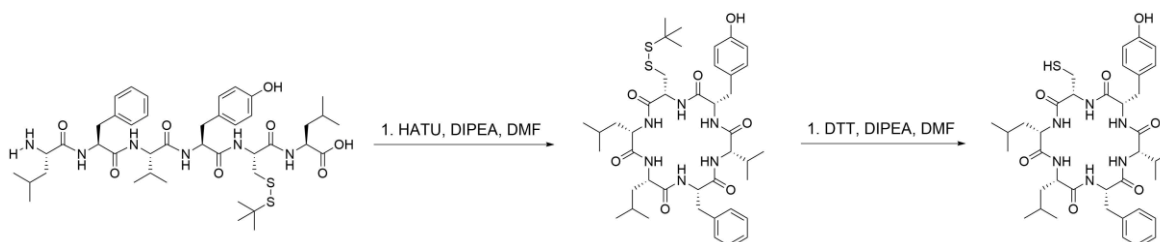
The linear sequence LFVYCL was synthesised from preloaded Leu-Wang resin using standard Fmoc SPPS, with the exception of the cysteine protecting group utilised. In place of the standard trityl protecting group an acid stable StBu group was used. As the StBu group is stable to TFA the thiol

remains protected after cleavage from the resin. This is necessary to prevent it from reacting in the cyclisation step, which would result in a mixture of products. Linear sequences were then cleaved from the resin using a TFA cleavage cocktail as shown in Scheme 7.



Scheme 7. Synthesis of linear LFVYCL using Fmoc SPPS, and subsequent cleavage from the Wang resin using a TFA cleavage cocktail.

Once the linear sequences were obtained, they next had to be cyclised. Efficient backbone cyclisation of linear peptide sequences shorter than eight amino acids can prove to be synthetically challenging, and is often dependent on the sequence being cyclised.¹⁹² In order to prevent intermolecular reactions occurring between linear peptides the cyclisation was carried out under high dilution (<5 mM), using the reagents EDC and Oxyma Pure. The final step in the synthesis of CLLFVY was to remove the StBu cysteine-protecting group, and this was done reductively using the reagent DTT, as shown in Scheme 8, to yield the desired peptide cyclo-CLLFVY.



Scheme 8. Cyclisation of LFVYCL and subsequent deprotection of cysteine butyl thiol group.

2.2 Recombinant Expression of HIF-1 α PAS-B Domain

In order to develop an *in vitro* assay for determining the binding affinity of CLLFVY, it was first necessary to obtain the HIF protein. As it had been previously demonstrated by Miranda *et al.* that CLLFVY binds to the PAS-B domain of HIF-1 α ,¹⁴⁶ it was decided that recombinant expression of the isolated PAS-B domain would provide an adequate model for further study. A study by Zhu *et al.* previously demonstrated the expression of this domain and so a similar approach was pursued.¹⁷ The HIF-1 α PAS-B domain (aa's 238-349) was expressed from a pET28a vector containing an N-terminal hexahistidine (HIS) tag (Figure 19), transformed into BL21 (DE3) Rosetta cells.

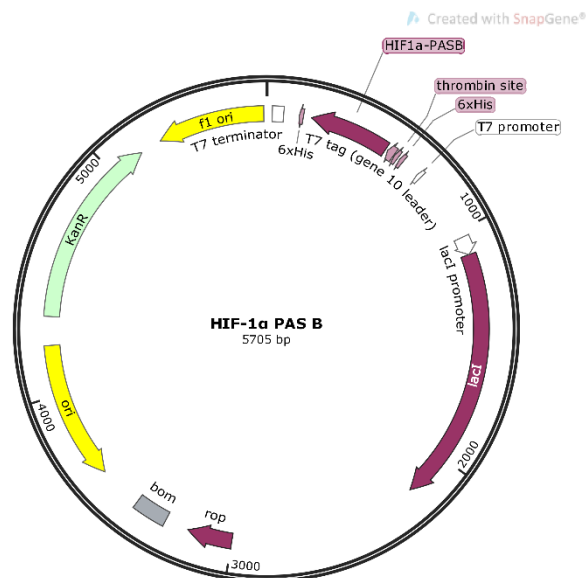


Figure 19. Plasmid map of the pET-28 vector used for expression of HIF-1 α PAS-B domain.

BL21 are the most common host for recombinant protein expression due to their lack of the proteases *lon* (4) and *ompT*, which helps prevent the degradation of the recombinantly expressed protein. The DE3 designation signifies that the strain contains the T7 RNA polymerase gene on its chromosome under the control of the *lac* promoter, this is crucial for expression of proteins from pET vectors which contain T7 promoters. The T7 promoter is selected over the *lac* promoter due to its superior expression activity, which can lead to the target protein accounting for 50 % of total cell protein.^{193,194} Lastly, BL21 (DE3) Rosetta cells were chosen as they contain a pRARE chloramphenicol resistant plasmid that enhances the expression of mammalian proteins containing codons that rarely occur in *E.coli*.

Protein expression was induced via IPTG from a culture of BL21 (DE3) Rosetta cells transformed with the HIF-1 α PAS-B pET-28 vector, which was incubated at 16 °C overnight. The cells were then lysed and the protein isolated from the soluble fraction via nickel affinity chromatography. The

protein was then purified via size exclusion chromatography (SEC) to yield a major peak at 79.2 mL, shown in Figure 20A. The expression was found to yield 2 mg/L of desired protein. Whilst this represents an increase in yield when compared to that obtained by Zhu et al (0.7 mg/L), it is still a low yield for over expression in *E. coli*, demonstrating the challenges of expressing HIF proteins. Analysis by SDS-PAGE showed that the protein fractions obtained from the S75 purification produced a single band at the expected mass. A small amount of protein was seen in the insoluble fraction suggesting that a small amount of the protein could be aggregating during expression.

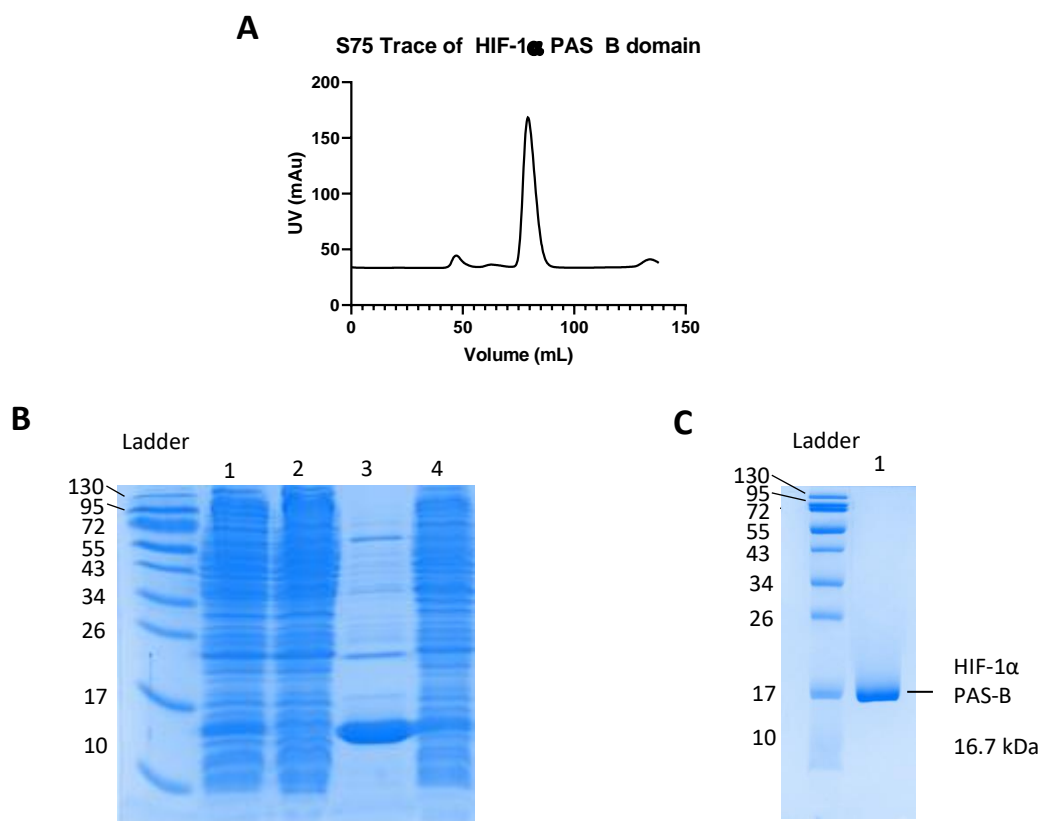


Figure 20. Analysis of expression of the HIF-1 α PAS-B domain. A) S75 trace of the HIF-1 α PAS-B domain. B) SDS-PAGE analysis of the expression of HIF-1 α PAS-B domain, 1: Insoluble fraction, 2: HIS-trap flow through, 3: HIS-trap elution, 4: Soluble fraction. C) SDS-PAGE analysis of the isolated HIF-1 α PAS-B domain.

The protein was also analysed via mass spectrometry, which showed two peaks upon deconvolution, one at a mass of 16577.0 Da and the other at 16756.1 Da (Figure 21). Whilst neither of these peaks correspond to the expected mass of the construct (16708 Da) they can both be explained by common modifications seen in recombinant proteins. The first peak at 16577.0 Da is equal to the mass of the construct minus methionine. This is almost certainly a result of N-terminal methionine excision, a co-translational process carried out by a class of enzymes called methionine aminopeptidases which is common in *E. coli*.¹⁹⁵ The second peak at 16756.1 Da shows a gain in mass of 178 Da, this is the expected mass gain upon HIS-tag

Chapter 2

gluconoylation a process thought to occur via the reaction of recombinant proteins with 6-phosphoglucono-1,5-lactone. This yields a gain of 258 Da which upon removal of a phosphate group by a phosphatase explains the additional 178 Da mass.¹⁹⁶

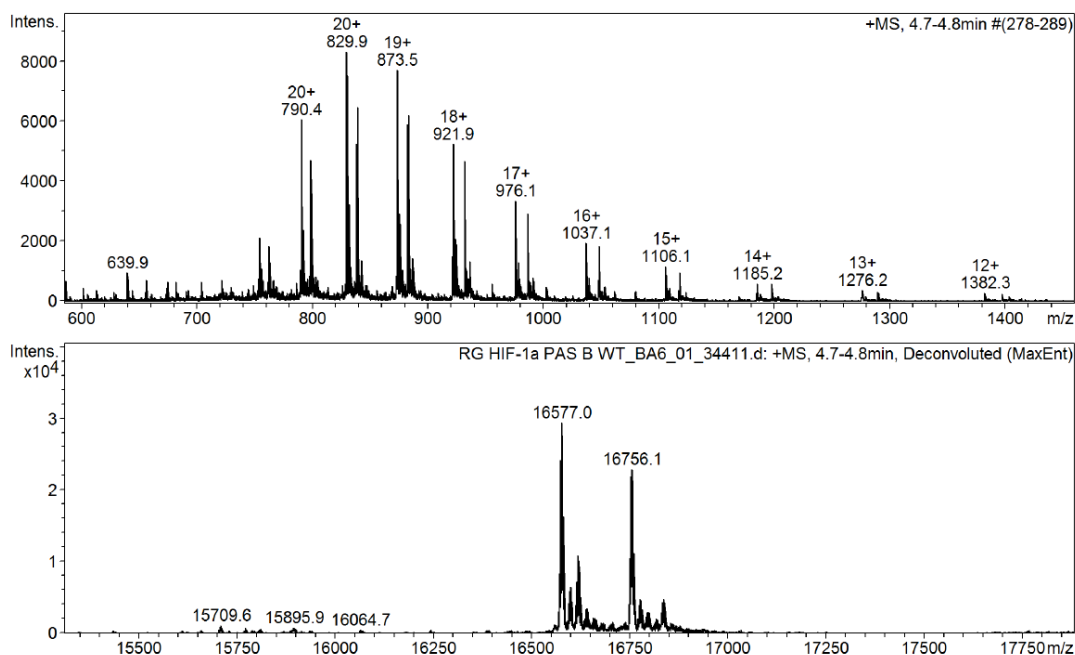


Figure 21. Mass spectrometry analysis of obtained HIF-1 α PAS-B domain. The mass spectrum shows the methionine cleaved major product at 16577.0 Da, and the gluconoylated protein at 16756.1.

2.3 Thermal Shift

In order to assess the effect of CLLFVY binding on protein stability, a thermal shift protocol was established. Thermal shift is a relatively simple biophysical method for the characterisation of protein-ligand interactions, which relies on the use of a fluorophore with a fluorescence intensity that is dependent upon its environment. By having a lower intensity in polar environments and higher in non-polar environments, the fluorophore is sensitive to the fold of the protein.¹⁹⁷ This fluorophore is mixed with the protein and ligand of interest in a multi-well plate, and the sample is then heated in increments with the fluorescence intensity measured. The heating causes the protein chain to start to unfold which exposes the hydrophobic residues which usually reside in the core of the protein, leading to an increase in fluorescence until the protein has completely denatured. The temperature at which the protein is halfway to being fully denatured is termed the melting temperature (T_m) and is used to quantify a proteins thermal stability. With continued heating the protein aggregates, shielding the dye from the hydrophobic regions, which coupled

with the increased dissociation rates at higher temperatures leads to a decrease in fluorescence intensity.¹⁹⁸ This process is outlined in Figure 22.

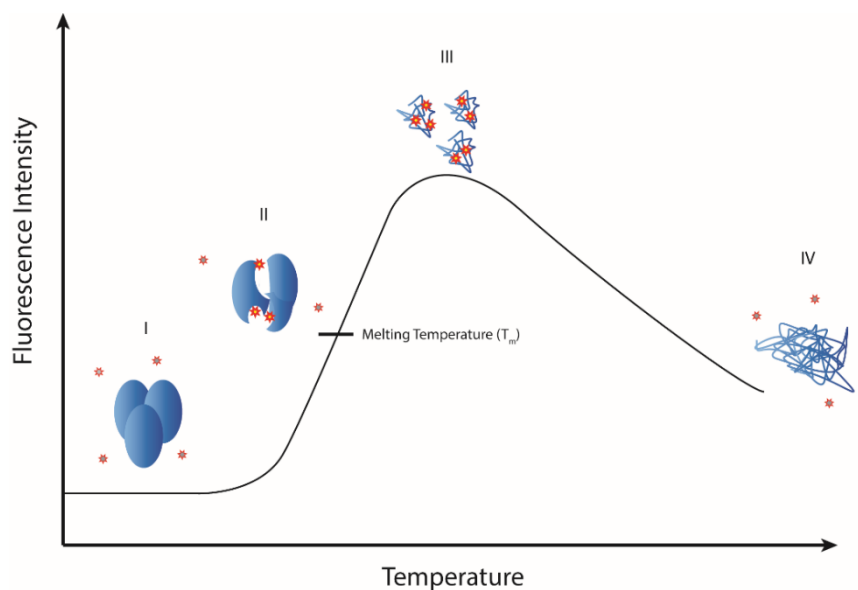


Figure 22. Outline of the principle of thermal shift. I) Protein is fully folded and so fluorescence intensity is low. II) With increasing temperature the protein unfolds, leading to the exposure of hydrophobic residues and an increase in fluorescence. III) The protein is fully unfolded leading to maximum fluorescence. IV) Protein aggregates leading to reduced access to hydrophobic environments and decreased fluorescence.

Due to the limited aqueous solubility of CLLFVY, DMSO is required in any assay buffer in order to reach suitable concentrations for *in vitro* analysis. To assess the proteins tolerance for different concentrations of DMSO, a thermal shift assay was conducted with the HIF-1 α PAS-B domain in buffers containing a range of DMSO. The melting curves for the different buffers are shown in Figure 23A. This data showed that the protein in buffer without DMSO had a melting temperature of 52.25 °C and that it was stable up to 10 % DMSO, with the melting temperature decreasing by 0.5 °C at 12.5 and 15 %, demonstrating a small decrease in stability. Therefore, all further assays were carried out in 10 % DMSO so that the maximum concentration of peptide could be used without compromising protein stability.

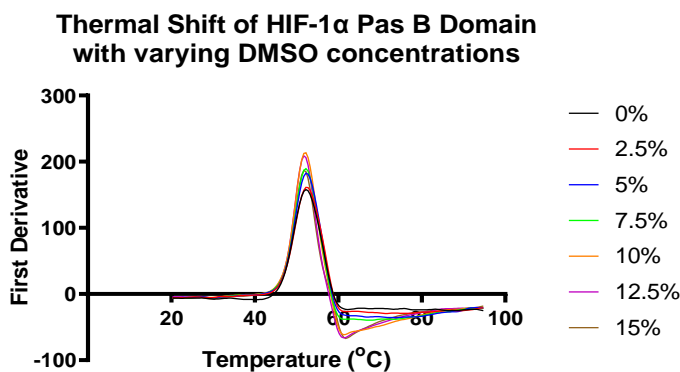


Figure 23. Thermal shift analysis of the HIF-1 α PAS-B domain in the presence of DMSO. First derivative analysis of a thermal shift assay assessing the effect of varying DMSO concentrations on the HIF-1 α PAS-B domain.

CLLFVY was found to be soluble in buffer with 10 % DMSO up to a concentration of 500 μ M. In order to assess the effect of the compound on the protein, a thermal shift assay was carried out with CLLFVY at concentrations of 500 and 250 μ M. The first derivatives of the resulting melt curves are shown in Figure 24A and the calculated melting temperatures in Figure 24B.

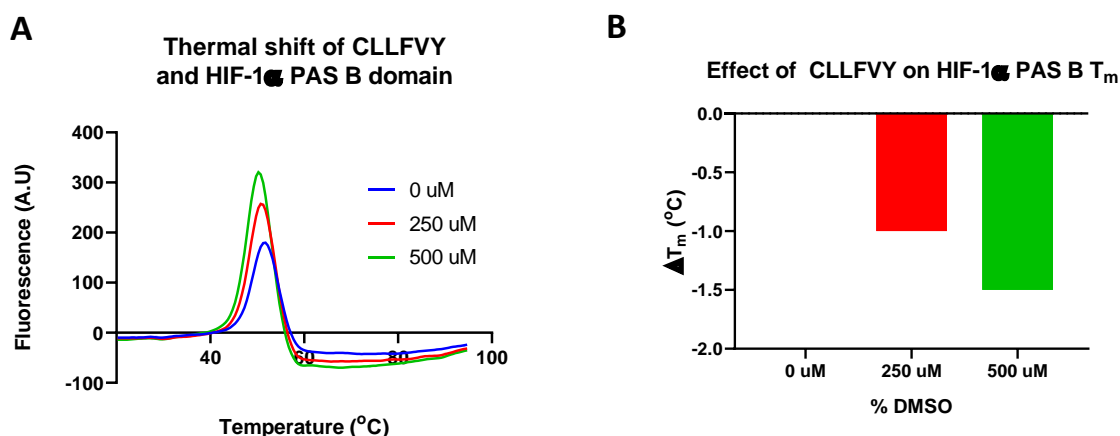


Figure 24. Thermal shift analysis of the HIF-1 α PAS-B domain in the presence of CLLFVY. First derivative analysis of a thermal shift assay assessing the effect of varying concentrations of cyclo-CLLFVY on the HIF-1 α PAS-B domain. B) Plot of the shift in melting temperature obtained under different cyclo-CLLFVY concentrations.

Interestingly CLLFVY at concentrations of 250 and 500 μ M were found to decrease the melting temperature of the HIF-1 α PAS-B domain by 1 and 1.5 $^{\circ}$ C respectively. Typically, a ligand binding event will stabilise the protein of interest and so a rise in melting temperature would be expected. There is evidence however that many ligands are capable of causing a decrease in protein melting temperature. One of the most common rationales for this effect is that the ligand preferentially binds to the unfolded protein, which leads to the unfolded protein becoming more energetically

favourable than the native fold, which is observed as a decreased melting temperature.¹⁹⁹ An alternative cause for a decreased melting temperature upon ligand binding is that the protein has multiple conformations, with the ligand preferentially binding to the conformer that has a lower melting temperature. PAS domains are well known to possess conformational plasticity which is often a result of 'open' and 'closed' conformations that are dependent on an external stimulus. This has been observed in the structurally similar aryl hydrocarbon receptor (AhR) and light oxygen-voltage-sensing (LOV) PAS-B domains, which adopt different conformations dependent upon external stimuli enabling their function as signalling domains.^{200–202} Therefore, it is possible that the HIF domain also possess multiple conformations to which CLLFVY has different affinities for, resulting in the observed thermal shift results. However, the small observed shifts in melting temperature and the multiple ways in which a decreased melting temperature can be interpreted means that an alternative biophysical technique is needed to more accurately determine the binding of CLLFVY to the HIF-1 α PAS-B domain.

2.4 The Development of an MST Assay

2.4.1 Introduction to Microscale Thermophoresis

MicroScale Thermophoresis (MST) is a biophysical technique for the quantification of bimolecular interactions, which utilises the movement of molecules across a thermal gradient, termed thermophoresis. By irradiating a sample with an IR laser a thermal gradient is produced, with molecules diffusing away from the induced hot spot (thermophoresis). In the case of fluorescently labelled molecules, the thermophoresis can be monitored as the change in fluorescence intensity at the point of irradiation. The fluorescent nature of MST enables highly sensitive measurements to be carried out with very low sample requirement. This was a major factor in deciding to pursue MST as the primary biophysical assay for investigating the binding of CLLFVY, as it only requires a small volume (3-4 μ L per capillary) of dilute protein (20 nM). This is an advantage when working with HIF proteins due to their aforementioned poor recombinant expression yields, with MST enabling multiple compounds to be screened from a single expression.

The rate at which molecules diffuse across the thermal gradient is dependent on a number of physical properties such as size, charge, and solvation, making MST sensitive to small changes to a proteins structure such as ligand binding. These changes in physical properties are accounted for by the Soret coefficient, which describes the balance between thermophoresis away from the hot spot and the diffusion of the molecules due to concentration gradients as shown in Equation 1 below.

$$S_T = \frac{D_T}{D}$$

Equation 1. Determination of the Soret coefficient. Where S_t is the Soret coefficient, D_t is the thermal diffusion coefficient and D is the diffusion coefficient.

The extent of thermophoresis of a sample can then be determined via the relation of the Soret coefficient to the concentration of molecules at hot and cold states as shown in Equation 2 below.

$$\frac{C_{hot}}{C_{cold}} = \exp(S_T \Delta T)$$

Equation 2. The extent of thermophoresis is calculated as the ratio between the concentration of molecules at higher temperatures and the concentration at low temperatures. Where C_{hot} is the concentration at high temperatures, C_{cold} is the concentration at low temperatures, S_t is the Soret coefficient and T is the temperature.

The concentration of the molecule of interest is measured by covalently attaching a fluorophore, the fluorescence intensity of which is then directly proportional to the concentration. As changes in fluorescence are usually quite small, they are often reported as F_{norm} values on a per mille scale and are calculated as shown in Equation 3.

$$\frac{F_{hot}}{F_{cold}} * 1000 = F_{norm}$$

Equation 3. Calculation of F_{norm} values. Where F_{hot} is the fluorescence intensity at high temperatures and F_{cold} is the fluorescence intensity at low temperatures.

By quantifying the change in normalised fluorescence, the binding constant for a ligand can be derived. This is achieved by expressing each condition as a superposition of the fully bound and unbound states as detailed in Equation 4.

$$F_{norm} = (1 - x)F_{norm}(unbound) + xF_{norm}(bound)$$

Equation 4. Expressing F_{norm} as a superposition of the fully unbound and fully bound states allows for the determination of the fraction bound and therefore the K_d . Where $F_{norm}(unbound)$ is the F_{norm} value at a fully unbound state, $F_{norm}(bound)$ is the F_{norm} value at a fully bound state and X is the fraction bound.

Re-arranging for x allows the determination of the fraction bound and thus the dissociation constant.^{203,204}

In order to determine binding constants, multiple capillaries containing differing concentrations of ligand are analysed. There are five phases to the analysis of a single capillary. First the initial fluorescence of the sample is measured, then a thermal gradient of between 2-6 °C is induced in the sample with an IR laser, which results in a rapid change in the fluorescence termed the T-jump. This is the result of direct effects of the temperature change on the fluorophore and is sensitive to the direct environment of the dye. As a result, changes in T-jump can be caused by ligands binding within close proximity to the dye or inducing a change in the protein which leads to a change in the dye's environment, and can be used for determining binding constants. The next phase is where the thermophoresis occurs, with a steady-state usually being reached within 20-30 seconds. The IR irradiation is then stopped and the molecules back-diffuse due to concentration gradients. These phases and a typical MST experiment are demonstrated in Figure 25.

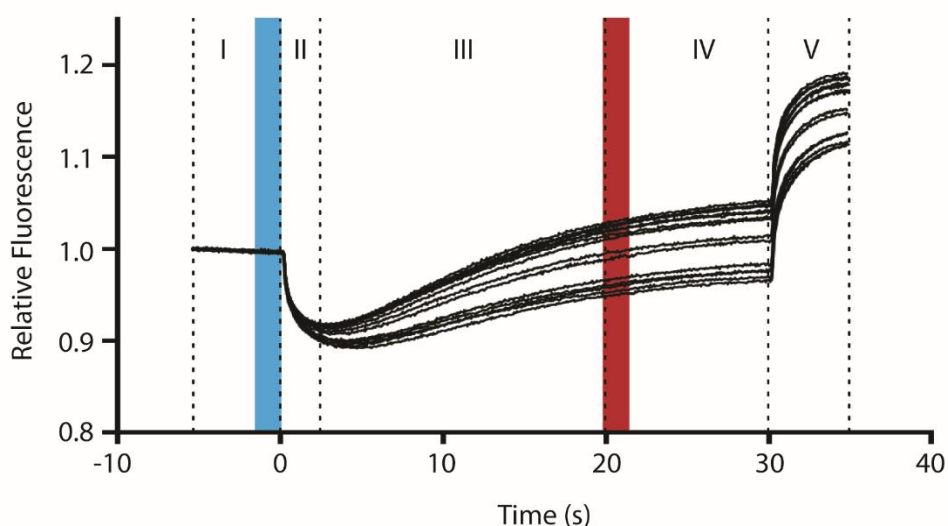
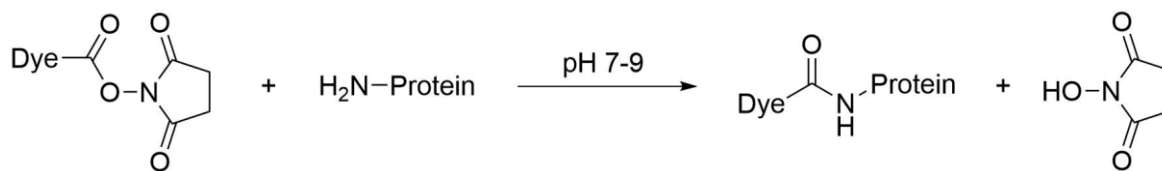


Figure 25 . Fluorescence over time of a typical MST experiment. I) Initial fluorescence. II) T-jump. III) Thermophoresis. IV) Steady-state. V) Back-diffusion. The values for F_{hot} are taken during the thermophoresis region of the MST trace (shown in red), with the F_{cold} value taken from the period of initial fluorescence (shown in blue).

2.4.2 Analysis of CLLFVY by MST

Before the affinity of CLLFVY could be determined, the HIF-1 α PAS-B domain first had to be fluorescently labelled. This was achieved using a cyanine dye containing an N-hydroxysuccinimide (NHS) moiety. The NHS group is reactive to amine groups present on lysine side chains and reacts via substitution at the activated ester position as depicted in Scheme 9.



Scheme 9. Labelling of protein for use in MST with an NHS-Cy5 dye.

Once the labelled HIF-1 α PAS-B domain was obtained it was incubated with a 16-point serial dilution of CLLFVY ranging from 500 μ M to 15 nM. Each concentration was then loaded into a capillary and analysed via MST and the normalised F_{norm} values plotted using GraphPad Prism. The data was plotted as percentage bound and the resulting curve for CLLFVY is shown in Figure 26.

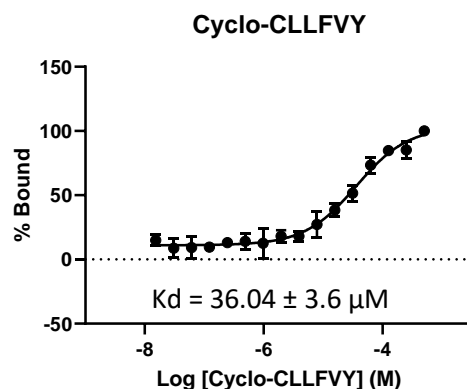


Figure 26. Cyclo-CLLFVY binding to HIF-1 α PAS-B domain as determined by MST. Error bars represent standard deviation $n=3$. Binding constants were determined by curve fitting in GraphPad Prism 8.

CLLFVY was found to bind to the HIF-1 α PAS-B domain with a K_d of $36.04 \pm 3.6 \mu\text{M}$. This is significantly weaker than the previously published values determined by Miranda *et al.* which were $1.3 \mu\text{M}$ by ELISA and 124 nM by ITC.¹⁴⁶ However, it is worth noting that both of these assays were carried out using an analogue of CLLFVY that was tagged with the cell penetrating peptide Tat (CGRKKRRQRRPPQ). It is therefore possible that the discrepancy in K_d values is due to interference by the Tat-tag. As this project aims to develop the cyclic peptide CLLFVY, the Tat-tag will not be investigated further and instead the K_d of CLLFVY alone will be optimised.

A number of control experiments were also carried out to ensure the specificity of CLLFVY binding. Three other proteins, HIF-2 α PAS-B, HIF-1 β PAS-B, and BSA were labelled with the same NHS-Cy5 dye and the binding of CLLFVY to each was assessed. The resulting MSTs are shown in Figure 27 and show that CLLFVY binds selectively to HIF-1 α PAS-B. Some binding affinity is seen for the HIF-2 α PAS-B domain, which is unsurprising given the high sequence homology of the two domains. None the less with an approximate K_d of $>300 \mu\text{M}$, CLLFVY still shows 10-fold higher affinity for the 1 α PAS-B domain which is in keeping with the observed selectivity in previous

assays.¹⁴⁶ With a functioning *in vitro* biophysical assay in hand, we next set out to probe the SAR and the importance of the amino acids in CLLFVY.

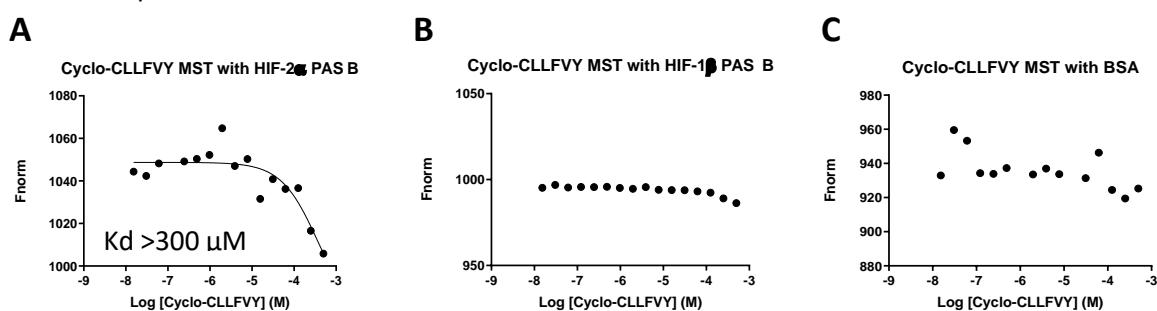


Figure 27. MST analysis of CLLFVY specificity. A) MST trace of cyclo-CLLFVY with the HIF-2 α PAS-B domain. B) MST trace of cyclo-CLLFVY with the HIF-1 β PAS-B domain. C) MST trace of cyclo-CLLFVY with BSA. Binding constants were determined by curve fitting in GraphPad Prism 8.

2.5 Alanine Scanning of CLLFVY

Alanine scanning is a powerful tool for probing the contribution of individual amino acids to a protein or peptides function. The replacement of individual amino acids with alanine removes the side chain functionality; any loss of binding affinity as a result of the substitution infers that the side chain was engaging in an interaction that contributes to the affinity of that sequence.²⁰⁵

Alanine is used in place of glycine as it still contains a chiral centre and so helps to maintain the secondary structure of the peptide. In order to determine which amino acids of CLLFVY were crucial for binding, six alanine scan derivatives were synthesised, sequentially replacing each amino acid with alanine. The binding of each alanine analogue was then assessed via MST against the HIF-1 α PAS-B domain with the results shown in Figure 28. Partial binding curves were obtained for five out of six of the derivatives, with the cysteine analogue not showing any binding activity. This suggests that the cysteine is the most critical residue of CLLFVY as all binding activity was lost with its replacement. The contribution of the other residues could not be determined accurately due to the limited solubility of the peptides preventing the use of higher concentrations. However, a reduction in affinity was seen for all substitutions suggesting all residues contribute to HIF binding. In order to probe the possible interaction of the cysteine side chain with the protein, it was decided that a series of derivatives containing similar chemical functionality would be synthesised.

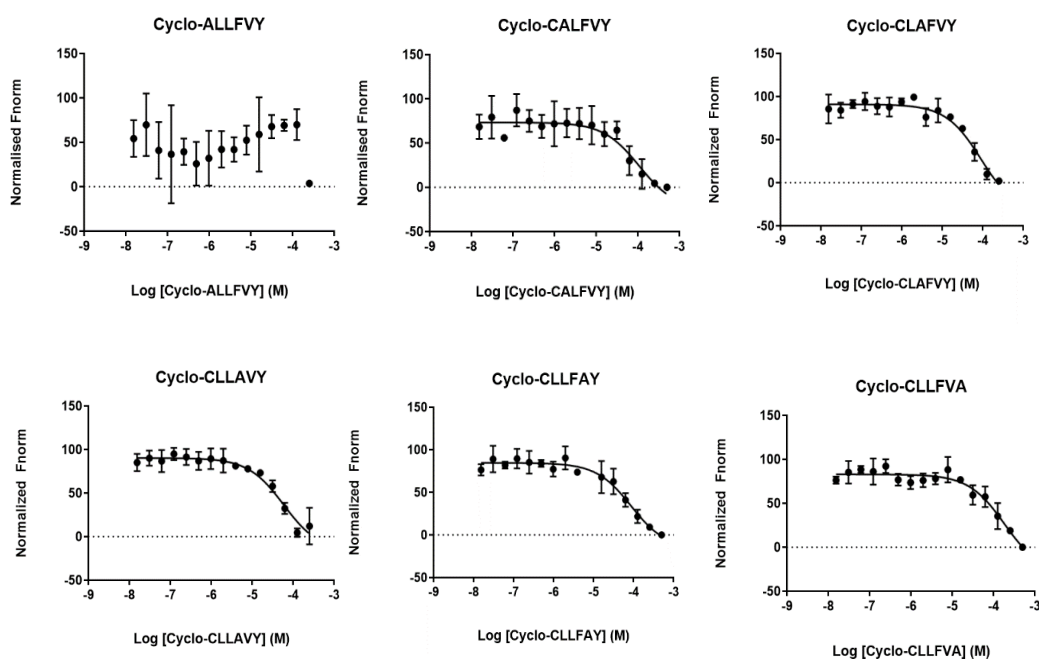


Figure 28. MST binding curves of the six CLLFVY alanine scan analogues against the HIF-1 α PAS-B domain. Error bars represent standard deviation ($n=3$). Binding constants were determined by curve fitting in GraphPad Prism 8.

2.6 Investigating Cysteine Analogues

To probe how the cysteine residue was interacting with the protein, seven analogues of the peptide were synthesised that incorporated similar functionalities. These included a homologated analogue as well as a peptide containing the opposite stereochemistry, D-cysteine. Serine and threonine were also included to probe whether the interaction is specific to sulfur, or if a different electronegative heteroatom would behave similarly. In addition, methionine and thioproline (Tp), which contain sulfur atoms as thioethers rather than free thiols, were included to probe the importance of the free thiol group. The structures of the chosen derivatives are shown in Figure 29. Once obtained, the compounds affinity for the HIF-1 α PAS-B domain was determined by MST. The affinity of each was determined and plotted in Figure 30.

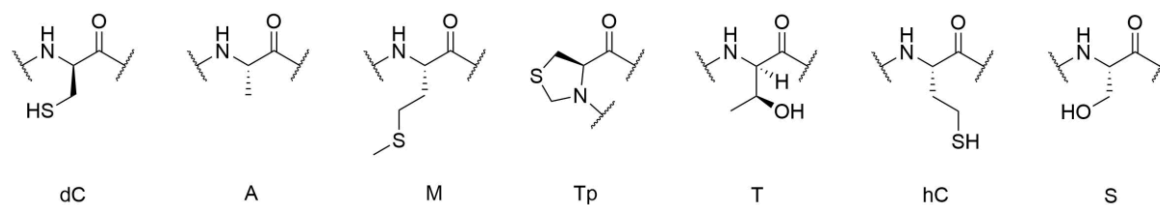


Figure 29. Structures of the seven amino acids that were substituted into the cysteine position of CLLFVY.

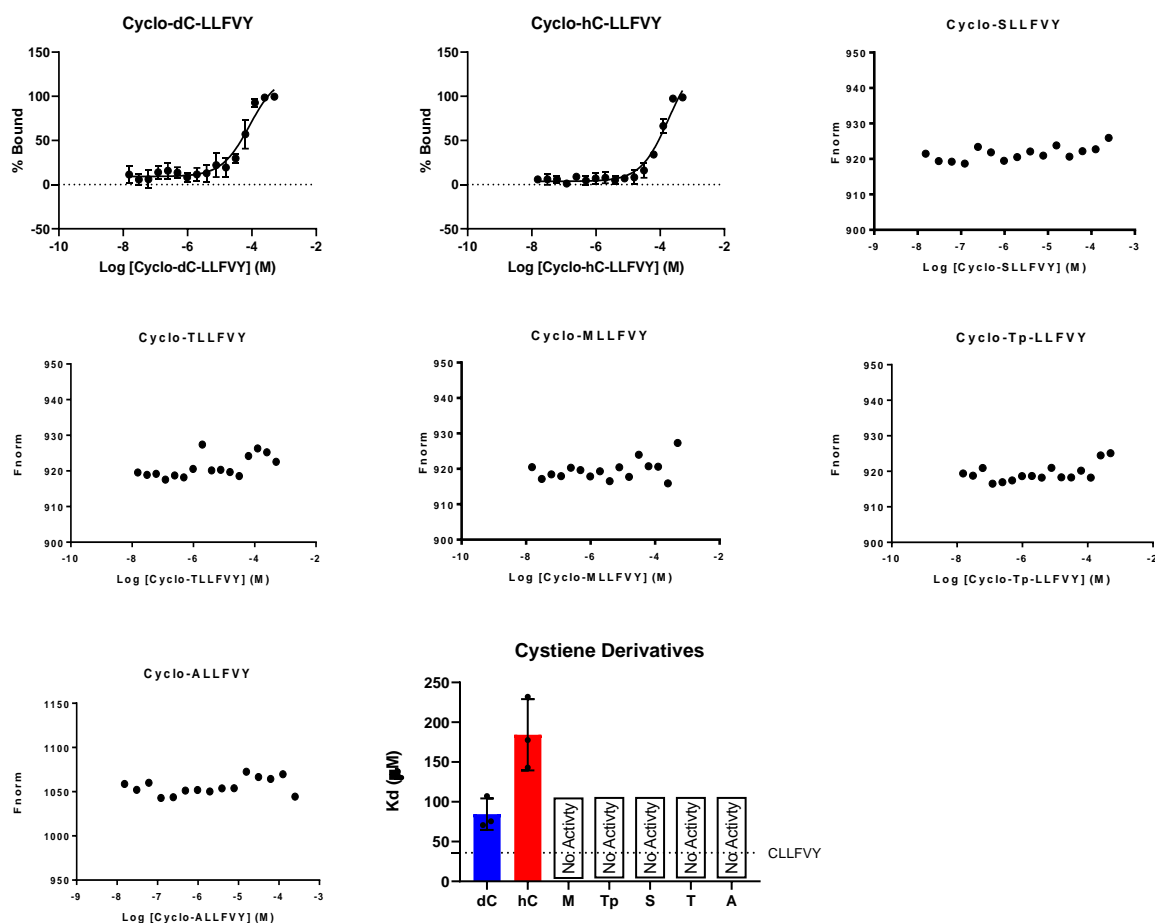


Figure 30. MST binding curves of cysteine derivatives and bar chart comparison of the obtained K_d values. Where present error bars represent standard deviation $n=3$. Binding constants were determined by curve fitting in GraphPad Prism 8.

The MSTs revealed that none of the analogues served as a replacement for cysteine, with the only derivatives retaining activity being homo-cysteine and D-cysteine. The error in these measurements was high, resulting from the curves not reaching saturation, but it is still clear that neither amino acid binds with greater affinity than cysteine. These results suggest that a free thiol group is necessary for interaction with HIF. The necessity of the free thiol and importance of the cysteine in the alanine scanning is suggestive of a covalent interaction being present via a disulphide linkage. This was deemed unlikely due to the presence of a 1000-fold excess of TCEP in the assay buffer when compared to the protein, which should reductively cleave any disulphides present. However, in order to be certain about this potential binding mechanism, we decided to probe the potential formation of a disulphide experimentally.

2.7 Investigating the Potential of Disulphide Mediated Binding

2.7.1 Dithiothreitol (DTT)

Our first approach in investigating potential disulphide bond formation was to evaluate the effect of the reducing agent present in the buffer. The buffers used so far have all utilised the reducing agent tris(2-Carboxyethyl) phosphine hydrochloride (TCEP), and whilst TCEP is generally considered more stable and more potent than the next most popular alternative, DTT,²⁰⁶ it was decided to investigate a reducing agent acting via a different mechanism.

A thermal shift experiment was first run to determine the tolerance of the HIF-1 α PAS-B domain for different concentrations of DTT. The results, shown below, indicate that the protein is stable in the presence of DTT up to a concentration of 2 mM, and therefore this concentration was used in the following experiments. The binding of CLLFVY to the HIF-1 α PAS-B domain in DTT buffer was analysed by MST and the results plotted Figure 31C.

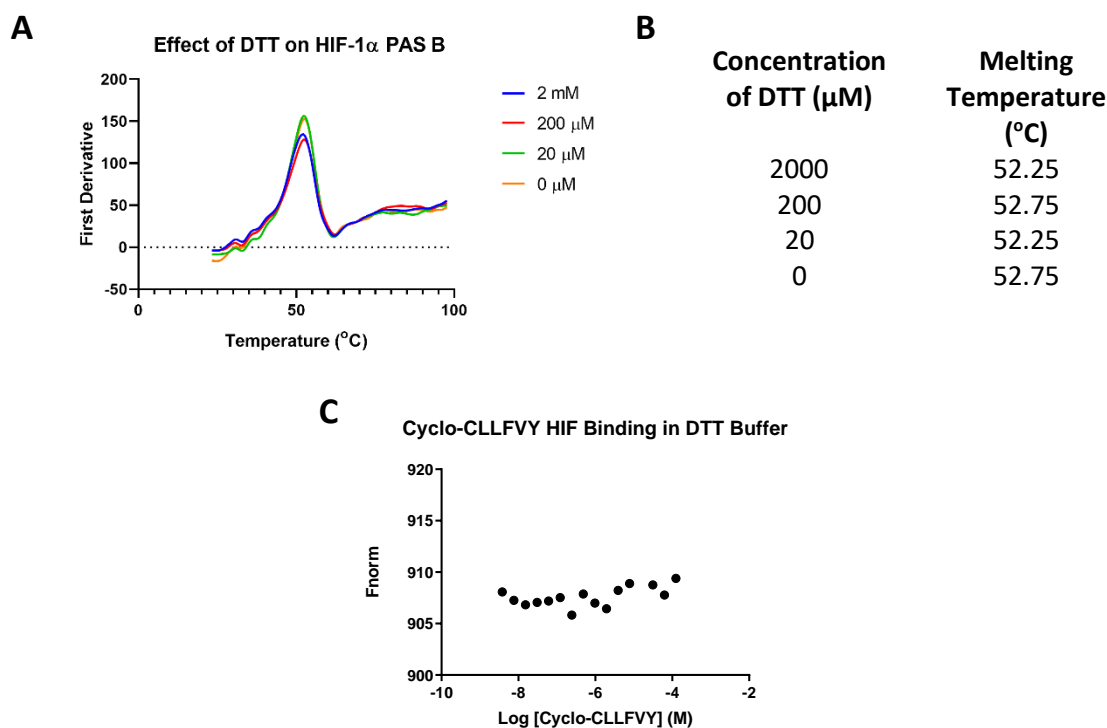


Figure 31. Assessing the impact of DTT on the HIF-1 α PAS-B domain and HIF-1 α PAS-B-CLLFVY binding. A) Thermal shift of varying DTT concentrations and HIF-1 α PAS-B domain. B) Table of melting temperatures obtained at varying DTT concentrations. C) MST of CLLFVY-HIF binding in DTT buffer (2 mM).

DTT was found to abolish all CLLFVY binding activity, suggesting that the peptides mode of action could be disulphide mediated. A control experiment of a serial dilution of DTT against the HIF-1 α

PAS-B was then run to investigate the effect of DTT alone, and the trace is shown in Figure 32. Surprisingly the serial dilution yielded a binding curve with a K_d of $68.4 \pm 6.9 \mu\text{M}$. Whilst this interaction is two times weaker than that of CLLFVY, DTT is present in the buffer at 30 times this K_d and so HIF will be fully bound to DTT in this buffer. Given that DTT binds the HIF-1 α PAS-B domain, it is not clear whether it is preventing CLLFVY binding by acting as reducing agent or as a competitive binder of the HIF protein. DTT binding to proteins is not unprecedented and has been previously reported to occur with a protein called pigpen, with the authors describing that it inhibited protein function in a non-disulphide specific manner.²⁰⁷ It is therefore possible that DTT could be binding to HIF in a similar manner, and whilst this is something that will be investigated in further work it is outside of the scope of this project. Given that DTT was found to bind the HIF-1 α PAS-B domain, it was reasoned that the HIF-1 α PAS-B domain may have an affinity for any free thiol, and therefore other thiol containing reducing agents were not investigated. Instead, the potential disulphide between CLLFVY and the protein was investigated using an alternative approach.

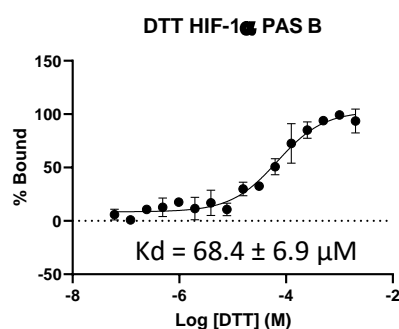


Figure 32. MST binding curve of DTT and the HIF-1 α PAS-B domain. Error bars represent standard deviation (n=3). Binding constants were determined by curve fitting in GraphPad Prism 8.

2.7.2 Cysteine Mutants

The alternative approach chosen was site-directed mutagenesis (SDM) of the HIF-1 α PAS-B domain. There are three cysteines in the HIF-1 α PAS-B domain that could potentially be involved in a disulphide interaction with the peptide: C255, C334 and C337. The locations of each in the PAS-B heterodimer crystal structure are shown in Figure 33. In this structure, only C334 is solvent exposed and directly involved with the PPI interface, with the other two buried internally. A reactivity study by Cardoso *et al.*, screening a library of 177 electrophilic compounds, revealed C255 to be the most reactive, with four compounds found to react preferentially at this position. The best of which, named 'Compound 5', was shown to reduce the affinity of protein dimerisation 10-fold.¹⁶ This is somewhat surprising given the position of C255 within the internal cavity of the protein, but does demonstrate that disulphide-linked covalent inhibitors of HIF can have

inhibitory activity. In order to assess whether CLLFVY was forming a disulphide with one of these cysteines three constructs were made, sequentially mutating each cysteine to an alanine. If the binding of CLLFVY was mediated by a disulphide to one of these cysteines then a significant loss of activity would be expected upon replacement of that cysteine with alanine. The HIF-1 α PAS-B pET-28a vector was used as a template, and site-directed mutagenesis (SDM) was carried out utilising primers containing the mismatched alanine codon GCT in place of the TGT cysteine codons.

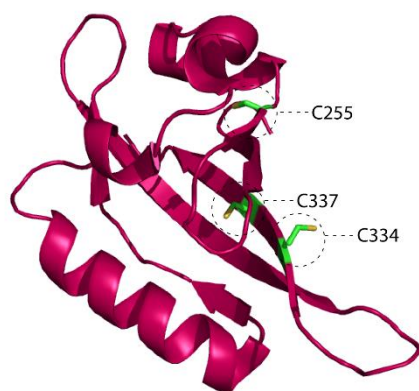


Figure 33. Location of the three cysteine residues to be mutated in the HIF-1 α PAS-B domain (PDB: 4H6J).¹¹⁸

The mutant proteins were expressed in BL21 (DE3) Rosetta cells using the same protocol as the wild type PAS-B, to yield the proteins C255A, C344A and C337A. The yield of all three proteins was poor which is likely attributable to their poor stability. This was further evidenced by the reduction in melting temperatures of all three mutant proteins when compared to the wild type protein (Figure 34). The C337A mutation yielded the greatest destabilisation, decreasing the proteins melting point by 7 °C when compared to the wild type protein. The C255A mutant resulted in a 4.5 °C drop in melting temperature and the C334A gave the smallest destabilisation of just 2 °C.

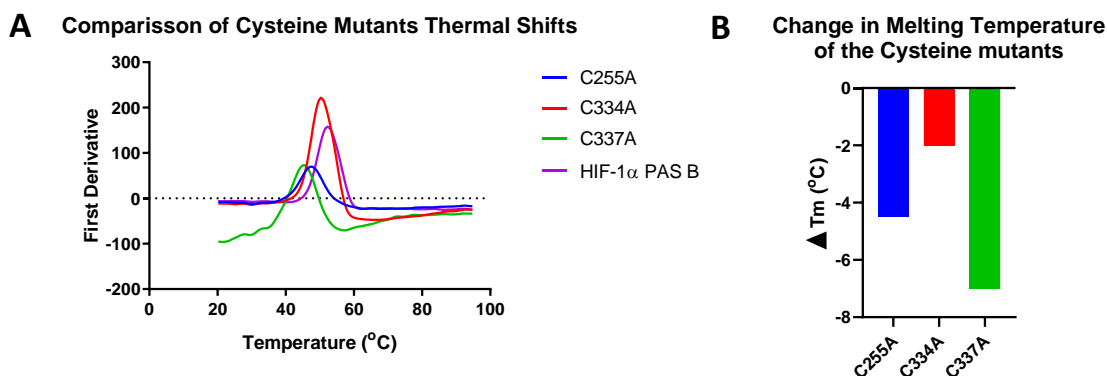


Figure 34. Thermal shift analysis of the HIF-1 α PAS-B cysteine mutants. A) Thermal shift of the cysteine mutants C255A, C334A and C337A. The wild type HIF is also shown for comparison. B) Plot of the change in melting temperature of the cysteine mutants when compared to the wild type protein.

Once obtained, the three cysteine mutants were labelled for MST using the same dye and procedure as for the wild type. MST analysis of CLLFVY binding to the three mutants was then carried out, with the results shown in Figure 35. The resulting binding curves show that CLLFVY binds to all three cysteine mutants. Whilst some loss of affinity is seen for C334A, this is most likely due to structural or stability changes caused by the mutation. If the loss of activity were the result of the loss of a disulphide interaction, we would expect to see the same loss of activity that was observed upon substituting cysteine in the CLLFVY alanine scan. Given that we do not see a return to the flat line that was observed for ALLFVY, it can be concluded that CLLFVY binding is not mediated by a disulphide interaction.

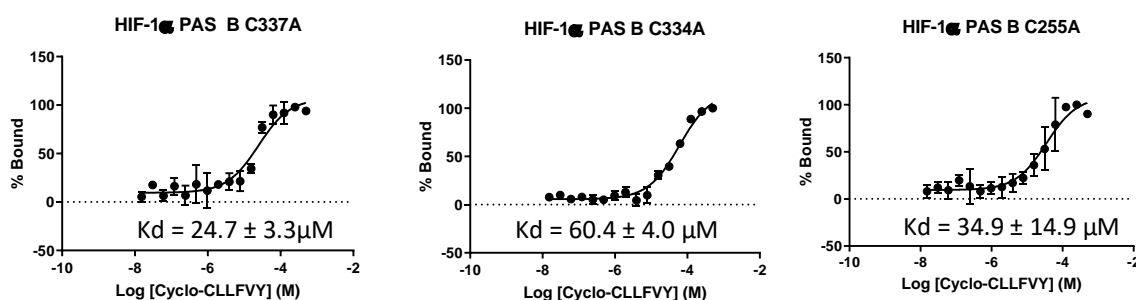


Figure 35. MST traces of CLLFVY binding to the mutant proteins C255A, C334A and C337A. Error bars represent standard deviation ($n=3$). Binding constants were determined by curve fitting in GraphPad Prism 8.

If CLLFVY binding is not mediated via a disulphide, then the thiol group must be contributing to binding via a different modality. Sulfur containing amino acids have been hypothesised to form interactions with aromatic rings since the 1980's, when it was noticed that these interactions occur more frequently than random chance would dictate.²⁰⁸ The S- π stacking interactions were also shown to have a preferred geometry, with planar facing orientations more favourable than edge-facing ones.²⁰⁹ Later studies also confirmed the formation of sulphur-aromatic interactions, both experimentally and computationally, this time measuring the energy of formation to be 1-3 kcal/mole.^{210,211} This type of interaction is now known to contribute to many biological interactions including protein-ligand interactions and protein tertiary structure.^{212,213}

The observed SAR from the cysteine derivatives is suggestive of this type of interaction as it is unique to sulfur containing amino acids, explaining why the oxygen containing serine and threonine gave poor binding affinities. The sensitivity of this type of interaction to geometry also corroborates the results, as changing the chain length or stereochemistry at this position also lead to a loss of activity. As the cysteine was found to be the most potent residue at this position it was left in place, and the SAR at a different position was investigated.

2.8 Investigating Phenylalanine Analogues

Given that the alanine scanning could not rank the importance of the remaining amino acids, it was decided that a position should be chosen empirically for further derivatisation. It was hoped that by creating further analogues the affinity of the compound for HIF could be increased to the point where another alanine scan could be carried out, with full curves obtained before hitting the solubility limit. The phenylalanine position was chosen for this, due to the greater diversity of available analogues, thus allowing for a more extensive probing of the SAR at this position. A set of 16 unnatural analogues of phenylalanine were chosen, with the structures shown in Figure 36.

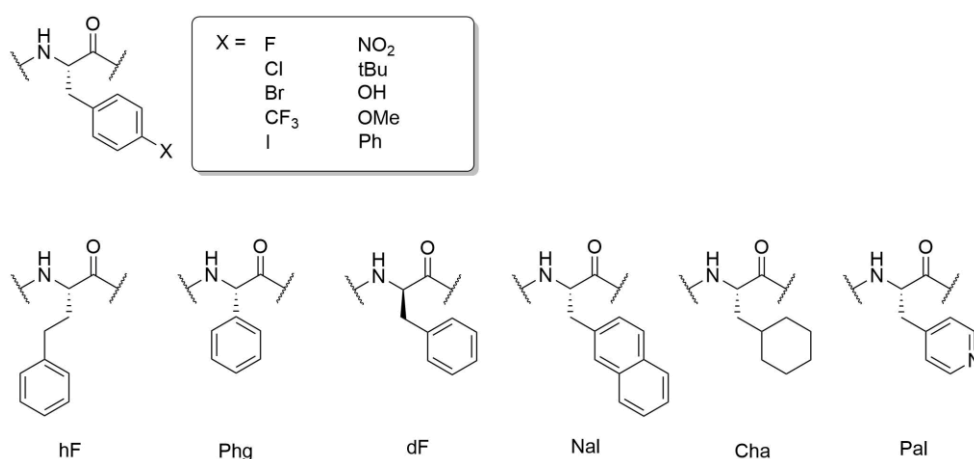


Figure 36. Structures of the amino acids substituted into the phenylalanine position of CLLFVY.

The linear sequences were synthesised from a Leu-Wang resin with standard Fmoc SPPS protocols to give the bulk resin LLCYV-NH₂. The resin was then split into syringes and the unnatural amino acids coupled and deprotected before the final linear sequences were cleaved from the resin using a TFA cleavage cocktail. Cyclisation was then attempted using EDC/ Oxyma Pure as before, however upon LCMS analysis some of the peptide sequences were found to have epimerised. This could be attributed to the changed position of cyclisation, with previous derivatives having been cyclised between the two leucine residues. In addition, the increased steric bulk of some of the unnatural analogues is likely to decrease the rate of cyclisation, and therefore increase the chance of racemisation.

Racemisation during the cyclisation step arises from preference of amide bonds to form an E isomer, and so adopting the Z isomer required for cyclisation is unfavourable. This can lead to linear sequences becoming very difficult to cyclise, especially when the peptide is an all L configuration sequence. Oxyma Pure was capable of suppressing racemisation of the coupling of the unnatural amino acids; however, this occurs on a much shorter timescale and so the potential for racemisation is smaller. The nature of the highly dilute cyclisation means that the reaction takes place over a longer time period and so a more potent suppressor of racemisation is

required. Two further classes of racemisation suppression reagents are phosphonium salts and aminium salts. Aminium derivatives were initially thought to react via a uronium intermediate. X-ray crystallography studies have since shown that these salts crystallise as aminium salts, and so it is now thought they react through a positively charged carbon. Aminium salts can therefore undergo nucleophilic attack from the N-terminus of the peptide, terminating coupling reactions. It is therefore important not to use these reagents in excess. Phosphonium based racemisation suppression reagents became popular with the discovery of the reagent PyBOP, which was notably devoid of the toxic by-products that were the major drawbacks of its predecessor BOP.²¹⁴ Phosphonium salts also offer an advantage over the above mentioned aminium salts as they do not interact with the amino group of the incoming amino acid and so there is no risk of the coupling step being blocked by this side reaction.

In order to try to prevent racemisation during cyclisation five different combinations of coupling reagents were investigated: HOBt/EDC, HOAt/EDC, Oxyma Pure/EDC, HATU, and PyBOP. Of these, two gave the desired peptide as a pure product (Figure 37), HATU and PyBOP. However, the phosphine by-product of the PyBOP coupling was found to co-elute with the peptide upon purification via RP-flash chromatography. Therefore, HATU was selected as the coupling reagent for all macrocyclisations from this point on.

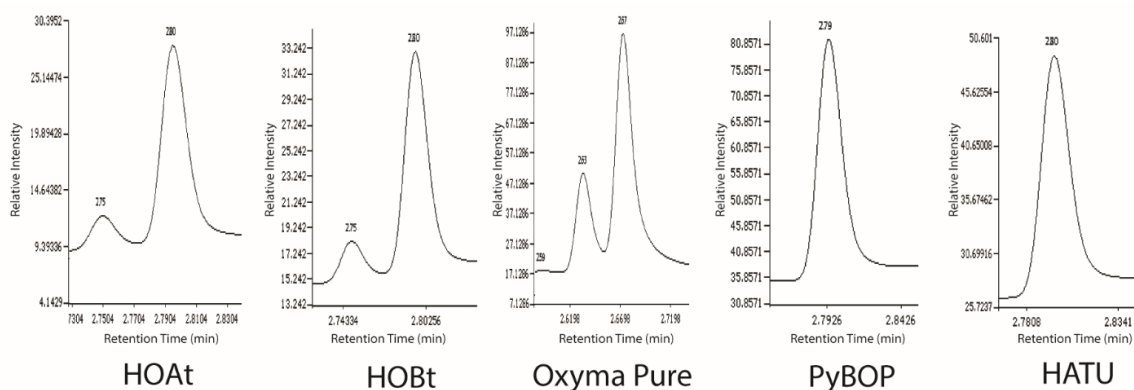


Figure 37. UV traces from LCMS analysis of cyclo-C(StBu)LL-Nal-VY cyclised using different coupling reagents. HOAt, HOBt and Oxyma Pure all showed two peaks of identical mass. PyBOP and HATU gave just one peak at the expected mass.

The ability of HATU to suppress racemisation has been linked to the position of the nitrogen in the pyridyl motif. The nitrogen atom in position seven has been hypothesised to participate in a neighbouring group effect, directing the C and N termini of the peptide together and thus speeding up the reaction.²¹⁵ This process is outlined in Figure 38.

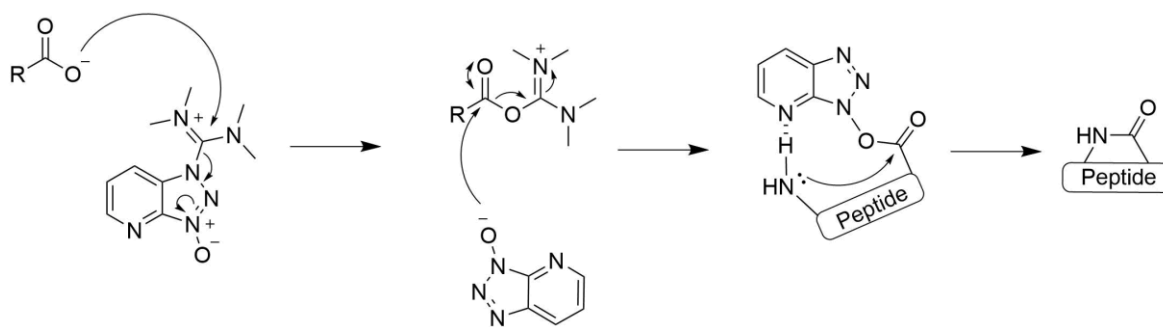


Figure 38. Proposed mechanism of the neighbouring group effect of HATU couplings.

Once the enantiopure derivatives had been successfully synthesised their affinity for the HIF-1 α PAS-B domain was assessed by MST. Single runs of 16 points were initially run in order to assess which analogues would be better binders than the parent peptide. Substitutions leading to a K_d lower than 10 μM were then analysed in triplicate in order to generate an accurate ranking of these derivatives. The obtained K_d values are summarised in Figure 39 and the resulting binding curves are shown in Figure 40.

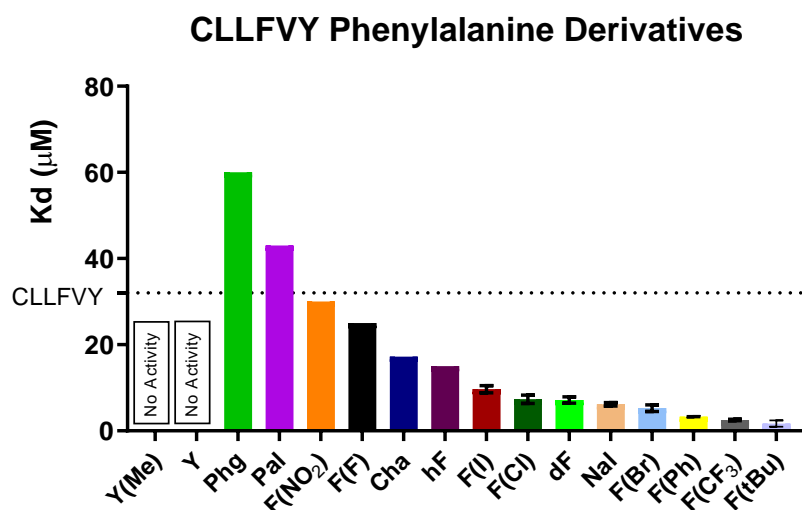


Figure 39. Comparison of the K_d values obtained for the phenylalanine derivatives. Where present error bars represent standard deviation ($n=3$).

The Y(Me), Y, and Pal derivatives gave the poorest affinities, with the Y and Y(Me) yielding incomplete binding curves. Both Y(Me) and Y contain electron-donating substituents, whilst the nitrogen atom in the 4-position of the pyridyl motif of Pal leads to a focussing of electron density onto the 4 position, and thus a more electron-poor aromatic system. The fact that these different electronic effects both led to a loss of activity is somewhat surprising. However, these substituents do all introduce lone pairs into the para position, and so the loss of activity could result from poor positioning of these lone pairs within the protein binding site.

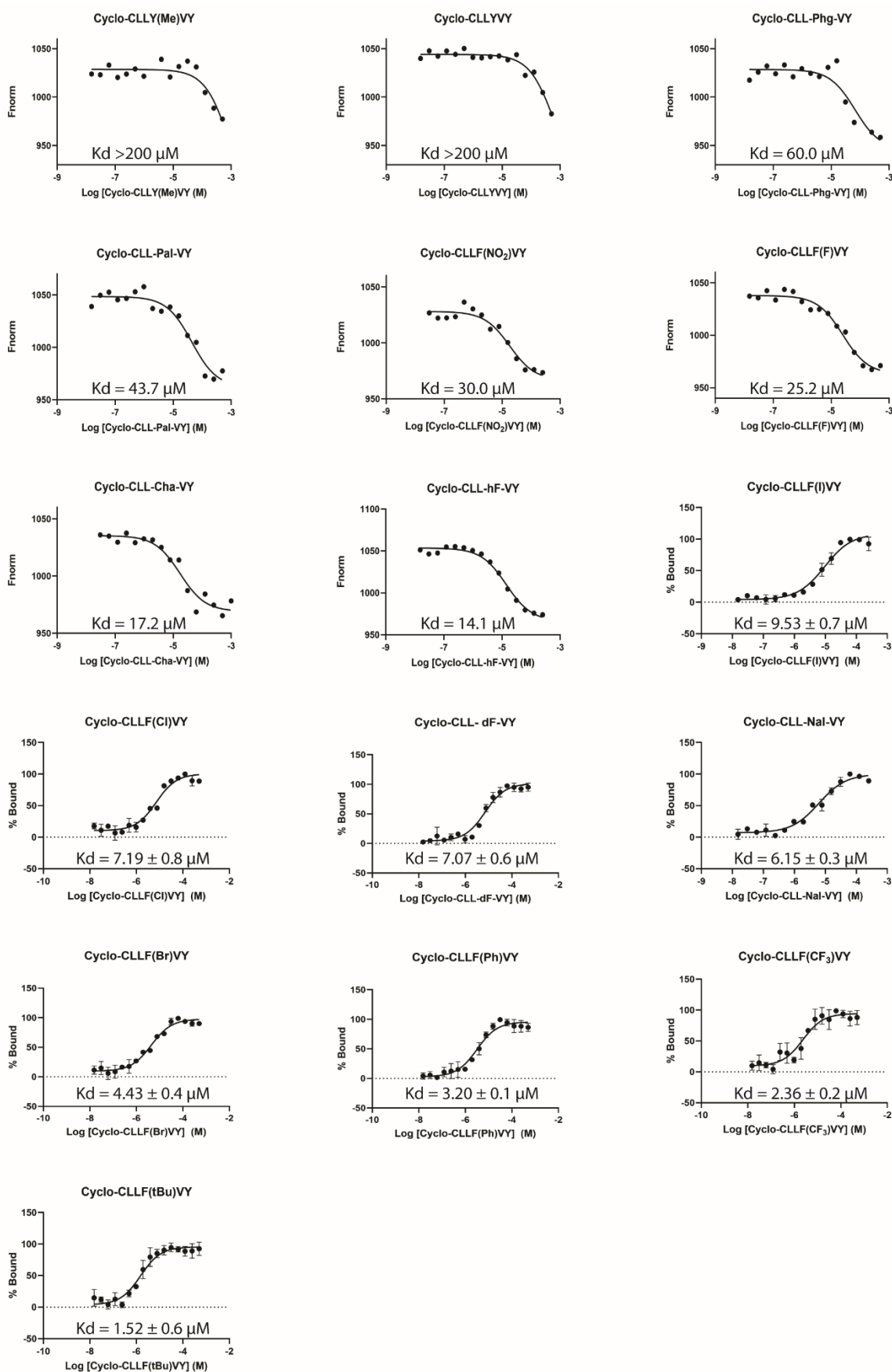


Figure 40. MST binding curves of the 16 phenylalanine derivatives of CLLFVY. Error bars represent standard deviation (n=3). Binding constants were determined by curve fitting in GraphPad Prism 8.

Chapter 2

The best affinities were observed for F(tBu), F(CF₃) and F(Ph), with F(tBu) yielding the greatest affinity of all the substitutions made with a K_d of $1.52 \pm 0.6 \mu\text{M}$, a 24-fold improvement over CLLFVY. All of these derivatives contain sterically bulky groups in the para position, showing that there is a high tolerance of steric bulk within the binding environment at this position. This is perhaps indicative of a hydrophobic pocket within the protein, which would also rationalise the poor tolerance of the more polar derivatives discussed above.

Another interesting result was the naphthyl derivative NaI, which gave a K_d of $6.15 \pm 0.3 \mu\text{M}$, a 6-fold improvement in affinity compared to the parent CLLFVY. This shows that extended aromatic systems are also well tolerated at this position. The positioning of the phenyl ring within its binding environment was also probed, and this revealed a trend upon changing the chain length between the peptide backbone and the phenyl ring. Elongating this chain via homo-phenylalanine yields a K_d approximately 2-fold lower than CLLFVY, whereas shortening the chain by a CH₂ unit with phenylglycine gave a K_d 2-fold higher than CLLFVY. This trend builds on the previously observed SAR that has revealed both an improved affinity for aromatic systems further from the backbone, as with NaI and F(Ph), and also a tolerance for bulky groups at this position, F(CF₃) and F(tBu). It is perhaps unsurprising then that extending the distance of the phenyl ring from the backbone leads to an increase in affinity and is likely a result of the more flexible linker enabling a more favourable orientation of the phenyl ring within its binding site.

Included in this set of derivatives is the para substitution of the halides series; F, Cl, Br and I. Para fluoro-phenylalanine was found to be the weakest binder with a K_d of $25.2 \mu\text{M}$, a modest improvement in binding over CLLFVY. Moving down the group to chlorine the K_d drops to $7.19 \pm 0.8 \mu\text{M}$, then drops to $4.43 \pm 0.4 \mu\text{M}$ for bromine before increasing to $9.53 \pm 0.7 \mu\text{M}$ for iodine. The electronic effect of the halogen series is to be both mesomerically donating and inductively withdrawing, with the strength of these effects diminishing when moving down the group. It has been demonstrated that electron donating groups such as OH and OMe are less tolerated at the para position than electron withdrawing groups such as F(NO₂) and F(CF₃). This could help rationalise the observed trend when moving down the halogen series, as the increasing affinity could be attributed to a weakening of the donating mesomeric effect due to the diminishing orbital overlap with the aromatic system. However, this does not explain the loss of activity upon reaching iodine.

Another characteristic of the halogens, excluding fluorine, is the presence of an electropositive region opposite the bond, termed a σ -hole. This region is larger in later halogens and can interact with Lewis bases.²¹⁶ This does correlate somewhat with the trend in observed affinities, with fluorine showing little effect on CLLFVY affinity whereas Cl and Br both have improvements. However, it is unlikely that this effect is responsible for the improvement in affinity, as once again

it would be expected that iodine would give an improved affinity over bromine. In addition, the derivative F(CF₃) gave the second best affinity of all derivatives ($2.36 \pm 0.2 \mu\text{M}$) despite displaying much more limited σ -hole character.²¹⁷

The importance of the aromaticity of the ring system was also investigated via the non-aromatic cyclohexyl alanine. This derivative showed a 2-fold improvement in K_d ($17.2 \mu\text{M}$) compared with the parent peptide, demonstrating that the aromatic system of the phenyl ring is not necessary for interacting with the protein. Further studies would be needed to validate whether the aromatic systems of other derivatives in this series are contributing to the peptides binding.

Lastly, the chirality at the phenylalanine position was also investigated by inverting the stereochemistry with D-phenylalanine. This was found to have a pronounced effect on the binding with a 13-fold increase in affinity (K_d = $7.07 \pm 0.6 \mu\text{M}$) observed for D-phenylalanine. This is a significant increase in activity, and could be due to an altered structure of the macrocycle as a whole as a result of the change in stereochemistry at this position. This is something that could be more thoroughly assessed in future work.

The most potent binder of this series was the para-t-butyl phenylalanine derivative ($1.52 \pm 0.6 \mu\text{M}$). However, concerns over the solubility of such a hydrophobic group, with visible precipitation observed at $500 \mu\text{M}$ with 10 % DMSO, meant that the second best derivative of the series, para-trifluoromethylphenylalanine ($2.36 \pm 0.2 \mu\text{M}$), was carried forward instead.

Having decided to move forward with the trifluoromethyl phenylalanine derivative, the stereochemistry of this position was also probed. This would show whether the gain in affinity observed previously from swapping the stereochemistry would be additive with that gained from the addition of the para CF₃ group. D-trifluoromethyl phenylalanine was incorporated into CLLFVY to yield CLL-dF(CF₃)-VY and its affinity assessed via MST, and the resulting binding curve is shown in Figure 41. The determined K_d was $2.86 \pm 0.63 \mu\text{M}$ which is within error of the L orientation of the F(CF₃) motif, and therefore it does not appear that these two have a synergistic effect. As such, the peptide CLL-F(CF₃)-VY was carried forward for further optimisation.

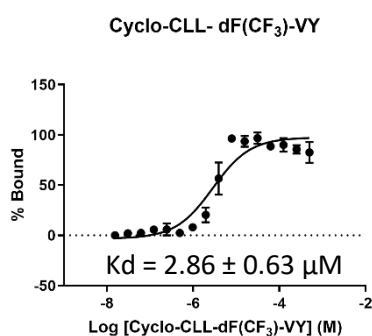


Figure 41. MST binding curve of CLL-dF(CF₃)-VY. Error bars represent standard deviation (n=3).

Binding constants were determined by curve fitting in GraphPad Prism 8.

2.9 CLLF(CF₃)VY Alanine Scan

With the K_d now in the single digit micromolar range, it was decided that the alanine scan could be repeated with the aim to obtain full binding curves. This should then enable the identification of key amino acid residues which can be used to guide the further development of CLLF(CF₃)VY. The five new alanine scanning analogues were synthesised via Fmoc-SPPS and their affinities assessed by MST against the HIF-1 α PAS-B domain. The results of this second alanine scan and a comparison of their K_d 's are shown in Figure 42. As expected the replacement of the cysteine and F(CF₃) residues led to a significant decrease in affinity. ALLF(CF₃)VY showed some weak binding activity, but as a full binding curve was not obtained, only a rough estimate of the K_d can be achieved ($>200 \mu\text{M}$). Replacement of the cysteine residue of CLLF(CF₃)VY with alanine led to a loss of activity that prevented a binding curve from being plotted (Figure 42). This corroborates the result of the alanine scan of CLLFVY which also showed the cysteine was critical for binding activity (section 2.5). The remaining four substitutions all yielded full binding curves with very similar affinities, all approximately 2-fold worse than the parent peptide. These results indicate that these amino acids are interacting with the HIF-1 PAS-B domain, but are contributing to binding less than the cysteine and F(CF₃). As these positions have been determined to have a lesser contribution to binding, it was reasoned that substitutions at these positions might introduce new interactions with HIF that would increase the peptides affinity, and so derivatives of the remaining positions were investigated.

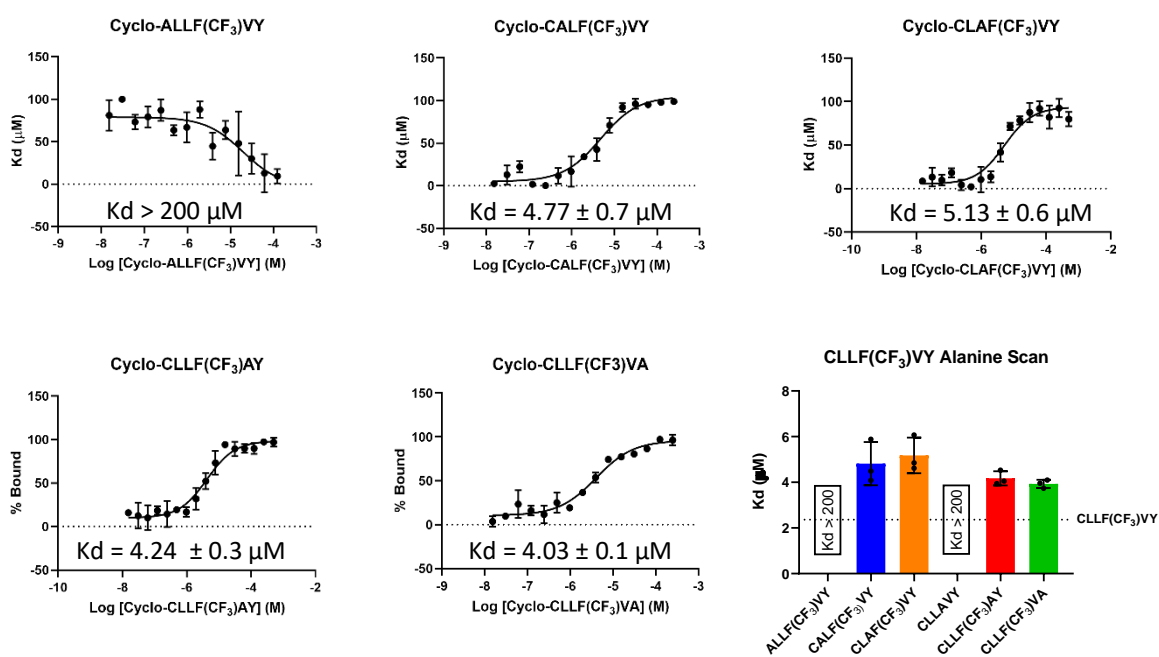


Figure 42. MST binding curves for the CLLF(CF₃)VY alanine scan, and comparison of obtained K_d values. Error bars represent standard deviation (n=3). Binding constants were determined by curve fitting in GraphPad Prism 8.

2.10 Tyrosine Analogues

The next residue investigated was tyrosine, and a series of 16 derivatives were synthesised using the same unnatural amino acids as were used for phenylalanine. The derivatives were assessed by MST and a comparison of their K_d values are shown in Figure 43, with the MST binding curves shown in Figure 44.

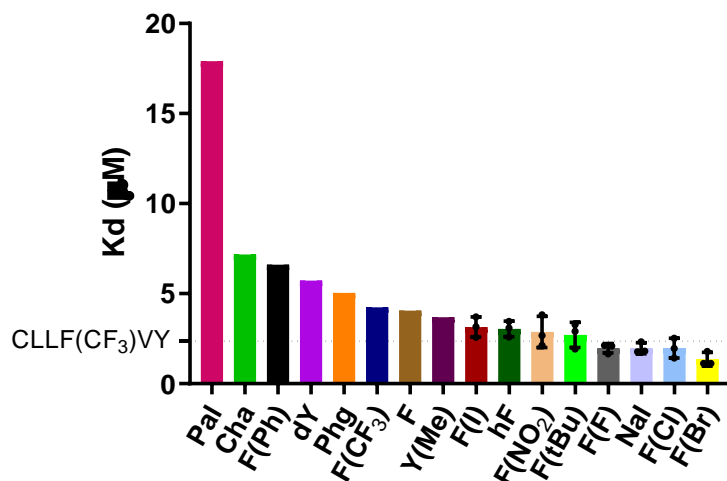


Figure 43. Comparison of the K_d values obtained for the tyrosine derivatives. Error bars represent standard deviation ($n=3$).

The pyridyl motif (Pal) showed a similar deleterious effect in this position as it did in the phenylalanine series, with an 8-fold loss in affinity observed. Notably, both the Y(Me) and F(NO₂) derivatives are far better tolerated at this position in the peptide, with Y(Me) leading to an approximate 1.5-fold loss in activity and F(NO₂) being within error of the parent peptide. This justifies the amino acid choice from the initial SICLOPPS screen, with the choice of tyrosine over phenylalanine reflecting the increased tolerance of polar atoms at this position.

The tyrosine position was shown to display an SAR significantly different from that of the phenylalanine position, with many of the substituents that gave the best affinities in the phenylalanine position now causing loss of binding activity. This is most evident in the F(Ph) which led to a three-fold loss of activity, compared to a ten-fold increase when in the phenylalanine position. This suggests that the phenyl rings of the tyrosine and phenylalanine have distinct binding modes within the protein-binding site. Similar results were seen with the F(CF₃) group which led to a two-fold loss of binding when in the tyrosine position. The larger F(tBu) substituent bound with a similar affinity to the parent peptide (CLLF(CF₃)VY). Again, we also see that the extended aromatic system of the naphthyl derivative (NaI) leads to an increase in affinity ($K_d = 1.92 \pm 0.2 \mu\text{M}$).

Chapter 2

The position of the phenyl ring was again probed using the phenylglycine and homophenylalanine derivatives. For these derivatives a similar trend was observed as for the phenylalanine position, with shortening of the linker via phenylglycine leading to a loss of activity ($5.04 \mu\text{M}$) when compared to CLLF(CF₃)VF ($4.06 \mu\text{M}$). In contrast, lengthening the linker with homophenylalanine led to an increase in affinity ($3.04 \pm 0.4 \mu\text{M}$). This is again most likely due to the increased flexibility allowing for a better orientation of the ring within its binding site.

The stereochemistry at this site was also investigated via the compound CLLF(CF₃)V-dY. This led to a two-fold loss of activity, and could be attributed to this orientation positioning the ring away from its binding site. Alternatively, the swap of stereochemistry could again be having an effect on the overall macrocycle structure, though given the relatively small change in binding activity this seems less likely than the inversion of stereochemistry at the phenylalanine position.

The para-substitution of the halogen series was also investigated, and it is this series that gave the best affinities of the tyrosine derivatives; CLLF(CF₃)VF(Cl) ($1.89 \pm 0.4 \mu\text{M}$), and CLLF(CF₃)VF(Br) ($1.35 \pm 0.3 \mu\text{M}$). The para-fluoro phenylalanine also led to an improvement in affinity with a K_d of $1.99 \pm 0.2 \mu\text{M}$, with only the para-iodo derivative leading to a small loss of activity at $3.06 \pm 0.5 \mu\text{M}$. Given the increase in affinity for the para-fluoro derivative, the increase in affinity for halogen substituents is very unlikely to be attributable to the previously discussed sigma effect as fluorine lacks a sigma hole. This is further evidenced by the poorer tolerance of the iodo derivative which displays the strongest sigma effect. The affinity of the derivatives increases moving down the halogen series with the exception of iodine. This could be due to its size preventing it from interacting with the protein in the same way as the other halogens, and indeed other substituents of similar size are also poorly tolerated (F(tBu) and F(CF₃)). If this were the case, then the increase in affinity could be attributed to the trends observed moving down the halogen series, including the increase in polarizability leading to greater Van der Waals interactions.

It is worth noting that the biggest improvement in this series F(Br) led to an improvement in K_d of $1 \mu\text{M}$, which corresponds to a change in free energy of binding of 1.38 kJ/mol (0.33 kcal/mol). This magnitude of energy change is most often associated with weak hydrophobic interactions, demonstrating that this substitution is an incremental improvement and is not a major component of compound binding.²¹³ There does not seem to be a clear consensus rationalising the changes observed in the 16 tyrosine derivatives. It is possible a wider variety of analogues with more varied chemistries may have a greater impact, and this could be investigated in future work. Through derivatisation at this position the affinity of the peptide has been improved two-fold and the best derivative (CLLF(CF₃)VF(Br) was carried forward (K_d = $1.35 \pm 0.3 \mu\text{M}$).

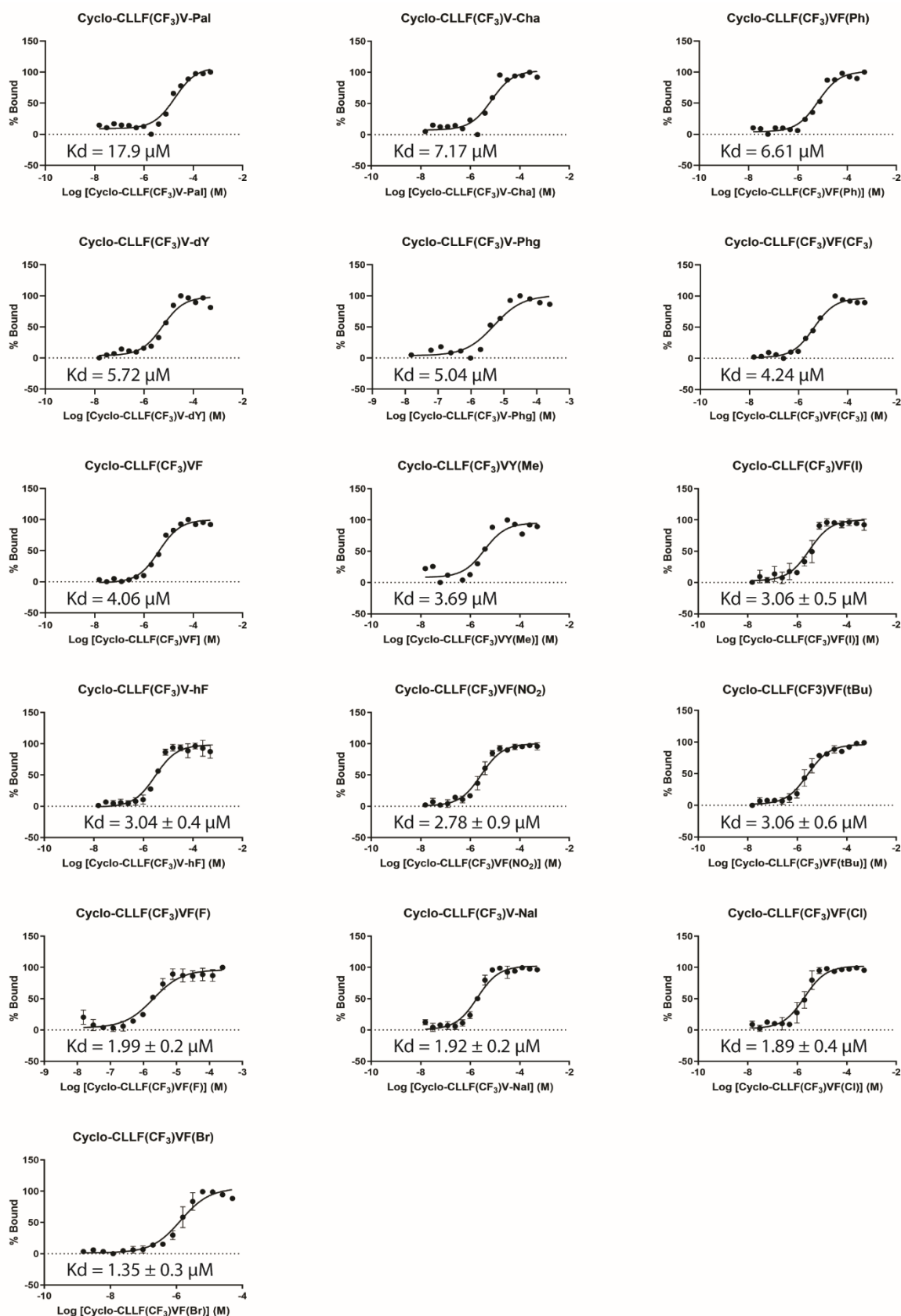


Figure 44. MST binding curves of the 16 tyrosine analogues of CLLF(CF₃)VY. Where present error bars represent standard deviation (n=3). Binding constants were determined by curve fitting in GraphPad Prism 8.

2.11 Valine Analogues

The valine position of CLLF(CF₃)VF(Br) was probed next. This required a different selection of unnatural analogues that more closely resembled the aliphatic side chain of the valine that was selected in the SICLOPPS screen. For this, the two other natural aliphatic amino acids (I,L) and a further five unnatural aliphatic amino acids (hL, NL, NV, Aib and Prop) were substituted in place of the valine. The structure of the substituted amino acids are shown in Figure 45. The analogues were synthesised using the same Fmoc SPPS protocols as before and their affinity for the HIF-1 α PAS-B domain determined by MST, with the resulting binding curves are shown in Figure 46.

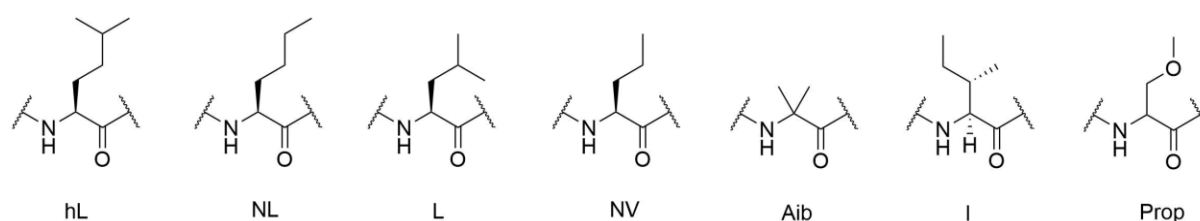


Figure 45. Structures of the amino acids substituted for valine.

Overall there was little improvement seen for the majority of the valine derivatives, with most peptides yielding a K_d of approximately 1 μ M. A small preference was seen for branched aliphatic side chains, with hL, L and I yielding the highest affinities of this series. Substitution of the valine for isoleucine yielded the greatest improvement with a 3-fold increase in affinity ($K_d = 473 \pm 246$ nM). This improvement is likely due to enhanced hydrophobic interactions of the isoleucine within the protein-binding site. Side chains with longer aliphatic chains also performed better than the parent peptide, though the improvement in affinity was modest. This is suggestive of weak hydrophobic interactions with the protein, which reinforces the results of the alanine scan that this position is not tightly engaged with the HIF protein.

2-amino-3-methoxypropionic acid (Prop) was also well tolerated despite containing the polar oxygen atom, with the obtained K_d being within error of the previous peptide CLLF(CF₃)VF(Br). This again demonstrates that the valine position is not contributing significantly to HIF binding.

The most prominent shift in affinity was observed with the NV derivative (CLLF(CF₃)-NV-F(Br)) which showed a 17-fold loss of affinity ($K_d = 22.45 \mu$ M). This is surprising given its structural similarity with the other analogues tested, with many only differing by a methyl group from the NV side chain. Given the similarity, it seems unlikely the loss of activity is due to the interaction of the side chain within the protein-binding site, and more likely due to a change in the peptides conformation. This could displace the cysteine and/or phenylalanine positions which are critical for the peptides affinity, leading to the observed loss of binding. Future work focussing on the impact of conformation on the peptides affinity for binding HIF will be needed to clarify this

result. As the isoleucine derivative gave the best affinity, this derivative was carried forward and the L2 position next probed.

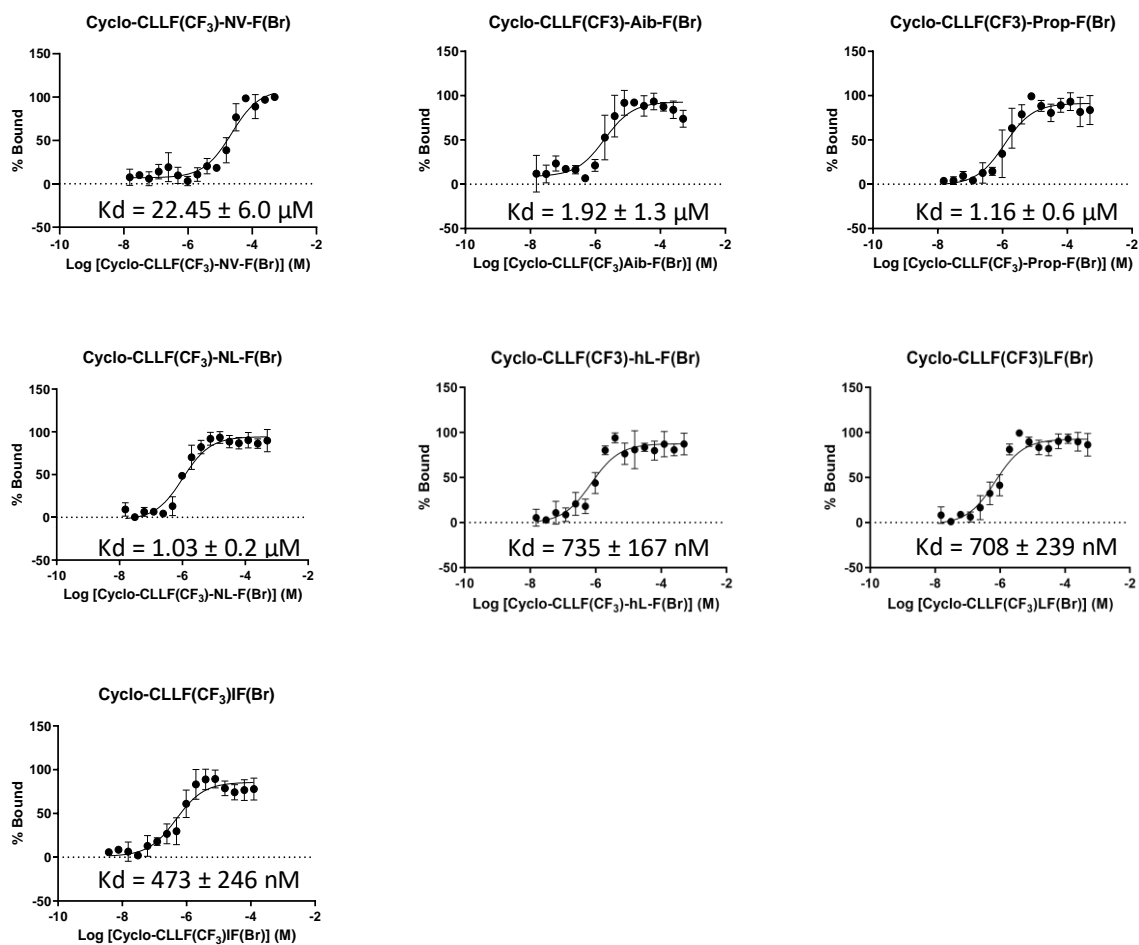


Figure 46. MST binding curves for the valine derivatives. Error bars represent standard deviation (n=3). Binding constants were determined by curve fitting in GraphPad Prism 8.

2.12 L2 Analogues

The same five unnatural analogues used in the valine series along with the natural amino acids I and V were then substituted for the L2 position (CLLF(CF₃)IF(Br)), and their binding determined by MST. The resulting binding curves are shown in Figure 47.

Upon running the MSTs, it was noted that the error of the binding curves in this derivative series has increased from those obtained for previous derivatives. The NL derivative in particular yielded a binding curve with very large error bars. The shape of this curve has also deviated away from the expected sigmoidal shape with a significant hook present at higher concentrations. The error in this measurement is significant enough that an accurate value for binding could not be determined. This could be because of a deteriorating solubility profile, though no clear precipitation was observed upon sample preparation. Alternatively, it is possible that the increasing hydrophobicity of the peptide has led to non-specific binding events that are affecting the MST results. It was therefore not possible to accurately determine the affinity of these derivatives at this point, and it was decided that the observed artefacts would be investigated experimentally.

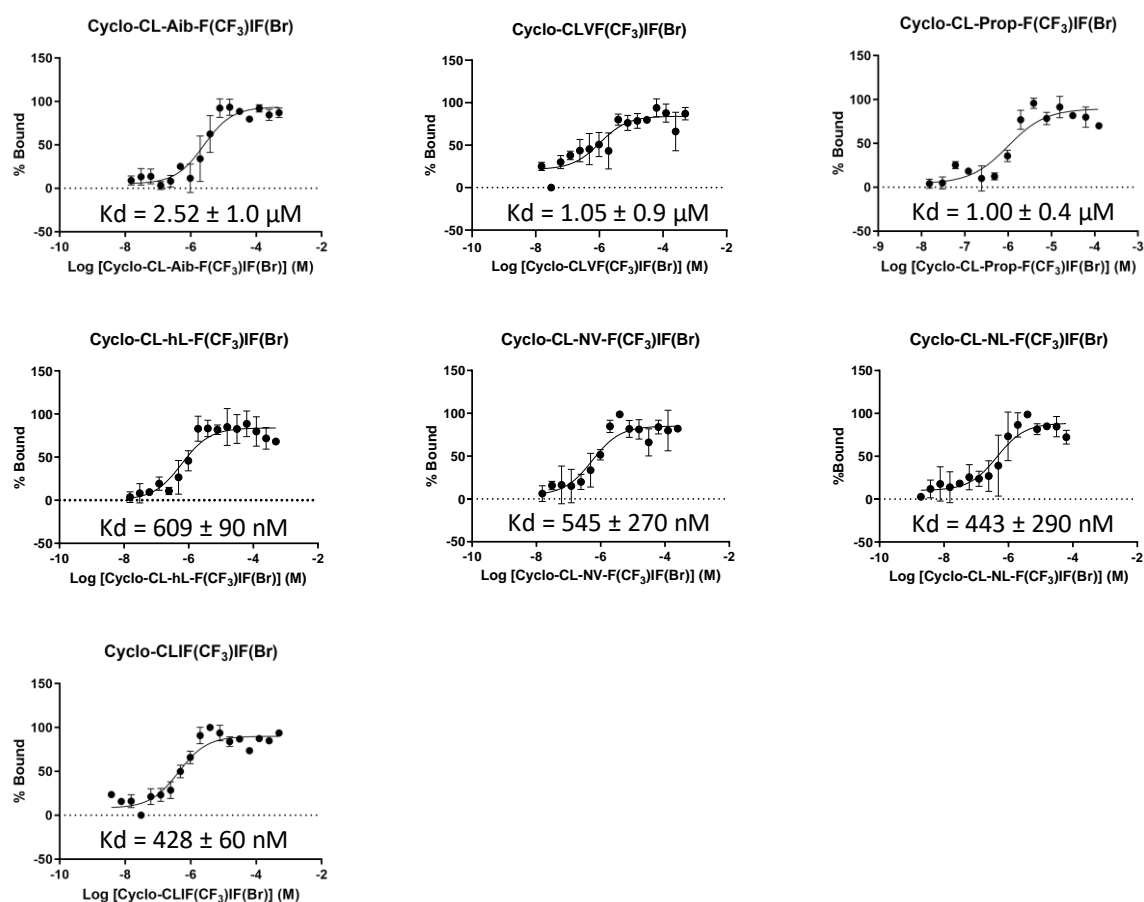


Figure 47. MST binding curves obtained for L derivatives. Error bars represent standard deviation (n=3). Binding constants were determined by curve fitting in GraphPad Prism 8.

2.13 Investigating Artefacts

To investigate the origin of the observed artefact in the peptide MSTs, an experiment was designed whereby the MST of the compound CL-NL-F(CF₃)IF(Br) was repeated in the presence of DTT (2mM). As has been previously demonstrated, DTT is a competitive binder of the HIF-1 α PAS-B domain, and so it was predicted that its addition should abolish the peptides normal binding activity to yield a flat line instead of a binding curve. If a secondary binding event were occurring, then a curve would still be obtained in the presence of DTT. The results of this experiment are shown in Figure 48.

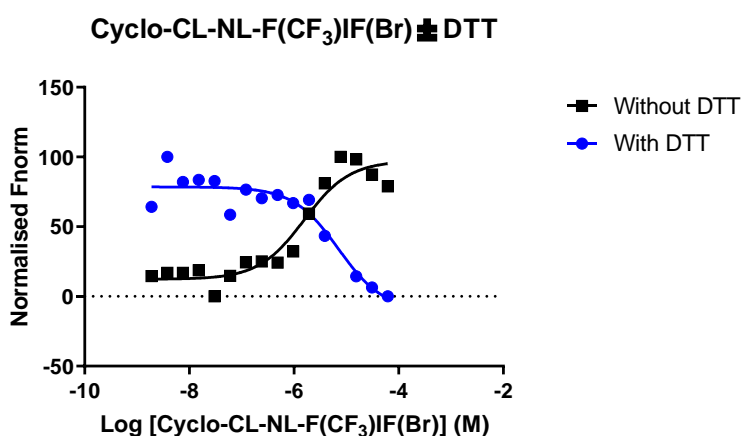


Figure 48. MST binding curves for CL-NL-F(CF₃)IF(Br) with and without DTT.

As predicted, the addition of DTT led to the abolishment of the typical binding activity and instead yielded a weaker curve in the opposite direction. Overlaying the curves obtained with and without DTT shows that this secondary curve is most likely the cause of the observed artefact in the L2 series. It can be seen that the superposition of the two curves will cause the observed drop in Fnorm for the top three concentrations. This is most likely the cause of the high error and deviation from a sigmoidal shape that was observed in many of the L derivative series. This secondary binding curve is most likely due to the increased hydrophobicity of the peptide causing non-specific binding at higher concentrations. However, there remains the possibility that the compound has been optimised to a secondary binding site on the HIF-1 α PAS-B domain. In order to investigate this possibility, the introduction of charged residues that would reduce the compounds hydrophobicity was investigated.

2.14 Introduction of Charged Residues

To investigate whether decreasing the hydrophobicity of the peptides would mitigate the secondary binding event, the introduction of charged residues at the L1 position was investigated. The introduction of these residues should improve the peptides aqueous solubility and therefore reduce any non-specific hydrophobic interactions. Two derivatives were synthesised containing either lysine or glutamic acid, therefore probing the impact of introducing both a positive and negative charge. The introduction of side chains containing carboxyl and amine functional groups presented an issue for the current method of peptide synthesis. This is because the side chain groups would be deprotected during the TFA resin cleave, leaving them capable of reacting during the peptide cyclisation step. To prevent this from occurring, a new synthetic strategy was devised whereby a 2-chlorotrityl chloride resin was used. This resin can be cleaved under milder conditions than Wang resin and so cleavage of the linear peptide can be achieved with the side chain protecting groups left intact.

The first step of this synthesis required loading of the 2-chlorotrityl chloride resin, which was achieved by swelling the resin in DCM before adding 1 eq. of the amino acid (Fmoc-Lys(Boc)-OH or Fmoc-Glu(OtBu)-OH), and 5 eq. of DIPEA. The resin loading is then determined by performing an Fmoc deprotection on an accurately weighed sample of the resin and measuring the absorbance of the resulting dibenzofulvene-piperidine adduct (Figure 49) at 301 nM.

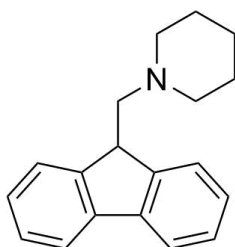
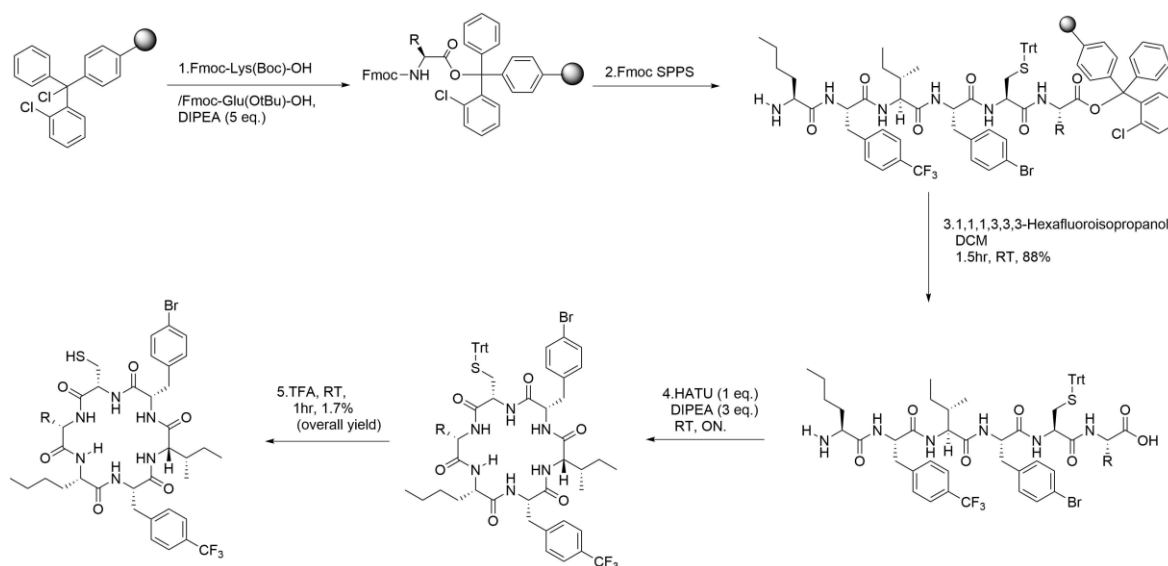


Figure 49. Dibenzofulvene-piperidine adduct that absorbs at a wavelength of 301 nm.

The desired resin bound linear peptide can then be synthesised using standard Fmoc SPPS, before being cleaved with 1,1,1,3,3,3-hexafluoroisopropanol. The protected linear peptide can then be cyclised under the previously described conditions to yield a protected cyclic which is globally deprotected with TFA as the final step. This synthesis is outlined in Scheme 10. As the side chain protecting groups are maintained after cleavage the cysteine StBu protecting group is no longer necessary, and instead the acid labile trityl protecting group is used to minimise the number of synthetic steps required.



Scheme 10. Synthesis of lysine and glutamic acid derivatives via protected linear sequences utilising a 2-chlorotrityl chloride resin.

Once synthesised the peptides were assessed by MST and the obtained traces are shown in Figure 50. Both peptides were found to bind with a similar affinity of approximately 1 μM . This is somewhat surprising given the oppositely charged side chains, and suggests that the side chains are orientated facing into the solvent rather than interfacing with the protein. Given the purpose of these derivatives was to solubilise the peptide, this orientation of the side chains into the solvent is probably advantageous, and indeed for both peptides the aforementioned tailing up of the MST traces is diminished. The glutamic acid derivative yields a curve with the expected sigmoidal shape with a K_d of $1.21 \pm 0.2 \mu\text{M}$. The lysine derivative has improved the curve shape and reduced the error. However, it is still showing some unusual points at the top plateau of the binding curve. As before there are two points which are tailing down at the highest concentrations tested (250, 500 μM). When these points are excluded, the K_d of this compound drops to $2.08 \pm 0.4 \mu\text{M}$ and this value likely better reflects the actual binding affinity of this compound.

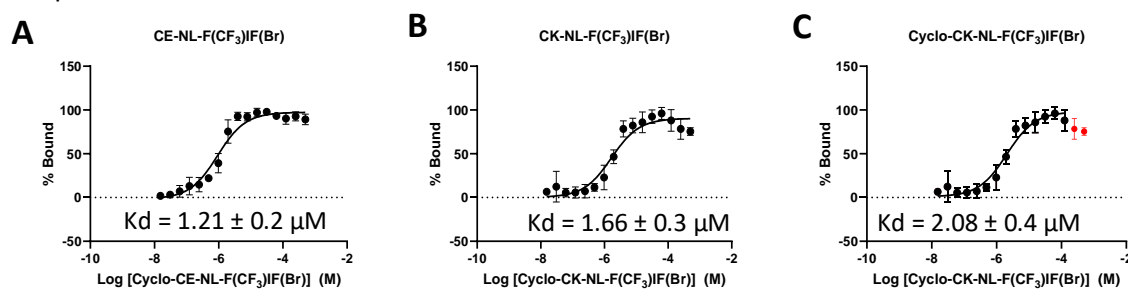


Figure 50. MST binding curves of the glutamic acid (A) and lysine derivatives (B). C) MST binding curve for CK-NL-F(CF₃)IF(Br) with the top two concentrations removed from the binding curve analysis (shown in red). Error bars represent standard deviation (n=3). Binding constants were determined by curve fitting in GraphPad Prism 8.

Chapter 2

The ability of the K and E amino acids to prevent non-specific binding was also probed using DTT. As has been previously demonstrated DTT acts as a competitive inhibitor of HIF and prevents the peptides major binding mode allowing the affinity of any secondary affects to be measured. The K and E derivatives were analysed by MST in a DTT (2 mM) buffer and the obtained binding curves are shown in Figure 51.

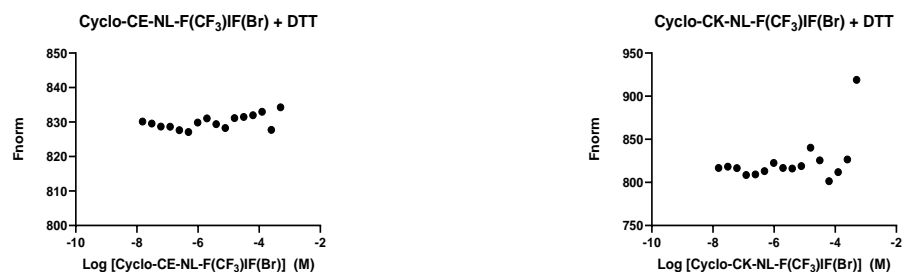


Figure 51. MST binding curves of CK-NL-F(CF₃)IF(Br) and CE-NL-F(CF₃)IF(Br) in the presence of DTT.

Both compounds showed very little activity in these experiments with the E derivative yielding a flat line, in keeping with the lack of any obvious secondary effects in the original binding curve. The K derivative also showed very little activity, with only the top concentration (500 μ M) having an increased Fnorm value. This reinforces what was seen in the original MST of this compound, where only the top concentrations deviated from the plotted binding curve.

These experiments demonstrate that the cyclic peptides CE-NL-F(CF₃)IF(Br) and CK-NL-F(CF₃)IF(Br) bind to the HIF-1 α PAS-B domain with affinities 30 times and 17 time better than the parent CLLFVY respectively. Moreover, they have demonstrated that the introduction of charged residues at the L1 position of CLLFVY are both well tolerated and able to diminish any non-specific binding effects as observed by MST. As it is not known at which point in the derivatisation of CLLFVY the secondary binding event was introduced, these peptides were selected as the lead compounds for use in further assay development, as it has been shown experimentally that they do not have secondary binding modes. The compounds developed here, (CE-NL-F(CF₃)IF(Br) and CK-NL-F(CF₃)IF(Br), represent a significant increase in binding affinity over the parent peptide and will be useful for the development of probes and additional biophysical assays.

2.15 Investigating HIF Isoform Specificity

Having probed the SAR at each position of CLLFVY and produced two lead compounds with improved affinities, experiments were next carried out to determine if the derivatisation of CLLFVY had affected its ability to selectively bind to the HIF-1 α isoform over HIF-2 α . This was probed using MST with the labelled HIF-2 α PAS-B domain, which was kindly provided by Dr Andy Ball. The MSTs were carried out as described for HIF-1 α and the resulting binding curves are shown in Figure 52 below.

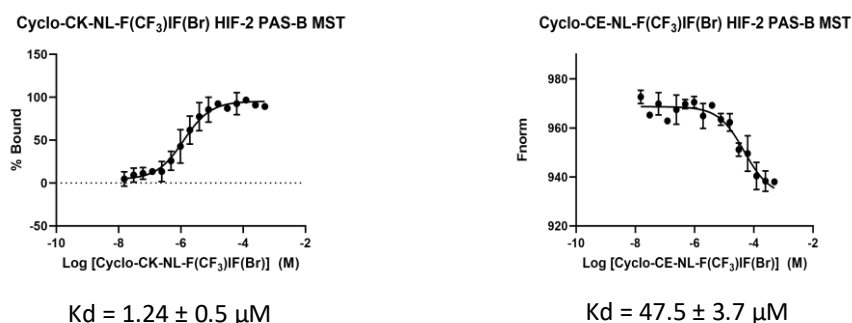


Figure 52. MST binding curves of CK-NL-F(CF₃)IF(Br) and CE-NL-F(CF₃)IF(Br) against the HIF-2 α PAS-B domain. Error bars represent standard deviation (n=3). Binding constants were determined by curve fitting in GraphPad Prism 8.

CK-NL-F(CF₃)IF(Br) bound to the HIF-2 α PAS-B domain with a K_d of $1.24 \pm 0.5 \mu\text{M}$, which is within error of the affinity obtained for HIF-1 α . Therefore, CK-NL-F(CF₃)IF(Br) has lost all selectivity between HIF- α isoforms. CE-NL-F(CF₃)IF(Br) bound to the HIF-2 α PAS-B domain with a K_d of $47.5 \pm 3.7 \mu\text{M}$. This represents a fifty-fold lower affinity when compared to HIF-1 α , and represents an increase in selectivity over the parent CLLFVY, which bound to HIF-1 α with approximately ten-fold higher affinity than HIF-2 α . The significant difference between these derivatives is suggestive of the charged residues influencing the selectivity of the peptide, with the positively charged K residue interacting favourably with the HIF-2 α PAS-B domain. Conversely, the negatively charged E residue leads to a fifty-fold loss of activity, which could be the result of a negatively charged residue present in the binding site of HIF-2 α that is not present in HIF-1 α . Experiments are currently underway to try to identify this amino acid which could provide valuable insight into the development of isoform specific inhibitors.

Chapter 2 Summary

In summary, Fmoc SPPS was used to synthesise the hit compound CLLFVY. An MST assay was subsequently developed that was capable of assessing the affinity of CLLFVY for the HIF-1 α PAS-B domain ($K_d = 36.0 \pm 3.6 \mu\text{M}$). This assay was then used to assess the importance of each residue of CLLFVY via alanine scanning. This revealed that the cysteine residue was the most critical residue for binding. A series of analogues demonstrated that it was the thiol side chain that was critical for binding, but it was revealed to not be mediated via a disulphide interaction. The phenylalanine position was then investigated, yielding the compound CLLF(CF₃)VY which bound with an affinity 15-fold better than CLLFVY ($2.36 \pm 0.2 \mu\text{M}$). Following a second alanine scan that did not reveal any other critical residues, an extensive SAR study was carried out probing the remaining positions of the CLLF(CF₃)VY peptide. These results highlighted a challenge in the development of inhibitors of PPIs, as the flat and featureless hot spots that are targeted often lead to increasing reliance on weaker Van der Waals forces that push compounds to become increasingly hydrophobic. This was evidenced by the observed second binding curve that was seen in the later derivatives. This non-specific binding event was mitigated via the introduction of charged residues leading to the development of the compounds CE-NL-F(CF₃)IF(Br) and CK-NL-F(CF₃)IF(Br). These peptides have improved the affinity of binding to the HIF-1 α PAS-B domain by 30 and 17-fold respectively. ($K_d = 1.21 \pm 0.2 \mu\text{M}$ and $2.08 \pm 0.4 \mu\text{M}$). The impact of the SAR study on isoform specificity was then investigated, which revealed that the selectivity was significantly affected by the charged residue E or K. CK-NL-F(CF₃)IF(Br) was found to act as a pan-HIF binder, whilst CE-NL-F(CF₃)IF(Br) was found to preferentially bind to HIF-1 α , with a 40-fold higher K_d for 1 α over 2 α .

Chapter 3 Development of Further *In Vitro* Assays

3.1 Enzyme-Linked Immunosorbent Assay

Up until this point all *in vitro* characterisation of CLLFVY, and derivatives thereof, have been carried out using MST. Whilst this biophysical technique has enabled the determination of the binding affinities of the peptides towards the HIF-1 α PAS-B domain, this does not necessarily translate to improved activity and a more potent inhibitor. In order to assess the compounds activity a different technique was required. Initially the use of enzyme-linked immunosorbent assay (ELISA) was explored for this purpose.

ELISAs are solid-phase immunoassays, commonly used for probing biomolecular interactions, and are generally considered as an easy and specific method for use as both a research and diagnostic tool.²¹⁸ The high specificity of this technique arises from the use of antibodies, which are capable of remaining selective even when the antigen of interest is only present at very low concentrations. There is now a plethora of commercially available antibodies for many different antigens, including antibodies that recognise common peptide tags. This is especially useful if the protein targets have been obtained via recombinant protein expression, as many proteins will already contain a targetable peptide tag. Whilst there are many types of ELISA, the technique which is best suited to monitoring PPIs is the indirect ELISA. Direct ELISAs use one antibody to recognise a plate bound antigen, whereas indirect ELISAs utilise a secondary antibody, with the primary antibody recognising the antigen and the secondary, often a polyclonal anti-species antibody, binding to the primary antibody. This secondary antibody increases the sensitivity of the assay and is also conjugated to a visualisation marker to enable the quantification of binding.²¹⁹

To run an indirect ELISA capable of assessing the interaction of HIF-1 α PAS-B and HIF-1 β PAS-B, the HIF-1 α PAS-B domain first had to be immobilised on a 96-well plate. This was achieved by incubating the HIF-1 α PAS-B domain in the wells of a clear Pierce nickel coated 96-well plate, which binds to the poly-histidine tag on the HIF-1 α PAS-B domain. The plate is then washed and a blocking step carried out via incubation with 2 % milk in PBST (PBS + Tween 20) to prevent non-specific interactions. The binding partner, HIF-1 β , is then added along with any ligand to be tested. If the ligand functions as an inhibitor of the dimerisation of the two proteins then HIF-1 β will not bind to HIF-1 α and so will be washed from the plate. However, if there is no ligand present or it does not function as an inhibitor, HIF-1 β will bind to HIF-1 α and therefore be linked to the plate. An antibody specific to HIF-1 β is then incubated with the plate-bound complex, in this case a mouse anti-FLAG antibody specific for the FLAG tag present on the N-terminus of HIF-1 β (DYKDDDDK). A secondary antibody, anti-mouse-horseradish peroxidase, is then added which

Chapter 3

specifically recognises the mouse anti-FLAG antibody and is conjugated to horseradish peroxidase (HRP). HRP is a metalloenzyme capable of oxidising various organic substrates when in the presence of hydrogen peroxide. Next, a reagent containing both hydrogen peroxide and 3,3',5,5'-tetramethylbenzidine (TMB) is added to each well. TMB is a chromogen, which yields a colour change upon oxidation and therefore functions as a colorimetric method for the detection of HRP, and so subsequently a quantitative method for the detection of HIF binding. Upon initial oxidation the TMB reagent yields a blue colour. A strong acid is then added to shift the maximum absorbance to 450 nm, and the magnitude of absorbance is then measured using a plate reader.^{220,221} This process is outlined in Figure 53.

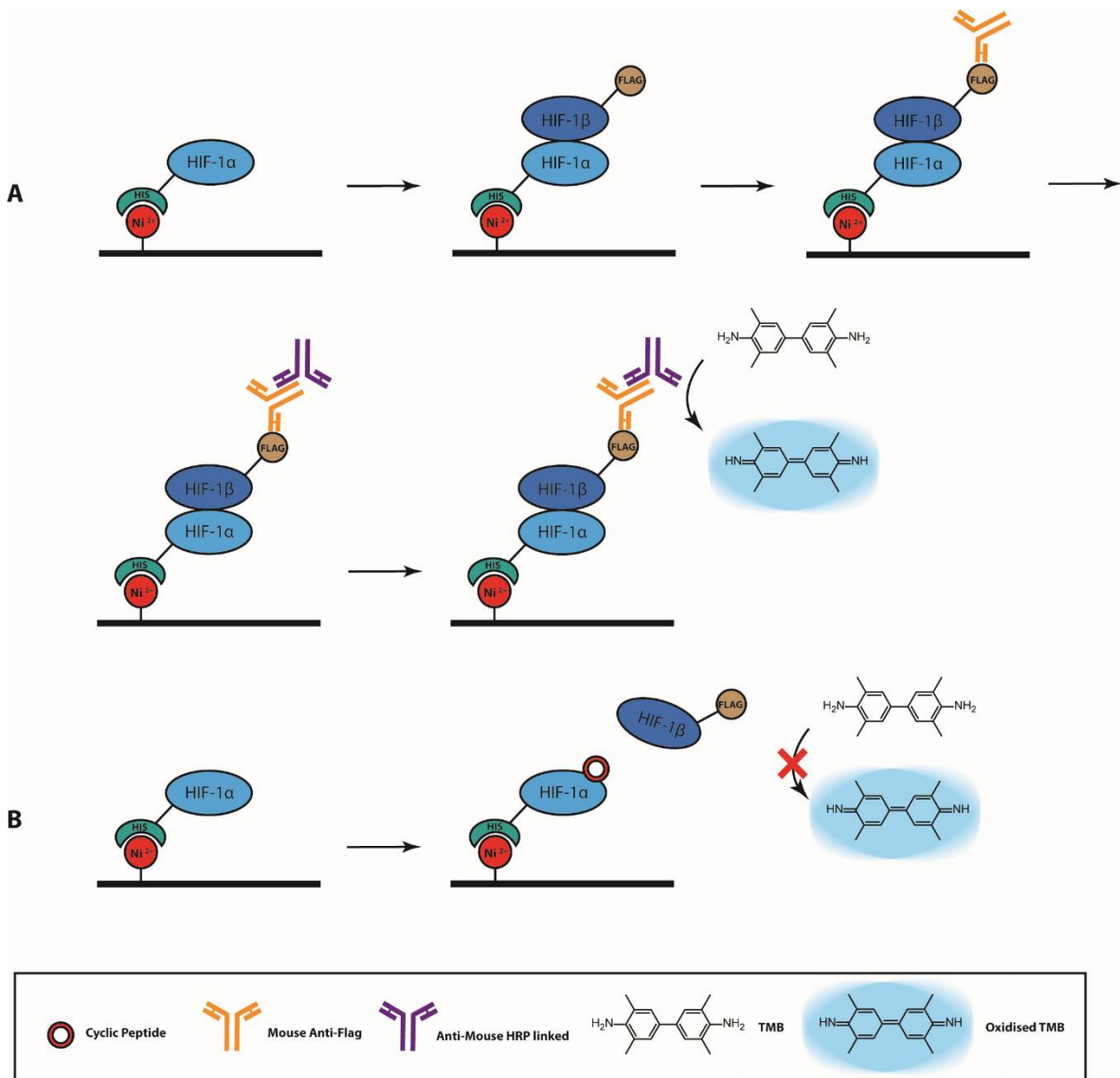


Figure 53. Overview of an ELISA assay. A) Illustration of the stages of an indirect ELISA assay for the assessment of HIF dimerisation without ligand present. B) Illustration of an indirect ELISA run in the presence of an inhibitor of HIF dimerisation, preventing the oxidation of the TMB dye and therefore no colour develops.

3.1.1 Expression of HIF-1 β PAS-B Domain

To assess if the peptides disrupt HIF-1 dimerisation, a different protein model is needed from that used in the MST assay, with one that contains both HIF-1 α and HIF-1 β now required. As CLLFVY has previously been shown by MST to bind to the PAS-B domain¹⁴⁶ it was deemed likely that it would act via the disruption of the HIF-1 α PAS-B-HIF-1 β PAS-B interaction, especially given the nature of PAS domains of similar proteins to mediate dimerisation interactions.⁴³ A previous model consisting of the PAS-B domains of both α and β was described by Cardoso *et al*, where they have reversed a salt bridge between the two domains to give a stable dimer, something that could not be achieved with the wild type domains.¹⁶ This mutant dimer was then used to yield a crystal structure which shows the two domains interacting via an anti-parallel beta-sheet contact (Figure 54A).¹⁶ This work builds on previous studies that identified four point mutations that increased the stability and solubility of the HIF-1 α PAS-B domain as well as decreasing its propensity to form homodimers.¹⁷ The four mutations include the reversed salt bridge, R245E, but also include E266H, R311H and S330L mutations, and were decided upon by comparing the HIF-1 α PAS-B structure with the more stable HIF-2 α PAS-B isoform. The location of these mutations are mapped onto the crystal structure of the HIF-heterodimer in Figure 54B. Given the improved solution characteristics and dimerisation properties of this mutated HIF domain, it was chosen for use in the ELISA. However, as the MSTs up until this point had been carried out on wild type HIF-

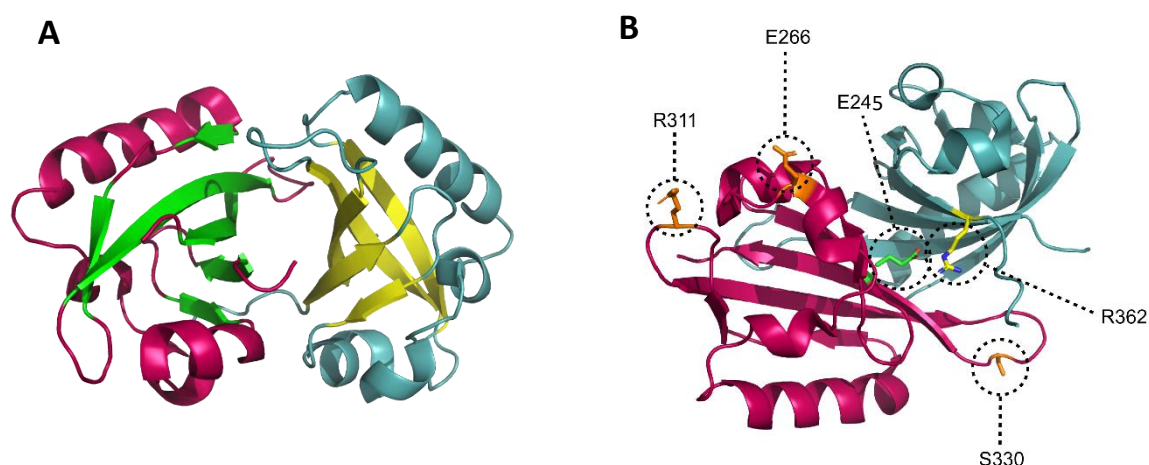


Figure 54. Analysis of the structure of the HIF-1 α PAS-B mutant. A) Crystal structure of HIF-1 α PAS-B (magenta) and HIF-1 β PAS-B (cyan) heterodimer with the anti-parallel beta sheet interaction highlighted green (α) and yellow (β) (adapted from PDB: 4H6J).¹¹⁸ B) Crystal structure of HIF-1 α PAS-B and HIF-1 β PAS-B heterodimer with highlighted mutations. Green: HIF-1 α R243E. Yellow: HIF-1 β E362R. Orange: Positions that have been mutated from structure shown, HIF-1 α R266E, R311H, S330L (adapted from PDB: 4H6J).¹¹⁸

1 α PAS-B domain, it was decided that the assay would also be attempted with the wild type α and β domains as well.

To utilise this new protein model, the HIF-1 β PAS-B domain and HIF-1 α PAS-B 4x mutant domain had to be obtained. The mutant HIF-1 α PAS-B domain was expressed by incorporating the gene encoding this protein into a pET-28 vector, which was transformed into *E. coli* BL21 (DE3) Rosetta cells and expressed under the same conditions as the wild type protein. As reported, the mutant HIF-1 α PAS-B domain was found to be much more amenable to recombinant expression, yielding 7.5 mg of protein per litre of expression. This is a ten-times higher yield than that obtained for the wild type protein (section 0). The protein was then analysed by thermal shift, which showed the increased stability of the mutant domain, which had a two-degree increase in melting temperature (54.25 $^{\circ}$ C) compared to the wild type (52.25 $^{\circ}$ C, Figure 23A). The first derivative of the thermal shift data is shown in Figure 55A. An aliquot of the mutant HIF-1 α PAS-B domain was then labelled as previously described (section 2.4.2) to enable its use in an MST assay. This enabled confirmation that the mutations had not affected the binding of CLLFVY, and a dose-response curve for a serial dilution of CLLFVY with the mutant HIF-1 α PAS-B domain is shown in Figure 55B. CLLFVY was found to bind to the mutant HIF domain with an affinity of $49.2 \pm 3.7 \mu\text{M}$, which is slightly weaker than the binding observed for the wild type ($36.04 \pm 3.6 \mu\text{M}$), representing a 1.3-fold loss of affinity. This small loss of activity should not impede the use of the mutant protein in the ELISA assay.

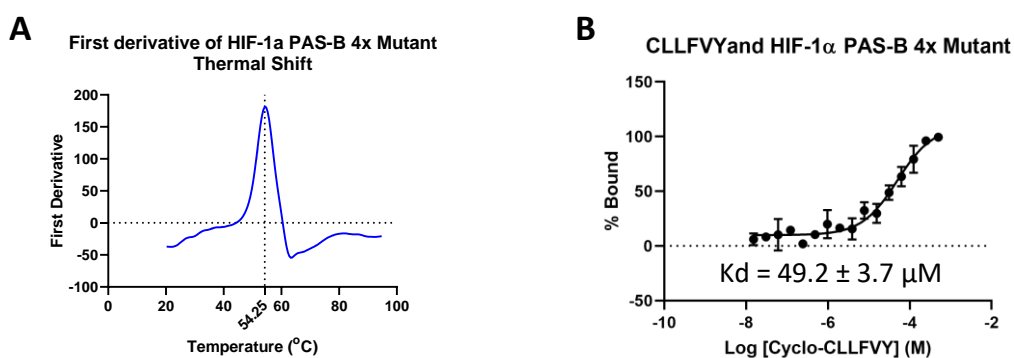


Figure 55. A) First derivative of thermal shift data obtained for the HIF-1 α PAS-B 4x mutant, with a melting temperature of 54.25 $^{\circ}$ C. B) Dose-response curve for CLLFVY binding to the HIF-1 α PAS-B 4x mutant. Error bars represent standard deviation (n=3). Binding constants were determined by curve fitting in GraphPad Prism 8.

The HIF-1 β PAS-B domains were obtained using a pET-28 vector with the gene coding for either the wild type or E362R mutant cloned in, after which the plasmids were transformed into *E. coli* BL21 (DE3) Rosetta cells and expressed following the same procedures as for the α domain. The construct also included two peptide tags, an N-terminal polyhistidine tag followed by a thrombin cleavage site, and then a FLAG tag. A map of the plasmid used is shown in Figure 56A. This

enabled easy purification after expression via the His-tag, but also enabled the removal of this tag post purification via incubation with thrombin. This is necessary to ensure the HIF-1 β PAS-B domains are not capable of binding to the Ni-plate used for the ELISA. The FLAG tag is included so that the β domain contains an epitope not present in the α domains, enabling specific recognition of this domain by the corresponding antibodies.

After the HIF-1 β PAS-B domains were expressed, the cells were lysed and the protein isolated from the soluble fraction via nickel affinity chromatography. The protein was then purified via SEC and protein-containing fractions were incubated with thrombin at 4 °C overnight. A reverse HIS-trap was then performed to remove any uncleaved protein as well as the cleaved HIS-tag. A further SEC purification was then carried out to yield the desired protein construct. The HIF-1 β PAS-B domains were found to express well under these conditions, yielding 11.8 mg/L of wild type HIF-1 β PAS-B and 12.9 mg/L of the E362R mutant. A thermal shift of the two HIF-1 β PAS-B constructs found that the wild type and mutant proteins both had melting temperatures of 51.25 °C (Figure 56B).

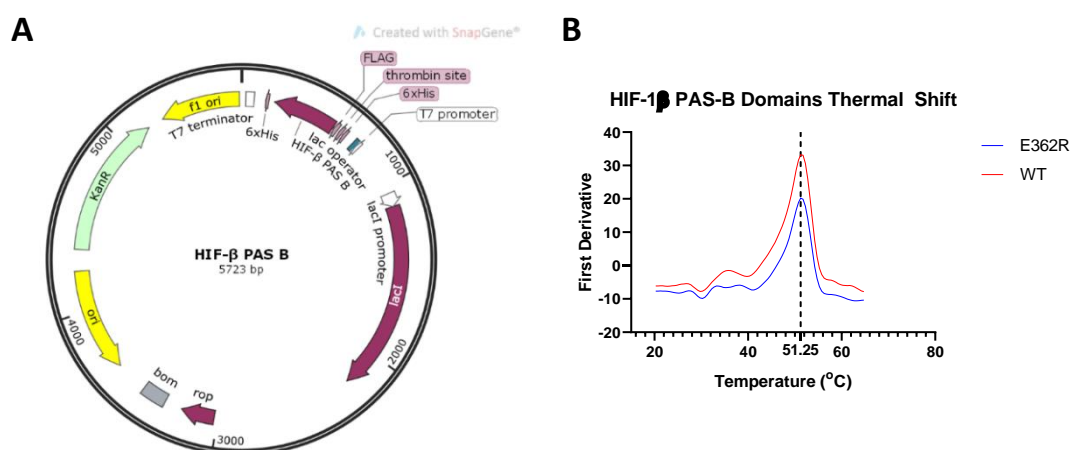


Figure 56. A) Plasmid map of the construct used for expression of HIF-1 β PAS-B domains. B) First derivative of the thermal shift data for the WT and mutant HIF-1 β PAS-B domains.

With α and β proteins expressed, the first experiment run was a positive control to demonstrate that the proteins could form a dimer in the ELISA assay. This was first attempted with the wild type PAS-B domains whereby the HIF-1 α PAS-B was bound to the plate and then a 12-point serial dilution of HIF-1 β was added to the wells. The relevant antibody incubations were then performed and 1-Step TMB Ultra Solution (100 μ L) was added to each of the wells. H₂SO₄ (2M, 100 μ L) was then added to each of the wells and the absorbance at 450 nm was plotted. This yielded the partial binding curve shown in Figure 57A. The relatively weak interaction between the domains led to an incomplete binding curve being obtained, with an approximate K_d of 50 μ M. This is a stronger interaction than the reported value, which is stated as being similar to that

of the wild-type 2α PAS-B domains ($K_d = 120 \mu\text{M}$).¹⁶ The next experiment utilised the mutant PAS-B domains, with the HIF-1 α 4x mutant bound to the plate and subsequent incubation with a serial dilution of HIF-1 β E362R protein. A full binding curve for this interaction was obtained with a K_d of $822 \pm 95 \text{ nM}$ (Figure 57B), representing an approximate 60-fold increase in affinity over the wild type proteins. This value is 10-fold weaker than the previously reported K_d ($83 \pm 20 \text{ nM}$) of the two domains as determined by ITC.¹⁶ This could be a result of the HIF-1 α domain being bound to the plate, and therefore blocking some of the binding interactions between the two proteins. However, as a full binding curve was obtained, an experiment was set up to determine whether the addition of CLLFVY would prevent the formation of the heterodimer.

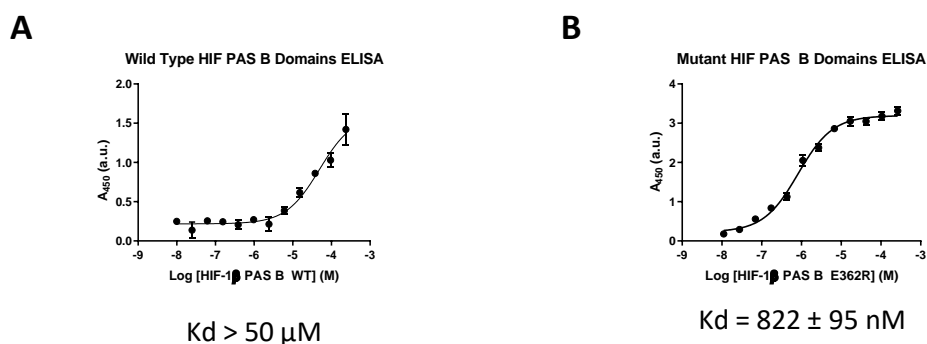


Figure 57. Assessment of dimer formation of HIF-1 α PAS-B and HIF-1 β PAS-B by ELISA. A) ELISA binding curve of wild type HIF-1 α PAS-B and wild type HIF-1 β PAS-B dimerisation. B) ELISA binding curve of the four times mutant HIF-1 α PAS-B and HIF-1 β PAS-B E362R dimerisation. Error bars represent standard deviation ($n=3$). Binding constants were determined by curve fitting in GraphPad Prism 8.

To assess the ability of CLLFVY to inhibit dimerisation, a serial dilution of CLLFVY ranging from $500 \mu\text{M}$ to 15 nM was added to plate bound HIF-1 α PAS-B in competition with the HIF-1 β PAS-B domain. The incubation with the antibodies and subsequent wash steps were carried out, and the extent of TMB oxidation was measured via the absorbance at 450 nm and plotted against the concentration of CLLFVY. The resulting dose-response is shown in Figure 58.

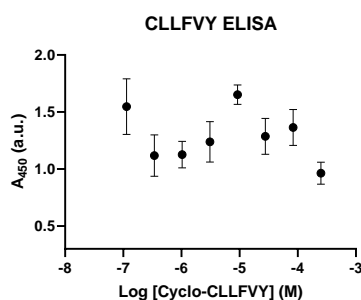


Figure 58. ELISA binding curve for CLLFVY inhibition of HIF-PAS-B dimerisation. Error bars represent standard deviation ($n=3$).

The addition of CLLFVY at even the highest concentration (500 μM) gave no change in the A_{450} , indicating that it is not disrupting dimer formation. There are a number of potential explanations for this result, including that the activity of CLLFVY is too low to be detected at this concentration range, or that the mutant heterodimer is not a valid model of HIF function. To address the first point, that the activity is too low to be detected, the optimised compounds CK-NL-F(CF₃)IF(Br) and CE-NL-F(CF₃)IF(Br) were also analysed by ELISA. These optimised peptides were found to bind the HIF-1 α PAS-B domain with affinities 17 and 30-fold higher than CLLFVY, respectively, by MST. If the activity of these compounds has also been improved, then it may be possible to measure their activity with the ELISA assay. The ELISA was repeated with both of the optimised compounds and the results are shown in Figure 59.

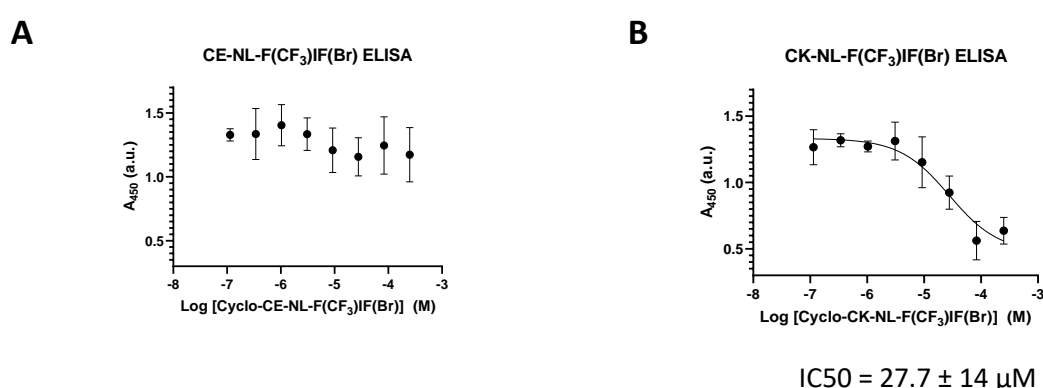


Figure 59. A) ELISA dose-response curve for CE-NL-F(CF₃)IF(Br). B) ELISA dose-response curve for CK-NL-F(CF₃)IF(Br). Error bars represent standard deviation (n=3). Binding constants were determined by curve fitting in GraphPad Prism 8.

Surprisingly, the results showed that CE-NL-F(CF₃)IF(Br) had no effect on the dimerisation of the HIF domains, whereas CK-NL-F(CF₃)IF(Br) did show inhibition with an IC₅₀ of 27.7 ± 14 μM . This indicates that the inhibition of the HIF PAS-B dimerisation is highly dependent on the charge present in L2 position of CLLFVY. The charged residues were previously thought to be orientated away from the HIF-1 α PAS-B domain given their similar affinity for the domain via MST. It is possible that in the heterodimer this position of the cyclic peptide can interact with the HIF-1 β PAS-B domain and therefore influence the peptides inhibitory activity. The IC₅₀ of CK-NL-F(CF₃)IF(Br) (27.7 ± 14 μM) is 10-fold worse than its affinity as determined by MST. This could be due to the plate bound nature of the ELISA interfering with the peptides binding, but it is also worth noting that the ELISA assay is measuring the ability of the compound to disrupt dimerisation, whereas the MST assay was only measuring the peptides ability to bind to the HIF-1 α PAS-B domain. If CLLFVY also has an activity 10-fold weaker than its affinity by MST then this would explain why a binding curve was not obtained in this assay, as inhibition would not occur until a higher concentration range than was used here. This result demonstrates that the SAR

studies carried out on CLLFVY have improved the peptides ability to inhibit HIF dimerisation as well as the peptides affinity for binding to HIF.

Another possibility for the discrepancy between affinity and activity could be that the chosen model is not biologically relevant. This could arise from the mutations made to the protein domains that, whilst enabling the formation of stable dimers, would not be present in the wild type proteins. At the time of writing no compounds have been identified using the HIF-1 PAS-B mutant heterodimer that have progressed into *in vivo* studies, and so its biological relevance is undetermined. However, comparisons can be drawn between the HIF-1 PAS-B model used here and the analogous HIF-2 PAS-B model.

First described by Scheuermann *et al*, it is the HIF-2 PAS-B heterodimer that was used to identify the mutations that can swap a crucial salt bridge, leading to a lower affinity interaction between the α and β domains.^{14,115} This model was screened with a small molecule library, which identified the compound THS-044 as capable of disrupting the dimerisation of these domains.¹¹⁵ This compound has since been developed into the compound PT2977 that is currently undergoing phase two clinical trials for the treatment of clear cell renal cell carcinoma.¹²⁰ This clearly demonstrates the potential of the HIF PAS-B heterodimer model as a useful tool in the development of HIF inhibitors. Whilst the PAS-B heterodimer model was used to identify the hit compound THS-044, further optimisation was carried out using an alternative protein model consisting of the bHLH, PAS-A and PAS-B domains of both HIF-1 α and HIF-1 β . This model was published in 2015 by Wu *et al*, and its structure was solved via x-ray crystallography of the dimer bound to DNA.¹⁸ The structures of both the HIF-1 and HIF-2 isoforms were solved, and in the case of HIF-2, two ligand bound structures were also described. Comparing this model with that of the PAS-B heterodimer reveals that the anti-parallel beta sheet interaction thought to be critical to HIF dimerisation in the PAS-B model is no longer present in the three-domain structure, as highlighted in Figure 60. It is therefore possible that the mutations introduced to induce dimerisation of the PAS-B domains in the old model have altered the way in which these domains would interact with respect to the wild type protein. Therefore, the use of the three-domain HIF model for *in vitro* assessment of CLLFVY and derivatives thereof was investigated.

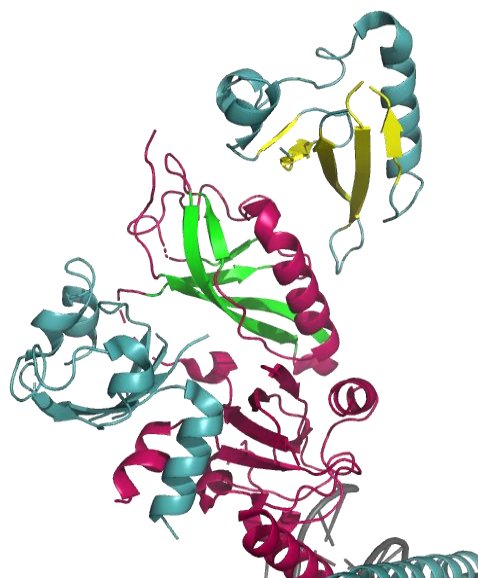


Figure 60. Structure of the HIF-1 three-domain heterodimer. Consisting of HIF-1 α (magenta) and HIF-1 β (cyan) with the orientation of the PAS-B beta sheets of HIF-1 α (green) and HIF-1 β (yellow) highlighted.

3.2 Expression of Three-Domain HIF Heterodimer

Before this model could be assessed, the three-domain protein first had to be obtained. This was achieved by co-expressing HIF-1 α bHLH-PAS-A-PAS-B (13-357) and HIF-1 β bHLH-PAS-A-PAS-B (82-464) in *E. coli* BL21 (DE3) Rosetta cells containing a pET-28a vector with the HIF-1 α gene and a pET-21d vector with the HIF-1 β gene. These plasmids are shown in Figure 61A and B.

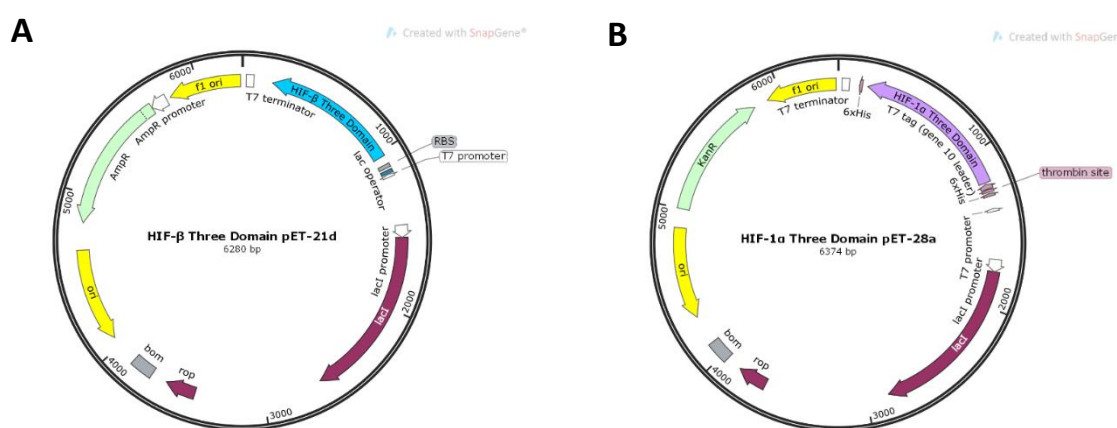


Figure 61. A) Plasmid map of the pET21-d vector containing the HIF-1 β three-domain gene. B) Plasmid map of the pET28-a vector containing the HIF-1 α three-domain gene.

These plasmids were chosen for their orthogonal antibiotic resistances, kanamycin and ampicillin respectively, which enables the retention of both plasmids in the cell. Once transformed, the cells were grown and incubated as for the HIF-1 α PAS-B domain. The cells were then lysed and the protein isolated from the cell lysate via a nickel affinity column. Notably only the HIF-1 α protein

contained a poly-histidine tag, and therefore any unbound HIF-1 β would be removed at this step. It was also necessary to perform a wash step with a salt solution (NaCl 1M) in order to elute prokaryotic DNA that had bound to the protein (Figure 62A). The protein was further purified via an S200 SEC which would also remove any monomeric protein. The fractions F1, F2 and F3 (Figure 62B) were analysed by SDS-PAGE (Figure 62C) and those containing protein at the expected masses (41.9 and 44.8 kDa for HIF-1 α and HIF-1 β respectively) were combined, with the yield of expression found to be 2.8 mg/ L. The protein was then concentrated and aliquots were frozen in liquid nitrogen before being stored at -80 °C.

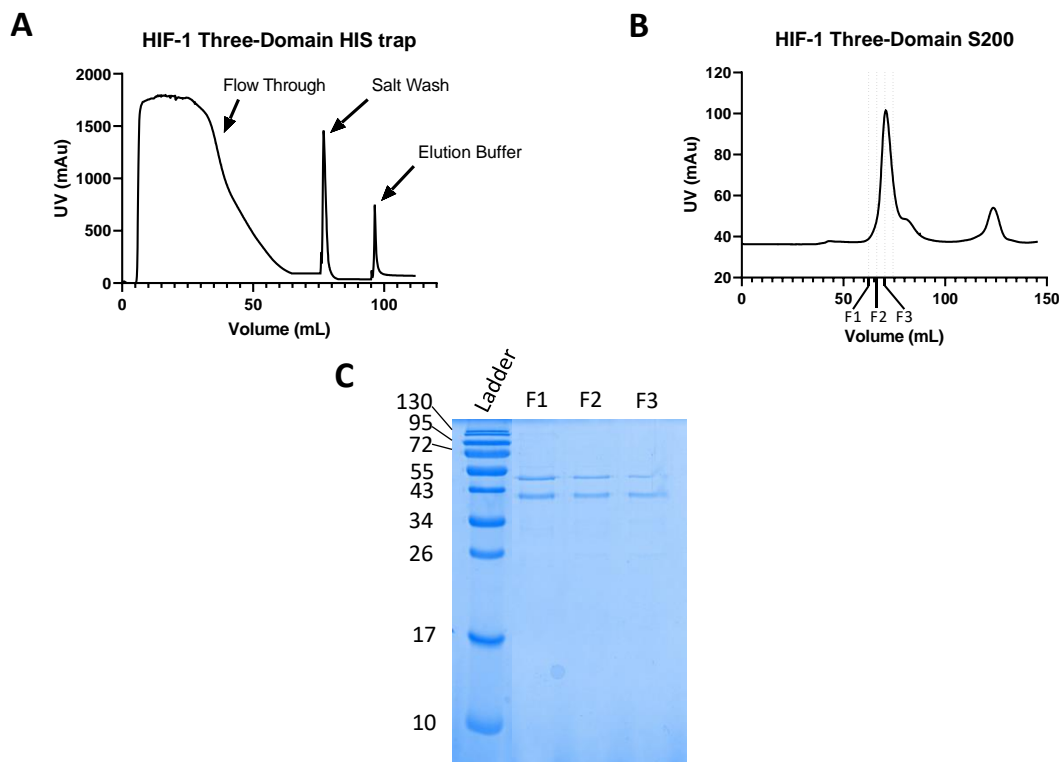


Figure 62. Analysis of the expression of the HIF three-domain protein. A) UV trace from the HIS trap purification of the HIF three-domain protein. B) UV trace from the S200 purification of the HIF three-domain protein. C) SDS-PAGE analysis of fractions obtained from the S200 purification.

A sample of the purified protein was then labelled for MST as previously described. To test whether the peptides would bind to this new HIF model an MST experiment was carried out with the parent peptide CLLFVY. A serial dilution ranging from 500 μ M to 15 nM was assessed using the same conditions as described for the HIF-1 α PAS-B MST experiments (section 2.4.2). The results of this experiment are shown in Figure 63, and show that there was no binding affinity between CLLFVY and the HIF three-domain model, even at 500 μ M.

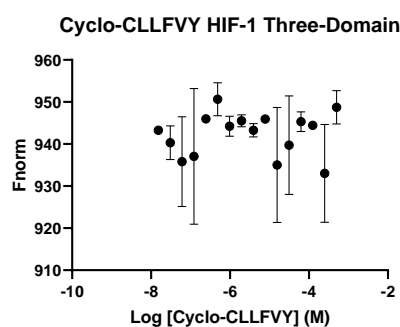


Figure 63. MST binding curve of CLLFVY against HIF three-domain. Error bars represent standard deviation (n=3).

The absence of binding affinity between CLLFVY and HIF three-domain could be a result of the new orientation of the HIF-1 α PAS-B domain within this model. However, it was reasoned that even if there was a small loss of activity between the two models then binding might not be seen for CLLFVY, which had a relatively weak affinity for the HIF-1 α PAS-B model of 36 μ M (Figure 26). Therefore the two optimised peptides, CE-NL-F(CF₃)IF(Br) and CK-NL-F(CF₃)IF(Br), were also tested. During the capillary scan, that is performed before the MST measurements are made, there was a dose-dependent change in the initial fluorescence for CE-NL-F(CF₃)IF(Br) as shown in Figure 64.

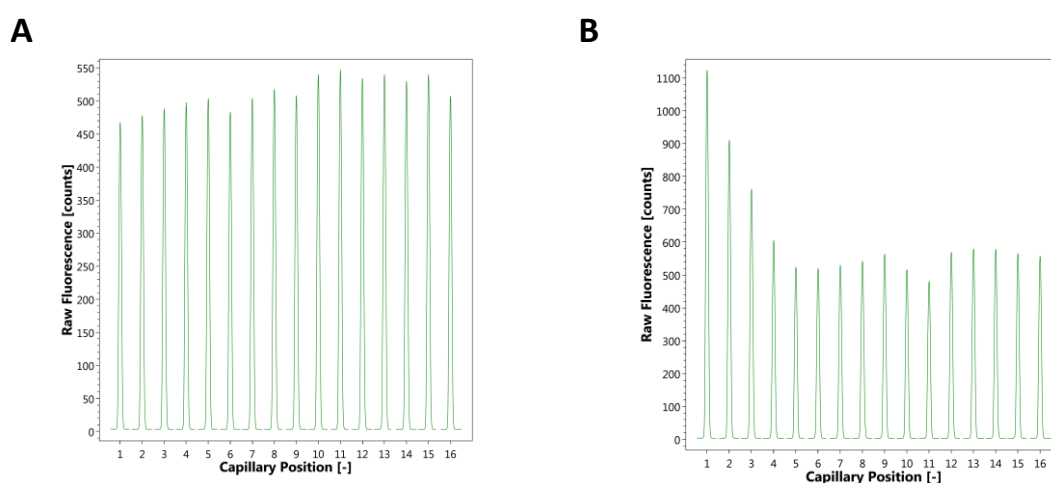


Figure 64. Capillary scans of initial fluorescence for CK-NL-F(CF₃)IF(Br) and CE-NL-F(CF₃)IF(Br) where 1 is the highest concentration (500 μ M) and 16 is the lowest (15 nM). A) Capillary scan of CK-NL-F(CF₃)IF(Br) showing that fluorescence does not deviate by more than $\pm 10\%$. B) Capillary scan of CE-NL-F(CF₃)IF(Br) showing a dose-dependent increase in fluorescence at high peptide concentrations.

The observed effect of the peptide on the fluorescence output of the dye can be due to a number of reasons. One explanation is that the peptides are binding to the protein near the site of labelling. This then causes a change in the dye's environment, and therefore affects its fluorescent

output. This effect is specific to ligand binding and leads to a change in the T-jump region of the MST that can be used to determine the affinity of ligand binding. Alternatively, a dose-dependent change in fluorescence output can also be non-specific, caused by ligand promoted aggregation of the protein. In order to determine whether the observed dose-dependent change is specific or not, an SDS-denaturation test (SD-test) was carried out. The four lowest and four highest concentrations of the peptides CE-NL-F(CF₃)IF(Br) and CK-NL-F(CF₃)IF(Br) were incubated with the HIF three-domain protein for 30 minutes at room temperature. The samples were then centrifuged at 15,000 rpm for 15 minutes. The supernatant (10 µL) was then mixed 1:1 with the SD mix (4 % SDS, 40 mM DTT) and incubated at 95 °C for 5 minutes. Centrifugation of the sample removes any aggregates that have formed during the incubation, and the SD mix followed by heating to 95 °C denatures the protein thus preventing any specific ligand binding from occurring. Therefore, if a ligand is causing a specific ligand binding effect, the SD-test should show equal fluorescence at all concentrations as the ligand can no longer bind and influence the dye. However, if the observed dose-dependent change in fluorescence were due to aggregation a dose-dependent drop in fluorescence would be observed in the SD-test. After treatment with the SD mix, a capillary scan was performed and the fluorescence of each concentration was plotted (Figure 65).

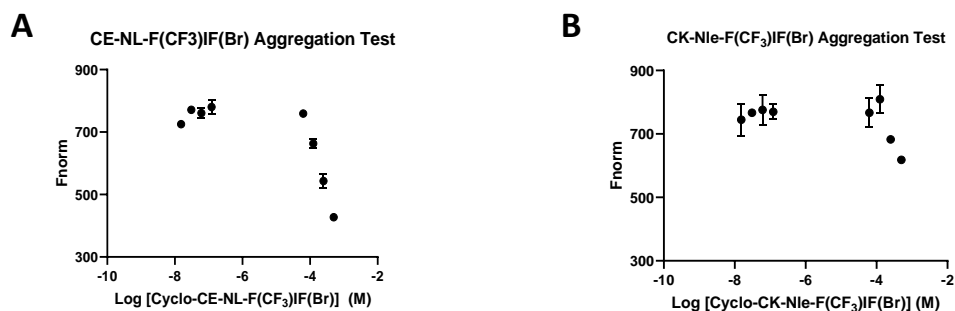


Figure 65. SD-Test of CK-NL-F(CF₃)IF(Br) and CE-NL-F(CF₃)IF(Br). A) CE-NL-F(CF₃)IF(Br) SD-Test, showing a loss of fluorescence at higher concentrations. B) CK-NL-F(CF₃)IF(Br) SD-Test, showing a small loss of fluorescence at higher concentrations. Error bars represent standard deviation (n=2).

Both peptides showed some aggregation at higher concentrations, however CE-NL-F(CF₃)IF(Br) was more prone to aggregation, with significant deviations away from the fluorescence observed at low peptide concentrations for the top three concentrations (Figure 65A). CK-NL-F(CF₃)IF(Br) also showed a loss of fluorescence at the highest concentration (500 µM) and a small deviation was also observed at 250 µM (Figure 65B). A loss of fluorescence during the SD-test is most likely attributable to the removal of aggregated protein during the centrifugation step. These results indicate that both peptides are causing aggregation of the labelled HIF-1 three-domain protein, but only at the highest concentrations used. As such, the MSTs of the peptides were carried out

with the concentrations showing aggregation excluded from the analyses to yield the dose-response curves in Figure 66. CE-NL-F(CF₃)IF(Br) was still found to deviate significantly from a sigmoidal binding curve at higher concentrations (Figure 66A). This is reminiscent of the superposition of two curves that was observed in the SAR studies of the L derivatives (section 2.13). Whilst this curve could be the result of non-specific binding, it is also possible that the second curve is caused by the dissociation of the HIF dimer. The first curve could be showing the compound binding to the complex, and the second the change in thermophoresis brought about by dissociation. Further experiments would be needed to verify this hypothesis. This assay could therefore not provide a reliable binding affinity for the CE-NL-F(CF₃)IF(Br) peptide. The CK-NL-F(CF₃)IF(Br) peptide did yield a full binding curve with a calculated affinity of $3.2 \pm 0.6 \mu\text{M}$ (Figure 66B). This is in very good agreement with the value obtained when using the PAS-B model ($2.08 \pm 0.4 \mu\text{M}$, Figure 50). These results validate that the peptides do bind to the HIF three-domain model. In addition, this result also validates that the use of the PAS-B alone model to improve compound affinity does translate to improved binding of a larger HIF model. The loss of activity moving from the PAS-B alone to three-domain model offers a potential explanation for the lack of binding seen for CLLFVY. As CLLFVY had a weaker affinity for the HIF-1 α PAS-B by MST, loss of activity in the three-domain MST could lead to the binding curve shifting beyond the solubility limit of the compound and so not be visible at the concentration ranges used.

The lack of a second binding curve for CK-NL-F(CF₃)IF(Br) could indicate it is not causing dissociation of the HIF dimer. The potential for secondary binding curves upon inhibition of dimerisation that cannot be deconvoluted from peptide binding indicates that this MST assay cannot be used for reliable assessment of the peptides. As it was shown that CK-NL-F(CF₃)IF(Br) will bind to this model, a different biophysical assay utilising this protein was investigated that would allow for the assessment of the peptides ability to inhibit HIF activity.

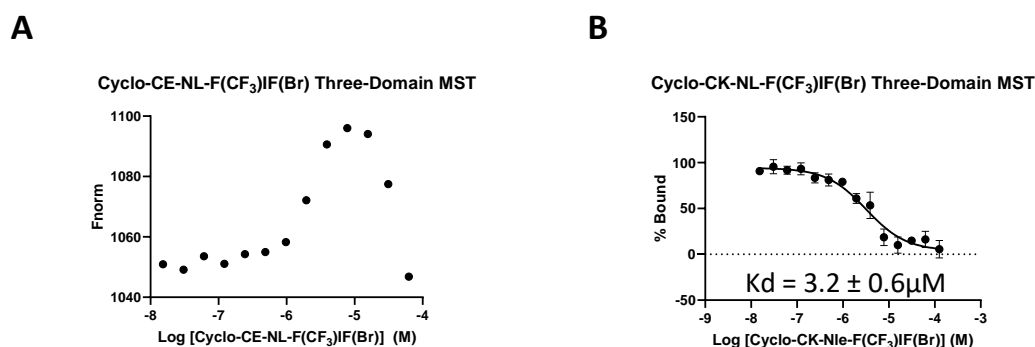


Figure 66. A) MST dose-response curve obtained for CE-NL-F(CF₃)IF(Br) with the HIF-1 three-domain protein. B) MST dose-response curve obtained for CK-NL-F(CF₃)IF(Br) with the HIF-1 three-domain protein. Error bars represent standard deviation (n=3). Binding constants were determined by curve fitting in GraphPad Prism 8.

3.3 Fluorescence Polarisation

Recent work by Fraydoon *et al.*, published in 2019, describes the use of the three-domain model for a fluorescence polarisation assay.¹¹⁹ This assay would be capable of determining a compounds ability to prevent HIF from binding DNA. The assays compatibility for use with CLLFVY and derivatives was therefore investigated. Fluorescence polarisation (FP) is a biophysical technique for monitoring changes in a molecules apparent molecular weight by measuring the change in polarisation of a fluorescent probes emitted light. This makes FP very useful for measuring binding events, which invariably lead to a change in molecular weight. The ability of FP to detect changes in molecular mass has been exploited in the investigation of many interactions, including protein-protein, protein-peptide, protein-nucleic acid and protein- small molecule.²²²⁻²²⁴ The basic principle is outlined in Figure 67 whereby a fluorescent sample is irradiated with a plane polarised light source. The degree of depolarisation of the emitted light is then dependent on the rotational diffusion of the fluorophore, which is in turn dependent on the molecular weight of the fluorescent molecule. An increase in the molecules apparent mass will result in a slower rotation speed, and therefore a higher degree of emitted plane polarised light.

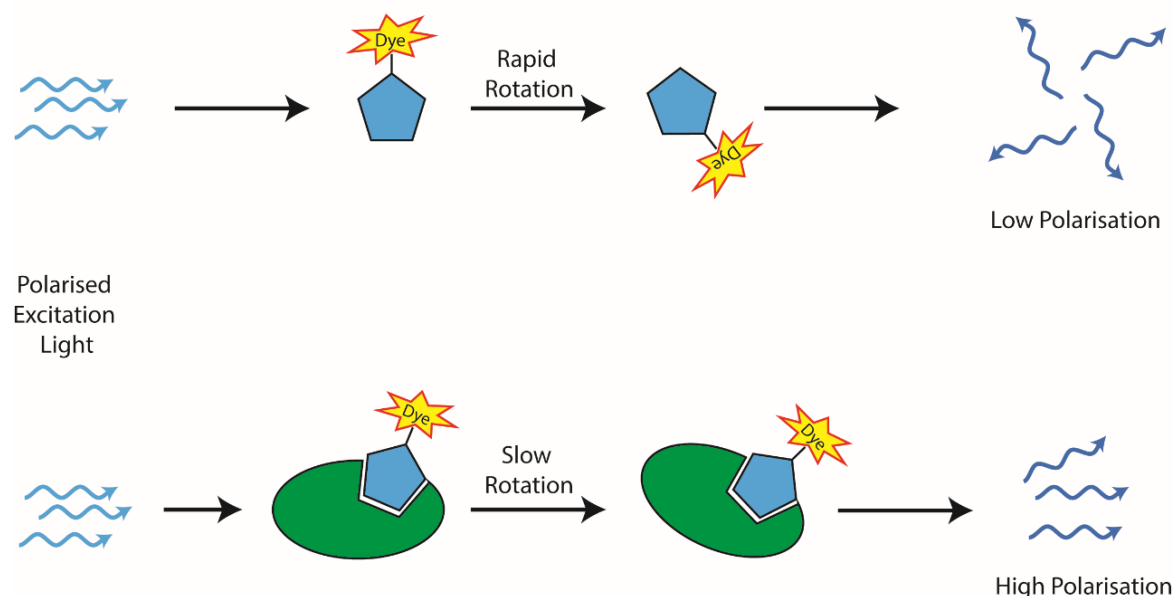


Figure 67. Outline of the principles of fluorescence polarisation. A) Fast rotation of small fluorescent molecules leads to rotational diffusion of polarised light. B) Upon a binding event the mass of the fluorophore is increased and the degree of rotational diffusion reduced, leading to emission of light that retains a high degree of polarisation.

The ratio between emitted polarised and unpolarised light is then measured and the FP signal calculated according to Equation 5, where I_{\parallel} is the intensity of light emitted parallel to the excitation light and I_{\perp} is the intensity of light emitted perpendicular to the excitation light. FP is then calculated as the difference between the perpendicular and parallel channels normalised by the total intensity of the emitted light.

$$FP = \frac{I_{\parallel} - I_{\perp}}{I_{\parallel} + I_{\perp}}$$

Equation 5. Calculation of fluorescence polarisation. Where I_{\parallel} is the intensity of light emitted parallel to the plane of polarisation and I_{\perp} is the intensity of light emitted perpendicular to the plane of polarisation.

3.3.1 Development of a HIF FP Assay

In order for the FP assay to work, a fluorescent probe is needed. To maximise the signal obtained in the FP assay, it is best to have a probe that has a lower relative molecular weight than its binding partner, as this will lead to a greater change in mass upon binding. Therefore, the HRE was chosen to be the fluorescent probe and was tagged with the fluorophore 6-carboxyfluorescein (6-FAM). This enables the direct measurement of DNA binding, with an increase in FP signal as the DNA is bound and its rotation slowed.¹¹⁹ The three-domain heterodimeric HIF-1 was obtained as previously described (section 3.2). The 6-FAM labelled forward DNA sequence 5'-GGCTGCGTACGTGCGGGTCGT-3' was annealed with the reverse strand 5'-ACGACCCGCACGTACGCAGCC-3' to yield the labelled duplex FAM-HRE. A serial dilution of protein was then added to the FAM-HRE (2 nM) and the fluorescence polarisation was measured on a plate reader to yield the dose-response curve shown in Figure 68A. The HIF-1 three domain protein was found to bind to the FAM-HRE with a K_d of 59.1 ± 4.4 nM, which is in good agreement with the value found by Fraydoon *et al.*, for HIF-2 binding to the FAM-HRE (65 nM).¹¹⁹ The next experiment was to determine the specificity of the assay by competing off the FAM-HRE with an unlabelled HRE. The unlabelled HRE was prepared by annealing the same reverse strand with an unlabelled forward strand, before a serial dilution of the unlabelled HRE ranging from 1 μ M to 0.4 nM was then prepared and added to the HIF-1 heterodimer (200 nM). FAM-HRE (2 nM) was then added and the FP signal again measured to yield the curve in Figure 68B. The competition assay yielded an IC_{50} for the unlabelled HRE of 15.7 ± 3.0 nM, showing that the protein preferentially

binds the unlabelled DNA. This result validates the assay for the detection of HIF-1-HRE binding activity.

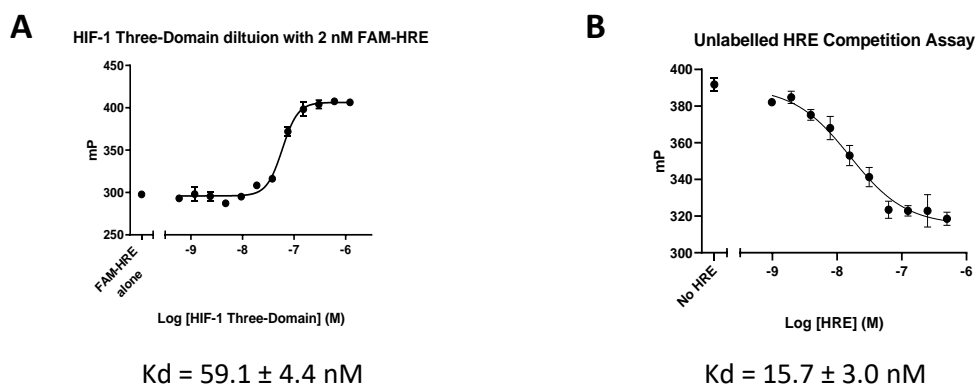


Figure 68. A) Dose-response curve of HIF three-domain protein with FAM-HRE. B) Dose-response curve from the competition assay between FAM-HRE and HRE. Error bars represent standard deviation ($n=3$). Binding constants were determined by curve fitting in GraphPad Prism 8.

With the assay in hand, the ability of CLLFVY to inhibit HIF-DNA binding was probed. This was achieved by serially diluting CLLFVY in DMSO before adding each concentration to the HIF-1 three-domain protein for a final concentration range of CLLFVY of 500 μ M to 15 nM. The FAM-HRE probe (2 nM) was then added and the fluorescence polarisation measured to yield the dose-response curve shown in Figure 69A. A dose-dependent inhibition of DNA binding was observed, and whilst the DNA was not completely disassociated at the highest concentration, a binding curve was fitted and was found to give an IC50 value of 76.5 ± 8.0 μ M. This is somewhat surprising given the lack of binding seen in the three-domain HIF MST assay (Figure 63). This could be caused by the protein labelling for MST; either the position of the dye could be blocking interactions of CLLFVY with the protein, or alternatively labelling could also cause a change in conformation of the protein, subsequently leading to the observed loss of binding in the MST assay. The IC50 value obtained in the FP assay is 2.1-fold worse than the affinity measured using the HIF-1 α PAS-B model MST assay (Figure 26). Whilst this is a significant difference in affinity, this result demonstrates that a ligand that binds to the HIF-1 α PAS-B model is capable of binding to the larger three-domain HIF model as well. This is also a crucial result as it shows that CLLFVY is capable of preventing the dimer from binding to the FAM-HRE. As a negative control, and to assess how the SAR thus far developed using the MST assay translates to activity, we next assessed the compound ALLFVY. According to the alanine scanning of CLLFVY, the cysteine was critical for binding (Figure 28). If ALLFVY does not bind, then no FP activity should be observed. This hypothesis was confirmed as a serial dilution of ALLFVY over the same concentration range of that used for CLLFVY showed no detectable activity, as shown in Figure 69B. This result

demonstrates the SAR observed using the HIF-1 α PAS-B domain MST assay does correlate with HIF inhibition.

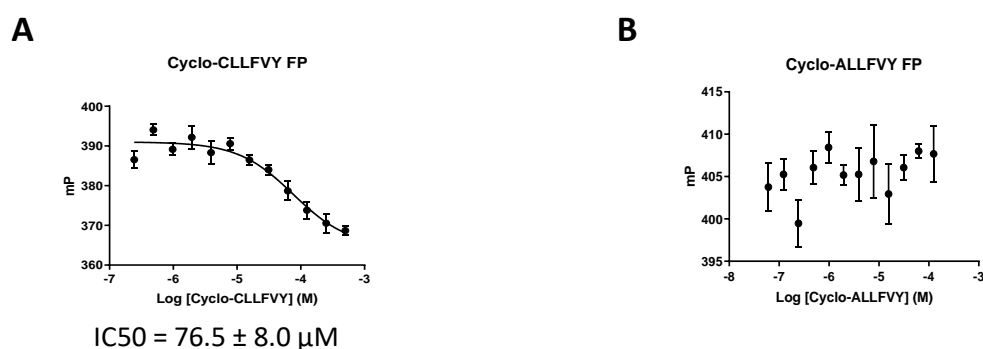


Figure 69. A) Dose-response curve of CLLFVY in the FAM-HRE FP assay. Dose-response curve of ALLFVY in the FAM-HRE FP assay. Error bars represent standard deviation ($n=3$). Binding constants were determined by curve fitting in GraphPad Prism 8.

As a loss of affinity by MST led to a loss of activity in the FP assay, it was predicted that the improvements made during the derivatisation of CLLFVY by MST lead to peptides that display better activity in the FP assay. This was first probed using the compounds CK-NL-F(CF₃)IF(Br) and CE-NL-F(CF₃)IF(Br). A serial dilution of each compound was incubated with the HIF three-domain heterodimer, the FAM-HRE probe (2 nM) was added and the FP measured as before to yield the dose-response curves shown in Figure 70.

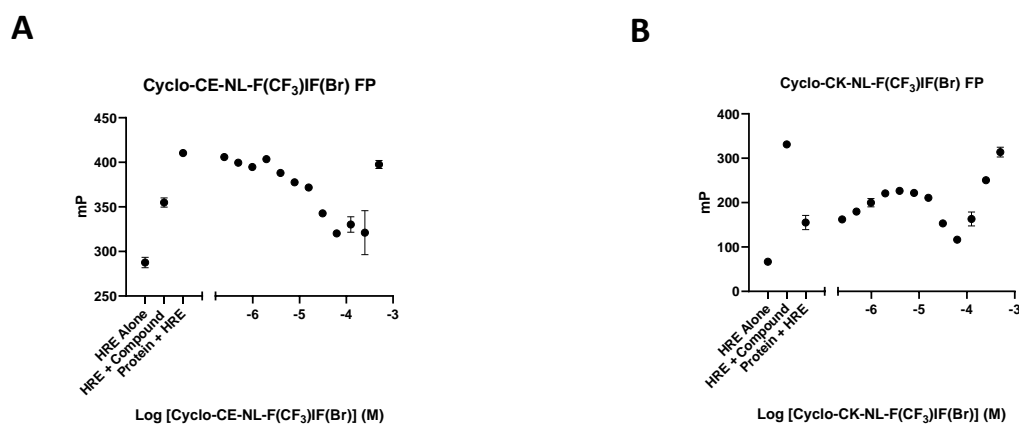


Figure 70. A) Dose-response curve for CE-NL-F(CF₃)IF(Br) in the FAM-HRE FP assay. B) Dose-response curve for CK-NL-F(CF₃)IF(Br) in the FAM-HRE FP assay. Error bars represent standard deviation ($n=3$).

Unfortunately, full binding curves could not be obtained for these two peptides as they were both found to bind to the FAM-HRE itself. This is evidenced by a high mP value for the control that contained the peptide at the highest concentration used (500 μ M) and the FAM-HRE probe

without protein. This was especially evident for CK-NL-F(CF₃)IF(Br) (Figure 70B) which showed a small increase in mP, the value then fell before rising again sharply. This is most likely due to the favourable interaction of the positively charged lysine side chain with the negative charges of the DNA phosphate groups. With the peptide binding to the HRE-HIF complex causing the initial rise in mP, the peptide then reaches a concentration where it starts to inhibit HIF HRE binding and so a drop in mP is observed. As the FAM-HRE is released into solution it is sequestered by more peptide causing the sharp rise at higher concentrations, with the top value matching that of the FAM-HRE + CK-NL-F(CF₃)IF(Br) control, strongly suggesting that the peptide is binding the FAM-HRE. To try to mitigate this effect, the ionic strength of the buffer was investigated. So far, all FP experiments have been carried out in a low salt buffer (20 mM) as was used by Fraydoon *et al.*¹⁸ To investigate whether DNA binding activity is still seen at higher salt concentrations a gradient was run from 4 to 400 mM NaCl, and the mP measured and plotted in Figure 71.

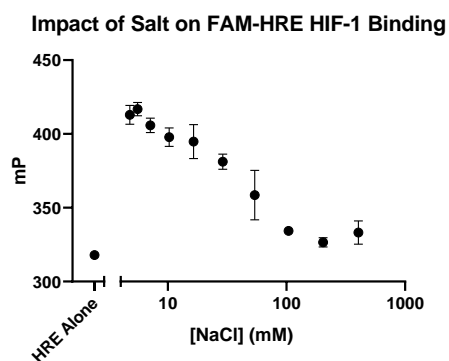


Figure 71. Effect of NaCl concentration on HIF-1 FAM-HRE binding as determined by fluorescence polarisation. Error bars represent standard deviation ($n = 3$).

Increasing concentrations of NaCl negatively affected the HIF-1 proteins ability to bind DNA, with a concentration of 200 mM completely abolishing binding. Even the relatively low concentrations used for previous assays were seen to have a deleterious effect on the measured mP. A small assay window may be achievable at higher concentrations up to approximately 50 mM. In order to assess the concentration of NaCl that would be needed to prevent CK-NL-F(CF₃)IF(Br) binding, a separate experiment was conducted whereby CK-NL-F(CF₃)IF(Br) (100 μ M) and FAM-HRE (2 nM) were added to buffers containing different concentrations of NaCl and the mP was measured. The same experiment was also carried out using CE-NL-F(CF₃)IF(Br) with the prediction that increasing NaCl should not affect its binding to the FAM-HRE as it is lacking the side chain positive charge. The results are shown in Figure 72. For CK-NL-F(CF₃)IF(Br) the interaction with the FAM-HRE is prevented at higher concentrations of NaCl, with a concentration of 200 mM inhibiting any detectable DNA binding (Figure 72A). This is considerably higher than the concentration needed to prevent HIF binding to DNA, and so increasing NaCl concentration is not a viable method for the reduction of non-specific DNA binding. For the sequence CE-NL-F(CF₃)IF(Br), increasing the

concentration of NaCl had little effect on its affinity for binding the FAM-HRE up to a concentration of 100 mM, with a slight increase observed from 100- 400 mM (Figure 72B).

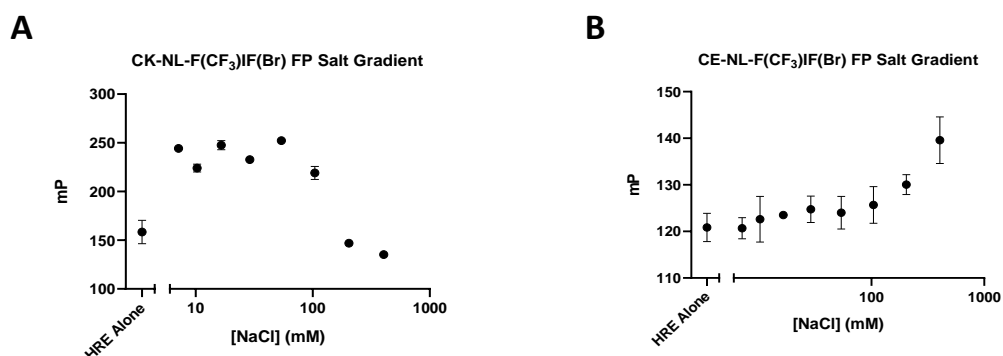


Figure 72. A) Effect of increasing NaCl concentration on CK-NL-F(CF₃)IF(Br) FAM-HRE binding. B) Effect of increasing NaCl concentration on CE-NL-F(CF₃)IF(Br) FAM-HRE binding.

Given that both peptides were found to bind to the FAM-HRE probe, the best peptide from the V derivative series (CLLF(CF₃)IF(Br)) was also investigated, with the assay being set up as before and the resulting dose-response curve shown in Figure 73. Gratifyingly, this derivative did yield a full binding curve. It was also shown that this peptide sequence did not bind to the HRE-FAM probe as evidenced by the low mP value obtained for the control containing FAM-HRE and CLLF(CF₃)IF(Br) (500 μM) without protein. This enabled the determination of an IC₅₀ value of $11.76 \pm 1.1 \mu\text{M}$, which represents a 6.5-fold improvement in activity over CLLFVY. This demonstrates that the approach taken of derivatisation guided by the compounds affinity for the HIF-1α PAS-B domain as determined by MST has yielded an improvement in the ability of the peptide to inhibit HIF DNA binding activity. This value is 25-fold higher than the affinity of CLLF(CF₃)IF(Br) as determined by MST ($473 \pm 246 \text{ nM}$, Figure 46). This discrepancy could arise from the difference in the protein models used, with the peptide having been optimised to a surface that is blocked in the three-domain model or interacting with amino acids whose positions are not consistent between the two models. Alternatively, the differences could arise from differences in the assays themselves, with the MST measuring the K_d of the compound, whereas FP determines an IC₅₀ value.

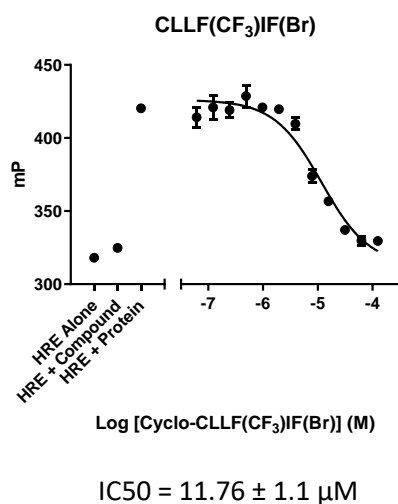


Figure 73. Dose-response curve for CLLF(CF₃)IF(Br) in the FAM-HRE FP assay. Error bars represent standard deviation (n=3). Binding constants were determined by curve fitting in GraphPad Prism 8.

3.4 Chapter 3 Summary

An assay capable of determining the activity of the developed peptide leads was pursued. Initially, efforts were made to develop an ELISA assay for the detection of inhibition of HIF heterodimer formation. A new HIF model was utilised with the expression of the 4x mutant HIF-1 α PAS-B domain, as well as two HIF-1 β PAS-B constructs. CLLFVY was shown to bind to the mutant HIF-1 α PAS-B domain with an affinity comparable to the wild-type protein ($49.2 \pm 3.7 \mu\text{M}$). The ELISA assay was shown to be functional with K_d values determined for both wild-type and mutant heterodimer formation ($50 \mu\text{M}$ and $822 \pm 95 \text{ nM}$ respectively). CLLFVY, CE-NL-F(CF₃)IF(Br) and CK-NL-F(CF₃)IF(Br) were then tested for inhibitory activity in the mutant HIF ELISA assay, with only CK-NL-F(CF₃)IF(Br) demonstrating any activity, with an IC₅₀ of $27.7 \pm 1.4 \mu\text{M}$. Future work will aim to elucidate the mechanism of charge dependence when disrupting the HIF PAS-B heterodimer. As only one peptide showed activity, a second assay was developed to corroborate this result.

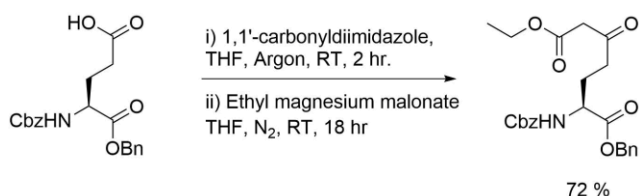
A larger HIF model, consisting of the bHLH-PAS A-PAS-B domains of both HIF-1 α and HIF-1 β , was expressed and the affinity of the peptides for this new protein model was probed using MST. This showed that CLLFVY had no detectable affinity for the labelled three-domain protein, and so the derivatives CE-NL-F(CF₃)IF(Br) and CK-NL-F(CF₃)IF(Br) were analysed. Both peptides were found to cause aggregation at higher concentrations, however excluding these points enabled dose-response curves to be plotted. CK-NL-F(CF₃)IF(Br) was found to bind with a K_d of $3.2 \pm 0.6 \mu\text{M}$ which is in good agreement with the HIF-1 α PAS-B MST value of $2.08 \pm 0.4 \mu\text{M}$. A value could not be determined for CE-NL-F(CF₃)IF(Br) due to a secondary event occurring.

The use of an FP assay based on work by Fraydoon *et al* was then developed to allow the assessment of the peptides ability to prevent HIF-DNA binding.¹¹⁹ The assay showed that the HIF three-domain was functionally active, binding to the fluorescently labelled FAM-HRE with a Kd of 76.5 ± 3.3 nM. The ability of the peptides to prevent FAM-HRE binding was then assessed, and CLLFVY was found to inhibit HIF FAM-HRE binding with an IC50 of 76.5 ± 8.0 μ M. Unfortunately, IC50 values could not be determined for CE-NL-F(CF₃)IF(Br) and CK-NL-F(CF₃)IF(Br) as they were found to bind to the FAM-HRE probe. Therefore the earlier derivative CLLF(CF₃)IF(Br) was also assessed and found to possess increased inhibitory activity over CLLFVY with an IC50 of 11.76 ± 1.1 μ M. This is a critical result as it demonstrates that the improvements in affinity during the CLLFVY SAR study translates to improved inhibitory activity.

Overall, three new *in vitro* assays have been developed that have demonstrated the SAR study of CLLFVY has successfully increased the peptides ability to inhibit HIF activity.

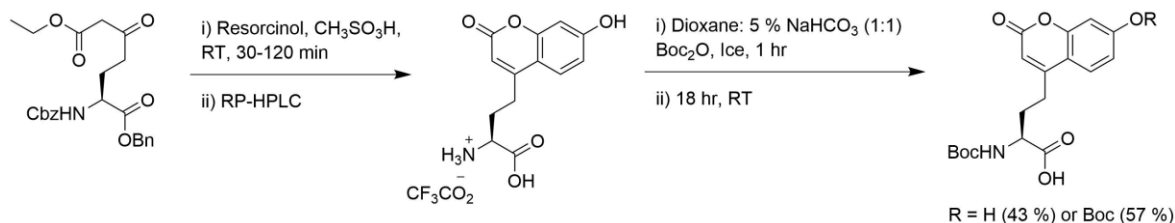
Chapter 4 Development of CLLFVY Fluorescent Probes

During the investigation of the SAR at the phenylalanine position it was noted that the extended aromatic derivative, Nal, was well tolerated, leading to a 6-fold increase in affinity compared to CLLFVY. The fluorescent molecule coumarin is structurally similar to Nal and has been demonstrated to be well suited for biological applications, and as such was considered for use in generating a fluorescent derivative of CLLFVY.^{225,226} The synthesis of a Boc-coumaryl amino acid would enable the incorporation of a coumaryl motif into CLLFVY via standard SPPS. The synthetic route followed was described by Brimble *et al.*, whereby the amino acid N- α -Cbz-L-glutamic acid α -benzyl ester is used as a starting point.²²⁷ The side chain carboxyl group is activated with carbonyl diimidazole (CDI) and reacted with ethyl magnesium malonate to yield the corresponding β -ketoester (Scheme 11).



Scheme 11. Activation of the glutamic acid side chain with CDI and subsequent reaction with ethyl magnesium malonate yields the desired β -ketoester.

Methane sulfonic acid and resorcinol are then added to the β -ketoester, which simultaneously deprotects the amino acid protecting groups and installs the coumaryl motif via a von Pechmann mechanism.²²⁸ The final step is to re-protect the N-terminus of the amino acid to enable its use in solid-phase peptide synthesis. This was achieved by reacting the unprotected amino acid with Boc anhydride, which yielded a mixture of amine protected and amine plus side chain protected coumaryl amino acid (Scheme 12). The di-Boc protected amino acid was then used in the synthesis of coumaryl containing peptides.



Scheme 12. Von Pechmann condensation of β -ketoester with resorcinol and simultaneous deprotection of protecting groups, followed by Boc protection of amino and side chain functionalities.

Once synthesised the amino acid was integrated into the phenylalanine position of CLLFVY, as well as the tyrosine position of the optimised compound CLLF(CF₃)VY, which also showed good tolerance of the Nal unnatural amino acid. The peptides were synthesised using standard SPPS protocols with the coumaryl amino acid added last. The Boc protecting groups were then removed during the resin cleavage with TFA. This yielded the two fluorescent peptides, CLL-hCou-VY and CLLF(CF₃)V-hCou, the structures of which are shown in Figure 74A and 75B, respectively.

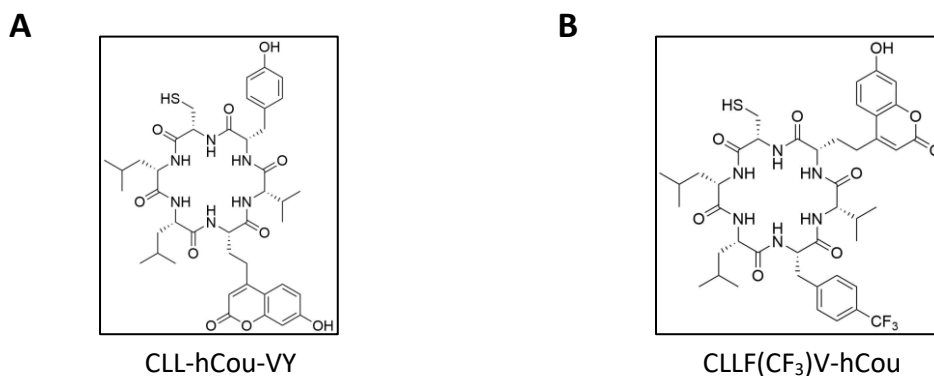


Figure 74. Structures of the fluorescent peptides CLL-hCou-VY (A) and CLLF(CF₃)V-hCou (B).

Once synthesised, the impact of the coumaryl motif on the affinity of binding to the HIF-1 α PAS-B domain was assessed by MST, with the resulting dose-response curves shown in Figure 75A and 76B. Interestingly CLL-hCou-VY showed a higher affinity towards the HIF-1 α PAS-B domain ($K_d = 10.6 \pm 1.2 \mu\text{M}$) than CLLF(CF₃)V-hCou ($K_d = 23.2 \pm 0.3 \mu\text{M}$) despite the latter containing the optimised F(CF₃) amino acid. Both of the peptides were found to have an improved K_d over the parent peptide CLLFVY, as predicted by the observed SAR when using the Nal derivative (section 2.8). The absorption-emission spectra of the coumarin fluorophore was also recorded for CLL-hCou-VY, which gave a maxima at 356 and 454 nm for absorbance and emission respectively (Figure 75C). These values are very close to the literature values of 4-methylumbelliferone, which the amino acid is derived from (365 and 445 nm).²²⁹

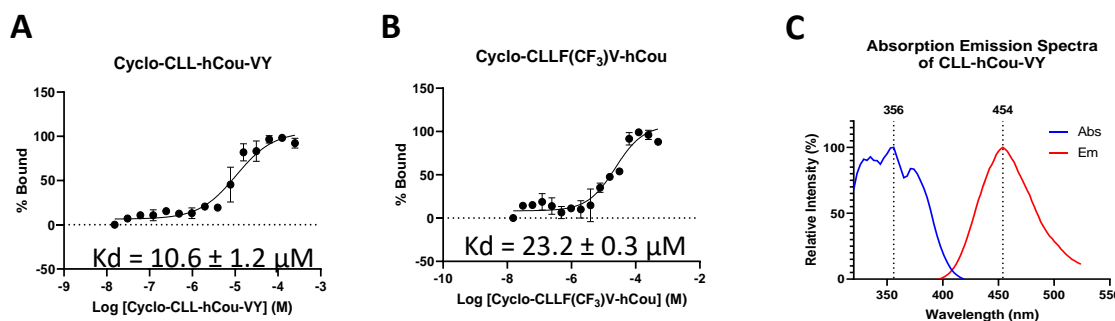


Figure 75. Analysis of the homo-coumaryl CLLFVY derivatives. MST dose-response curves of CLL-hCou-VY (A) and CLLF(CF₃)V-hCou against the HIF-1 α PAS-B protein (B). Errors bars represent standard deviation (n=3). C) Absorption-emission spectra for CLL-hCou-VY showing maxima at 356 and 454 nm. Binding constants were determined by curve fitting in GraphPad Prism 8.

Having obtained a fluorescent cyclic peptide that was shown to retain HIF binding activity, the potential of these peptides for use in mammalian cells was investigated. All cellular experiments were performed by Cyrielle Doigneaux. The first experiment conducted was to assess the passive cell permeability of the compound. This was achieved by dosing HeLa cells with the relevant peptide at a concentration of 50 μM in 1% final DMSO concentration in the cell media. The cells were then incubated at 37 °C overnight before being observed by confocal microscopy using a 405 nm UV laser and PMT detection. Both CLL-hCou-VY and CLLF(CF₃)V-hCou were observed in the cell cytoplasm as shown in Figure 76A and 77B, respectively.

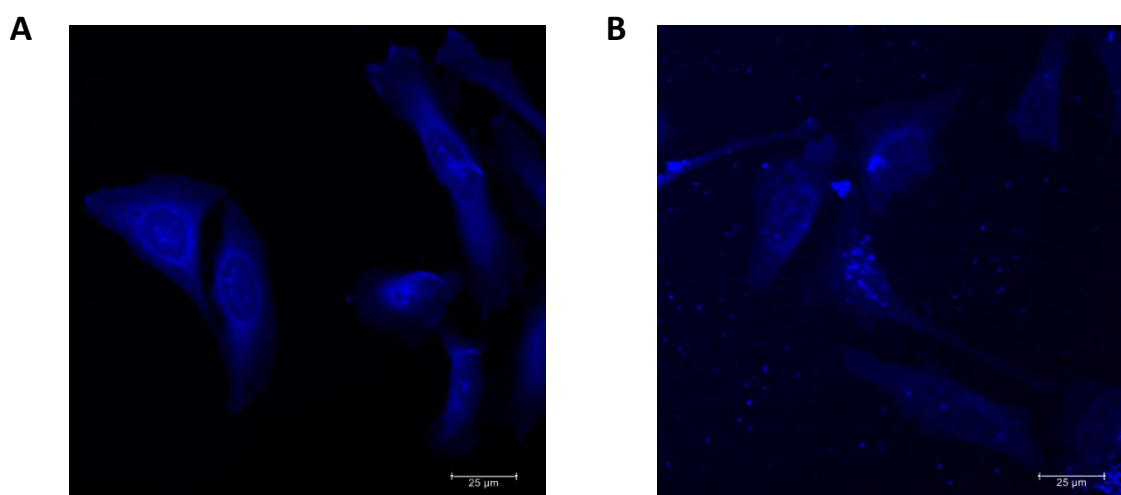


Figure 76. A) HeLa cells dosed with CLL-hCou-VY (50 μM) B) HeLa cells dosed with CLLF(CF₃)V-hCou (50 μM). Images acquired by Dr Cyrielle Doigneaux.

The parent compound CLLFVY had previously been shown to require a cell penetrating peptide in order to be internalised by mammalian cell lines, and so it is somewhat surprising that the fluorescent derivative CLL-hCou-VY was seen to be internalised into cells at concentrations that allow clear visualisation.¹⁴⁶ Both compounds were found to be permeable to some extent, however CLLF(CF₃)V-hCou also exhibited some aggregation outside of the cells (Figure 76B). As such, further experiments were therefore carried out with CLL-hCou-VY.

As CLL-hCou-VY was shown to be cell permeable, an experiment was next designed to assess the effect of the compound on the location of HIF-1 α within the cell. A control experiment was first run to determine where HIF is localised within the cell in both normoxia and when treated with the compound desferrioxamine (DFX). DFX is an iron chelator that prevents the hydroxylation of the HIF- α subunit by PHDs, causing an accumulation of HIF-1 α independent of oxygen concentration and therefore inducing increased HIF transcriptional activity.^{230,231} DFX is thus commonly used to mimic hypoxic environments.

HeLa cells plated onto glass coverslips were either mock treated (culture media only) or treated with DFX (100 μ M) for 16 hours. The cells were then subjected to an immunofluorescence assay aiming at localising HIF-1 α . This was carried out by incubating the previously fixed and permeabilised cells with a primary antibody, anti-HIF-1 α , and secondary antibody coupled to an Alexa568 fluorophore to enable visualisation. Cells were then mounted on a glass slide using DAPI-containing media to enable nuclei visualisation. The cells were then imaged using a confocal microscope. With detection at 580-700 nm for visualisation of the Alexa568 probe, and a wavelength of 405 nm for visualisation of the DAPI nuclear stain. The resulting images are shown in Figure 77.

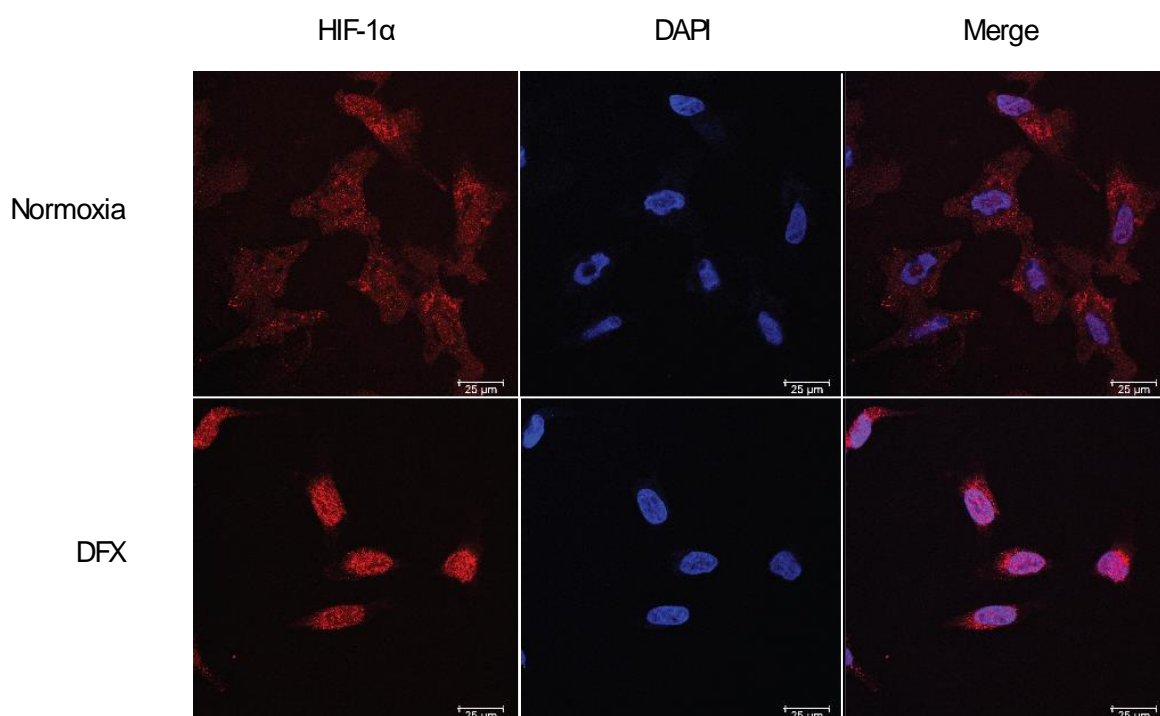


Figure 77. Immunofluorescence of HIF-1 α in HeLa cells with and without DFX treatment. Cells were also treated with nuclear DAPI stain and an overlay of the two channels is also shown. Images acquired by Dr Cyrielle Doigneaux.

As expected, the fluorescence intensity associated with HIF-1 α was lower in normoxia than in DFX-treated cells and the protein displayed homogenous distribution throughout the cell. However, when treated with DFX, the HIF-1 α immunofluorescence signal was found to be much stronger as the protein is no longer degraded, and can thus accumulate within the cell. It is also seen to translocate to the nucleus as would be expected under the mimicked hypoxic conditions consistent with its function as a transcription factor.

An experiment was then set up with the cells dosed with CLL-hCou-VY at 50 μ M in normoxia both with and without DFX-treatment. Both the peptide (blue) and HIF-1 α (red) were visualised (Figure 78).

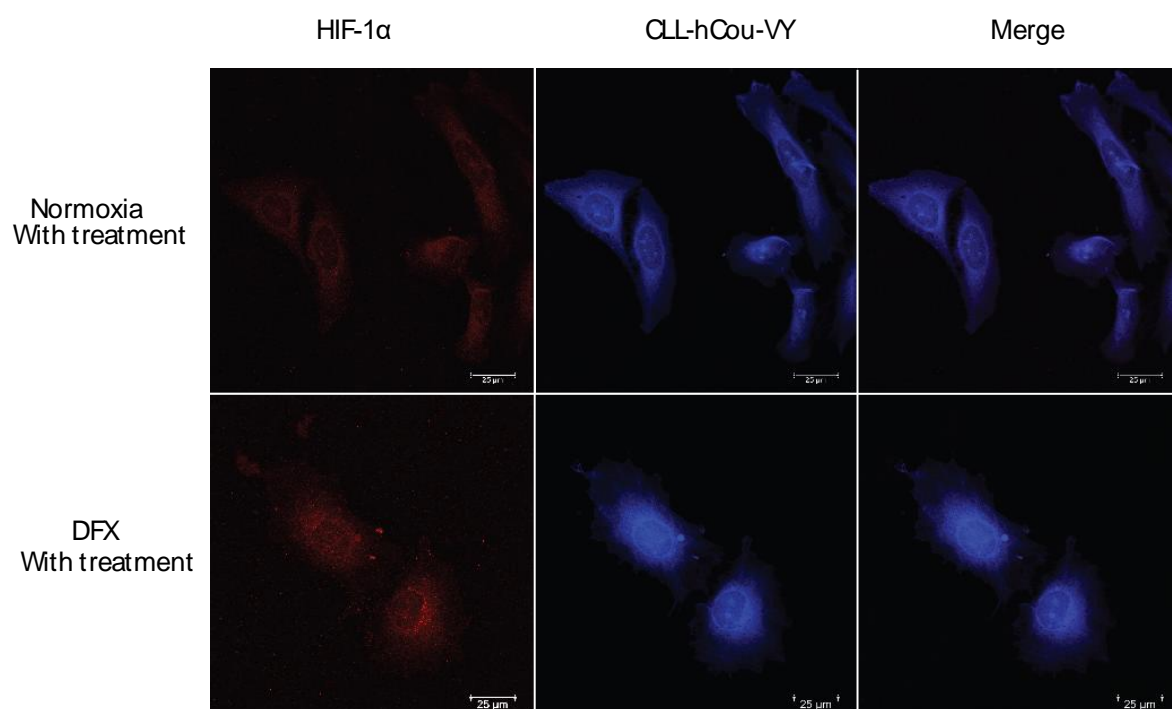


Figure 78. Immunofluorescence of HIF-1 α in HeLa cells with and without CLL-hCou-VY treatment. An overlay of the two images is also shown. Images acquired by Dr Cyrielle Doigneaux.

The cell pictures reveal that the cells treated with CLL-hCou-VY show not only reduced levels of HIF-1 α under DFX conditions (Figure 78) compared to untreated cells (Figure 77), but also that the HIF-1 α is no longer present in the nucleus (Figure 78). The reduction of total HIF-1 α levels is somewhat surprising given that CLLFVY has previously been reported to not affect HIF-1 α protein levels by western blot analysis.¹⁴⁶ The absence of HIF-1 α fluorescence in the nucleus of the cells under DFX-treated conditions demonstrates that HIF cannot be transcriptionally active, though determining the exact mechanism by which HIF has been excluded from the nucleus will require further work.

This is a good example of how fluorescent derivatives of CLLFVY could be helpful in determining the mechanisms of action of this class of compound. These probes however, are still in their infancy and considerable further work is needed before they will be able to be fully utilised. Further experiments already planned include the use of concurrent HIF-1 β immunofluorescence in order to assess the co-localisation of the two proteins upon treatment with the peptides. Efforts for the live imaging of the treatment are also underway and it is hoped that this will reveal at which point the compounds are intercepting HIF-1 α , and whether it is prevented from translocating to the nucleus or whether the domain does translocate but when inactive is ejected back into the cytoplasm.

Chapter 5 Determining the Structural Basis of CLLFVY Binding

5.1 Introduction to Protein X-ray Crystallography

Developing a structural understanding of CLLFVY binding would enable a more rational approach to its further development, and as such would be an invaluable resource for future studies. The two main techniques for probing protein structure are NMR and X-ray crystallisation. Previous experiments within the Tavassoli group have demonstrated that the HIF proteins do not express under minimal media conditions, removing the option of protein observed NMR experiments via isotopically labelled protein. In addition, the poor solubility of the peptide ligands limits the applicability of ligand observed NMR techniques such as Saturation Transfer Difference (STD) and Water-Ligand Observation with Gradient Spectroscopy (waterLOGSY). Given these limitations, it was decided that the interaction of HIF and CLLFVY would be probed using X-ray crystallography.

Since 1912 the predictable manner in which an ordered crystal will diffract X-rays has been used to gain insight into the structure of that crystal.^{232,233} This is possible as the distance between planes of atoms in a crystal lattice are of the same order as the wavelength of the X-rays used, and this relationship is defined by Bragg's law Equation 6.

$$2d \sin \theta = n\lambda$$

Equation 6. Bragg's law. Where d is the distance between two planes, λ is the wavelength of the incident X-ray and n is an integer.

The first example of X-rays being used to solve a protein structure came in 1958 when Kendrew *et al.* solved the structure of sperm whale myoglobin to a resolution of 6 Å, which they quickly improved upon to yield a structure at 2 Å resolution.^{234,235} At this resolution, the backbone of a protein can be clearly seen and the position of most side chains can be solved, making X-ray crystallography an incredibly powerful technique for probing protein structure. Since this first structure was published, innovations such as synchrotron radiation and cryocrystallography have led the field of X-ray crystallography to revolutionise structural biology.^{236,237} With these advances structures can be obtained quicker and The Protein Data Bank (PDB, a central repository for protein structures) at the time of writing now contains over 142,000 protein structures, 90 % of which were determined by X-ray crystallography.^{238,239}

Before a structure can be solved, the protein must be obtained in crystalline form. Whilst there have been multiple innovations in protein crystallisation, this is still the limiting step when trying

to gain structural information. The process of obtaining crystals that will diffract well can require a long process, that in the case of novel proteins is largely a matter of trial and error. In comparison to small molecules, protein crystals have a large solvent component, which accounts for an average of 50 % of the crystal.²⁴⁰ Whilst the presence of this solvent leads to weaker diffraction patterns it also allows the determination of a protein structure that resembles what the protein would actually look like in solution. The large solvent content also has the additional benefit that it creates channels within the crystal lattice, which can be exploited for the use of crystal soaking.²⁴¹ Soaking of small molecules into a protein crystal allows the screening of multiple ligands with only one set of conditions needed for growing the protein crystals. In comparison, co-crystallisation of a ligand-protein complex often requires completely new precipitant conditions.²⁴² To obtain protein crystals a concentrated solution of the protein of interest is concentrated until supersaturation is achieved. In practise this can be achieved using multiple techniques; two of the most prevalent being sitting drop and hanging drop set-ups which are illustrated in Figure 79.²⁴³

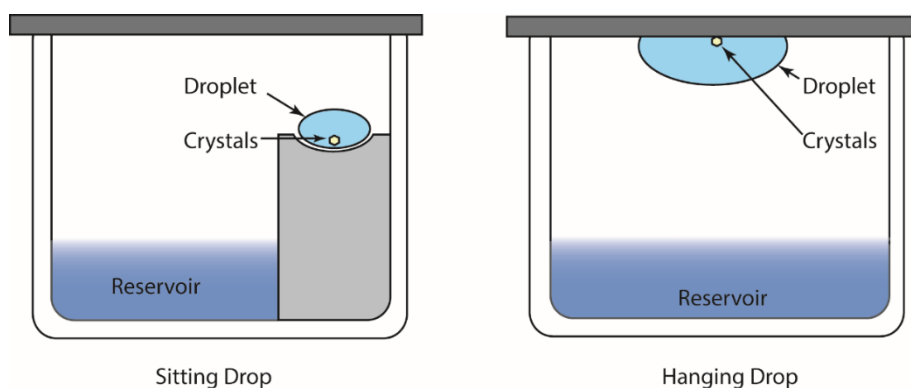


Figure 79. Illustration of a sitting drop and hanging drop crystallisation set up.

These utilise a reservoir containing a 'mother liquor', which contains the precipitants (buffer, salts, PEG, organic liquid) and a small droplet (1-5 μL) consisting of the protein solution mixed with the mother liquor. As the system reaches equilibrium, water is lost from the droplet into the reservoir restoring the concentration of protein to that of the stock solution and concentrating the precipitant. In order for crystals to form, the precipitant and protein concentrations must increase to the point where nucleation can occur, after which a supersaturated solution is required for these critical nuclei to grow.²⁴⁴ A phase diagram demonstrating this process is shown in Figure 80.

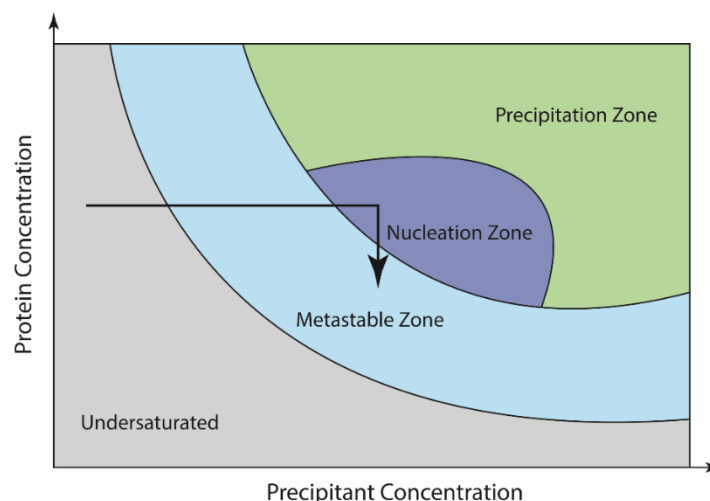


Figure 80. 2D visualisation of the solubility of protein as a function of precipitant concentration.

In the under-saturated zone, no crystal growth can occur. The remaining three zones represent saturated protein solutions, the metastable zone is where crystal growth can occur and the nucleation zone is where spontaneous nucleation can occur. At higher concentrations only amorphous precipitation is observed.

Once the crystalline protein has been obtained, it can be analysed by x-ray diffraction. The protein crystals are picked and cooled in liquid nitrogen which helps minimise radiation damage during data collection.²⁴⁵ The resulting non-destructive interference, as governed by Bragg's law, results in a diffraction pattern in reciprocal space. The diffraction pattern is the combination of multiple reflections, which are all dependent on three factors: the direction, amplitude and phase of the diffracted wave. During an x-ray diffraction analysis, the position and intensities of the reflections are measured, and in turn can be used to determine the direction and amplitude of the diffracted wave. The position of each reflection can then be used to calculate the diffraction angle θ and the distance between the lattice planes (d_{hkl}), which in turn allows the assignment of Miller indexes (h,k,l) to each reflection plane and unit cell parameters to be determined.²⁴⁶

The intensity of the reflection can be used to determine the amplitude of the diffracted waves. Whilst the intensity can be affected by other factors, the factor that relates to the crystal lattice is called the structure factor (F_{hkl}) and is calculated using the Fourier sum shown in Equation 7.²⁴⁷

$$F(hkl) = \sum_{j=1}^N f_j \exp[2\pi i(hu_j + kv_j + lw_j)]$$

Equation 7. Calculation of the structure factor. Where the position of the j^{th} atom is given by the coordinates u_j, v_j, w_j with the scattering factor f_j .

In order to generate an electron density map that correlates to the position of each atom in the lattice, the structure factors of each reflection needs to be expressed as the sum of contributions of electron density from each volume element of the unit cell. This can be achieved by performing a Fourier transform on Equation 8 to calculate $\rho(xyz)$, the electron density within the unit cell, which is used to generate an electron density map.²⁴⁷

$$\rho(xyz) = \frac{1}{V} \sum_{hkl} F(hkl) \exp[-2\pi i(hx + ky + lz)]$$

Equation 8. Calculation of electron density within the unit cell. Where $\rho(xyz)$ is the electron density at point (x, y, z) , V is the volume of the unit cell and $F(hkl)$ is the structure factor.

Whilst the amplitudes ($|F(hkl)|$) can be calculated from the intensity of the diffracted x-ray, the phases of those x-rays ($\phi(hkl)$) is not determined experimentally and so must be determined via a different method. A number of techniques for phase determination have been developed as the field of x-ray crystallography has evolved, such as isomorphous replacement, anomalous scattering, and molecular replacement.²⁴⁸ However, as there are existing crystal structures of HIF the method of determining phases used in this work will be molecular replacement.

As previous HIF structures have been solved, they can be used as models for solving the obtained diffraction data in a process called molecular replacement. This is a facile way to obtain the unknown phase information, whereby a known protein structure is used as a template for the collected data. Automated algorithms then use rotational and translational parameters to fit the data to the selected model.²⁴⁹ In order to obtain a structure in good agreement with the electron density map it is crucial to pick a suitable model for replacement. The model should have high sequence similarity, a good level of completeness and low root mean squared deviation (R.M.S.D) from the protein of interest. To assess how well the modelled data matches the measured diffraction an R-factor is calculated as described in Equation 9. This R-factor compares the agreement between the structure factor amplitudes from the model with those in the observed data set, and is also used in further refinement cycles.

$$R = \frac{\sum |F_{obs}| - |F_{calc}|}{\sum |F_{obs}|}$$

Equation 9. R-factor calculation for assessing how well the current model fits to the observed diffraction data. $|F_{obs}|$ and $|F_{calc}|$ represent the observed and calculated amplitudes.

5.2 Crystallising HIF

As has been discussed previously there are only two existing crystal structures of HIF-1 that have been solved, the heterodimeric HIF-1 α PAS-B HIF-1 β PAS-B structure, and the DNA bound bHLH-PAS domain heterodimer, both of which present potential difficulties for crystallisation of HIF in the presence of an inhibitor of HIF dimerisation. The existing structures rely on HIF-1 α being bound to HIF-1 β , and so binding of an inhibitor of this interaction should prevent heterodimer formation. In addition, this problem is exacerbated in the case of the bHLH-PAS domains model as this structure was crystallised bound to DNA, and as has been demonstrated CLLFVY, and derivatives thereof, prevent DNA binding in the HRE-FP assay.

A structure containing only the HIF-1 α protein would circumvent these problems. Given that the majority of the SAR studies on CLLFVY utilised the HIF-1 α PAS-B domain, and this domain was shown to bind the compounds with the highest affinity, it was decided a PAS-B alone structure would be pursued. The necessity for a protein model that is stable at high concentrations and available in high quantity meant that the mutant HIF-1 α PAS-B domain (HIF-1 α PAS-B 4x) was selected.

A collaboration between a previous PhD student (Dr Andy Foster) and GlaxoSmithKline led to the discovery of conditions that yielded protein crystals with the HIF-1 α PAS-B domain alone. The conditions were a 1:1 mixture of HIF-1 α PAS-B (4.0 mg/mL in HIF buffer) and a solution of sodium succinate (pH 7). The process also required the enzyme thrombin to be added to each of the wells to carry out an *in situ* thrombin cleavage. The HIF-1 α PAS-B contains a thrombin cleavage site enabling the removal of the proteins HIS tag, which may be impeding crystallisation.

The conditions were replicated within our lab in a hanging drop set up. Sodium succinate (0.8 M, pH 7, 500 μ L) was added to the reservoir and a mixture of this mother liquor and a solution of the HIF-1 α PAS-B 4x mutant in HIF buffer was pipetted onto a glass slide. Three droplets were added to the slide in ratios of 1:3, 1:1 and 3:1, HIF-1 α PAS-B 4x: mother liquor. Each droplet then had thrombin (10 % V/V) added. The slide was then attached to a hanging drop well with vacuum grease and incubated at 18 $^{\circ}$ C for seven days. The resulting crystals were then imaged and are shown in Figure 81. The conditions produced small crystals of poor quality, and are therefore

unlikely to diffract well. Therefore, the possibility of using these crystals as seeds for further experiments was explored.



Figure 81. Image of HIF-1 α PAS-B crystals obtained using previously discovered crystallisation conditions.

5.2.1 Seeding Experiments

Crystal seeding is a technique that tackles one of the most prominent challenges of protein crystallography, that is that the conditions that are most ideal for nucleation (high levels of supersaturation) are not the same as the best conditions for crystal growth (low level supersaturation). Seeding involves taking previously nucleated crystals and transferring them to a less saturated protein solution which maintains the nucleated crystal in the metastable zone (Figure 80), allowing for steady crystal growth that results in larger protein crystals.²⁵⁰ To improve the size and quality of the HIF-1 α PAS-B 4x protein crystals, seeding experiments were attempted. This involved crushing the poorly formed crystals from the previous experiment, and making a range of seed stocks at different dilutions (1:10, 1:100, 1:1000 and 1:10,000 seed stock: mother liquor). A hanging drop plate was then set up with a reservoir of succinic acid (0.8M, pH7, 500 μ L), and droplets of HIF-1 α PAS-B 4x: mother liquor in a ratio of 1:3, 1:1 and 3:1 were added to 4 glass slides. Each slide then had a different concentration of seed stock added to each droplet (0.5 μ L). After seven days, crystals were seen to be growing, predominately in the 3:1 droplets. These droplets were imaged under a polarised microscope and the images are shown in Figure 82.

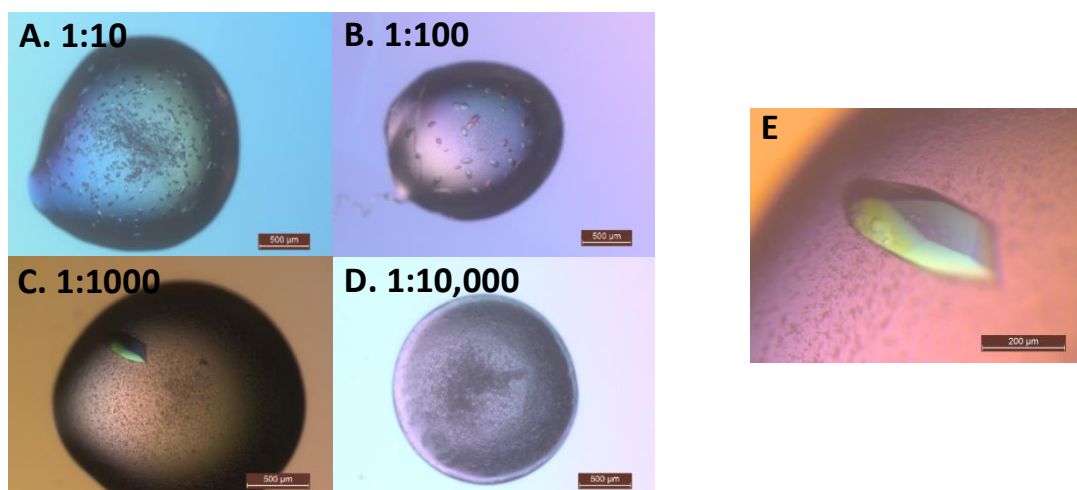


Figure 82. Images of crystals obtained from seeding experiments. A) Image of droplet containing 10x diluted seed stock. B) Image of droplet containing 100x diluted seed stock. C) Image of droplet containing 1000x diluted seed stock. D) Image of droplet containing 10,000x diluted seed stock. E) Zoomed in image of the crystal obtained in C.

At a dilution of 1:10 the seed stock was too concentrated and so many nucleation sites were present, leading to the growth of many small crystals. At a dilution of 1:100, the higher dilution means that there were fewer crystals but of a better size. The best results were seen at a dilution of 1:1000 where one much larger crystal (400 μm) was observed. Beyond this, the 10,000 times diluted seed stock did not appear to have any nucleated material as the protein was seen to precipitate rather than crystallise. The largest crystal (Figure 82E) was picked with a cryo-loop and placed into a droplet of 30 % glycerol in mother liquor as a cryo-protectant, before being flash frozen in liquid nitrogen and stored at $-196\text{ }^{\circ}\text{C}$. The crystal was then analysed at Diamond Light Source on beamline I04. An image of the loaded crystal and resulting diffraction pattern are shown in Figure 83.

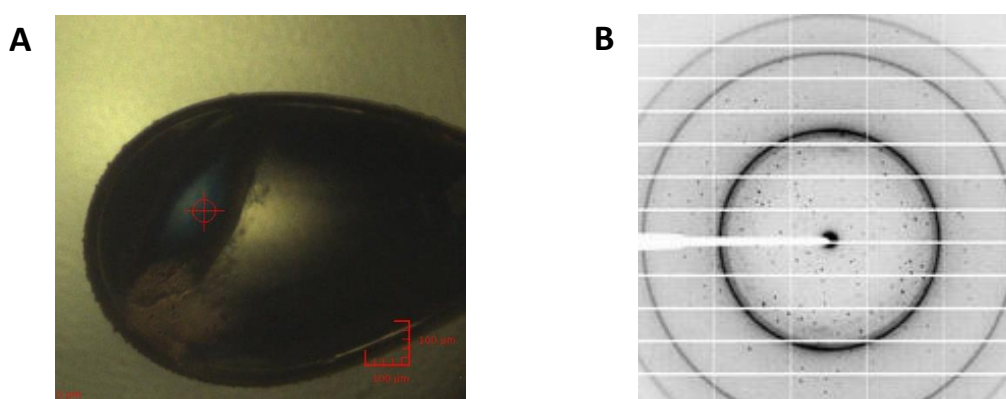


Figure 83. A) Loop containing a crystal of HIF-1 α PAS-B. Loop size 0.4-0.5 mm, beam size 20 x 20 μm . B) Associated diffraction pattern from 12 keV synchrotron radiation.

Chapter 5

A high-resolution dataset was collected from the crystal at a wavelength of 0.9159 Å. The collected data was automatically processed and the diffraction statistics are given in Table 1.

Table 1. Diffraction statistics obtained for the HIF-1 α PAS-B crystal on the I04 beamline at Diamond Light Source. Highest resolution shell shown in parentheses.

Diffraction Statistics	HIF-1 α PAS-B
Space Group	P 31 2 1
a, b, c (Å)	39.2, 39.2, 117.54
α , β , γ (°)	90, 90, 120
Wavelength (Å)	0.9159
Resolution (Å)	58.77-1.58 (1.58-1.61)
No. of reflections	239,559
No. of unique reflections	15,132
R_{merge}	0.073 (2.007)
R_{meas}	0.075 (2.182)
$I/\sigma I$	14.9 (1.1)
CC1/2	0.99 (0.6)
Completeness (%)	99.9

The statistics of the obtained data set show that the maximum resolution achieved was 1.58 Å and that the data had a very good level of completeness. The data set was then solved using molecular replacement with the Phenix Phaser-MR program using the HIF-1 α model extracted from the HIF-1 PAS-B heterodimer structure, PDB:4H6J.^{251,252} This produced a model in good agreement with the electron density map which was further refined via Phenix-refine software and manual adjustments made using WinCoot.^{251,253} The final statistics for this model are given in Table 2 and are summarised using the Phenix polygon tool in Figure 84. This tool enables validation of the final structure by comparing the obtained statistical parameters with all other structures in the PDB of a similar resolution. The polygon analysis showed that all parameters for the HIF-1 α PAS-B model were within the ranges for structures submitted at similar resolutions.

Table 2. Refinement statistics for the HIF-1 α PAS-B data set.

	HIF-1 α PAS-B
No. of reflections	7596
R_{work}/R_{free}	0.22/ 0.28
No. atoms:	
Protein	796
Water	49
B-factors:	
Protein	38.9
Water	49
R.M.S deviations:	
Bond lengths (\AA)	0.007
Bond angles (\AA)	0.94

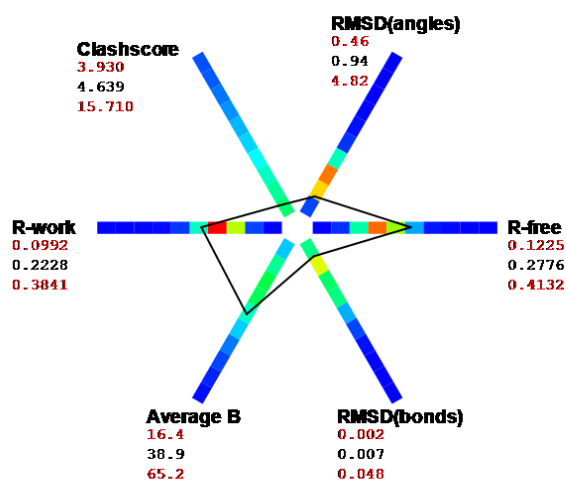


Figure 84. Polygon plot of refinement statistics for the HIF-1 α PAS-B data set. Values in red are upper and lower limits of these parameters found in the PDB. Refinement was carried out using the Phenix suite of software.

The solved structure reveals that the HIF-1 α PAS-B domain crystallises as a homodimer with the two domains interfacing via an anti-parallel beta sheet interaction that resembles that of the HIF-1 PAS-B heterodimer (Figure 85).¹⁶ The interaction is mediated by six polar contacts between the two domains. The first interaction is a complex formed between the E243, Y254, D256 and R258, with hydrogen bonds formed between the R258 and E243 and between Y254 and D256. Interestingly, a further interaction is observed between D256 and R258, forming a complex

between the four amino acids (Figure 85A). The next interaction is a hydrogen bond between the carbonyl oxygen of E318 of one domain and the hydroxyl proton of T327 on the other (Figure 85B). It is also of note that the E243 is the mutated residue necessary for formation of a stable heterodimer and would be an arginine in the wild type protein. These interactions are then reflected down a line of symmetry to give the six interdomain interactions.

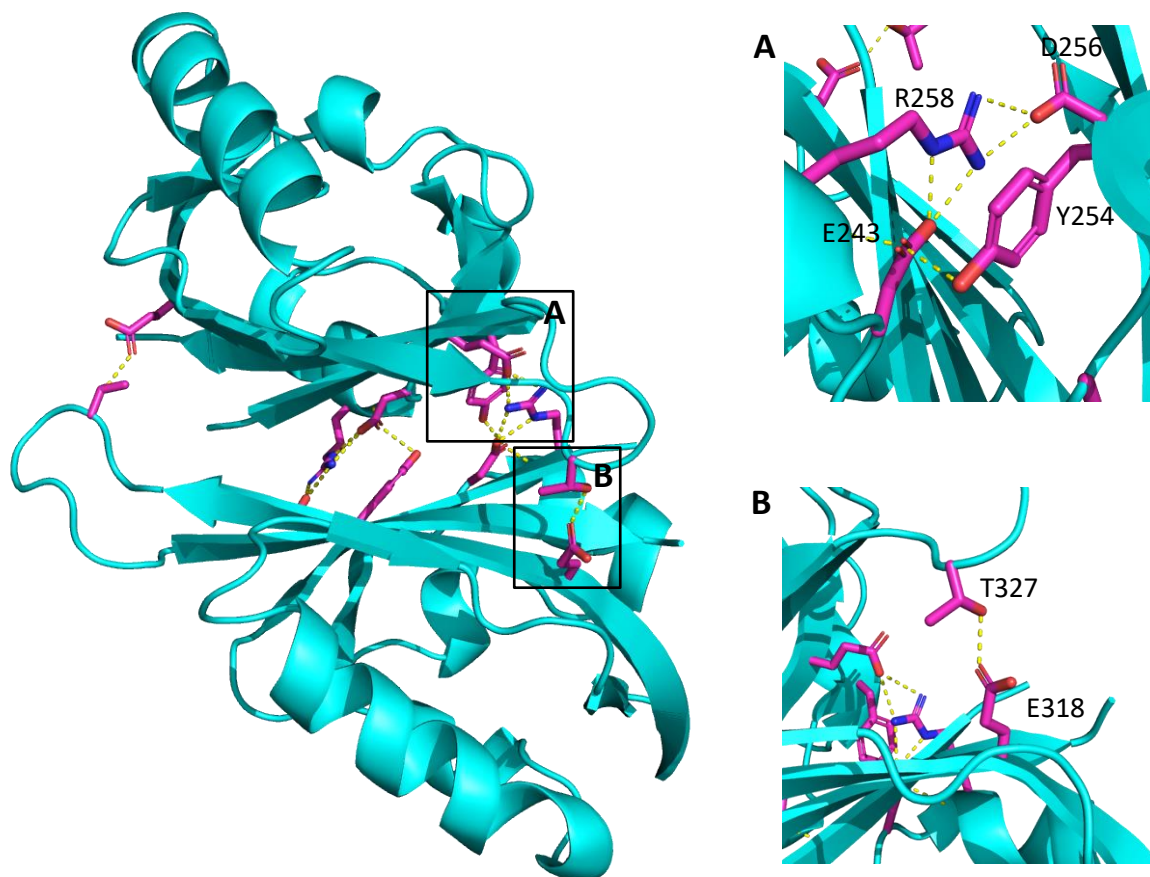


Figure 85. Protein structure obtained upon solving the obtained data set. The HIF-1 α PAS-B was seen to form a homodimeric structure mediated by polar contacts between six residues of each domain. A) Complex formed between the E243, and R258 of one domain and Y254 and D256 of the other domain. B) Hydrogen bond between the carbonyl oxygen of E318 of one domain and the hydroxyl proton of T327 of the other domain.

The structure also reveals the presence of three structural water molecules, which is consistent with the HIF-1 α PAS-B domain in the heterodimeric model (Figure 86).¹⁶

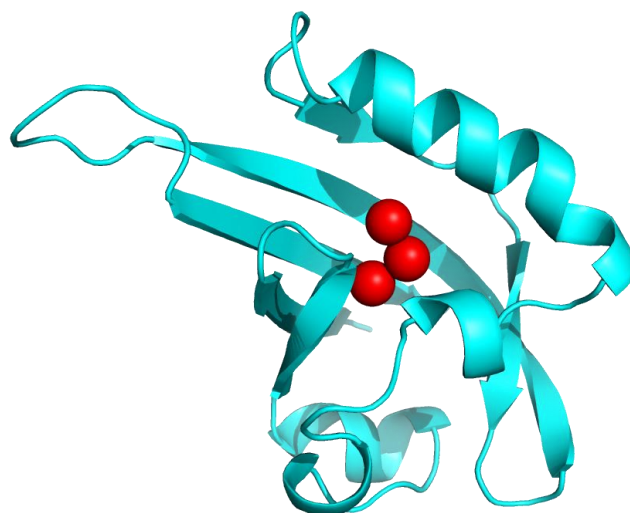


Figure 86. Structural waters present in the internal cavity of HIF-1 α PAS-B domain.

5.2.2 Soaking HIF-1 α PAS-B crystals with CLLFVY

Having obtained an apo structure of the HIF-1 α PAS-B domain we next sought to use this model to obtain a structural understanding of CLLFVY binding. With conditions for the formation of HIF-1 α PAS-B crystals in hand, the use of these crystals in soaking experiments was explored. Several hanging drop wells were set up as previously described (section 5.2.1) which produced a number of protein crystals. CLLFVY (0.5 μ L 50 mM 100% DMSO or 0.5 μ L 500 μ M 20 % DMSO) was then added to drops containing crystals. The crystals were then left to soak for 0.5, 3 or 24 hours. All of the crystals that had 0.5 μ L, 50 mM CLLFVY in 100 % DMSO added were found to dissolve. Of the crystals that had 0.5 μ L 500 μ M CLLFVY in 20 % DMSO added, those that were soaked for 24 hours dissolved but the 3 hour and 0.5 hour soaked crystals remained intact. These were transferred to a droplet of 30 % glycerol before being frozen and stored in liquid nitrogen. Analysis of the soaked crystals was carried out as previously detailed (section 5.2.1) on the IO4 beamline at Diamond Light Source. The crystals yielded high quality data sets that once solved revealed that no ligand was present. There are numerous reasons that could explain the absence of CLLFVY in the crystals, including; CLLFVY is too large to permeate the crystal lattice, the conditions used in soaking were sub-optimal, or that the homodimer is blocking the binding site of the peptide. Further experiments would be needed to optimise this approach and investigate different soaking conditions, perhaps at lower concentrations over longer time periods.

5.2.3 Co-Crystallisation of HIF-1 α PAS-B and CLLFVY

In addition to the attempts at soaking the HIF crystals with the peptides, the possibility of co-crystallisation was also investigated. This was initially carried out using similar conditions as those used to obtain the apo structure. HIF-1 α PAS-B was incubated for 1 hour at room temperature

with CLLFVY in either 1:1 or 1:5 molar ratios. The samples were then centrifuged at 10,000 rpm to remove any precipitate and the protein-peptide mixture was added to droplets in a hanging drop plate. Three droplets were created per well with ratios of 1:1, 1:3 or 3:1 peptide-protein: mother liquor. The mother liquor used varied per well and covered a range of succinic acid concentrations (0.5-1.0 M) at both pH 7 and pH 8. In addition, several wells were set up using the exact conditions used to obtain the apo structure (0.8 M succinic acid, pH 7) including addition of the 1:1000 seed stock (0.5 μ L per droplet) and thrombin (10 % V/V). Almost all of these conditions led to the formation of precipitate with no visible crystallisation, with the exception of the seeded wells where one droplet contained a small crystal that was too small to obtain a data set from (Figure 87). Given that this small crystal was only seen in the seeded droplet, in the same conditions as that used for the apo structure, it seems likely that this crystal is also unbound HIF-1 α PAS-B, though this cannot be confirmed unless a crystal large enough to diffract is obtained.

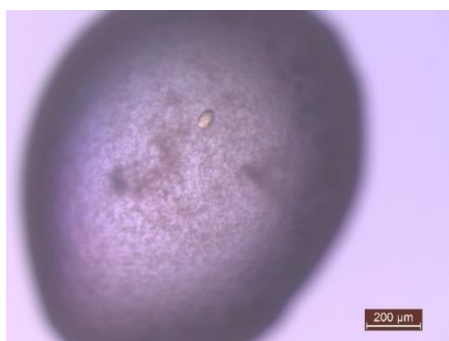


Figure 87. Image of a poor quality protein crystal obtained during the attempted co-crystallisation of HIF-1 α PAS-B with CLLFVY. Conditions used were succinic acid (0.8 M, pH 7), with seed stock (0.5 μ L, 1:1000 HIF-1 α PAS-B).

These results demonstrate that the addition of CLLFVY in DMSO inhibits the formation of crystalline protein in these conditions. In order to search for more amenable conditions for co-crystallisation, a screen of two buffer libraries (Hampton research HR2-110 and HR2-112) containing 98 different buffer conditions was carried out. Sitting drop plates were used with HIF-1 α PAS-B (4.0 mg/ mL) and buffer mixed in droplets in 1:1, 1:3 and 3:1 ratios. The plates were then incubated at 18 $^{\circ}$ C for seven days and inspected under a microscope for any crystal formation. Upon inspection, no crystals were seen to have formed in any of the wells. This could be as a result of the poor solubility of the peptide causing precipitation to occur in many of the wells. Future experiments could investigate the use of some of the CLLFVY derivatives created in this work to try and circumvent this issue. It may also be that a wider array of conditions for co-crystallisation still need to be explored and work is currently underway to examine this possibility. In addition, it is also possible that this protein model is not suitable for obtaining peptide-bound structures and future work will also explore the use of different protein models.

5.2.4 Optimisation of crystallisation conditions

One of the drawbacks of the current method for crystal growth is the requirement of an *in-situ* thrombin cleave of the protein. In order to obtain a ligand-bound structure a simpler workflow would be advantageous. To simplify the workflow, and allow for more efficient optimisation of crystallisation conditions, the potential of cleaving the protein with thrombin before crystallisation was investigated. This was achieved by incubating the HIF-1 α PAS-B 4x protein with thrombin (1 unit/ mg) at 4 °C overnight. The protein was then purified via a Ni²⁺ affinity column to remove the cleaved HIS-tag and then further purified by an S75 size-exclusion column (SEC). The purified protein was then analysed by mass spectrometry (Figure 88) which showed the presence of several different masses. This is likely a result of the thrombin non-specifically cleaving at secondary sites in the protein, something which has been previously reported as an issue with this enzyme.²⁵⁴ Notably, the expected mass of 14,814 Da is missing from this spectrum. The sequence of the different truncations are shown in Figure 89. Interestingly, whilst the thrombin is cleaving beyond the recognition site (LVPRGS), the predominant species with a mass of 13151.9 has the amino acid sequence closest to that of the solved HIF-1 α PAS-B structure. Given the higher temperature and longer time period of the *in-situ* thrombin cleavage used during crystallisation, it is not unreasonable to assume that the protein is cleaved multiple times to yield the same truncated protein species. It would then stand to reason that this smaller protein is more amenable to crystallisation than the uncleaved HIF-1 α PAS-B domain. Given the evidence that this truncated HIF model is needed for crystallisation, a new protein construct was designed to express the model without need for an *in-situ* enzymatic cleavage.

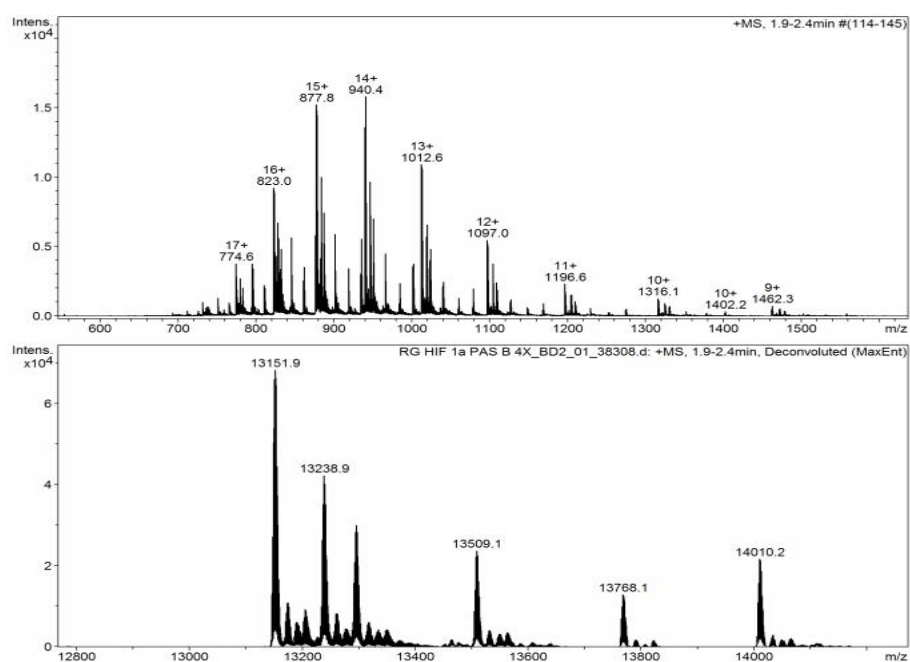


Figure 88. Mass spectrum obtained from thrombin cleaved HIF-1 α PAS-B domain.

Chapter 5

```

HIF-1aPAS-B      MGSSHHHHHSSGLVPRGSHMASMTGGQQMGRGSMDSKTFLEHSLDMKFSYCDERITELMGYHPEELLGRSIYEYHYAL
14010.2          -----GGQQMGRGSMDSKTFLEHSLDMKFSYCDERITELMGYHPEELLGRSIYEYHYAL
13768.1          -----QMGRGSMDSKTFLEHSLDMKFSYCDERITELMGYHPEELLGRSIYEYHYAL
13509.1          -----GRGSMDSKTFLEHSLDMKFSYCDERITELMGYHPEELLGRSIYEYHYAL
13238.9          -----GSMDSKTFLEHSLDMKFSYCDERITELMGYHPEELLGRSIYEYHYAL
13151.9          -----SMDSKTFLEHSLDMKFSYCDERITELMGYHPEELLGRSIYEYHYAL
HIF-1aCrystal    -----KTFLEHSLDMKFSYCDERITELMGYHPEELLGRSIYEYHYAL
                  *****

HIF-1aPAS-B      DSDHLTKTHHDMFTKGQVTTGQYRMLAKHGGYVWVETQATVIYNTKNLQPQCIVCVNYVVSIGIIQHD
14010.2          DSDHLTKTHHDMFTKGQVTTGQYRMLAKHGGYVWVETQATVIYNTKNLQPQCIVCVNYVVSIGIIQHD
13768.1          DSDHLTKTHHDMFTKGQVTTGQYRMLAKHGGYVWVETQATVIYNTKNLQPQCIVCVNYVVSIGIIQHD
13509.1          DSDHLTKTHHDMFTKGQVTTGQYRMLAKHGGYVWVETQATVIYNTKNLQPQCIVCVNYVVSIGIIQHD
13238.9          DSDHLTKTHHDMFTKGQVTTGQYRMLAKHGGYVWVETQATVIYNTKNLQPQCIVCVNYVVSIGIIQHD
13151.9          DSDHLTKTHHDMFTKGQVTTGQYRMLAKHGGYVWVETQATVIYNTKNLQPQCIVCVNYVVSIGIIQHD
HIF-1aCrystal    DSDHLTKTHHDMFTKGQVTTGQYRMLAKHGGYVWVETQATVIYNTKNLQPQCIVCVNYVVSIGIIQHD
                  *****

```

Figure 89. Amino acid sequences corresponding to the different peaks observed in the mass spectrum of the thrombin cleaved HIF-1 α PAS-B domain. A comparison with the amino acid sequence of the solved HIF-1 α PAS-B homodimer shows that the protein with mass 13,151 displays the highest sequence homology.

5.2.5 Development of a new protein model for crystallisation

As the mass spectrometry analysis had shown that the protein species forming crystals is most likely a shorter sequence than was being expressed, a new protein model was designed to express only the protein sequence observed in the solved structure (HIF-1 α PAS-B₍₂₃₈₋₃₄₂₎). This protein would still need an affinity tag to enable purification of the expressed protein, which would have to be cleaved to prevent it from interfering in the crystallisation process. Given the promiscuity of the thrombin cleavage, the use of a different enzyme for cleaving the affinity tag was investigated. For this, the high specificity enzyme Tobacco Etch Virus (TEV) protease was selected. This required the design of a new vector with the appropriate recognition sequence at the point of cleavage. For this the gene coding for the HIF-1 α PAS-B₍₂₃₈₋₃₄₂₎ protein was cloned into a pETM11 vector (Figure 90).

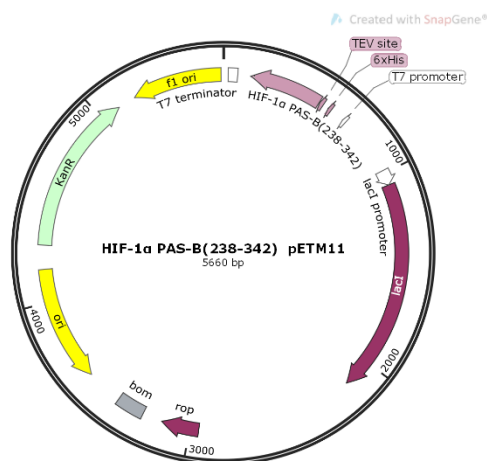


Figure 90. Plasmid map of pETM11 vector containing HIF-1 α PAS-B₍₂₃₈₋₃₄₂₎ and a TEV cleavage site to enable removal of the N-terminal HIS-tag.

The plasmid was transformed into BL21 DE3 Rosetta cells and expressed as previously described (section 0). The purified protein was then incubated with TEV (1 unit/ mg) at 4 °C overnight. The protein was then purified via a Ni²⁺ affinity column to remove the cleaved HIS-tag and TEV, and then further purified by an S75 SEC. The purified protein was then analysed by mass spectrometry (Figure 91).

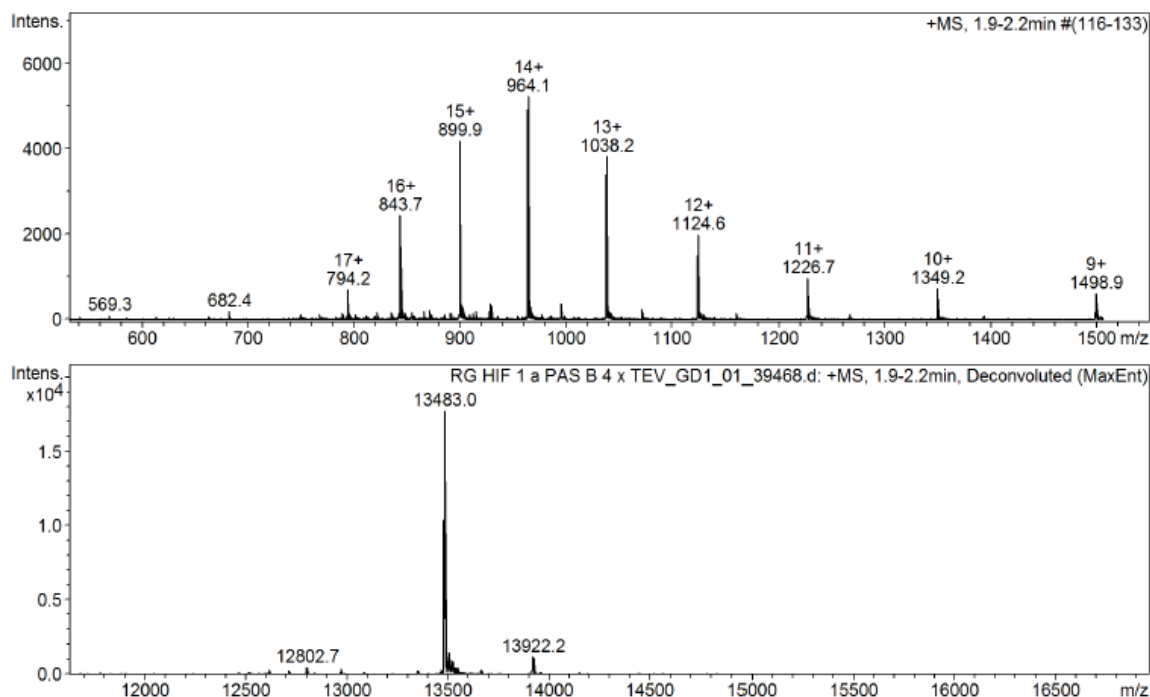


Figure 91. Mass spectrum obtained for TEV cleaved HIF-1 α PAS B domain.

Gratifyingly, the majority of the obtained protein was a single species with the expected mass of 13,483 Da. A comparison of the HIF-1 α PAS-B₍₂₃₈₋₃₄₂₎ and cleaved HIF-1 α PAS-B₍₂₃₈₋₃₄₂₎ protein sequences is shown in Figure 92.

```

HIF-1apETM11  MKHHHHHPMSDYDIPTTENLYFQGAMAMDSKTF LSEHSLDMKFSYCDERITELMGYHPPEELLGRSIEY EYHALDSDHLT
HIF-1aTEV      -----GAMAMDSKTF LSEHSLDMKFSYCDERITELMGYHPPEELLGRSIEY EYHALDSDHLT
                *****

HIF-1apETM11  KTHHDMFTKGQVTTGQYRMLAKHGGYVWVETQATVIYNTKNLQPQCIVCVNYVVS GIIQHD
HIF-1aTEV      KTHHDMFTKGQVTTGQYRMLAKHGGYVWVETQATVIYNTKNLQPQCIVCVNYVVS GIIQHD
                *****

```

Figure 92. Sequences of HIF-1 α PAS-B₍₂₃₈₋₃₄₂₎ and cleaved HIF-1 α PAS-B₍₂₃₈₋₃₄₂₎ protein.

Having obtained the cleaved HIF-1 α PAS-B₍₂₃₈₋₃₄₂₎, an attempt was made to crystallise the protein using the previously discovered conditions. The protein was added to a droplet in a hanging drop set up in a 3:1 ratio with succinic acid (0.8 M, pH7) and incubated at 18 °C for seven days. The resulting crystals are shown in Figure 93.

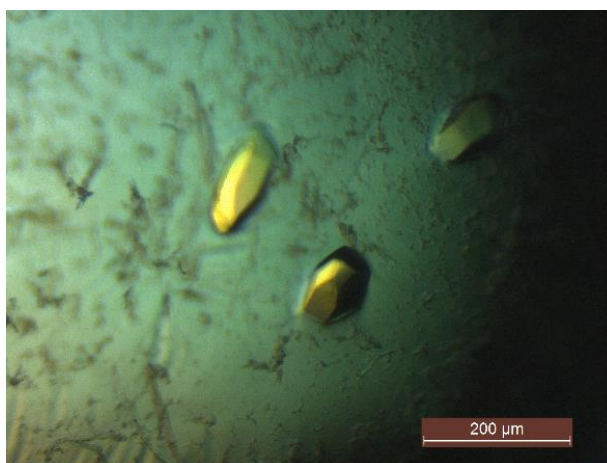


Figure 93. Crystals of TEV-cleaved HIF-1α PAS-B₍₂₃₈₋₃₄₂₎ protein.

Whilst some small crystals were seen to form, they were significantly smaller than those obtained in Figure 82E. This not unexpected given the change in crystallisation conditions brought about by the omission of thrombin and the new, shorter protein model. These small crystals show that the new protein model is amenable to crystallisation, and with some optimisation it is hoped that diffraction quality crystals will be obtained that can then be used in ligand based experiments.

5.3 Chapter 5 Summary

Conditions for the crystallisation of the HIF-1α PAS-B protein have been optimised via the use of seeding to enable the growth of crystals suitable for data collection. A high quality data set for the diffraction of a HIF-1α PAS-B crystal was obtained and solved to yield a novel HIF structure that shows that the HIF-1α PAS-B domain crystallises as a homodimer. The proteins were found to interface via an anti-parallel beta-sheet interaction analogous to the HIF-1/2 PAS-B heterodimers.¹⁶ The interaction contained 6 polar contacts between the two HIF-1α PAS-B domains, and each domain was also found to have an internal cavity filled with three structural water molecules.

The conditions for crystallisation of the HIF-1α PAS-B homodimer were then used to grow crystals for use in soaking experiments. A range of different concentrations of CLLFVY and DMSO were soaked for either 0.5, 3 or 24 hours. Diffraction of crystals from the 0.5 and 3-hour time points revealed that no ligand was present in the crystals. Co-crystallisation was also attempted utilising two commercial buffer screens (Hampton research HR2-110 and HR2-112) but no crystals were observed to form.

The role of thrombin in the crystallisation of the HIF-1α PAS-B was then investigated. Thrombin was found to cleave the protein beyond the recognition site to yield a truncated HIF model. Inspection of the resulting cleaved protein revealed that the shortened sequence was similar to

that used in previous HIF-1 α PAS B crystal models, leading to the hypothesis that these smaller proteins are more amenable to crystallisation.¹⁶ A new protein construct was therefore designed that contained a TEV cleavage site preceding the shortened protein sequence. An expression protocol was then developed that yielded the HIF-1 α PAS-B₍₂₃₈₋₃₄₂₎ model without the need for an *in-situ* enzymatic cleavage. Initial experiments have shown that this new protein model does crystallise, though further optimisation will be needed to yield crystals suitable for diffraction.

Chapter 6 Conclusions

As the master regulator of the cellular response to hypoxia, HIF is responsible for upregulating hundreds of genes that enable cells to adapt to drops in the concentration of environmental oxygen. The prevalence of hypoxia in the cores of cancerous tumours has made HIF a key target for therapeutic intervention. Whilst a number of indirect approaches have been developed to inhibit HIF activity, there is still a lack of compounds that are capable of directly targeting HIF. Compounds capable of directly binding to HIF would be invaluable to not only minimise off target effects, but also as use as probes for furthering our understanding of HIF activity. The most obvious target for inhibition of HIF activity is to inhibit the dimerisation of the two HIF subunits, HIF-1 α and HIF-1 β , but the development of inhibitors targeting this interaction has proven challenging. Typical small molecule inhibitors struggle to target the comparatively featureless surfaces of PPIs, which has led to the development of new scaffolds such as cyclic peptides. Cyclic peptides are better positioned to tackle these challenging targets, as they have larger surface areas and a secondary structure that enables them, to mimic the interacting protein surface.

The previously discovered cyclic peptide CLLFVY was shown to be the first compound capable of inhibiting HIF-1 α dimerisation in an isoform specific manner.¹⁴⁶ CLLFVY was discovered utilising a SICLOPPS library coupled with an RTHS that enabled the identification of peptides capable of inhibiting the dimerisation of HIF-1 α and HIF-1 β . CLLFVY was tagged with a cell-penetrating peptide and shown to inhibit the expression of HIF-1 regulated genes by qPCR without affecting HIF-2 specific genes.

This project aimed to improve the binding affinity of CLLFVY through an *in vitro* approach, via the systematic incorporation of unnatural amino acids. The improved compound was then to be used in the development of new *in vitro* assays that would probe the compounds activity as well as provide new tools for the development of the next generation of HIF inhibitors. The use of CLLFVY as a fluorescent probe for use in cell studies was also to be investigated, as well as trying to determine the structural basis for CLLFVY binding.

6.1 Exploring the Structure Activity Relationship of CLLFVY

CLLFVY was synthesised via an fmoc-SPPS strategy and the HIF-1 α PAS-B domain was obtained via recombinant expression. The affinity of CLLFVY for the protein was then assessed via MST and was found to have a K_d of $36.0 \pm 3.6 \mu\text{M}$. An alanine scan of the peptide was then carried out which revealed that the cysteine was contributing most significantly to binding of the HIF-1 α PAS-B domain. Derivatisation at this position revealed that a free thiol group was necessary for

binding. Despite requiring the presence of a free thiol group, CLLFVY does not bind via a covalent disulphide linkage. This was proven using mutant HIF constructs that were synthesised via SDM of the HIF-1 α PAS-B domain, swapping the three cysteines present for alanine (C255A, C334A and C337A). There was no significant loss of affinity between the mutants and CLLFVY, demonstrating that the interaction is not disulphide-mediated. The remaining positions of CLLFVY were then investigated with sequential replacement of each residue with a variety of analogues that enabled the SAR of CLLFVY to be probed. This led to the development of multiple peptides with greatly improved K_d values compared to the parent peptide. Upon derivatisation of the L2 position artefacts were observed to appear in the MST data, and an investigation was conducted to identify the cause of this artefact. The use of the competitive inhibitor DTT revealed that a secondary binding event was occurring, and the hypothesis proposed that the compounds had become sufficiently hydrophobic that they were starting to bind non-specifically to the HIF-1 α PAS-B domain. This was investigated via the incorporation of charged residues at the L1 position of the peptide. Incorporation of a glutamic acid or lysine residue abolished the secondary binding event, allowing for an accurate determination of the peptides K_d values as $1.21 \pm 0.2 \mu\text{M}$ and $2.08 \pm 0.4 \mu\text{M}$ respectively, representing a 30 and 17-fold improvement in affinity when compared to CLLFVY. Interestingly, the introduction of these charged residues was found to have a significant impact on the selectivity of the sequence. The positively charged CK-NL-F(CF₃)IF(Br) was found to bind to both the HIF-1 α PAS-B domain as well as the HIF-2 α PAS-B domain, whereas CE-NL-F(CF₃)IF(Br) bound preferentially to HIF-1 α with a 40-fold higher K_d for 1 α over 2 α .

These peptides achieved the aim of improving the affinity of CLLFVY for the HIF-1 α PAS-B domain, with the lead peptide CE-NL-F(CF₃)IF(Br) being 30 times more potent. This compound will therefore be a valuable tool for the development of future assays and inhibitors.

6.2 Development of *In Vitro* HIF Assays

This project also investigated the use of other *in vitro* techniques to further characterise the peptides ability to inhibit HIF. Initial attempts were made to probe the activity of the peptides via an ELISA assay that utilised the PAS-B domains of both HIF-1 α and HIF-1 β . The wild type proteins were shown to have weak affinity for one another and so mutant PAS-B domains were expressed. These were subsequently shown to have a 50-fold higher affinity for dimer formation. This model was then used to determine that the lead peptide, CK-NL-F(CF₃)IF(Br), was capable of inhibiting HIF dimerisation with an IC₅₀ value of $27.7 \pm 1.4 \mu\text{M}$. Surprisingly, neither the parent CLLFVY or the most potent peptide, CE-NL-F(CF₃)IF(Br), were found to inhibit HIF dimerisation by ELISA. An alternative protein model was chosen in order to corroborate this result, and the three-domain HIF-1 protein was obtained recombinantly. Initial tests via MST with the new protein model

revealed that CLLFVY had no affinity for the new protein model, but that both optimised peptide sequences did. However, an accurate K_d could only be determined for CK-NL-F(CF₃)IF(Br) ($K_d = 3.2 \pm 0.6 \mu\text{M}$) as a secondary binding event was seen during the analysis of CE-NL-F(CF₃)IF(Br). This could potentially be the result of the HIF dimer breaking apart upon treatment with CE-NL-F(CF₃)IF(Br).

As it was not possible to separate out the multiple events observed in the MST, an FP assay was utilised that was capable of detecting whether or not HIF was bound to DNA, thus enabling the activity of the peptides to be determined without a dependence on the dimer staying intact. In this assay, the parent CLLFVY was seen to be capable of preventing HIF from binding to DNA with an IC₅₀ of $76.5 \pm 8.0 \mu\text{M}$. Whilst partial activity was observed for both optimised peptides, they were also both observed to bind to the DNA at high concentrations, preventing an accurate determination of their IC₅₀ values. Therefore, the earlier derivative CLLF(CF₃)IF(Br) was also assessed and found to possess increased inhibitory activity with an IC₅₀ of $11.76 \pm 1.1 \mu\text{M}$. Critically, this showed that the improvement in affinity during the CLLFVY SAR study translates into improved inhibitory activity.

6.3 Development of a Fluorescent CLLFVY Derivative

In the pursuit of a fluorescent derivative of CLLFVY a coumaryl amino acid was synthesised. It was then incorporated into CLLFVY and CLLF(CF₃)VY to yield the peptides CLL-hCouVY and CLLF(CF₃)V-hCou. MST analysis of the two fluorescent peptides showed that they both displayed a good affinity for the HIF-1 α PAS-B domain, with K_d values of $10.6 \pm 1.2 \mu\text{M}$ and $23.2 \pm 0.3 \mu\text{M}$ respectively, both of which represent an improvement in affinity compared to CLLFVY ($36.04 \pm 3.6 \mu\text{M}$). HeLa cells treated with the two peptides showed that both compounds were visualised in the cytoplasm, however CLLF(CF₃)VY was observed to form aggregates. Treatment of HeLa cells with CLL-hCou-VY in the presence and absence of DFX was then investigated and the location of HIF-1 α within these cells was probed via immunofluorescence. These experiments revealed a loss of HIF-1 α associated fluorescence signal upon treatment with CLL-hCou-VY. In addition, HIF-1 α associated fluorescence was no longer observed in the nucleus, suggesting it is no longer transcriptionally active.

6.4 Investigating the Structural Basis of CLLFVY Binding

This project also aimed to elucidate the structural basis of CLLFVY binding via x-ray crystallography. Whilst pursuing this goal, conditions for the crystallisation of the HIF-1 α PAS-B domain were optimised and seeding experiments yielded large crystals. The crystals were used to

Chapter 6

obtain a high quality data set that was solved to reveal that the HIF-1 α PAS-B domain crystallised as a homodimer, representing a novel HIF structure. Attempts were made to obtain a ligand bound structure by both soaking and co-crystallisation. Whilst these attempts have thus far been unsuccessful they provide the groundwork for future efforts. Removal of the *in situ* thrombin cleave has also been achieved via the design of a new protein construct enabling the cleavage of the protein HIS-tag via TEV during the purification workflow. The cleaved HIF-1 α PAS-B was shown to form small crystals, and future work will be needed to re-optimize the crystallisation conditions to enable the growth of crystals of a suitable size for data collection.

Chapter 7 Future Directions

During the SAR study of CLLFVY, the stereochemistry at the phenylalanine position was demonstrated to have a significant impact on the affinity of the peptide, with D-phenylalanine binding with an affinity 12-fold better than the L analogue. This suggests that the conformation of the peptide plays an important role in its binding to the HIF-1 α PAS-B domain. To investigate the importance of the peptide's conformation, NMR techniques such as NOESY could be used to characterise the solution structures of the peptides synthesised in this work. This can then be used to identify the preferred conformation of the peptides for HIF binding.

Another consideration is that the optimised peptides have not yet been tested in mammalian cell lines. CLLFVY was demonstrated to have HIF-1 inhibitory activity in a number of cell based assays, including quantitative PCR, proximity ligation assays and a HRE-luciferase reporter assay.¹⁴⁶ As the lead peptides developed in this project have been shown to have improved HIF inhibitory activity *in vitro*, future experiments should be conducted to test whether they also perform better in a cellular environment.

During the SAR study of CLLFVY, it was discovered that the HIF-1 α PAS-B domain appeared to bind preferentially to compounds with a free thiol moiety, including to the reducing agent DTT. This knowledge could be used to screen a library of compounds containing thiol groups to identify other ligands capable of binding to HIF. It is also possible that the thiol group is key to the inhibition of HIF activity. This could be probed using the three-domain HIF FP assay whereby a selection of compounds containing thiol groups could be assessed for inhibitory activity. In addition, as DTT was shown to compete with CLLFVY for binding to the HIF-1 α PAS-B domain, it is reasonable to assume they are binding to a similar site within the protein. If a ligand bound crystal structure could be obtained for a HIF-DTT complex, then the site of binding, and reason for HIFs apparent thiophilicity, could be elucidated. The greatly improved solubility and smaller size of DTT makes it an ideal candidate for future crystal soaking experiments.

One of the most interesting results of this project is the impact of charge on the isoform selectivity of the peptides CE-NL-F(CF₃)IF(Br) and CK-NL-F(CF₃)IF(Br). The positively charged CK-NL-F(CF₃)IF(Br) binding to both HIF-1 α and HIF-2 α , whereas CE-NL-F(CF₃)IF(Br) was shown to possess 40-fold selectivity for binding to HIF-1 α . Future experiments should focus on probing the cause of these differences. As previously stated, developing isoform specific inhibitors of HIF has been especially challenging given the high sequence homology of the HIF- α isoforms, and so an understanding of how this selectivity arises would enable the development of HIF-1 α or HIF-2 α specific inhibitors. This task would be greatly helped by the acquisition of structural data to elucidate the interaction of the two peptides with the different protein isoforms.

Chapter 7

Future studies will also continue to pursue a crystal structure of CLLFVY bound to HIF. The novel HIF structure detailed in this project will provide a foundation for the continued development of soaking experiments, as well as providing the potential for cross-seeding experiments for the discovery of co-crystallisation conditions. A structure of the CLLFVY bound HIF protein would be an invaluable resource for the rational development of future HIF inhibitors.

Chapter 8 Experimental

8.1 Equipment and Materials

All reagents were purchased from Fisher Scientific (UK) or Sigma-Aldrich (UK) unless otherwise specified. Restriction enzymes and microbiology reagents were purchased from New England Biolabs Inc, Fisher Scientific (UK) or Promega and were used as directed by the manufacturer. Primers and oligonucleotides were synthesised by Integrated DNA technologies (IDT, UK). DNA sequencing was carried out by Eurofins MWG operon (Germany). PCR was carried out in a BIO-RAD T100 thermal cycler. DNA purification was carried out using GeneJet Plasmid Miniprep kit or GeneJet PCR Purification kit. DNA and protein concentrations were determined using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, USA). Protein extinction coefficients were estimated using ExPASy ProtParam tool. Bacterial cell cultures were incubated in a Thermo Scientific MaxQ 6000 shaking incubator. Optical densities of cell cultures were determined using a Cary 100 bio UV-Visible spectrophotometer (Agilent Technologies, UK). Refrigerated centrifugation <4000 rpm was carried out using a Heraeus Biofuge Primo R centrifuge (Thermo Electron Cororation, UK). Centrifugation at speeds >4000 rpm was carried out using a Heraeus Multifuge X3 FR system (Thermo Scientific, UK).

Agarose and SDS-PAGE gels were run using a BioRad Power Pac Basic, and the resulting gels analysed using a BIO-RAD Universal Hood II Gel Doc XR system. Protein purification was carried out using an ÄKTA Prime Plus FPLC system. A 1mL HisTrap Column (GE Healthcare) was used for purification of His-tagged proteins. A HiLoad 16/600 Superdex 75 or 200 pg column (GE Healthcare) was used for size exclusion chromatography. Buffer pH was determined using an Accument AB150 pH meter (Fisher Scientific).

Nickel coated 96-well plates were purchased from Thermo Fisher Scientific. FP and ELISA measurements were carried out on a CLARIOstar plate reader (BMG Labtech, UK). MST measurements were carried out using a Monolith NT.115 and premium coated capillaries (NanoTemper Technologies, Germany). Thermal shift measurements were made using a BioRad CFX Connect Real Time PCR System.

8.2 Preparation of Materials

8.2.1 Lysogeny Broth (LB)

LB Miller powder (25 g) was suspended in deionised water (1 L) and autoclaved for 20 minutes at 115 °C. The solution was then left to cool to room temperature before use.

8.2.2 LB Agar plates

LB agar powder (8.8 g) was dissolved in deionised water (250 mL) and autoclaved at 121 °C for 20 minutes on the liquid media cycle. The solution was then allowed to cool to 50 °C before addition of the relevant antibiotics. The liquid agar (25 mL) was then poured into petri dishes and left to solidify. The plates were then dried in an incubator at 37 °C for 30 minutes before use.

8.2.3 Super optimal culture (SOC) media

Sterile LB medium (50 mL) was supplemented with MgCl₂ (10 mM), MgSO₄ (20 mM) and glucose solution 0.2 % (W/V). The solution was mixed and then filtered through a 0.22 µm filter before being stored in 10 mL aliquots at 4 °C.

8.2.4 Antibiotics

Antibiotic stock solutions were made using freshly autoclaved deionised water with the exception of chloramphenicol, which was dissolved in ethanol. Solutions were made to the stock concentrations shown in Table 3 then filtered through a 0.22 µm sterile filter (Millipore, UK) and stored at 4 °C.

Table 3. Stock and working concentrations of antibiotics used.

Antibiotic	Stock Concentration (mg/mL)	Working concentration (µg/mL)
Kanamycin	50	50
Chloramphenicol	35	35
Carbenicillin	100	100

8.2.5 Preparation of chemically competent cells

A small amount of frozen BL21 Rosetta DE3 cells were added to LB broth (10 mL), containing chloramphenicol at the working concentration given, and incubated at 37 °C overnight. A 1 % subculture was taken and added to LB broth (25 mL) again containing chloramphenicol and incubated at 37 °C until OD₆₀₀ reached 0.6. The culture was then centrifuged at 3100 rpm for 15 minutes at 4 °C. The supernatant was discarded and the pellet re-suspended in TFB 1 Buffer prepared as described in Table 4 (25 mL). This was again centrifuged at 3100 rpm for 15 min at 4 °C and the supernatant again discarded. The pellet was then re-suspended in TFB 2 buffer prepared as described in Table 5 (1 mL). The solution was divided into 100 µL aliquots and flash frozen in dry ice before being stored at -80 °C.

Table 4. TFB I buffer composition

Chemical	Quantity	Final Concentration
Potassium acetate	0.59 g	30 mM
Rubidium Chloride	2.42 g	100 mM
Calcium Chloride	0.29 g	10 mM
Manganese chloride	2.00 g	50 mM
Glycerol	30 mL	15 % V/V

The solution was then adjusted to pH 5.8 with acetic acid and made up to 200 mL with autoclaved deionised water. The solution was filtered and 10 mL aliquots were frozen at -80 °C until needed.

Table 5. TFB II buffer composition

Chemical	Quantity	Final Concentration
MOPS	0.21 g	10 mM
Rubidium Chloride	0.12 g	75 mM
Calcium Chloride	1.10 g	10 mM
Glycerol	30 mL	15 % V/V

The solution was then adjusted to pH 6.5 with NaOH and made up to 100 mL with autoclaved deionised water. The solution was filtered and 1 mL aliquots were frozen at -80 °C until needed.

8.2.6 Agarose gels

1 % agarose gels were prepared by dissolving agarose powder (1 g) in Tris-Acetate EDTA (TAE) (100 mL) by heating in a microwave until dissolution was complete. The mixture was allowed to

cool and 2 drops of ethidium bromide (0.625 mg/mL) added. The mixture was stirred with gently swirling before being poured into a mould and allowed to set for 20-30 minutes.

Table 6. TAE buffer composition

Reagent	Quantity	Final Concentration
Tris base	242 g	2 M
Glacial acetic acid	57.1 mL	1 M
EDTA	14.6 g	50 nM
Deionised water	Up to 1L	-

8.2.7 Preparation of SDS-PAGE gels

Reagents shown in Table 7 were combined adding tetramethylethylenediamine (TEMED) and ammonium persulfate last to initiate polymerisation, and the mixture loaded between glass plates. The gel was levelled via the addition of isopropyl alcohol (1 mL) and left to set for 30 minutes. The isopropyl alcohol was then removed and stacking gel, prepared as described in Table 8, poured on top of the resolving gel. The desired comb was then added and the stacking gel allowed to set for ~20 minutes. The gels were then either used immediately or stored at 4 °C wrapped in a paper towel soaked in deionised water in a sealed bag.

Table 7. SDS-PAGE running gel (15 %) composition

Reagent	Volume (mL)
Water	5.4
1.5 M Tris-Base pH 8.8	3.75
10 % SDS (w/v)/ H ₂ O	0.15
Acrylamide/ Bisacrylamide (40 %/ 0.8 % w/v)	5.6
10 % APS (w/v) H ₂ O	0.075
TEMED	0.01
dH ₂ O	5.4

Table 8. SDS-PAGE stacking gel composition

Reagent	Volume (mL)
0.5 M Tris-Base (pH 6.8)	1.25
10 % SDS (w/v)/ H ₂ O	0.05
Acrylamide/ Bisacrylamide (40 %/ 0.8 %) (w/v)	0.63
10 % APS (w/v) H ₂ O	0.025

TEMED	0.005
dH ₂ O	3.1

8.2.8 SDS-PAGE loading buffer

Table 9. SDS-PAGE loading buffer (2x) composition

Reagent	Concentration
Tris-HCl	100 mM
Sodium dodecyl sulfate solution (10 % w/v)	4 % v/v
Bromophenol blue	0.04 %
EDTA	25 mM
Glycerol	40 % v/v
Dithiothreitol (DTT)	25 mM

8.2.9 SDS-PAGE running buffer

Table 10. SDS-PAGE running buffer composition

Reagent	Quantity
Tris Base	3.02 g
Glycine	18.8 g
Sodium dodecyl sulfate solution (20 % w/v)	5 mL
dH ₂ O	Up to 1 L

8.2.10 Coomassie protein stain solution

Table 11. Coomassie protein staining solution composition

Reagent	Quantity
Coomassie Brilliant Blue	1.0 g
Acetic acid	100 mL
MeOH	400 mL
dH ₂ O	500 mL

8.2.11 Coomassie de-staining solution**Table 12. Coomassie de-staining solution composition**

Reagent	Quantity
Acetic acid	100 mL
MeOH	200 mL
dH ₂ O	700 mL

8.2.12 Protein buffers**Table 13. HIF buffer**

Reagent	Quantity	Final concentration
NaCl	8.76 g	150 mM
Tris Base	6.06 g	50 mM
TCEP	0.286 g	1 mM
Glycerol	50 mL	5 % (v/v)
dH ₂ O	Up to 1 L	-

Reagents were combined to give the final concentrations stated above. The pH was then adjusted to 8.

Table 14. HIS-trap protein loading buffer

Reagent	Quantity	Final Concentration
Imidazole	0.068 g	20 mM
HIF buffer	Up to 50 mL	-

Table 15. HIS-trap protein elution buffer

Reagent	Quantity	Final Concentration
Imidazole	0.851 g	250 mM
HIF buffer	Up to 50 mL	-

Table 16. Labelling buffer

Reagent	Quantity	Final Concentration
NaHCO ₃	0.420 g	100 mM
NaCl	0.438 g	150 mM
dH ₂ O	Up to 50 mL	-

Table 17. HIF MST Assay Buffer

Reagent	Quantity	Final Concentration
Tween 20 (10 %) (v/v H ₂ O)	250 µL	0.05 % (v/v)
HIF buffer	Up to 50 mL	-

Table 18. HIF FP Assay Buffer

Reagent	Quantity	Final concentration
NaCl	0.584 g	20 mM
Tris Base	1.21 g	20 mM
dH ₂ O	Up to 500 mL	

Reagents were combined to give the final concentrations stated above. The pH was then adjusted to 8.

8.2.13 Annealing buffer

Table 19. Annealing buffer

Reagent	Final concentration
NaCl	50 mM
Tris Base	10 mM
EDTA	1 mM

The buffer was then adjusted to pH 8.

8.3 General biology procedures

8.3.1 Polymerase Chain Reaction (PCR)

PCR mixtures were prepared as described in Table 20. The solution was mixed thoroughly before adding the DNA polymerase. PCR reaction were carried out in a thermocycler using the programme in Table 21.

Table 20. Composition of PCR reaction mixture

	Volume (µL)	Final Concentration (µM)
Polymerase buffer (5x or 10x)	5 or 10	1x
dNTP solution (10 mM)	1	200
Forward primer	1	0.3
Reverse primer	1	0.3

DNA Template	1	Variable (≈ 0.2 ng/ μ L)
DNA polymerase	0.5	
Distilled water	Up to 50	

Table 21. PCR programme

Cycle	Denaturation	Annealing	Extension
1	95 °C for 3 minutes	-	-
2-34	95 °C for 3 minutes	X °C for 30 seconds	72 °C for Y minutes
35	-	-	72 °C for 5 minutes

X represents the annealing temperature, which is a variable dependent on primers. The annealing temperature used was calculated using the NEB melting temperature calculator. The time Y is calculated as 30 seconds per 500bp being amplified.

8.3.2 Colony PCR

Colonies were picked from agar plates and diluted in LB media (10 μ L). A master mix for the colonies and a water control were then prepared as described in Table 22. The colony solutions (1 μ L) or water (1 μ L) were then added to aliquots of the master mix (10 μ L) and PCR run as detailed in Table 21. Successful colonies were identified via agarose gel electrophoresis.

Table 22. Colony PCR master mix

	Volume (μ L)	Final Concentration (μ M)
GoTaq polymerase buffer (5x)	20	1x
dNTP solution (10 mM)	2	200
Forward primer (10 μ M)	3	0.3
Reverse primer (10 μ M)	3	0.3
DNA Template	1	Variable (≈ 0.2 ng/ μ L)
DNA polymerase	0.5	
Distilled water	Up to 50	

8.3.3 Plasmid Purification

Cells of the relevant *E. coli* strain were taken from a frozen stock (-80 °C) and grown in LB media (10 mL), containing the appropriate antibiotics, overnight at 37 °C. The resulting culture was then

centrifuged at 4000 rpm for 15 minutes at 4 °C. The plasmid was then extracted from the resulting pellet using a Thermo Scientific GeneJET plasmid Miniprep kit.

8.3.4 Transformation of chemically competent cells

Plasmid (5 µL) was added to of chemically competent cells (100 µL) and then incubated on ice for 30 minutes. The cells were then heat shocked at 42 °C for 30 seconds before being transferred to 895 µL of SOC medium and incubated at 37 °C for 1 hour. 100 µL of the culture was then transferred to 10 mL of LB medium containing chloramphenicol and kanamycin and was incubated at 37 °C overnight.

8.3.5 Restriction Digests

Digestion of plasmids and PCR products was carried out by mixing the DNA (1 µg) with the appropriate buffer and enzyme then incubating as per the manufacturer's instructions. Digested vectors were then treated with Shrimp Alkaline Phosphatase (New England Biolabs) as per the manufacturer's instructions. The digested products were then purified using a Thermo Scientific GeneJET PCR purification kit.

8.3.6 Ligations

The digested vector was mixed with either 3 or 5 molar equivalents of digested insert and 10 x T7 ligase buffer (1.5 µL) was added. T7 DNA ligase (1 µL) (Promega) was then added and the mixture made up to 15 µL with dH₂O. The ligation was then carried out for 1 hour at RT after which the solution (5 µL) was directly transformed into chemically competent cells.

8.3.7 Site-Directed Mutagenesis

PCR of the HIF-1α PAS-B pET-28a vector was performed using Pfu polymerase (Promega) and a forward and reverse primer containing the desired mutagenesis site from those shown below. The PCR product was then incubated with DpnI (1 unit) for 1 hour at RT to remove the template. The products were the purified using a Thermo Scientific GeneJET PCR purification kit. The purified product was then transformed into chemically competent DH5α cells.

C255A SDM F: GAAATTTTCTTATGCTGATGAAAGAATTAC

C255A SDM R: GTAATTCTTTCATCAGCATAAGAAAATTC

C334A SDM F: ATTCTCAACCACAGGCCATTGTATGTGTGA

Chapter 8

C334A SDM R: TCACACATACAATGGCCTGTGGTTGAGAAT

C337A SDM F: ACAGTGCATTGTAGCTGTGAATTACGTTGT

C337A SMF R: ACAACGTAATTCACAGCTACAATGCACTGT

8.3.8 Running SDS-PAGE gels

SDS-PAGE gels were made as detailed in section 8.2.7. Samples were then prepared by mixing SDS-PAGE loading buffer (8.2.8) (1:1) to a total volume of 20 μ L. The samples were then heated to 95 °C for 10 minutes in a PCR cycler. A ladder was added to the first lane and the samples loaded into subsequent wells. The gel was then run at 180 V for 90 minutes using a Bio Rad 165-8000 filled with SDS-PAGE running buffer (8.2.9). The gel was then stained by applying Coomassie stain solution (8.2.10) and leaving on a rocker for 30 minutes. The gel was then rinsed with water and de-stained via application of the de-stain solution (8.2.11) with rocking for 2 hours. The gels were then washed with water before being visualized using a Bio Rad Chemidoc MP imaging system.

8.3.9 Annealing HRE oligonucleotides

The 6-FAM-labelled forward strand: 5'-GGCTGCGTACGTGCGGGTCGT-3' (10 μ M) and unlabelled reverse strand: 5'-ACGACCCGCACGTACGCAGCC-3' (10 μ M) were mixed 1:1 in annealing buffer (final volume 50 μ L). The mixture was then heated to 95 °C for 2 minutes and a gradient carried out in a thermocycler reducing the temperature to 20 °C over 45 minutes.

8.3.10 Construction of the HIF-1 α PAS-B₍₂₃₈₋₃₄₂₎ vector

The HIF-1 α PAS-B₍₂₃₈₋₃₄₂₎ insert was amplified via PCR (section 8.3.1) using Q5 DNA polymerase from the pET28a HIF-1 α PAS-B mutant plasmid and the following primers:

Forward 5'-GAGTGCGCCGCAAGCTTGTGACGG-3'

Reverse 5'-GTTGTTCCATGGCGATGGATAGCAAGACTTTCCTC-3'

The purified insert and the pETM11 vector were then digested using NOT1 and NCO1 restriction enzymes as described in section 8.3.5. The purified digestion products were ligated as described in section 8.3.6. The plasmid was then transformed into *E. coli* DH5 α cells and grown on LB agar under kanamycin control.

8.4 Protein procedures

8.4.1 Expression and Purification of HIF-1 α PAS-B Domain

A 1 % (v/v) subculture of transformed cells was added to 1 L of LB broth supplemented with chloramphenicol and kanamycin. This culture was then incubated with shaking at 200 rpm at 37 °C until an OD₆₀₀ of 0.6 was achieved. IPTG (0.1 mM) was then added to induce protein expression and the culture incubated at 16 °C, for 18 hours. The cells were then pelleted by centrifugation at 4000 rpm for 15 minutes. The cells were then re-suspended in protein loading buffer (5 mL/g) and a crushed Pierce protease inhibitor tablet was added to the suspension. Cells were lysed via sonication (12 cycles 10 seconds on 20 seconds off) whilst on ice. The suspension was then centrifuged at 14,500 rpm for 1 hour. The pellet was discarded and the supernatant filtered through a 0.22 μ m sterile filter (Millipore UK). The supernatant was then loaded onto a 1 mL His Trap column and equilibrated in protein loading buffer. The bound protein was then eluted using protein elution buffer and the protein containing fractions combined. The combined fractions were then loaded onto an S75 size exclusion column pre-equilibrated in HIF buffer. The buffer exchanged and purified protein was eluted in HIF buffer and either, aliquoted and flash frozen in liquid nitrogen before being stored at -80 °C, or carried forward for labelling.

8.4.2 Expression and Purification of HIF-1 Three Domain

A 1 % (v/v) subculture of cells co-transformed with plasmids containing HIF-1 α bHLH-PAS A-PAS-B and HIF-1 β bHLH-PAS A-PAS-B were added to 2 L of LB broth supplemented with chloramphenicol, kanamycin and carbenicillin. This culture was then incubated with shaking at 200 rpm at 37 °C until an OD₆₀₀ of 0.6 was achieved. IPTG (0.2 mM) was then added to induce protein expression and the culture incubated at 18 °C, for 18 hours. The cells were then pelleted by centrifugation at 4000 rpm for 15 minutes. The cells were then re-suspended in protein loading buffer (5 mL/g) and a crushed Pierce protease inhibitor tablet and DNase were added to the suspension. Cells were lysed via sonication (12 cycles 10 seconds on 20 seconds off) whilst on ice. The suspension was then centrifuged at 14,500 rpm for 1 hour. The pellet was discarded and the supernatant filtered through a 0.22 μ m sterile filter (Millipore UK). The supernatant was then loaded onto a 1 mL His Trap column and equilibrated in protein loading buffer. A wash step of NaCl (1M) was then carried out to elute any non-specifically bound DNA. The bound protein was then eluted using protein elution buffer and the protein containing fractions combined. The combined fractions were then loaded onto an S200 size exclusion column pre-equilibrated in HIF buffer. The buffer exchanged and purified protein was eluted in HIF buffer and either, aliquoted and flash frozen in liquid nitrogen before being stored at -80 °C, or carried forward for labelling.

8.4.3 Protein labelling

A 150 μL aliquot of the protein fraction obtained from the S75 was taken and buffer exchanged into labelling buffer using a PD SpinTrap G-25 column. The concentration of protein was adjusted to 100 μL at 20 μM using labelling buffer and 100 μL of 60 μM NT-NHS647 dye in DMSO was added to the mixture. It was then left to incubate at room temperature in the dark for 30 minutes. The dyed protein was then purified and simultaneously buffer exchanged into assay buffer on a PD MiniTrap G-25 column to yield dyed protein. The dyed protein was separated into 10 μL aliquots flash frozen in liquid nitrogen and stored at $-80\text{ }^{\circ}\text{C}$.

8.4.4 Thrombin/ TEV Cleavage

Following purification by size exclusion chromatography the protein was concentrated to a volume $<5\text{ mL}$ and the concentration determined on the Nanodrop ND-1000. 10 units of enzyme were added for each mg of protein and the mixture then incubated on a rocker at $4\text{ }^{\circ}\text{C}$ overnight. The protein solution was then loaded onto a HisTrap column and the cleaved protein eluted with loading buffer. This fraction was then further purified by size exclusion chromatography before being concentrated to the required concentration, aliquoted and flash frozen in LN_2 then stored at $-80\text{ }^{\circ}\text{C}$.

8.5 Assay protocols

8.5.1 Thermal shift

A master mix (64 μL) was made for each condition consisting of the protein of interest (15 μM) and reagent to be tested, in HIF buffer. When peptides were to be tested DMSO (10 % v/v) was also added). SYPRO orange dye (5000x) (Life Technologies) was diluted 1:9 in HIF buffer and then added to the master mix (1 μL). The master mixes (3 x 20 μL) were then pipetted into a BioRad Hard Shell Clear-bottomed 96 well plate. The assay was then conducted in a BioRad CFX Connect Real Time PCR system with a temperature gradient run from $20\text{-}95\text{ }^{\circ}\text{C}$ in $0.5\text{ }^{\circ}\text{C}$ increments. The fluorescence intensity of the dye was measured and the data plotted in GraphPad Prism 8 software. Melting temperatures were determined by generating the first derivative of the obtained curves and then finding the X value which gave the highest fluorescence reading.

8.5.2 HIF PAS-B MSTs

An aliquot of labelled HIF-1 α PAS-B was thawed on ice and diluted to 50 nM with HIF MST assay buffer, before being centrifuged at 12,000 rpm for 15 minutes. The compound was then serially

diluted in DMSO to give 10 x stocks. These were then diluted 1:5 in HIF MST assay buffer to give 2 x stocks, before finally being mixed 1:1 (v/v) with the labelled protein to give a final concentration of 25 nM protein and a final DMSO content of 10 %. Samples were loaded into NanoTemper® Monolith NT.115 premium treated capillaries and measurements were carried out using a Monolith NT.115 system at 25 °C using 50 % MST power and 50 % LED power. Data was analysed by NanoTemper® analysis software. Binding curves were fitted using GraphPad Prism 8 software.

8.5.3 Enzyme-Linked Immunosorbent Assay

To each well of a clear Pierce nickel coated 96-well plate (Thermo Scientific), purified His₆-HIF-1 α PAS-B protein in HIF buffer was added (25 μ L, 0.1 μ M), the plate was then incubated for 1 hr at RT with rocking. The wells were then washed 2 x with HIF buffer (150 μ L) and 1 x with PBS Tween (0.05 % Tween 20, 150 μ L). 2 % milk in PBS Tween (150 μ L) was then added to each well and incubated for 1 hr at RT with rocking as a blocking step. The wells were then washed 2 x PBS Tween (150 μ L) and 1 x with HIF buffer (150 μ L). For HIF-B dimerisation a serial dilution of HIF-1 β PAS-B ranging from 250 μ M to 11 nM in HIF buffer (100 μ L) was then added. For CLLFVY inhibition a serial dilution of CLLFVY ranging from 500 μ M to 15 nM in 100 % DMSO (10 μ L) was added at the same time as HIF-1 β PAS-B (10 μ M, 90 μ L) for a 10 % final concentration of DMSO. The plate was then incubated at RT for 1 hr with rocking. The wells were then washed 2 x with HIF buffer (150 μ L) and 1 x with PBS Tween (150 μ L). A 1:1000 dilution of mouse anti-flag antibody in 2 % milk PBS Tween (50 μ L) was then added to each well and the plate incubated for a further hour at RT with rocking. The wells were then washed with 3 x PBS Tween (150 μ L) with 5 minutes incubation per wash. A 1:6000 dilution of sheep anti mouse linked to horse radish peroxidase in 2 % milk PBS Tween (50 μ L) was then added and a further incubation of 1 hr at RT with rocking carried out. The plate was then washed as before and 1-Step TMB Ultra Solution (Thermo Scientific) (100 μ L) was added to each well and incubated for 15 minutes. H₂SO_{4(aq)} (2M, 100 μ L) was then added to each well and the absorbance at 450 nm determined on a BMG Clariostar plate reader.

8.5.4 HIF Three Domain MSTs

An aliquot of labelled HIF-1 three-domain protein was thawed on ice and diluted to 50 nM with HIF MST assay buffer, before being centrifuged at 12,000 rpm for 15 minutes. The compound was then serially diluted in DMSO to give 10 x stocks. These were then diluted 1:5 in HIF MST assay buffer to give 2 x stocks, before finally being mixed 1:1 (v/v) with the labelled protein to give a final concentration of 25 nM protein and a final DMSO content of 10 %. Samples were incubated for 30 minutes at room temperature before being loaded into NanoTemper® Monolith NT.115 premium treated capillaries. Measurements were carried out using a Monolith NT.115 system at

Chapter 8

25 °C using 50 % MST power and 50 % LED power. Data was analysed by NanoTemper® analysis software. Binding curves were fitted using GraphPad Prism 8 software.

8.5.5 Fluorescence Polarization Assays

HIF-1 three-domain protein was thawed on ice and diluted to 150 nM in HIF FP buffer. Compounds for analysis were serially diluted in DMSO to give 10x stocks. The compounds (10 µL, 100 % DMSO) were then added to the protein (80 µL). The samples were then incubated at room temperature for 30 minutes. FAM-HRE (10 µL, 20 nM) was then added and the samples loaded into a 384 well solid black non-binding microplate (Corning, USA). The fluorescence polarization of each well was then measured using gain and focus settings optimized by the used software set to a signal of 400 mP on a control well.

8.5.6 MST SD-Test

The three highest and three lowest concentrations (5000, 2500, 1250 µM and 600, 300, 150 nM) of the peptide to be tested (2 µL, 100 % DMSO) were added to the HIF-1α PAS-B domain (18 µL) and the samples centrifuged at 15,000 rpm for 10 minutes. The supernatant of each (10 µL) was then added 1:1 with the SD mix (4% SDS, 40 mM DTT) and the mixtures incubated at 95 °C for 5 minutes. The samples were then loaded into NanoTemper® Monolith NT.115 premium treated capillaries and the initial fluorescence of each measured at 50 % LED power on a Monolith NT.115 system.

8.6 Chemical procedures

8.6.1 Wang Resin SPPS

Wang resins pre-loaded with Fmoc amino acids were swollen in DMF for 10 minutes with agitation via argon flow. Fmoc deprotection was achieved via the addition of piperidine 20 % (v/v) in DMF. Fmoc protected amino acids were then coupled to the resin via the addition of a solution of the amino acid (3 eq.), Oxyma Pure (3 eq.) and DIC (3 eq.) in DMF and leaving the mixture with agitation for 1 hour. The resin was then washed with 3 x DMF, 3 x DCM and 3 x Et₂O and coupling confirmed via Kaiser test. Deprotection of the Fmoc group was again achieved via the addition of piperidine 20 % (v/v) in DMF with agitation for 30 minutes. The resin was washed as previously described and deprotection confirmed via Kaiser Test. These steps were repeated until the desired resin bound sequence was achieved.

8.6.2 Cleavage of Wang Resin

Cleavage of the linear sequence and simultaneous deprotection of side chain protecting groups was achieved using a cleavage cocktail of TFA (4.75 mL), H₂O (0.125 mL) and TIS (0.125 mL) and stirring for 2 hours at RT. The resin was then removed via filtration and the TFA removed under reduced pressure. The peptide was then precipitated with ice cold (-80 °C) diethyl ether. The ether was decanted and the peptide dried under reduced pressure.

8.6.3 2-Chlorotrityl chloride resin Loading

2-Chlorotrityl chloride resin (1 eq.) was placed in a sinter funnel and DCM (10 mL) added. The resin was then agitated via argon flow for 15 minutes to swell the resin. The DCM was then removed from the funnel and the Fmoc amino acid to be loaded (1 eq.) was dissolved in DCM (10 mL) and added to the resin. DIPEA (3 eq.) was then added and the resin agitated via argon flow for 2 hours. MeOH (1 mL) was then added to the funnel to end cap any unreacted sites and the resin agitated with argon flow for a further 15 minutes. The reaction mixture was then removed under reduced pressure and the resin washed with 3x DCM, 3x DMF and 3x Et₂O. The resin was then dried under reduced pressure.

8.6.4 Determining Fmoc-AA resin loading

Approximately 10 mg of dry resin was accurately weighed into a glass vial and DMF (800 µL) added. The resin was then left to swell for 15 minutes. Piperidine (200 µL) was then added and the resin left for a further 15 minutes. An aliquot of the reaction mixture (100 µL) was then diluted with DMF (900 µL). The resin loading was then determined from the UV absorbance of the dibenzofulvene-piperidine adduct at a wavelength of 301 nm as determined using a NanoDrop ND-1000 Spectrophotometer (NanoDrop Technologies, USA).

8.6.5 2-Chlorotrityl chloride resin SPPS

Pre-loaded 2-chlorotrityl chloride resins were deprotected with 20 % piperidine in DMF. The linear sequence was then synthesised as described for Fmoc-Wang resin synthesis.

8.6.6 2-Chlorotrityl chloride resin cleavage

To yield protected linear peptides, linear bound 2-chlorotrityl chloride resin was transferred to a vial and 1,1,1,3,3,3-hexafluoroisopropanol (HFIP) (5 mL) and DCM (5 mL) added. The mixture was then stirred at RT for 2 hrs. The resin was then removed via filtration and the solvent removed

Chapter 8

under reduced pressure. The peptide was then precipitated using ice cold (-80 °C) diethyl ether. The ether was decanted and the peptide dried under reduced pressure.

8.6.7 Cyclisation via EDC/ Oxyma Pure

Linear peptides were dissolved in DMF (1 mL/ mg) and EDC (3 eq.) and Oxyma pure (3eq.) were added. The mixture was then stirred at RT overnight. Excess DMF was removed under reduced pressure and the crude solution purified via reverse phase flash chromatography.

8.6.8 Cyclisation Via HATU

Linear peptides were dissolved in DMF (1 mL/ mg) and HATU (3 eq.) DIPEA (3 eq.) and HOAt (1eq.) were added. The mixture was then stirred at RT overnight. Excess DMF was removed under reduced pressure and the crude solution purified via reverse phase flash chromatography.

8.6.9 Cys(StBu) deprotection

Cys(StBu) containing sequences were deprotected by dissolving the cyclic peptide in DMF (1 mL) and adding DTT (10 eq.) and DIPEA (10 eq.) the mixture was then stirred for 30 minutes at RT. The mixture was then filtered through a 0.22 µm PTFE syringe filter and purified via reverse-phase HPLC.

8.6.10 Deprotection of protected cyclic peptides

Protected cyclic peptides were dissolved in TFA cleavage cocktail (TFA (4.75 mL), H₂O (0.125 mL) and TIS (0.125 mL)) and stirred at room temperature for 1 hour. Excess TFA was then removed under reduced pressure.

8.6.11 Kaiser free amine test

After each amino acid coupling step and Fmoc deprotection step, a few beads were transferred into a vial and 10 drops of KCN (0.02 mM) in pyridine and 3 drops of ninhydrin (0.30 M) in ethanol were added. The mixture was then heated to 130 °C for 2 minutes and the presence or absence of blue colour determined if the step was complete.

8.6.12 Purification via HPLC

Peptides were manually injected as solutions in either DMF or ACN via a Waters Flex inject system into a RP-HPLC system consisting of a Waters 1525 binary pump coupled to a Waters 2998

photodiode array detector. Peptides were then purified using a Waters X Select CSH prep C18 5 μ m OBD 19 x 250 mm column. A binary solvent system consisting of solution A (0.1 % TFA / water) and B (0.1 % TFA / MeCN) was run at the gradient shown below and elution monitored at 220 nm (peptide backbone). Fractions were collected automatically when the UV trace was in excess of a set threshold. Like fractions were combined and the ACN removed under reduced pressure. The sample was the lyophilised to yield the desired product.

% A	% B	Time (min)
95	5	0
10	90	0-15
10	90	15-20

8.6.13 Analytical HPLC

Analytical HPLCs were carried out on an Agilent 1260 Infinity II HPLC system running a binary solvent system consisting of solution A (0.1 % TFA / water) and B (0.1 % TFA / MeCN). Peptides were injected onto a Poroshell 120 EC-C18 column (2.7 μ m particle size, 3.0 x 100mm) and the gradient shown below run at a flow rate of 0.625 mL/ min. UV absorbance was recorded at both 220 and 280 nm.

% A	% B	Time (min)
95	5	0
0	100	0-20
0	100	20-22

8.6.14 Purification via Flash Chromatography

A Biotage Isolera One system equipped with a SNAP Ultra 30 g HP-Sphere C18 25 μ m column was used. A binary solvent system consisting of solution A (0.1 % TFA / water) and B (0.1 % TFA / MeCN) was run at 80 % A 20 % B for 3CV at 25 mL / min the peptide was then injected directly onto the column as a solution in either DMF or ACN. A gradient was then run with the solvent steps shown below. Fractions were collected automatically when UV_{220nm} exceeded 40 mAu. Collected fractions were analysed via mass spectrometry to confirm their identity and the desired fractions were concentrated under reduced pressure and then lyophilised to yield the desired product.

% A	% B	Number of column volumes
80	20	2
10	90	7
10	90	2

8.6.15 ESI⁺MS

Samples were analysed using a Waters (Manchester, UK) TQD mass spectrometer equipped with a triple quadrupole analyser. Samples were introduced to the mass spectrometer *via* an Acquity H-Class quaternary solvent manager (with TUV detector at 254nm, sample and column manager). Ultra-performance liquid chromatography was undertaken *via* a Waters BEH C18 column (50 mm x 2.1mm 1.7µm).

Gradient 20% acetonitrile (0.2% formic acid) to 100% acetonitrile (0.2% formic acid) in five minutes at a flow rate of 0.6 mL/min. Low-resolution mass spectra were recorded using positive ion electrospray ionisation.

8.6.16 High Resolution Mass Spectrometry (HRMS)

Samples were analysed using a MaXis (Bruker Daltonics, Bremen, Germany) mass spectrometer equipped with a Time of Flight (TOF) analyser. Samples were introduced to the mass spectrometer via a Dionex Ultimate 3000 autosampler and uHPLC pump. Gradient 20% acetonitrile (0.2% formic acid) to 100% acetonitrile (0.2% formic acid) in five minutes at 0.6 mL/min. Column, Acquity UPLC BEH C18 (Waters) 1.7 micron 50 x 2.1 mm. High resolution mass spectra were recorded using positive ion electrospray ionisation.

8.6.17 ¹H and ¹³C NMR

NMR spectra were recorded on Bruker AVII400 or Bruker AVIIHD400 FT-NMR spectrometers in the indicated solvent at 298 K. Chemical shifts for proton and carbon spectra are reported on the delta scale in ppm and were referenced to residual solvent references or internal TMS reference.

8.7 HeLa cell experiments

8.7.1 Assessment of peptide permeability

Treatment and imaging was carried out by Dr Cyrielle Doigneaux using the following protocol. 5.0 x 10⁴ HeLa cells were plated on glass coverslips in 35 mm dishes in a final volume of 1.5 mL of

DMEM + 10% FBS per well. The cells were then incubated overnight at 37 °C. The media was then removed and the cells were rinsed with cold PBS (1 mL). 1.5 mL of fresh warm media containing CLL-hCou-VY or CLLF(CF₃)V-hCou each at 50 µM (1% final DMSO concentration) or media containing 1% DMSO only (as control) was added to the cells. When a treatment with DFX was carried out, DFX was added to the cell media at a final concentration of 100 µM. The cells were subsequently incubated overnight at 37 °C. Following incubation, the media was removed and the cells washed with cold PBS (three times 1mL). The cells were then fixed by incubating them in 1 mL of PBS containing 4% of PFA for 10 minutes at room temperature. Following this the slides were then dried in the dark and the coverslips mounted on slides with DAPI-free mounting medium. The slides were sealed with nail polish and stored at -20 °C. Fixed cells were imaged using a Nikon plan apochromat 63x oil-immersion objective on a Leica SP8 inverted scanning confocal microscope. A 405 nm solid-state laser line and PMT detection was utilised for visualisation of coumarin.

8.7.2 HIF-1 α immunofluorescence experiments

Treatment and imaging was carried out by Dr Cyrielle Doigneaux using the following protocol. 5.0 x 10⁴ HeLa cells were plated on glass coverslips in 35 mm dishes in a final volume of 1.5 mL of DMEM + 10% FBS per well. The cells were then incubated overnight at 37 °C. The media was then removed and the cells were rinsed with cold PBS (1 mL). 1.5 mL of fresh warm media containing CLL-hCou-VY or CLLF(CF₃)V-hCou each at 50 µM (1% final DMSO concentration) or media containing 1% DMSO only (as control) was added to the cells. When a treatment with DFX was carried out, DFX was added to the cell media at a final concentration of 100 µM. The cells were subsequently incubated overnight at 37 °C. Following incubation, the media was removed and the cells washed with cold PBS (three times 1mL). The cells were then fixed by incubating them in 1 mL of PBS containing 4% of PFA for 10 minutes at room temperature. The solution was then removed and the cells were washed with PBS (3 x 1 mL) and were subsequently permeabilized using 1 mL of 0.5 % TX-100 in PBS (room temperature for 15 minutes). The permeabilization solution was then removed and the cells rinsed with PBS (3 x 1 mL). 1 mL of blocking buffer (1 % BSA and glycine (22.52 mg/mL) in PBST (PBS + 0.1 % Tween-20)) was then added to each well and the plate was incubated at room temperature for 1 hour. The blocking solution was then removed and 1.5 mL of anti-HIF-1 α antibody (NB100-449, Novus biologicals) diluted 1:100 in 1% BSA/PBST was added. The plate was then incubated overnight at 4 °C in a humidified chamber. The primary antibody solution was then removed and the cells washed with PBS (3 x 1 mL with 5 minute incubation on a rocker). 1 mL of a solution of secondary antibody anti-rabbit coupled to Alexa 568 (A11011, Thermofisher) at a 1:1000 dilution in 0.2 % BSA/PBST was then added to each well and

the plate was incubated for 1 hour at room temperature. The secondary antibody solution was then removed and the cells washed with PBS (3 x 1mL with 5 minute incubation on a rocker). The slides were then dried in the dark and the coverslips mounted on slides with DAPI-free mounting medium. The slides were sealed with nail polish and stored at -20 °C. Fixed cells were imaged using a Nikon plan apochromat 63x oil-immersion objective on a Leica SP8 inverted scanning confocal microscope. A 561 nm solid-state laser line and PMT detection at 580-700 nm was used for visualisation of the Alexa568 probe, and a wavelength of 405 nm with PMT detection was utilised for visualisation of coumarin.

8.8 Crystallography

8.8.1 Crystallisation of HIF-1 α PAS-B 4x using Dr Andrew Fosters conditions

HIF-1 α PAS-B 4x domain was expressed as previously described (section 8.4.1) and concentrated to 8.0 mg /mL in HIF buffer. A solution of succinic acid (0.8 M, pH7, 500 μ M) was then prepared. The two solutions were then mixed in 1:1, 1:3 and 3:1 ratios to yield 3 droplets on glass slides. Each slide was then suspended over a reservoir of the mother liquor (succinic acid (0.8 M, pH7, 500 μ L)) in a 24 well VDX plate (Hampton Research) sealed with vacuum grease. The plate was then incubated at 18 °C for seven days and imaged using a polarised light source. Crystals for data collection were picked using a 0.4-0.5 mm cryo-loop and placed into a solution of 30 % glycerol in mother liquor for approximately 5 seconds before being frozen in liquid nitrogen. The crystals were then stored in liquid nitrogen until data collection. Crystals were analysed at Diamond Light Source on beamline I04 using a wavelength of 0.9159 Å, and beam size of 20 x 20 μ m. The structure of the HIF-1 α PAS-B homodimer was solved by molecular replacement with the program Phaser-MR from the phenix software suite, using the HIF-1 PAS-B heterodimer (PDB: 4H6J) as the search model.^{16,251,252} Further manual model building was carried out using the software Coot with refinement by phenix.refine.^{251,253} All structural figures were prepared using PyMol (The PyMol Molecular graphics system, Version 2.0, Schrödinger, LLC).

8.8.2 Crystal soaking experiments

Crystals were grown in a hanging drop set up as previously described (section 8.8.1). A solution of CLLFVY (0.5 μ L, 50 mM 100 % DMSO or 0.5 μ L, 500 μ M, 20 % DMSO) was then added to the droplets containing crystals and the mixture left to soak for either 0.5, 3 or 24 hours. Intact crystals were then picked with a cryo-loop and transferred to a droplet containing 30 % glycerol in mother liquor before being flash frozen in liquid nitrogen.

8.8.3 Co-Crystallisation experiments

8.8.3.1 Succinic acid conditions

HIF-1 α PAS-B 4x (8.0 mg/mL) was incubated with CLLFVY (10 % DMSO) in either 1:1 or 1:5 molar ratios, for 1 hour. The sample was then centrifuged at 10,000 rpm for 15 minutes. Droplets were then set up on a glass slide in ratios of 1:1, 1:3 and 3:1 Protein-CLLFVY: succinic acid (0.8 M, pH7) and 1000x seed stock (0.5 μ L) added to each droplet. The slide was then suspended over a reservoir of succinic acid (0.8 M, pH7, 500 μ L) in a hanging drop set up. The plate was incubated for 7 days at 18 °C before being visualised using a polarised light source.

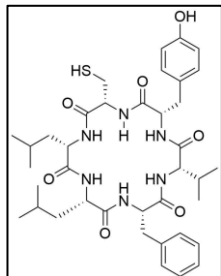
8.8.3.2 Broad screen conditions

HIF-1 α PAS-B 4x (8.0 mg/mL) was incubated with CLLFVY (10 % DMSO) in a 1:1 molar ratio. A 96-well sitting drop plate then had the buffers from the HR2-110 and HR2-112 screens (50 μ L) (Hampton research) added to the reservoir of each well. Droplets in 1:1, 1:3 and 3:1 ratios of protein-CLLFVY:reservoir were then added to the sitting drop wells and the plate sealed. The plate was then incubated for 7 days at 18 °C and visualised under a polarised light source.

8.9 Compound Characterisation

8.9.1 CLLFVY

Cyclo-CLLFVY

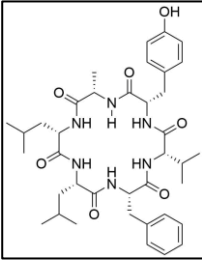
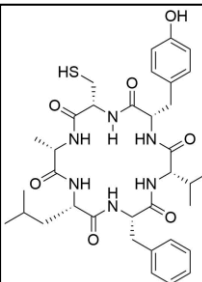
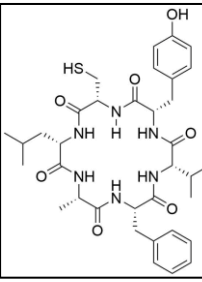
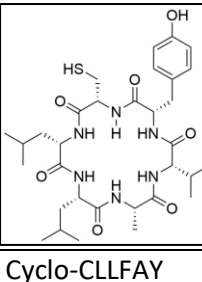
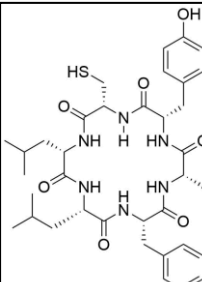


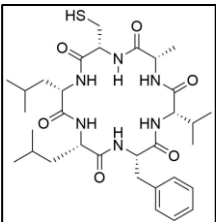
The linear resin bound peptide Resin-LCYVFL was synthesised using Wang resin SPPS as outlined in section 8.6.1. The linear peptide was then cleaved from the resin as described in section 8.6.2. The linear peptide was then cyclised using EDC and Oxyma Pure (section 8.6.7), before being purified via reverse-phase flash chromatography (section 8.6.14), to yield the StBu protected cyclic peptide C(StBu)LLFVY. The StBu group was then removed as

described in section 8.6.9 and the peptide purified via reverse-phase HPLC (section 8.6.12), like fractions were combined and concentrated under reduced pressure, then lyophilised to yield a white powder (11.8 mg, 8.0 % overall yield). Analytical HPLC $R_t = 5.290$ min (93 % purity). m/z (ESI⁺): 739.8 [M+H]⁺ (100 %), 1479.2 [2M+H]⁺ (50 %). HRMS (ESI⁺) found: 739.3833, [M+H]⁺ calculated: 739.3847.

8.9.2 CLLFVY Alanine Scan Derivatives

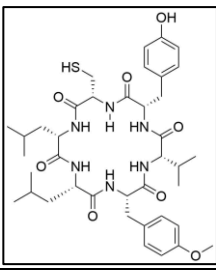
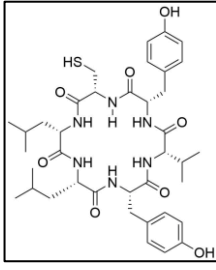
Linear resin bound peptides were synthesised using Wang resin SPPS as outlined in section 8.6.1. The linear sequences were then cleaved from the resin as described in section 8.6.2. The linear peptides were then cyclised using EDC and Oxyma Pure (section 8.6.7) before being purified via reverse-phase flash chromatography (section 8.6.14). StBu protected peptides were then deprotected (section 8.6.9), and the final cyclic peptides were purified via reverse-phase HPLC (section 8.6.12), the product containing fractions were combined and concentrated under reduced pressure, then lyophilised to yield the desired cyclic peptides.

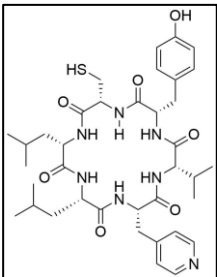
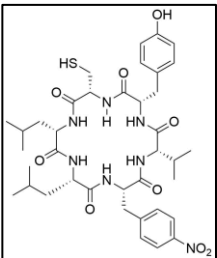
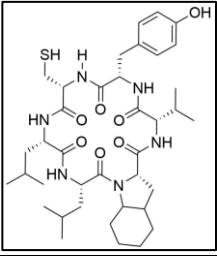
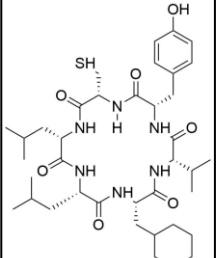
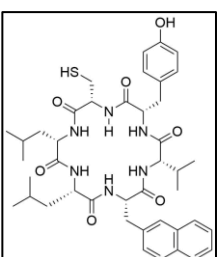
Compound	Yield (mg, %)	Analytical HPLC R _t (min)	Analytical HPLC Purity (%)	m/z (ESI ⁺)	HRMS (ESI ⁺) Calculated	HRMS (ESI ⁺) Found
Cyclo-ALLFVY 	13.7, 13.3	13.177	94	707.7 [M+H] ⁺ (100 %), 1437.2 [2M+H] ⁺ (50 %)	707.4127	707.4132
Cyclo-CALFVY 	18.7, 18.5	11.847	92	697.8 [M+H] ⁺ (100 %), 1416.3 [2M+Na] ⁺ (35 %)	697.3367	697.3378
Cyclo-CLAFVY 	24.9, 24.6	11.972	95	697.7 [M+H] ⁺ (100 %), 1395.5 [2M+H] ⁺ (50 %)	697.3378	697.3386
Cyclo-CLLAVY 	14.7, 15.2	11.356	99	663.7 [M+H] ⁺ (100 %), 1348.3 [2M+Na] ⁺ (40 %)	663.3534	663.3539
Cyclo-CLLFAY 	16.2, 11.4	12.824	97	711.7 [M+H] ⁺ (100 %), 1422.3 [2M+H] ⁺ (70 %)	711.3534	711.3536

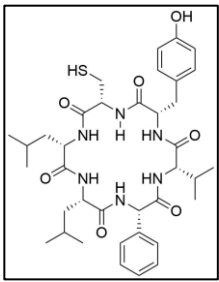
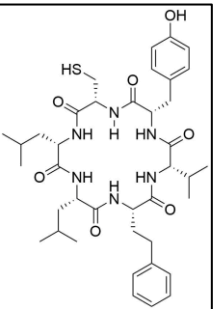
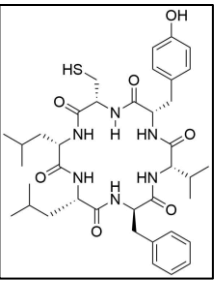
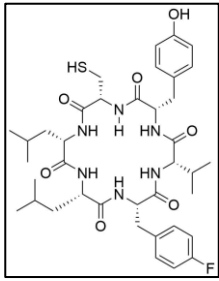
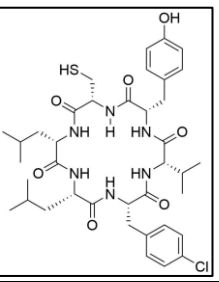
Cyclo-CLLFVA	8.0, 6.19	13.612	94	647.7 [M+H] ⁺ (100 %), 1316.3 [2M+Na] ⁺ (50 %)	647.3585	647.3576
						

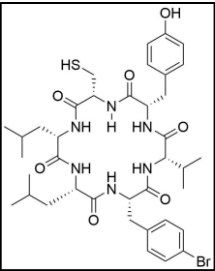
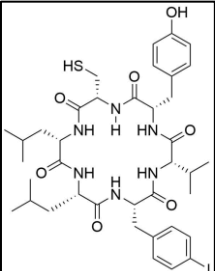
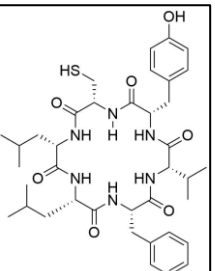
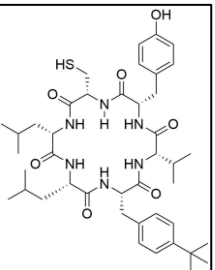
8.9.3 CLLFVY Phenylalanine derivatives

The linear resin bound sequence resin-LLCYV-NH₂ was synthesised using Wang resin SPPS (8.6.1). The resin was then split equally into 10 syringes and swelled in a small amount of DMF. A solution of unnatural amino acid (1 eq.), Oxyma pure (3 eq.) and DIC (3 eq.) was added and the syringes placed on a rocker for 1 hour. The reaction mixtures were then removed via vacuum filtration and the resins washed with DMF x 2, DCM x 2 and Et₂O x 1. Fmoc protected amino acids were then deprotected via the addition of piperidine 20 % in DMF and rocking for 30 minutes. The resins were then washed as before. The linear sequences were then cleaved from the resin as described in section 8.6.2, and cyclised using HATU, HOAt and DIPEA as described in section 8.6.8. The StBu group was then removed (section 8.6.9) and the peptide purified via reverse-phase HPLC (section 8.6.12), like fractions were combined and concentrated under reduced pressure, then lyophilised to yield the desired cyclic peptides.

Compound	Yield (mg, %)	Analytical HPLC R _t (min)	Analytical HPLC Purity (%)	m/z (ESI ⁺)	HRMS (ESI ⁺) Calculated	HRMS (ESI ⁺) Found
Cyclo-CLLY(Me)VY	3.7, 3.3	13.419	90	769.8 [M+H] ⁺ (100 %), 791.8 [M+Na] ⁺ (25 %)	769.3953	769.3944
						
Cyclo-CLLYVY	25.1, 22.2	11.830	99	755.7 [M+H] ⁺ (100 %)	755.3797	755.3791
						

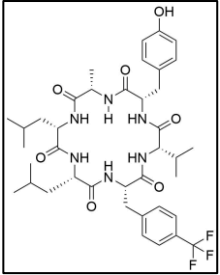
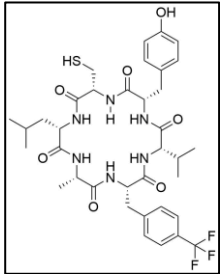
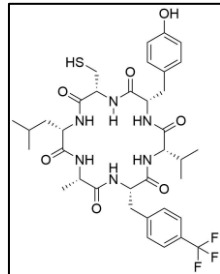
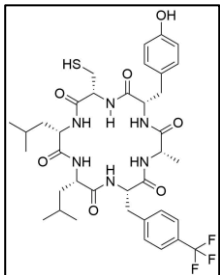
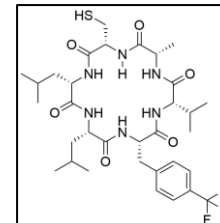
<p>Cyclo-CLL-PAL-VY</p> 	16.2, 15.1	9.721	89	740.8 [M+H] ⁺ (100 %)	740.3800	740.3817
<p>Cyclo-CLLF(4-NO₂)VY</p> 	11.5, 10.1	13.580	93	784.8 [M+H] ⁺ (100 %)	784.3698	784.3695
<p>Cyclo-CLL-Oic-VY</p> 	2.9, 2.7	14.282	92	743.8 [M+H] ⁺ (100 %)	743.4160	743.4151
<p>Cyclo-CLL-Cha-VY</p> 	3.0, 2.9	15.188	99	745.8 [M+H] ⁺ (100 %), 1491.6 [2M+H] ⁺ (90 %)	745.4317	745.4309
<p>Cyclo-CLL-Nal-VY</p> 	14.0, 8.9	15.317	92	789.7 [M+H] ⁺ (100 %), 811.6 [M+Na] ⁺ (35 %)	789.4004	789.3997

Cyclo-CLL-Phg-VY 	22.3, 15.4	12.903	87	725.7 [M+H]⁺ (100 %), 1450.1 [2M+H]⁺ (45 %)	725.3691	725.3687
Cyclo-CLL-hF-VY 	29.5, 19.3	14.175	87	753.7 [M+H]⁺ (100 %), 1149.6 [2M+H]⁺ (45 %)	753.4004	753.3991
Cyclo-CLL-dF-VY 	17.9, 14.8	13.746	94	739.8 [M+H]⁺(100 %)	739.3847	739.3846
Cyclo-CLL(F)VY 	11.9, 10.9	13.905	92	757.8 [M+H]⁺ (100 %)	757.3753	757.3751
Cyclo-CLL(F)Cl)VY 	8.5, 7.1	14.519	94	773.7 [M+H]⁺ (100 %), 795.7 [M+Na]⁺ (30 %)	773.3458	773.3452

Cyclo-CLLF(Br)VY 	12.1, 7.4	14.657	90	819.5 [M+H]⁺ (100 %)	817.2953	817.2926
Cyclo-CLLF(I)VY 	15.8, 9.1	15.144	90	865.7 [M+H]⁺ (100 %)	865.2814	865.2803
Cyclo-CLLF(CF ₃)VY 	29.4, 18.2	14.840	90	807.7 [M+H]⁺ (100 %), 829.7 [M+Na]⁺ (60 %)	807.3721	807.3720
Cyclo-CLLF(tBu)VY 	16.6, 10.4	16.384	93	795.87 [M+H]⁺ (100 %)	795.4473	795.4449

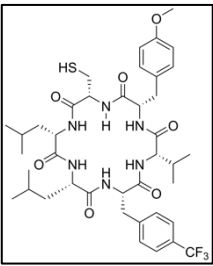
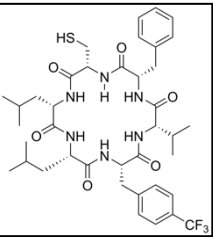
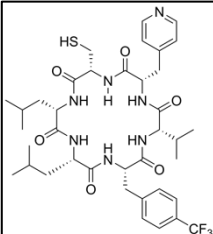
8.9.4 Trifluoromethylphenylalanine alanine scan derivatives

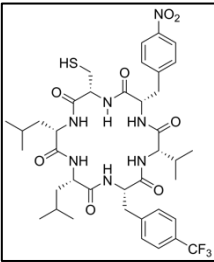
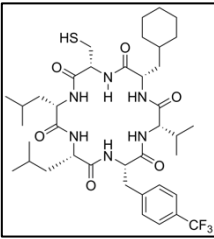
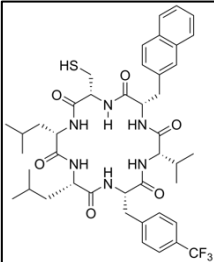
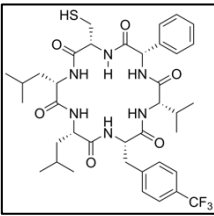
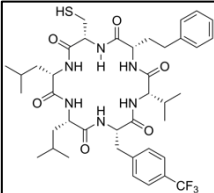
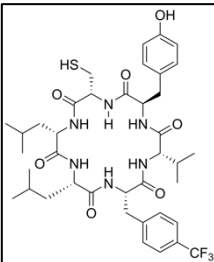
Linear peptides were synthesised using Wang resin SPPS (section 8.6.1) and cleaved as detailed in section 8.6.2. The linear sequences were then cyclised using HATU, HOAt and DIPEA (section 8.6.8) and purified via reverse-phase flash chromatography (section 8.6.14). StBu protected peptides were then deprotected as described in section 8.6.9 and the cyclic peptides were then purified via reverse-phase HPLC (section 8.6.12), the product containing fractions were combined and concentrated under reduced pressure, then lyophilised to yield the desired cyclic peptides.

Compound	Yield (mg, %)	Analytical HPLC R _t (min)	Analytical HPLC Purity (%)	m/z (ESI ⁺)	HRMS (ESI ⁺) Calculated	HRMS (ESI ⁺) Found
Cyclo-ALLF(CF ₃)VY 	9.6, 12.4	14.255	81	775.6 [M+H] ⁺ (100 %), 797.6 [M+Na] ⁺ (20 %)	775.4001	775.4001
Cyclo-CALF(CF ₃)VY 	13.6, 17.8	16.382	97	765.5 [M+H] ⁺ (100 %), 787.4 [M+Na] ⁺ (20 %)	765.3252	765.3244
Cyclo-CLAF(CF ₃)VY 	9.8, 12.8	13.564	89	765.5 [M+H] ⁺ (100 %)	765.3252	765.3242
Cyclo-CLLF(CF ₃)AY 	13.6, 17.5	13.931	91	779.5 [M+H] ⁺ (100 %)	779.3408	779.3424
Cyclo-CLLF(CF ₃)VA 	6.7, 9.4	13.931	91	715.5 [M+H] ⁺ (100 %), 1452.7 [2M+Na] ⁺ (55 %)	715.3453	715.3450

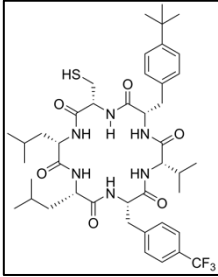
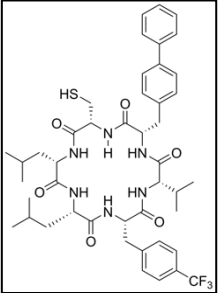
8.9.5 CLLF(CF₃)VY Tyrosine Derivatives

The linear resin bound sequence resin-LC-NH₂ was synthesised using Wang resin SPPS (section 8.6.1). The resin was then split equally into 10 syringes and swelled in a small amount of DMF. Standard SPPS protocols (section 8.6.1) were then followed to yield the resin bound linear sequence resin-LCXVF(CF₃)L where X is the unnatural amino acid. The linear sequences were then cleaved from the resin as described in section 8.6.2, and cyclised using HATU, HOAt and DIPEA as described in section 8.6.8. The StBu group was then removed (section 8.6.9) and the peptide purified via reverse-phase HPLC (section 8.6.12), like fractions were combined and concentrated under reduced pressure, then lyophilised to yield the desired cyclic peptides.

Compound	Yield (mg, %)	Analytical HPLC R _t (min)	Analytical HPLC Purity (%)	<i>m/z</i> (ESI ⁺)	HRMS (ESI ⁺) Calculated	HRMS (ESI ⁺) Found
CLLF(CF ₃)VY(Me) 	9.2, 5.6	16.557	99	821.7 [M+H] ⁺ (100 %)	821.3878	821.3866
CLLF(CF ₃)VF 	62.7, 39.7	16.769	95	791.6 [M+H] ⁺ (100 %), 813.6 [M+Na] ⁺ (20 %)	791.3772	791.3779
CLLF(CF ₃)V-Pal 	24.7, 15.6	12.770	97	792.6 [M+H] ⁺ (100 %)	792.3725	792.3716

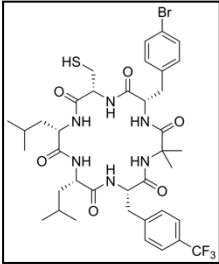
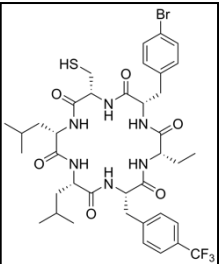
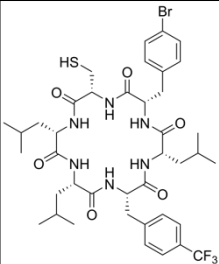
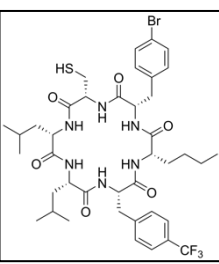
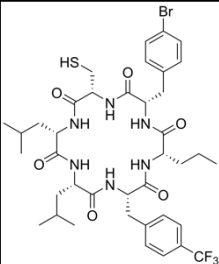
CLLF(CF ₃)VF(NO ₂) 	20.0, 12.0	12.877	99	836.6 [M+H] ⁺ (100%)	836.3623	836.3616
CLLF(CF ₃)V-Cha 	16.2, 10.2	18.111	97	797.7 [M+H] ⁺ (100%)	797.4242	797.4242
CLLF(CF ₃)V-Nal 	31.9, 19.0	17.959	99	841.6 [M+H] ⁺ (100%)	841.3929	841.3919
CLLF(CF ₃)V-Phg 	16.1, 10.4	16.544	92	777.7 [M+H] ⁺ (100%)	777.3616	777.3627
CLLF(CF ₃)V-hF 	2.7, 1.7	17.200	90	806.7 [M+H] ⁺ (100%)	805.3929	805.3933
CLLF(CF ₃)V-dY 	12.1, 7.5	14.683	97	807.7 [M+H] ⁺ (100%)	807.3721	807.3720

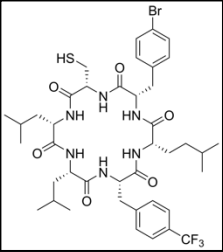
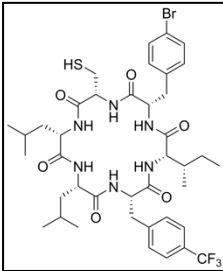
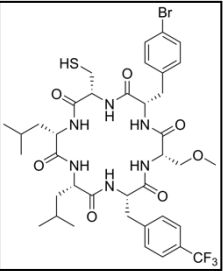
<p>CLLF(CF₃)VF(F)</p> <p>The structure shows a complex molecule with a central core of two isopropyl groups connected by a sulfur atom. This core is linked via amide bonds to a chain of four nitrogen atoms. The chain is terminated by a pentafluorophenyl group (CF₃) and a p-fluorophenyl group (F).</p>	28.4, 17.6	16.665	95	809.7 [M+H] ⁺ (100%)	809.3678	809.3672
<p>CLLF(CF₃)VF(Cl)</p> <p>The structure is identical to the first row, but the terminal phenyl ring is a p-chlorophenyl group (Cl).</p>	28.4, 17.2	17.376	99	825.6 [M+H] ⁺ (100%)	825.3382	825.3382
<p>CLLF(CF₃)VF(Br)</p> <p>The structure is identical to the first row, but the terminal phenyl ring is a p-bromophenyl group (Br).</p>	8.3, 4.8	17.580	99	871.6 [M+H] ⁺ (100%)	869.2881	869.2877
<p>CLLF(CF₃)VF(I)</p> <p>The structure is identical to the first row, but the terminal phenyl ring is a p-iodophenyl group (I).</p>	19.8, 10.8	17.893	91	917.7 [M+H] ⁺ (100%)	917.2739	917.2728
<p>CLLF(CF₃)VF(CF₃)</p> <p>The structure is identical to the first row, but the terminal phenyl ring is a p-(trifluoromethyl)phenyl group (CF₃).</p>	31.6, 18.4	16.284	99	859.7 [M+H] ⁺ (100%)	859.3646	859.3640

CLLF(CF ₃)VF(tBu) 	24.1, 14.2	17.333	99	847.7 [M+H] ⁺ (100 %), 869.7 [M+Na] ⁺ (60 %)	847.4398	847.4391
CLLF(CF ₃)VF(Ph) 	29.1, 16.8	18.351	98	867.7 [M+H] ⁺ (100 %), 889.7 [M+Na] ⁺ (30 %)	867.4085	867.4078

8.9.6 CLLF(CF₃)VF(Br) Valine Derivatives

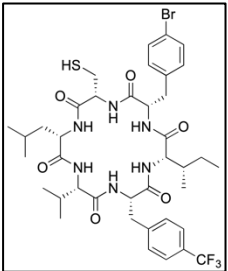
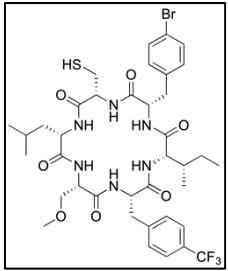
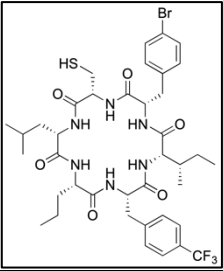
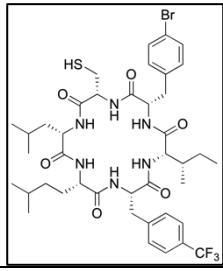
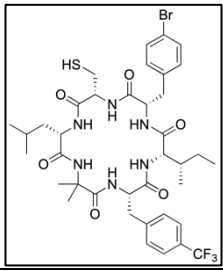
Fmoc-Leu-Wang resin (0.1mmol) was added to 8 syringes and the Fmoc group removed via exposure to 20 % piperidine in DMF. A solution of amino acid (1 eq.), Oxyma pure (3 eq.) and DIC (3 eq.) were added and the syringes placed on a rocker for 1 hour. The reaction mixtures were then removed via vacuum filtration and the resins washed with DMF x 2, DCM x 2 and Et₂O x 1. Fmoc protected amino acids were then deprotected via the addition of piperidine 20 % in DMF and being placed on a rocker for 30 minutes. The resins were then washed as before. These steps were repeated until the desired linear sequence was synthesised. The linear sequences were then cleaved from the resin using a TFA cleavage cocktail (section 8.6.2) and cyclised using HATU and DIPEA (section 8.6.8) excess DMF was removed under reduced pressure and the crude cyclic peptide used without further purification. The StBu group was then removed (section 8.6.9) and the peptide purified via reverse-phase HPLC (section 8.6.12), like fractions were combined and concentrated under reduced pressure, then lyophilised to yield the desired cyclic peptides.

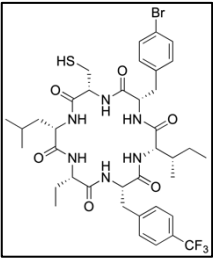
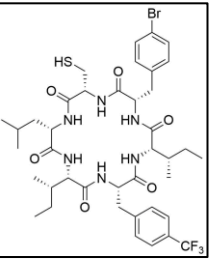
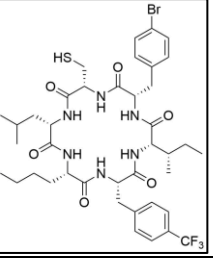
Compound	Yield (mg, %)	Analytical HPLC R _t (min)	Analytical HPLC Purity (%)	m/z (ESI ⁺)	HRMS (ESI ⁺) Calculated	HRMS (ESI ⁺) Found
CLLF(CF ₃)-Aib-F(Br)	3.3, 3.86	16.835	89	857.4 [M+H] ⁺ (100 %)	877.2546	877.2460
						
CLLF(CF ₃)-Abu-F(Br)	32.9, 38.5	16.609	99	857.4 [M+H] ⁺ (100 %), 879.5 [M+Na] ⁺ (40 %)	877.2546	877.2537
						
CLLF(CF ₃)LF(Br)	25.9, 29.4	17.873	99	885.5 [M+H] ⁺ (100 %)	905.2859	905.2863
						
CLLF(CF ₃)-NL-F(Br)	15.2, 17.2	18.107	95	885.5 [M+H] ⁺ (100 %), 905.4 [M+Na] ⁺ (60 %)	883.3034	883.3034
						
CLLF(CF ₃)-NV-F(Br)	23.8, 27.4	17.340	98	893.5 [M+Na] ⁺ (100 %), 869.5 [M+H] ⁺ (90 %)	869.2877	869.2873
						

CLLF(CF ₃)-hL-F(Br) 	33.3, 18.6	18.577	99	899.7 [M+H] ⁺ (100 %)	897.3190	897.3166
CLLF(CF ₃)IF(Br) 	14.2, 8.0	18.149	99	885.6 [M+H] ⁺ (100 %), 907.6 [M+Na] ⁺ (40 %)	883.3034	883.3020
CLLF(CF ₃)-Prop- F(Br) 	14.2, 8.0	16.822	98	873.5 [M+H] ⁺ (100 %), 893.4 [M+Na] ⁺ (99 %)	871.2670	871.2660

8.9.7 CLLF(CF₃)IF(Br) Leucine Derivatives

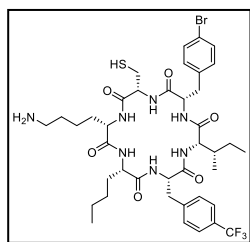
Fmoc-Leu-Wang resin (0.1mmol) was added to 8 syringes and the Fmoc group removed via exposure to 20 % piperidine in DMF. A solution of amino acid (1 eq.), Oxyma pure (3 eq.) and DIC (3 eq.) were added and the syringes placed on a rocker for 1 hour. The reaction mixtures were then removed via vacuum filtration and the resins washed with DMF x 2, DCM x 2 and Et₂O x 1. Fmoc protected amino acids were then deprotected via the addition of piperidine 20 % in DMF and being placed on a rocker for 30 minutes. The resins were then washed as before. These steps were repeated until the desired linear sequence was synthesised. The linear sequences were then cleaved from the resin using a TFA cleavage cocktail (section 8.6.2) and cyclised using HATU and DIPEA (section 8.6.8) excess DMF was removed under reduced pressure and the crude cyclic peptide used without further purification. The StBu group was then removed (section 8.6.9) and the peptide purified via reverse-phase HPLC (section 8.6.12), like fractions were combined and concentrated under reduced pressure, then lyophilised to yield the desired cyclic peptides.

Compound	Yield (mg, %)	Analytical HPLC R _t (min)	Analytical HPLC Purity (%)	m/z (ESI ⁺)	HRMS (ESI ⁺) Calculated	HRMS (ESI ⁺) Found
CLVF(CF ₃)IF(Br) 	0.4, 0.5	18.116	97	871.6 [M+H] ⁺ (100%), 890.6 [M+Na] ⁺ (45 %)	869.2877	869.2850
CL-Prop-F(CF ₃)IF(Br) 	6.1, 7.0	17.001	95	873.5 [M+H] ⁺ (100 %)	871.2670	871.2670
CL-NV-F(CF ₃)IF(Br) 	6.7, 7.7	17.729	99	871.6 [M+H] ⁺ (100 %)	869.2877	869.2853
CL-hL-F(CF ₃)IF(Br) 	16.2, 18.1	18.838	99	899.7 [M+H] ⁺ (100 %)	897.3190	897.3166
CL-Aib-F(CF ₃)IF(Br) 	24.7, 28.9	17.737	96	857.6 [M+H] ⁺ (100 %)	855.2721	855.2721

CL-Abu-F(CF ₃)IF(Br) 	18.5, 21.6	16.316	91	857.6 [M+H] ⁺ (100 %)	855.2710	855.2721
CLI F(CF ₃)IF(Br) 	4.5, 5.1	18.566	96	885.6 [M+H] ⁺ (100 %)	883.3034	883.3020
CL-NL-F(CF ₃)IF(Br) 	12.5, 14.2	18.310	99	885.6 [M+H] ⁺ (100 %)	883.3034	883.3021

8.9.8 Polar Derivatives

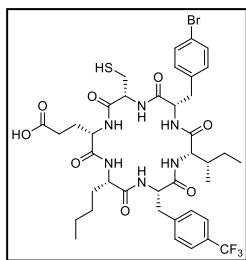
CK-NL-F(CF₃)IF(Br)



2-chlorotriptyl chloride resin was loaded with Fmoc-Lys(Boc) as detailed in section 8.6.3 and the resin loading determined (section 8.6.4). The linear sequence K(Boc)C(trt)F(Br)IF(CF₃)-NL was then synthesised via Fmoc SPPS (section 8.6.5). The linear sequence was then cleaved from the resin using HFIP, leaving the protecting groups intact (section 8.6.6). The protected

linear was then cyclised with HATU (section 8.6.8) to yield the protected cyclic peptide that was immediately globally deprotected with a TFA cleavage cocktail (section 8.6.10). The crude product was then dissolved in DMF before being purified via reverse-phase HPLC (section 8.6.13) to yield the desired cyclic peptide sequence cyclo-CK-NL-F(CF₃)IF(Br) as a white powder (51.2 mg, 28.5 %). Analytical HPLC R_t = 14.309 min (90.0 % purity). m/z (ESI⁺): 900.5 [M+H]⁺ (100 %). HRMS (ESI⁺) found: 898.3061, [M+H]⁺ calculated: 898.3143.

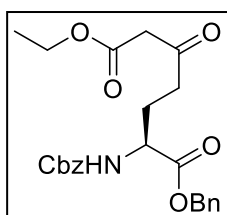
CE-NL-F(CF₃)IF(Br)



2-chlorotrityl chloride resin was loaded with Fmoc-Glu(OtBu) as detailed in section 8.6.3 and the resin loading determined (section 8.6.4). The linear sequence E(OtBu)C(trt)F(Br)IF(CF₃)-NL was then synthesised via Fmoc SPPS (section 8.6.5). The linear sequence was then cleaved from the resin using HFIP, leaving the protecting groups intact (section 8.6.6). The protected linear was then cyclised with HATU (section 8.6.8) to yield the protected cyclic peptide that was immediately globally deprotected with a TFA cleavage cocktail (section 8.6.10). The crude product was then dissolved in DMF before being purified via reverse-phase HPLC (section 8.6.13) to yield the desired cyclic peptide sequence cyclo-CE-NL-F(CF₃)IF(Br) as a white powder (5.7 mg, 3.2 %). Analytical HPLC R_t = 15.749 min (97 % purity). m/z (ESI⁺): 901.4 [M+H]⁺ (100 %). HRMS (ESI⁺) found: 899.2605, [M+H]⁺ calculated: 899.2619.

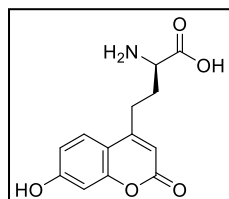
8.9.9 Coumarin Synthesis

1-Benzyl 7-ethyl (S)-2-(((benzyloxy)carbonyl)amino)-5-oxoheptanedioate



Cbz-Glu-OBn (2.00 g, 5.39 mmol, 1 eq.) was dissolved in THF (25 mL) and CDI (0.96 g, 5.92 mmol, 1.1 eq.) was added. The mixture was then stirred at RT for 2 hours under argon. Ethyl magnesium malonate (0.95 g, 2.96 mmol, 0.55 eq.) was then added and the mixture stirred under argon for a further 16 hours. The product was extracted with diethyl ether (3 x 50 mL) and washed with NaHCO₃ (3 x 50 mL), water (3 x 50 mL) and brine (3 x 50 mL). The organic layer was then dried with anhydrous MgSO₄ and concentrated under reduced pressure to yield a clear oil. The crude product was purified by silica column chromatography with an eluent of 1:1 PE:EtOAc to yield the β -ketoester (2.24 g, 94.1 %) as a white solid. m/z (ESI⁺): 464.4 [M+Na]⁺ (100 %). ¹H NMR (400 MHz, CDCl₃) δ ppm 1.28 (t, J =7.09 Hz, -OCH₂CH₃, 3 H) 1.90 - 2.28 (m, NHCHCH₂CH₂, 2 H) 2.46 - 2.70 (m, NHCHCH₂CH₂, 2 H) 3.39 (s, COCH₂COO, 2 H) 4.17 (q, J =7.09 Hz, -OCH₂CH₃, 2 H) 4.43 (td, J =8.01, 5.26 Hz, NHCHCH₂CH₂, 1 H) 5.12 (s, PhCH₂CONH, 2 H) 5.19 (s, COOCH₂Ph, 2 H) 5.44 (br d, J =8.07 Hz, NHCHCH₂CH₂, 1 H) 7.28 - 7.39 (m, ArH, 10 H). Spectral data matches literature ²²⁷

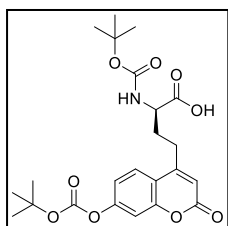
(S)-2-Amino-4-(7-hydroxy-2-oxo-2H-chromen-4-yl)butanoic acid (H₂N-hCou-OH)



The β -ketoester (1.0 g, 2.27 mmol, 1 eq.) was added slowly to a solution of resorcinol (1.25 g, 11.35 mmol, 5 eq.) in methanesulfonic acid (5 mL). The solution was stirred at RT for 2 hours before being diluted with ice-cold ether (300 mL). The precipitate was isolated via filtration and then dissolved in water before being filtered again and lyophilised. The resulting residue was purified via RP-chromatography to give the coumaryl amino acid (0.170 g, 28.5 %) as an off-white solid. m/z

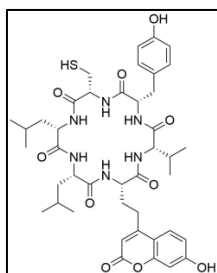
(ESI⁺): 264.2 [M+H]⁺ (100 %). ¹H NMR (400 MHz, DMSO-*d*₆) δ ppm 7.65 (d, *J* = 8.80 Hz, 1 H) 6.83 (dd, *J* = 8.68, 2.32 Hz, 1 H) 6.75 (d, *J* = 2.45 Hz, 1 H) 6.12 (s, 1 H) 3.85 (br t, *J* = 5.99 Hz, 1 H) 2.80 - 2.97 (m, 2 H) 2.00 - 2.19 (m, 2 H). Spectral data matches literature ²²⁷

(S)-2-((Tert-butoxycarbonyl)amino)-4-(7-((tert-butoxycarbonyl)oxy)-2-oxo-2H-chromen-4-yl)butanoic acid (Boc-hCou(Boc)-OH)

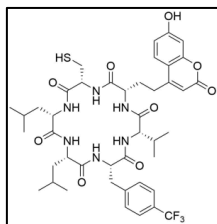


The coumaryl amino acid (0.4 g, 1.5 mmol, 1 eq.) was dissolved in 5% aqueous NaHCO₃/dioxane (1:1 v/v) (40 mL) and placed in an ice bath. Boc anhydride (3.27 g, 15 mmol, 10 eq.) was then added and the reaction stirred on ice for 1 hour before being warmed up to RT over 16 hours. The mixture was then acidified to pH 3 with 10% (w/v) citric acid_(aq) (60mL) and the resulting solution extracted with EtOAc (3 x 75mL). The extract was then washed with water (3 x 100 mL) and brine (3 x 100 mL) before being dried over anhydrous MgSO₄ and concentrated under reduced pressure. The resulting oil was purified via RP-chromatography to yield a mixture of the mono-boc coumaryl amino acid (17.4 mg, 3.2 %) and di-boc coumaryl amino acid and (379.4, 54.6 %). *m/z* (ESI⁺): 408.3 [M -tBu +H]⁺ (100 %). ¹H NMR (400 MHz, DMSO-*d*₆) δ ppm 7.92 (d, *J*=8.80 Hz, 1 H), 7.38 (d, *J*=2.45 Hz, 1 H), 7.32 (d, *J*=8.07 Hz, 1 H), 7.26 (dd, *J*=8.68, 2.32 Hz, 1 H), 6.34 (s, 1 H), 3.86 - 4.10 (m, 1 H), 2.87 (br t, *J*=7.58 Hz, 2 H), 1.85 - 2.10 (m, 2 H), 1.52 (s, 9 H), 1.41 (s, 9 H).

CLL-hCou-VY



Cyclo-CLL-hCou-VY was obtained as a white solid (9.4 mg, 11.2 %). Analytical HPLC *R*_t = 10.990 min (91 % purity). *m/z* (ESI⁺): 837.5 [M+H]⁺ (100 %), 859.5 [M+Na]⁺ (100 %). HRMS (ESI⁺) found: 836.3769, [M+H]⁺ calculated: 836.3779.

CLLF(CF₃)V-hCou

Cyclo-CLLF(CF₃)V-hCou was obtained as a white solid (6.6 mg, 7.4 %).

Analytical HPLC R_t = 11.133 min (87 % purity). *m/z* (ESI⁺): 889.6 [M+H]⁺ (100 %). HRMS (ESI⁺) found: 889.3768, [M+H]⁺ calculated: 889.3776.

All compound raw data as well as protein characterisation data can be found on the attached CD.

Bibliography

- 1 G. L. Semenza, Oxygen sensing, homeostasis, and disease, *N. Engl. J. Med.*, 2011, **365**, 537–547.
- 2 N. V. Iyer, L. E. Kotch, F. Agani, S. W. Leung, E. Laughner, R. H. Wenger, M. Gassmann, J. D. Gearhart, A. M. Lawler, A. Y. Yu and G. L. Semenza, Cellular and developmental control of O₂ homeostasis by hypoxia-inducible factor 1 α , *Genes Dev.*, 1998, **12**, 149–162.
- 3 W. X. Hong, M. S. Hu, M. Esquivel, G. Y. Liang, R. C. Rennert, A. McArdle, K. J. Paik, D. Duscher, G. C. Gurtner, H. P. Lorenz and M. T. Longaker, The Role of Hypoxia-Inducible Factor in Wound Healing, *Adv. Wound Care*, 2014, **3**, 390–399.
- 4 G. L. Semenza, Perspectives on Oxygen Sensing Minireview, *Cell*, 1999, **98**, 281–284.
- 5 B. Muz, P. de la Puente, F. Azab and A. K. Azab, The role of hypoxia in cancer progression, angiogenesis, metastasis, and resistance to therapy, *Hypoxia*, 2015, **11**, 83–92.
- 6 S. R. McKeown, Defining normoxia, physoxia and hypoxia in tumours - Implications for treatment response, *Br. J. Radiol.*, 2014, **87**, 1–12.
- 7 M. Höckel, K. Schlenger, B. Aral, M. Mitze, U. Schäffer and P. Vaupel, Association between tumor hypoxia and malignant progression in advanced cancer of the uterine cervix, *Cancer Res.*, 1996, **56**, 4509–4515.
- 8 A. M. Jubb, F. M. Buffa and A. L. Harris, Assessment of tumour hypoxia for prediction of response to therapy and cancer prognosis, *J. Cell. Mol. Med.*, 2010, **14**, 18–29.
- 9 A. C. Koong, E. Y. Chen and A. J. Giaccia, Hypoxia Causes the Activation of Nuclear Factor κ B through the Phosphorylation of I κ B α on Tyrosine Residues, *Cancer Res.*, 1994, **54**, 1425–1430.
- 10 K. A. Seta, Z. Spicer, Y. Yuan, G. Lu and D. E. Millhorn, Responding to hypoxia: lessons from a model cell line, *Sci. STKE*, 2002, **2002**, 1–17.
- 11 D. J. Manalo, A. Rowan, T. Lavoie, L. Natarajan, B. D. Kelly, S. Q. Ye, J. G. N. Garcia, G. L. Semenza, J. G. N. Garcia, G. L. Semenza, J. G. N. Garcia and G. L. Semenza, Transcriptional regulation of vascular endothelial cell responses to hypoxia by HIF-1, *Blood*, 2005, **105**, 659–669.
- 12 G. L. Wang, B.-H. H. Jiang, E. A. Rue and G. L. Semenza, Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS heterodimer regulated by cellular O₂ tension, *Proc. Natl. Acad. Sci. U. S. A.*, 1995, **92**, 5510–5514.
- 13 G. N. Masoud and W. Li, HIF-1 α pathway: Role, regulation and intervention for cancer therapy, *Acta Pharm. Sin. B*, 2015, **5**, 378–389.
- 14 P. B. Card, P. J. A. Erbel and K. H. Gardner, Structural basis of ARNT PAS-B dimerization: Use of a common beta-sheet interface for hetero- and homodimerization, *J. Mol. Biol.*, 2005, **353**, 664–677.
- 15 T. H. Scheuermann, D. R. Tomchick, M. Machius, Y. Guo, R. K. Bruick and K. H. Gardner, Artificial ligand binding within the HIF2 α PAS-B domain of the HIF2 transcription factor, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 450–455.
- 16 R. Cardoso, R. Love, C. L. Nilsson, S. Bergqvist, D. Nowlin, J. Yan, K. K. C. Liu, J. Zhu, P. Chen, Y. L. Deng, H. J. Dyson, M. J. Greig and A. Brooun, Identification of Cys255 in HIF-1 α as a novel site for development of covalent inhibitors of HIF-1 α /ARNT PasB domain protein-

Bibliography

- protein interaction, *Protein Sci.*, 2012, **21**, 1885–1896.
- 17 J. Zhu, M. Martinez-Yamout, R. Cardoso, J. Yan, R. A. Love, N. Grodsky, A. Brooun and H. J. Dyson, Homodimerization of the PAS-B domains of hypoxia-inducible factors, *J. Phys. Chem. B*, 2012, **116**, 6960–6965.
- 18 D. Wu, N. Potluri, J. Lu, Y. Kim and F. Rastinejad, Structural integration in hypoxia-inducible factors, *Nature*, 2015, **524**, 303–308.
- 19 D. R. Mole, C. Blancher, R. R. Copley, P. J. Pollard, J. M. Gleadle, J. Ragousis and P. J. Ratcliffe, Genome-wide association of hypoxia-inducible factor (HIF)-1 α and HIF-2 α DNA binding with expression profiling of hypoxia-inducible transcripts, *J. Biol. Chem.*, 2009, **284**, 16767–16775.
- 20 K. W. Schulte, E. Green, A. Wilz, M. Platten and O. Daumke, Structural Basis for Aryl Hydrocarbon Receptor-Mediated Gene Activation, *Structure*, 2017, **25**, 1025–1033.
- 21 M. H. Hefti, K. J. François, S. C. De Vries, R. Dixon and J. Vervoort, The PAS fold: A redefinition of the PAS domain based upon structural prediction, *Eur. J. Biochem.*, 2004, **271**, 1198–1208.
- 22 A. Moglich, R. A. Ayers and K. Moffat, Structure and Signaling Mechanism of Per-ARNT-Sim Domains, *Structure*, 2009, **17**, 1282–1294.
- 23 J. Yang, L. Zhang, P. J. A. Erbel, K. H. Gardner, K. Ding, J. A. Garcia and R. K. Bruick, Functions of the Per/ARNT/Sim domains of the hypoxia-inducible factor, *J. Biol. Chem.*, 2005, **280**, 36047–36054.
- 24 Y. V. Liu, J. H. Baek, H. Zhang, R. Diez, R. N. Cole and G. L. Semenza, RACK1 Competes with HSP90 for Binding to HIF-1 α and is Required for O₂-independent and HSP90 Inhibitor-induced Degradation of HIF-1 α , *Mol. Cell.*, 2007, **25**, 207–217.
- 25 M. E. Hubbi, W. Luo, J. H. Baek and G. L. Semenza, MCM Proteins Are Negative Regulators of Hypoxia-Inducible Factor 1, *Mol. Cell.*, 2011, **42**, 700–712.
- 26 B. H. Jiang, J. Z. Zheng, S. W. Leung, R. Roe and G. L. Semenza, Transactivation and inhibitory domains of Hypoxia-inducible factor 1 α : Modulation of transcriptional activity by oxygen tension, *J. Biol. Chem.*, 1997, **272**, 19253–19260.
- 27 J. L. Ruas, U. Berchner-Pfannschmidt, S. Malik, K. Gradin, J. Fandrey, R. G. Roeder, T. Pereira and L. Poellinger, Complex regulation of the transactivation function of hypoxia-inducible factor-1 α by direct interaction with two distinct domains of the creb-binding protein/p300, *J. Biol. Chem.*, 2010, **285**, 2601–2609.
- 28 C. J. Hu, A. Sataur, L. Wang, H. Chen and M. C. Simon, The N-Terminal Transactivation Domain Confers Target Gene Specificity of Hypoxia-inducible Factors HIF-1 α and HIF-2 α , *Mol. Biol. Cell*, 2007, **18**, 986–994.
- 29 S. J. Freedman, Z. Y. J. Sun, F. Poy, A. L. Kung, D. M. Livingston, G. Wagner and M. J. Eck, Structural basis for recruitment of CBP/p300 by hypoxia-inducible factor-1 α , *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 5367–5372.
- 30 E. Berra, D. Roux, D. E. Richard and J. Pouyssegur, Hypoxia-inducible factor-1 α (HIF-1) escapes O₂-driven proteasomal degradation irrespective of its subcellular localization: Nucleus or cytoplasm, *EMBO Rep.*, 2001, **2**, 615–620.
- 31 A. C. R. Epstein, J. M. Gleadle, L. A. McNeill, K. S. Hewitson, J. O'Rourke, D. R. Mole, M. Mukherji, E. Metzen, M. I. Wilson, A. Dhanda, Y. M. Tian, N. Masson, D. L. Hamilton, P. Jaakkola, R. Barstead, J. Hodgkin, P. H. Maxwell, C. W. Pugh, C. J. Schofield and P. J. Ratcliffe, C. elegans EGL-9 and mammalian homologs define a family of dioxygenases that

- regulate HIF by prolyl hydroxylation, *Cell*, 2001, **107**, 43–54.
- 32 M. Hirsilä, P. Koivunen, V. Günzler, K. I. Kivirikko and J. Myllyharju, Characterization of the human prolyl 4-hydroxylases that modify the hypoxia-inducible factor, *J. Biol. Chem.*, 2003, **278**, 30772–30780.
- 33 R. K. Bruick and S. L. McKnight, A conserved family of prolyl-4-hydroxylases that modify HIF, *Science*, 2001, **294**, 1337–1340.
- 34 K. Tanimoto, Mechanism of regulation of the hypoxia-inducible factor-1 alpha by the von Hippel-Lindau tumor suppressor protein, *EMBO J.*, 2000, **19**, 4298–4309.
- 35 L. E. Huang, J. Gu, M. Schau and H. F. Bunn, Regulation of hypoxia-inducible factor 1 α is mediated by an O₂-dependent degradation domain via the ubiquitin-proteasome pathway, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 7987–7992.
- 36 T. A. F. Cardote, M. S. Gadd and A. Ciulli, Crystal Structure of the Cul2-Rbx1-EloBC-VHL Ubiquitin Ligase Complex, *Structure*, 2017, **25**, 901–911.
- 37 D. Lando, D. J. Peet, J. J. Gorman, D. A. Whelan, M. L. Whitelaw and R. K. Bruick, FIH-1 is an asparaginyl hydroxylase enzyme that regulates the transcriptional activity of hypoxia-inducible factor, *Genes Dev.*, 2002, **16**, 1466–1471.
- 38 P. C. Mahon, K. Hirota and G. L. Semenza, FIH-1: a novel protein that interacts with HIF-1 α and VHL to mediate repression of HIF-1 transcriptional activity, *Genes Dev.*, 2001, **15**, 2675–2686.
- 39 J. M. Elkins, K. S. Hewitson, L. A. McNeill, J. F. Seibel, I. Schlemminger, C. W. Pugh, P. J. Ratcliffe and C. J. Schofield, Structure of factor-inhibiting hypoxia-inducible factor (HIF) reveals mechanism of oxidative modification of HIF-1 α , *J. Biol. Chem.*, 2003, **278**, 1802–1806.
- 40 P. Koivunen, M. Hirsilä, V. Günzler, K. I. Kivirikko and J. Myllyharju, Catalytic Properties of the Asparaginyl Hydroxylase (FIH) in the Oxygen Sensing Pathway Are Distinct from Those of Its Prolyl 4-Hydroxylases, *J. Biol. Chem.*, 2004, **279**, 9899–9904.
- 41 I. P. Stolze, Y. M. Tian, R. J. Appelhoff, H. Turley, C. C. Wykoff, J. M. Gleadle and P. J. Ratcliffe, Genetic analysis of the role of the asparaginyl hydroxylase factor inhibiting hypoxia-inducible factor (HIF) in regulating HIF transcriptional target genes, *J. Biol. Chem.*, 2004, **279**, 42719–42725.
- 42 D. Wu, X. Su, N. Potluri, Y. Kim and F. Rastinejad, NPAS1-ARNT and NPAS3-ARNT crystal structures implicate the bHLH-PAS family as multi-ligand binding transcription factors, *Elife*, 2016, **5**, 1–15.
- 43 D. Wu, N. Potluri, Y. Kim and F. Rastinejad, Structure and Dimerization Properties of the Aryl Hydrocarbon Receptor PAS-A Domain, *Mol. Cell. Biol.*, 2013, **33**, 4346–4356.
- 44 E. Berra, D. E. Richard, E. Gothié and J. Pouyssegur, HIF-1-dependent transcriptional activity is required for oxygen-mediated HIF-1 α degradation, *FEBS Lett.*, 2001, **491**, 85–90.
- 45 D. M. Stroka, T. Burkhardt, I. Desbaillets, R. H. Wenger, D. A. H. Neil, C. Bauer, M. Gassmann and D. Candinas, HIF-1 is expressed in normoxic tissue and displays an organ-specific regulation under systemic hypoxia, *FASEB J.*, 2001, **15**, 2445–2453.
- 46 Y. V. Liu and G. L. Semenza, RACK1 vs. HSP90: Competition for HIF-1 α degradation vs. stabilization, *Cell Cycle*, 2007, **6**, 656–659.
- 47 M. V. Blagosklonny, W. G. An, L. Y. Romanova, J. Trepel, T. Fojo and L. Neckers, P53 Inhibits Hypoxia-Inducible Factor-Stimulated Transcription, *J. Biol. Chem.*, 1998, **273**, 11995–

Bibliography

- 11998.
- 48 R. Ravi, B. Mookerjee, Z. M. Bhujwalla, C. H. Sutter, D. Artemov, Q. Zeng, L. E. Dillehay, A. Madan, G. L. Semenza and A. Bedi, Regulation of tumor angiogenesis by p53-induced degradation of hypoxia-inducible factor 1 α , *Genes Dev.*, 2000, **14**, 34–44.
- 49 M. Y. Koh, R. Lemos, X. Liu and G. Powis, The hypoxia-associated factor switches cells from HIF-1 α to HIF-2 α dependent signaling promoting stem cell characteristics, aggressive tumor growth and invasion, *Cancer Res.*, 2011, **71**, 4015–4027.
- 50 M. Y. Koh, B. G. Darnay and G. Powis, Hypoxia-Associated Factor, a Novel E3-Ubiquitin Ligase, Binds and Ubiquitinates Hypoxia-Inducible Factor 1, Leading to Its Oxygen-Independent Degradation, *Mol. Cell. Biol.*, 2008, **28**, 7081–7095.
- 51 W. Luo, J. Zhong, R. Chang, H. Hu, A. Pandey and G. L. Semenza, Hsp70 and CHIP selectively mediate ubiquitination and degradation of hypoxia-inducible factor (HIF)-1 α but not HIF-2 α , *J. Biol. Chem.*, 2010, **285**, 3651–3663.
- 52 A. Lyberopoulou, I. Mylonis, G. Papachristos, D. Sagris, A. Kalousi, C. Befani, P. Liakos, G. Simos and E. Georgatsou, MgcRacGAP, a cytoskeleton regulator, inhibits HIF-1 transcriptional activity by blocking its dimerization, *Biochim. Biophys. Acta - Mol. Cell Res.*, 2013, **1833**, 1378–1387.
- 53 B. Van De Sluis, M. Vooijs, E. Burstein, B. Van De Sluis, X. Mao, Y. Zhai, A. J. Groot and J. F. Vermeulen, COMMD1 disrupts HIF-1 α / β dimerization and inhibits human tumor cell invasion, *J. Clin. Invest.*, 2010, **120**, 2119–2130.
- 54 A. Kalousi, I. Mylonis, A. S. Politou, G. Chachami, E. Paraskeva and G. Simos, Casein kinase 1 regulates human hypoxia-inducible factor HIF-1, *J. Cell Sci.*, 2010, **123**, 2976–2986.
- 55 I. M. Yasinska and V. V. Sumbayev, S-nitrosation of Cys-800 of HIF-1 α protein activates its interaction with p300 and stimulates its transcriptional activity, *FEBS Lett.*, 2003, **549**, 105–109.
- 56 N. V. Chaika, T. Gebregiworgis, M. E. Lewallen, V. Purohit, P. Radhakrishnan, X. Liu, B. Zhang, K. Mehla, R. B. Brown, T. Caffrey, F. Yu, K. R. Johnson, R. Powers, M. A. Hollingsworth and P. K. Singh, MUC1 mucin stabilizes and activates hypoxia-inducible factor 1 alpha to regulate metabolism in pancreatic cancer, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 13787–13792.
- 57 S. C. Land and A. R. Tee, Hypoxia-inducible factor 1 α is regulated by the mammalian target of rapamycin (mTOR) via an mTOR signaling motif, *J. Biol. Chem.*, 2007, **282**, 20534–20543.
- 58 W. Luo, H. Hu, R. Chang, J. Zhong, M. Knabel, R. O’Meally, R. N. Cole, A. Pandey and G. L. Semenza, Pyruvate kinase M2 is a PHD3-stimulated coactivator for hypoxia-inducible factor 1, *Cell*, 2011, **145**, 732–744.
- 59 Z. Chen, X. Liu, Z. Mei, Z. Wang and W. Xiao, EAF2 Suppresses Hypoxia-Induced Factor 1 Transcriptional Activity by Disrupting Its Interaction with Coactivator CBP/p300, *Mol. Cell. Biol.*, 2014, **34**, 1085–1099.
- 60 T. Kietzmann, D. Mennerich and E. Y. Dimova, Hypoxia-Inducible Factors (HIFs) and phosphorylation: Impact on stability, localization, and transactivity, *Front. Cell Dev. Biol.*, 2016, **4**, 1–14.
- 61 I. Mylonis, G. Chachami, M. Samiotaki, G. Panayotou, E. Paraskeva, A. Kalousi, E. Georgatsou, S. Bonanou and G. Simos, Identification of MAPK phosphorylation sites and their role in the localization and activity of hypoxia-inducible factor-1 α , *J. Biol. Chem.*, 2006, **281**, 33095–33106.

- 62 H. Suzuki, A. Tomida and T. Tsuruo, Dephosphorylated hypoxia-inducible factor 1 α as a mediator of p53-dependent apoptosis during hypoxia, *Oncogene*, 2001, **20**, 5779–5788.
- 63 D. Flugel, A. Gorchach, C. Michiels and T. Kietzmann, Glycogen Synthase Kinase 3 Phosphorylates Hypoxia-Inducible Factor 1 and Mediates Its Destabilization in a VHL-Independent Manner, *Mol. Cell. Biol.*, 2007, **27**, 3253–3265.
- 64 X. Liu, Z. Chen, C. Xu, X. Leng, H. Cao, G. Ouyang and W. Xiao, Repression of hypoxia-inducible factor α signaling by Set7-mediated methylation, *Nucleic Acids Res.*, 2015, **43**, 5081–5098.
- 65 P. Hill, D. Shukla, M. G. B. Tran, J. Aragonés, H. T. Cook, P. Carmeliet and P. H. Maxwell, Inhibition of hypoxia inducible factor hydroxylases protects against renal ischemia-reperfusion injury, *J. Am. Soc. Nephrol.*, 2008, **19**, 39–46.
- 66 F. M. Mowat, U. F. O. Luhmann, A. J. Smith, C. Lange, Y. Duran, S. Harten, D. Shukla, P. H. Maxwell, R. R. Ali and J. W. B. Bainbridge, HIF-1 α and HIF-2 α are differentially activated in distinct cell populations in retinal ischaemia, *PLoS One*, 2010, **5**, 1–9.
- 67 L. Holmquist-Mengelbier, E. Fredlund, T. Löfstedt, R. Noguera, S. Navarro, H. Nilsson, A. Pietras, J. Vallon-Christersson, Å. Borg, K. Gradin, L. Poellinger and S. Pålman, Recruitment of HIF-1 α and HIF-2 α to common target genes is differentially regulated in neuroblastoma: HIF-2 α promotes an aggressive phenotype, *Cancer Cell*, 2006, **10**, 413–423.
- 68 C.-J. Hu, L.-Y. Wang, L. A. Chodosh, B. Keith and M. C. Simon, Differential Roles of Hypoxia-Inducible Factor 1 (HIF-1) and HIF-2 in Hypoxic Gene Regulation, *Mol. Cell. Biol.*, 2003, **23**, 9361–9374.
- 69 N. Takeda, E. L. O’Dea, A. Doedens, J. W. Kim, A. Weidemann, C. Stockmann, M. Asagiri, M. C. Simon, A. Hoffmann and R. S. Johnson, Differential activation and antagonistic function of HIF- α isoforms in macrophages are essential for NO homeostasis, *Genes Dev.*, 2010, **24**, 491–501.
- 70 B. L. Krock, N. Skuli and M. C. Simon, Hypoxia-Induced Angiogenesis: Good and Evil, *Genes and Cancer*, 2011, **2**, 1117–1133.
- 71 S. Parangi, M. O’Reilly, G. Christofori, L. Holmgren, J. Grosfeld, J. Folkman and D. Hanahan, Antiangiogenic therapy of transgenic mice impairs de novo tumor growth, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, **93**, 2002–2007.
- 72 D. M. McDonald and P. L. Choyke, Imaging of angiogenesis: From microscope to clinic, *Nat. Med.*, 2003, **9**, 713–725.
- 73 M. A. Konerding, E. Fait and A. Gaumann, 3D microvascular architecture of pre-cancerous lesions and invasive carcinomas of the colon, *Br. J. Cancer*, 2001, **84**, 1354–1362.
- 74 M. W. Dewhirst, H. Kimura, S. Rehmus, R. D. Braun, D. Papahadjopoulos, K. Hong and T. W. Secomb, Microvascular studies on the origins of perfusion-limited hypoxia, *Br. J. Cancer*.
- 75 F. Jung, L. A. Palmer, N. Zhou and R. A. Johns, Hypoxic regulation of inducible nitric oxide synthase via hypoxia inducible factor-1 in cardiac myocytes, *Circ. Res.*, 2000, **86**, 319–325.
- 76 V. H. Haase, Regulation of erythropoiesis by hypoxia-inducible factors, *Blood Rev.*, 2013, **27**, 41–53.
- 77 I. Papandreou, R. A. Cairns, L. Fontana, A. L. Lim and N. C. Denko, HIF-1 mediates adaptation to hypoxia by actively downregulating mitochondrial oxygen consumption, *Cell Metab.*, 2006, **3**, 187–197.

Bibliography

- 78 M. S. Ullah, A. J. Davies and A. P. Halestrap, The plasma membrane lactate transporter MCT4, but not MCT1, is up-regulated by hypoxia through a HIF-1 α -dependent mechanism, *J. Biol. Chem.*, 2006, **281**, 9030–9037.
- 79 G. Zhang, Y. Zhang, D. Dong, F. Wang, X. Ma, F. Guan and L. Sun, MCT1 regulates aggressive and metabolic phenotypes in bladder cancer, *J. Cancer*, 2018, **9**, 2492–2501.
- 80 G. Bellot, R. Garcia-Medina, P. Gounon, J. Chiche, D. Roux, J. Pouyssegur and N. M. Mazure, Hypoxia-Induced Autophagy Is Mediated through Hypoxia-Inducible Factor Induction of BNIP3 and BNIP3L via Their BH3 Domains, *Mol. Cell. Biol.*, 2009, **29**, 2570–2581.
- 81 H. Zhang, M. Bosch-Marce, L. A. Shimoda, S. T. Yee, H. B. Jin, J. B. Wesley, F. J. Gonzalez and G. L. Semenza, Mitochondrial autophagy is an HIF-1-dependent adaptive metabolic response to hypoxia, *J. Biol. Chem.*, 2008, **283**, 10892–10903.
- 82 Z. Chen, Y. Li, H. Zhang, P. Huang and R. Luthra, Hypoxia-regulated microRNA-210 modulates mitochondrial function and decreases ISCU and COX10 expression, *Oncogene*, 2010, **29**, 4362–4368.
- 83 S. Y. Chan, Y. Y. Zhang, C. Hemann, C. E. Mahoney, J. L. Zweier and J. Loscalzo, MicroRNA-210 Controls Mitochondrial Metabolism during Hypoxia by Repressing the Iron-Sulfur Cluster Assembly Proteins ISCU1/2, *Cell Metab.*, 2009, **10**, 273–284.
- 84 O. Warburg, F. Wind and E. Negelein, The Metabolism of Tumors in the Body, *J. Gen. Physiol.*, 1927, **8**, 519–530.
- 85 G. L. Semenza, HIF-1 mediates the Warburg effect in clear cell renal carcinoma, *J. Bioenerg. Biomembr.*, 2007, **39**, 231–234.
- 86 H. Lu, R. A. Forbes and A. Verma, Hypoxia-inducible factor 1 activation by aerobic glycolysis implicates the Warburg effect in carcinogenesis, *J. Biol. Chem.*, 2002, **277**, 23111–23115.
- 87 M. Y. Mboge, B. P. Mahon, R. McKenna and S. C. Frost, Carbonic anhydrases: Role in pH control and cancer, *Metabolites*, 2018, **8**, 1–31.
- 88 S. Y. Liao, O. N. Aurelio, K. Jan, J. Zavada and E. J. Stanbridge, Identification of the MN/CA9 protein as a reliable diagnostic biomarker of clear cell carcinoma of the kidney, *Cancer Res.*, 1997, **57**, 2827–2831.
- 89 S. Y. Liao, C. Brewer, J. Závada, J. Pastorek, S. Pastorekova, A. Manetta, M. L. Berman, P. J. DiSaia and E. J. Stanbridge, Identification of the MN antigen as a diagnostic biomarker of cervical intraepithelial squamous and glandular neoplasia and cervical carcinomas, *Am. J. Pathol.*, 1994, **145**, 598–609.
- 90 M. A. Proescholdt, C. Mayer, M. Kubitzka, T. Schubert, S.-Y. Liao, E. J. Stanbridge, S. Ivanov, E. H. Oldfield, A. Brawanski and M. J. Merrill, Expression of hypoxia-inducible carbonic anhydrases in brain tumors, *Neuro. Oncol.*, 2005, **7**, 465–475.
- 91 A. Giatromanolaki, M. I. Koukourakis, E. Sivridis, J. Pastorek, C. C. Wykoff, K. C. Gatter and A. L. Harris, Expression of hypoxia-inducible carbonic anhydrase-9 relates to angiogenic pathways and independently to poor outcome in non-small cell lung cancer, *Cancer Res.*, 2001, **61**, 7992–7998.
- 92 N. M. Chau, P. Rogers, W. Aherne, V. Carroll, I. Collins, E. McDonald, P. Workman and M. Ashcroft, Identification of novel small molecule inhibitors of hypoxia-inducible factor-1 that differentially block hypoxia-inducible factor-1 activity and hypoxia-inducible factor-1 α induction in response to hypoxic stress and growth factors, *Cancer Res.*, 2005, **65**, 4918–4928.
- 93 A. Rapisarda, B. Uranchimeg, D. A. Scudiero, M. Selby, E. A. Sausville, R. H. Shoemaker and

- G. Melillo, Identification of small molecule inhibitors of hypoxia-inducible factor 1 transcriptional activation pathway, *Cancer Res.*, 2002, **62**, 4316–4324.
- 94 M. Xia, K. Bi, R. Huang, M. H. Cho, S. Sakamuru, S. C. Miller, H. Li, Y. Sun, J. Printen, C. P. Austin and J. Inglese, Identification of small molecule compounds that inhibit the HIF-1 signaling pathway, *Mol. Cancer*, 2009, **8**, 1–13.
- 95 P. E. Jung, D. Kong, E. J. Park and A. G. Stephen, Echinomycin, a Small-Molecule Inhibitor of Hypoxia-Inducible Factor-1 DNA-Binding Activity, *Am. Assoc. Cancer Res.*, 2005, **65**, 9047–9055.
- 96 K. A. Lee, D. Z. Qian, S. Rey, H. Wei, J. O. Liu and G. L. Semenza, Anthracycline chemotherapy inhibits HIF-1 transcriptional activity and tumor-induced mobilization of circulating angiogenic cells, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 2353–2358.
- 97 A. L. Kung, S. D. Zabludoff, D. S. France, S. J. Freedman, E. A. Tanner, A. Vieira, S. Cornell-Kennon, J. Lee, B. Wang, J. Wang, K. Memmert, H.-U. U. Naegeli, F. Petersen, M. J. Eck, K. W. Bair, A. W. Wood and D. M. Livingston, Small molecule blockade of transcriptional coactivation of the hypoxia-inducible factor pathway, *Cancer Cell*, 2004, **6**, 33–43.
- 98 B. B. Lao, I. Grishagin, H. Mesallati, T. F. Brewer, B. Z. Olenyuk and P. S. Arora, In vivo modulation of hypoxia-inducible signaling by topographical helix mimetics, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 7531–7536.
- 99 G. M. Burslem, H. F. Kyle, P. Prabhakaran, A. L. Breeze, T. A. Edwards, S. L. Warriner, A. Nelson and A. J. Wilson, Synthesis of highly functionalized oligobenzamide proteomimetic foldamers by late stage introduction of sensitive groups, *Org. Biomol. Chem.*, 2016, **14**, 3782–3786.
- 100 H. J. Choi, B. J. Song, Y. D. Gong, W. J. Gwak and Y. Soh, Rapid degradation of hypoxia-inducible factor-1 α by KRH102053, a new activator of prolyl hydroxylase 2, *Br. J. Pharmacol.*, 2008, **154**, 114–125.
- 101 H. L. Shan, H. S. Dong, Y. S. Chun, K. L. Myung, M. S. Kim and J. W. Park, A novel mode of action of YC-1 in HIF inhibition: Stimulation of FIH-dependent p300 dissociation from HIF-1 α , *Mol. Cancer Ther.*, 2008, **7**, 3729–3738.
- 102 A. Field-Smith, G. J. Morgan and F. E. Davies, Bortezomib (VelcadeTM) in the treatment of multiple myeloma, *Ther. Clin. Risk Manag.*, 2006, **2**, 271–279.
- 103 D. H. Shin, Y. S. Chun, D. S. Lee, L. E. Huang and J. W. Park, Bortezomib inhibits tumor adaptation to hypoxia by stimulating the FIH-mediated repression of hypoxia-inducible factor-1, *Blood*, 2008, **111**, 3131–3136.
- 104 N. J. Mabeesh, D. E. Post, M. T. Willard, B. Kaur, E. G. Van Meir, J. W. Simons and H. Zhong, Geldanamycin induces degradation of hypoxia-inducible factor 1 α protein via the proteosome pathway in prostate cancer cells, *Cancer Res.*, 2002, **62**, 2478–2482.
- 105 L. Xiang, D. M. Gilkes, P. Chaturvedi, W. Luo, H. Hu, N. Takano, H. Liang and G. L. Semenza, Ganetespib blocks HIF-1 activity and inhibits tumor growth, vascularization, stem cell maintenance, invasion, and metastasis in orthotopic mouse models of triple-negative breast cancer, *J Mol Med*, 2014, **92**, 151–164.
- 106 E. Hur, H. H. Kim, S. M. Choi, J. H. Kim, S. Yim, H. J. Kwon, Y. Choi, D. K. Kim, M. O. Lee and H. Park, Reduction of hypoxia-induced transcription through the repression of hypoxia-inducible factor-1 α /aryl hydrocarbon receptor nuclear translocator DNA binding by the 90-kDa heat-shock protein inhibitor radicicol, *Mol. Pharmacol.*, 2002, **62**, 975–982.
- 107 B. H. Jiang, E. Rue, G. L. Wang, R. Roe and G. L. Semenza, Dimerization, DNA binding, and

Bibliography

- transactivation properties of hypoxia- inducible factor 1, *J. Biol. Chem.*, 1996, **271**, 17771–17778.
- 108 J.-E. Park, D. Kong, R. Fisher, J. Cardellina, R. H. Shoemaker and G. Melillo, Targeting the PAS-A Domain of HIF-1 α for Development of Small Molecule Inhibitors of HIF-1, *Cell Cycle*, 2006, **5**, 1847–1853.
- 109 K. Lee, H. Zhang, D. Z. Qian, S. Rey, J. O. Liu and G. L. Semenza, Acriflavine inhibits HIF-1 dimerization, tumor growth, and vascularization., *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 17910–17915.
- 110 B. Obstoy, M. Salaun, L. Veresezan, R. Sesbou  , P. Bohn, F. X. Boland and L. Thiberville, Safety and performance analysis of acriflavine and methylene blue for in vivo imaging of precancerous lesions using fibered confocal fluorescence microscopy (FCFM): an experimental study, *BMC Pulm. Med.*, 2015, **15**, 1–10.
- 111 C. J. Lee, C. H. Yue, Y. Y. Lin, J. C. Wu and J. Y. Liu, Antitumor activity of acriflavine in human hepatocellular carcinoma cells, *Anticancer Res.*, 2014, **34**, 3549–3556.
- 112 A. Mangraviti, T. Raghavan, F. Volpin, N. Skuli, D. Gullotti, J. Zhou, L. Asnaghi, E. Sankey, A. Liu, Y. Wang, D. H. Lee, N. Gorelick, R. Serra, M. Peters, D. Schriefer, F. Delaspre, F. J. Rodriguez, C. G. Eberhart, H. Brem, A. Olivi and B. Tyler, HIF-1 α - Targeting Acriflavine Provides Long Term Survival and Radiological Tumor Response in Brain Cancer Therapy, *Sci. Rep.*, 2017, **7**, 1–13.
- 113 P. Zargar, E. Ghani, F. Jalali Mashayekhi, A. Ramezani and E. Eftekhar, Acriflavine enhances the antitumor activity of the chemotherapeutic drug 5-fluorouracil in colorectal cancer cells, *Oncol. Lett.*, 2018, **15**, 10084–10090.
- 114 C. Lee, C. Yue, Y. Y. Lin, Y. Y. Lin, S. Kao, J. Liu and Y. Chen, Antitumor Activity of Acriflavine in Lung Adenocarcinoma Cell Line A549, *Anticancer Res.*, 2014, **34**, 6467–6472.
- 115 T. H. Scheuermann, D. R. Tomchick, M. Machius, Y. Guo, R. K. Bruick and K. H. Gardner, Artificial ligand binding within the HIF2 α PAS-B domain of the HIF2 transcription factor, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 450–455.
- 116 T. H. Scheuermann, Q. Li, H.-W. W. Ma, J. Key, L. Zhang, R. Chen, J. A. Garcia, J. Naidoo, J. Longgood, D. E. Frantz, U. K. Tambar, K. H. Gardner and R. K. Bruick, Allosteric inhibition of hypoxia inducible factor-2 with small molecules, *Nat. Chem. Biol.*, 2013, **9**, 271–276.
- 117 W. Chen, H. Hill, A. Christie, M. S. Kim, E. Holloman, A. Pavia-Jimenez, F. Homayoun, Y. Ma, N. Patel, P. Yell, G. Hao, Q. Yousuf, A. Joyce, I. Pedrosa, H. Geiger, H. Zhang, J. Chang, K. H. Gardner, R. K. Bruick, C. Reeves, T. H. Hwang, K. Courtney, E. Frenkel, X. Sun, N. Zojwalla, T. Wong, J. P. Rizzi, E. M. Wallace, J. A. Josey, Y. Xie, X. J. Xie, P. Kapur, R. M. McKay and J. Brugarolas, Targeting renal cell carcinoma with a HIF-2 antagonist, *Nature*, 2016, **539**, 112–117.
- 118 H. Cho, X. Du, J. P. Rizzi, E. Liberzon, A. A. Chakraborty, W. Gao, I. Carvo, S. Signoretti, R. Bruick, J. A. Josey, E. M. Wallace and W. G. Kaelin Jr, On-Target Efficacy of a HIF2 α Antagonist in Preclinical Kidney Cancer Models, *Nature*, 2016, **539**, 107–111.
- 119 D. Wu, X. Su, J. Lu, S. Li, B. L. Hood, S. Vasile, N. Potluri, X. Diao, Y. Kim, S. Khorasanizadeh and F. Rastinejad, Bidirectional modulation of HIF-2 activity through chemical ligands, *Nat. Chem. Biol.*, 2019, **15**, 367–376.
- 120 R. Xu, K. Wang, J. P. Rizzi, H. Huang, J. A. Grina, S. T. Schlachter, B. Wang, P. M. Wehn, H. Yang, D. D. Dixon, R. M. Czerwinski, X. Du, E. L. Ged, G. Han, H. Tan, T. Wong, S. Xie, J. A. Josey and E. M. Wallace, 3-[(1 S ,2 S ,3 R)-2,3-Difluoro-1-hydroxy-7-methylsulfonylindan-4-yl]oxy-5-fluorobenzonitrile (PT2977), a Hypoxia-Inducible Factor 2 α (HIF-2 α) Inhibitor for

- the Treatment of Clear Cell Renal Cell Carcinoma , *J. Med. Chem.*, 2019, **62**, 6876–6893.
- 121 F. Corrêa, J. Key, B. Kuhlman and K. H. Gardner, Computational Repacking of HIF-2 α Cavity Replaces Water-Based Stabilized Core, *Structure*, 2016, **24**, 1918–1927.
- 122 A. L. Garner and K. D. Janda, Protein-Protein Interactions and Cancer: Targeting the Central Dogma, *Curr. Top. Med. Chem.*, 2010, **11**, 258–280.
- 123 J. A. Wells and C. L. McClendon, Reaching for high-hanging fruit in drug discovery at protein-protein interfaces, *Nature*, 2007, **450**, 1001–1009.
- 124 A. Whitty and G. Kumaravel, Between a rock and a hard place?, *Nat. Chem. Biol.*, 2006, **2**, 112–118.
- 125 T. Clackson and J. A. Wells, A hot spot of binding energy in a hormone-receptor interface, *Science*, 1995, **267**, 383–386.
- 126 Y. Wang, R. Coulombe, D. R. Cameron, L. Thauvette, M. J. Massariol, L. M. Amon, D. Fink, S. Titolo, E. Welchner, C. Yoakim, J. Archambault and P. W. White, Crystal Structure of the E2 Transactivation Domain of Human Papillomavirus Type 11 Bound to a Protein Interaction Inhibitor, *J. Biol. Chem.*, 2004, **279**, 6976–6985.
- 127 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, In Vivo Activation of the p53 Pathway by Small-Molecule Antagonists of MDM2, *Science*, 2004, **303**, 844–848.
- 128 T. Oltersdorf, S. W. Elmore, A. R. Shoemaker, R. C. Armstrong, D. J. Augeri, B. A. Belli, M. Bruncko, T. L. Deckwerth, J. Dinges, P. J. Hajduk, M. K. Joseph, S. Kitada, S. J. Korsmeyer, A. R. Kunzer, A. Letai, C. Li, M. J. Mitten, D. G. Nettesheim, S. C. Ng, P. M. Nimmer, J. M. O'Connor, A. Oleksijew, A. M. Petros, J. C. Reed, W. Shen, S. K. Tahir, C. B. Thompson, K. J. Tomaselli, B. Wang, M. D. Wendt, H. Zhang, S. W. Fesik and S. H. Rosenberg, An inhibitor of Bcl-2 family proteins induces regression of solid tumours, *Nature*, 2005, **435**, 677–681.
- 129 M. R. Arkin, M. Randal, W. L. DeLano, J. Hyde, T. N. Luong, J. D. Oslob, D. R. Raphael, L. Taylor, J. Wang, R. S. McDowell, J. A. Wells and A. C. Braisted, Binding of small molecules to an adaptive protein-protein interface, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 1603–1608.
- 130 C. Wilson and M. Arkin, Small-Molecule Inhibitors of IL-2/IL-2R: Lessons Learned and Applied, *Curr. Top. Microbiol. Immunol.*, 2011, **348**, 25–29.
- 131 A. Groß, C. Hashimoto, H. Sticht and J. Eichler, Synthetic peptides as protein mimics, *Front. Bioeng. Biotechnol.*, 2016, **3**, 1–16.
- 132 S. H. Joo, Cyclic peptides as therapeutic agents and biochemical tools, *Biomol. Ther.*, 2012, **20**, 19–26.
- 133 G. M. Pauletti, S. Gangwar, T. J. Siahaan, J. Aubé and R. T. Borchardt, Improvement of Oral Peptide Bioavailability: Peptidomimetics and prodrug strategies, *Adv. Drug Deliv. Rev.*, 1997, **27**, 235–256.
- 134 K. Fosgerau and T. Hoffmann, Peptide therapeutics: Current status and future directions, *Drug Discov. Today*, 2015, **20**, 122–128.
- 135 L. Di, Strategic Approaches to Optimizing Peptide ADME Properties, *AAPS J.*, 2015, **17**, 134–143.
- 136 L. Gentilucci, R. De Marco and L. Cerisoli, Chemical Modifications Designed to Improve Peptide Stability: Incorporation of Non-Natural Amino Acids, Pseudo-Peptide Bonds, and Cyclization, *Curr. Pharm. Des.*, 2010, **16**, 3185–3203.
- 137 H. John, E. Maronde, W. Forssmann, M. Meyer and K. Adermann, N-terminal acetylation

Bibliography

- protects glucagon-like peptide GLP-1-(7-34)-amide from DPP-IV-mediated degradation retaining cAMP- and insulin-releasing capacity., *Eur. J. Med. Res.*, 2008, **13**, 73–78.
- 138 L. H. Brinckerhoff, V. V. Kalashnikov, L. W. Thompson, G. V. Yamshchikov, R. A. Pierce, H. S. Galavotti, V. H. Engelhard and C. L. Slingsluff, Terminal modifications inhibit proteolytic degradation of an immunogenic MART-1 peptide: Implications for peptide vaccines, *Int. J. Cancer*, 1999, **83**, 326–334.
- 139 A. Sharman and J. Low, Vasopressin and its role in critical care, *Continuing Education in Anaesthesia, Critical Care and Pain*, 2008, **8**, 134–137.
- 140 H. Agersø, L. S. Larsen, A. Riis, U. Lövgren, M. O. Karlsson and T. Senderovitz, Pharmacokinetics and renal excretion of desmopressin after intravenous administration to healthy subjects and renally impaired patients, *Br. J. Clin. Pharmacol.*, 2004, **58**, 352–358.
- 141 D. Tedesco and L. Haragsim, Cyclosporine: A Review, *J. Transplant.*, 2012, 1–7.
- 142 F. R. Bruniera, F. M. Ferreira, L. R. M. Saviolli, M. R. Bacci, D. Feder, M. D. L. G. Pedreira, M. A. S. Peterlini, L. A. Azzalis, V. B. C. Junqueira and F. L. A. Fonseca, *The use of vancomycin with its therapeutic and adverse effects: A review*, 2015, vol. 19.
- 143 F. Weis, F. Vogt, R. Sodian and A. Beiras-fernandez, Daptomycin: a novel lipopeptide antibiotic against Gram-positive pathogens, *Infect. Drug Resist.*, 2010, **3**, 95–101.
- 144 D. C. Sadowski, Use of octreotide in the acute management of bleeding esophageal varices, *Can. J. Gastroenterol.*, 1997, **11**, 339–343.
- 145 M. C. Smith and J. E. Gestwicki, Features of protein–protein interactions that translate into potent inhibitors: topology, surface area and affinity, *Expert Rev. Mol. Med.*, 2012, **14**, 1–20.
- 146 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, A cyclic peptide inhibitor of HIF-1 heterodimerization that inhibits hypoxia signaling in cancer cells, *J. Am. Chem. Soc.*, 2013, **135**, 10418–10425.
- 147 Y. Hayashi, J. Morimoto and H. Suga, In Vitro Selection of Anti-Akt2 Thioether-Macrocyclic Peptides Leading to Isoform-Selective Inhibitors, *ACS Chem. Biol.*, 2012, **7**, 607–613.
- 148 G. Guidotti, L. Brambilla and D. Rossi, Cell-Penetrating Peptides: From Basic Research to Clinics, *Trends Pharmacol. Sci.*, 2017, **38**, 406–424.
- 149 C. A. Lipinski, F. Lombardo, B. W. Dominy and P. J. Feeney, Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings, *Adv. Drug Deliv. Rev.*, 1997, **23**, 3–25.
- 150 M. Boehm, K. Beaumont, R. Jones, A. S. Kalgutkar, L. Zhang, K. Atkinson, G. Bai, J. A. Brown, H. Eng, G. H. Goetz, B. R. Holder, B. Khunte, S. Lazzaro, C. Limberakis, S. Ryu, M. J. Shapiro, L. Tylaska, J. Yan, R. Turner, S. S. F. Leung, M. Ramaseshan, D. A. Price, S. Liras, M. P. Jacobson, D. J. Earp, R. S. Lokey, A. M. Mathiowetz and E. Menhaji-Klotz, Discovery of Potent and Orally Bioavailable Macrocyclic Peptide-Peptoid Hybrid CXCR7 Modulators, *J. Med. Chem.*, 2017, **60**, 9653–9663.
- 151 W. M. Hewitt, S. S. F. F. Leung, C. R. Pye, A. R. Ponkey, M. Bednarek, M. P. Jacobson and R. S. Lokey, Cell-permeable cyclic peptides from synthetic libraries inspired by natural products, *J. Am. Chem. Soc.*, 2015, **137**, 715–721.
- 152 D. S. Nielsen, H. N. Hoang, R. J. Lohman, T. A. Hill, A. J. Lucke, D. J. Craik, D. J. Edmonds, D. A. Griffith, C. J. Rotter, R. B. Ruggeri, D. A. Price, S. Liras and D. P. Fairlie, Improving on Nature: Making a Cyclic Heptapeptide Orally Bioavailable, *Angew. Chemie - Int. Ed.*, 2014,

- 53**, 12059–12063.
- 153 N. Favalli, G. Bassi, J. Scheuermann and D. Neri, DNA-encoded chemical libraries – achievements and remaining challenges, *FEBS Lett.*, 2018, **592**, 2168–2180.
- 154 G. P. Smith, Filamentous fusion phage: Novel expression vectors that display cloned antigens on the virion surface, *Science*, 1985, **228**, 1315–1317.
- 155 S. Sattar, N. J. Bennett, W. X. Wen, J. M. Guthrie, L. F. Blackwell, J. F. Conway and J. Rakonjac, Ff-nano, short functionalized nanorods derived from Ff (f1, fd or M13) filamentous bacteriophage, *Front. Microbiol.*, 2015, **6**, 1–13.
- 156 L. M. Mullen, S. P. Nair, J. M. Ward, A. N. Rycroft and B. Henderson, Phage display in the study of infectious diseases, *Trends Microbiol.*, 2006, **14**, 141–147.
- 157 M. A. McLafferty, R. B. Kent, R. C. Ladner and W. Markland, M13 bacteriophage displaying disulfide-constrained microproteins, *Gene*, 1993, **128**, 29–36.
- 158 K. T. O’Neil, R. H. Hoess, S. A. Jackson, N. S. Ramachandran, S. A. Mousa and W. F. DeGrado, Identification of novel peptide antagonists for GPIIb/IIIa from a conformationally constrained phage peptide library, *Proteins Struct. Funct. Bioinforma.*, 1992, **14**, 509–515.
- 159 B. A. Desimmie, M. Humbert, E. Lescrinier, J. Hendrix, S. Vets, R. Gijsbers, R. M. Ruprecht, U. Dietrich, Z. Debyser and F. Christ, Phage display-directed discovery of LEDGF/p75 binding cyclic peptide inhibitors of HIV replication, *Mol. Ther.*, 2012, **20**, 2064–2075.
- 160 S. C. Meyer, T. Gaj and I. Ghosh, Highly selective cyclic peptide ligands for neutravidin and avidin identified by phage display, *Chem. Biol. Drug Des.*, 2006, **68**, 3–10.
- 161 N. C. Wrighton, F. X. Farrell, R. Chang, A. K. Kashyap, F. P. Barbone, L. S. Mulcahy, D. L. Johnson, R. W. Barrett, L. K. Jolliffe and W. J. Dower, Small peptides as potent mimetics of the protein hormone erythropoietin, *Science*, 1996, **273**, 458–463.
- 162 P. Timmerman, J. Beld, W. C. Puijk and R. H. Meloen, Rapid and quantitative cyclization of multiple peptide loops onto synthetic scaffolds for structural mimicry of protein surfaces, *ChemBioChem*, 2005, **6**, 821–824.
- 163 C. Heinis, T. Rutherford, S. Freund and G. Winter, Phage-encoded combinatorial chemical libraries based on bicyclic peptides, *Nat. Chem. Biol.*, 2009, **5**, 502–507.
- 164 D. S. Wilson, A. D. Keefe and J. W. Szostak, The use of mRNA display to select high-affinity protein-binding peptides, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 3750–3755.
- 165 R. W. Roberts, Totally in vitro protein selection using mRNA-protein fusions and ribosome display, *Curr. Opin. Chem. Biol.*, 1999, **3**, 268–273.
- 166 R. Wang, S. w. Cotten and R. Liu, mRNA Display Using Covalent Coupling of mRNA to Translated Proteins, *Methods Mol. Biol.*, 2012, **805**, 87–100.
- 167 H. Saito, D. Kourouklis and H. Suga, An in vitro evolved precursor tRNA with aminoacylation activity, *EMBO J.*, 2001, **20**, 1797–1806.
- 168 M. Ohuchi, H. Murakami and H. Suga, The flexizyme system: a highly flexible tRNA aminoacylation tool for the translation apparatus, *Curr. Opin. Chem. Biol.*, 2007, **11**, 537–542.
- 169 T. Passioura and H. Suga, A RaPID way to discover nonstandard macrocyclic peptide modulators of drug targets, *Chem. Commun.*, 2017, **53**, 1931–1940.
- 170 Y. Yamagishi, I. Shoji, S. Miyagawa, T. Kawakami, T. Katoh, Y. Goto and H. Suga, Natural product-like macrocyclic N-methyl-peptide inhibitors against a ubiquitin ligase uncovered

Bibliography

- from a ribosome-expressed de novo library, *Chem. Biol.*, 2011, **18**, 1562–1570.
- 171 J. Morimoto, Y. Hayashi and H. Suga, Discovery of macrocyclic peptides armed with a mechanism-based warhead: Isoform-selective inhibition of human deacetylase SIRT2, *Angew. Chemie - Int. Ed.*, 2012, **51**, 3423–3427.
- 172 K. Iwasaki, Y. Goto, T. Katoh and H. Suga, Selective thioether macrocyclization of peptides having the N-terminal 2-chloroacetyl group and competing two or three cysteine residues in translation, *Org. Biomol. Chem.*, 2012, **10**, 5783–5786.
- 173 Y. Goto, H. Murakami and H. Suga, Initiating translation with D-amino acids, *RNA*, 2008, **14**, 1390–1398.
- 174 C. P. Scott, E. Abel-Santos, M. Wall, D. C. Wahnnon and S. J. Benkovic, Production of cyclic peptides and proteins in vivo, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 13638–13643.
- 175 A. Tavassoli and S. J. Benkovic, Split-intein mediated circular ligation used in the synthesis of cyclic peptide libraries in *E. coli*, *Nat. Protoc.*, 2007, **2**, 1126–1133.
- 176 T. S. Young, D. D. Young, I. Ahmad, J. M. Louis, S. J. Benkovic and P. G. Schultz, Evolution of cyclic peptide protease inhibitors, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 11052–11056.
- 177 A. Tavassoli and S. J. Benkovic, Genetically selected cyclic-peptide inhibitors of AICAR transformylase homodimerization, *Angew. Chemie - Int. Ed.*, 2005, **44**, 2760–2763.
- 178 A. R. Horswill, S. N. Savinov and S. J. Benkovic, A systematic method for identifying small-molecule modulators of protein-protein interactions, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 15591–15596.
- 179 E. L. Osher, F. Castillo, N. Elumalai, M. J. Waring, G. Pairaudeau and A. Tavassoli, A genetically selected cyclic peptide inhibitor of BCL6 homodimerization, *Bioorg. Med. Chem.*, 2018, **26**, 3034–3038.
- 180 E. K. Leitch, N. Elumalai, M. Fridén-Saxin, G. Dahl, P. Wan, P. Clarkson, E. Valeur, G. Pairaudeau, H. Boyd and A. Tavassoli, Inhibition of low-density lipoprotein receptor degradation with a cyclic peptide that disrupts the homodimerization of IDOL E3 ubiquitin ligase, *Chem. Sci.*, 2018, **9**, 5957–5966.
- 181 A. L. Male, F. Forafonov, F. Cuda, G. Zhang, S. Zheng, P. C. F. Oyston, P. R. Chen, E. D. Williamson and A. Tavassoli, Targeting *Bacillus anthracis* toxicity with a genetically selected inhibitor of the PA/CMG2 protein-protein interaction, *Sci. Rep.*, 2017, **7**, 1–9.
- 182 C. N. Birts, S. K. Nijjar, C. A. Mardle, F. Hoakwie, P. J. Duriez, J. P. Blaydes and A. Tavassoli, A cyclic peptide inhibitor of C-terminal binding protein dimerization links metabolism with mitotic fidelity in breast cancer cells, *Chem. Sci.*, 2013, **4**, 3046–3057.
- 183 A. Tavassoli, SICLOPPS cyclic peptide libraries in drug discovery, *Curr. Opin. Chem. Biol.*, 2017, **38**, 30–35.
- 184 I. N. Mistry and A. Tavassoli, Reprogramming the Transcriptional Response to Hypoxia with a Chromosomally Encoded Cyclic Peptide HIF-1 Inhibitor, *ACS Synth. Biol.*, 2017, **6**, 518–527.
- 185 R. B. Merrifield, Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide, *J. Am. Chem. Soc.*, 1963, **85**, 2149–2154.
- 186 A. A. El-Faham and F. Albericio, Peptide Coupling Reagents, More than a Letter Soup, *Chem. Rev.*, 2011, **111**, 6557–6602.
- 187 J. M. Humphrey and A. R. Chamberlin, Chemical synthesis of natural product peptides: Coupling methods for the incorporation of noncoded amino acids into peptides, *Chem.*

- Rev.*, 1997, **97**, 2243–2266.
- 188 F. Albericio, R. Chinchilla, D. J. Dodsworth and C. Nájera, New trends in peptide coupling reagents, *Org. Prep. Proced. Int.*, 2001, **33**, 203–303.
- 189 S.-Y. Y. Han and Y.-A. A. Kim, Recent development of peptide coupling reagents in organic synthesis, *Tetrahedron*, 2004, **60**, 2447–2467.
- 190 A. El-Faham and F. Albericio, Synthesis and Application of N-Hydroxylamine Derivatives as Potential Replacements for HOBt, *European J. Org. Chem.*, 2009, **2009**, 1499–1501.
- 191 R. Subirós-Funosas, R. Prohens, R. Barbas, A. El-Faham and F. Albericio, Oxyma: An efficient additive for peptide synthesis to replace the benzotriazole-based HOBt and HOAt with a lower risk of explosion, *Chem. - A Eur. J.*, 2009, **15**, 9394–9403.
- 192 C. J. White and A. K. Yudin, Contemporary strategies for peptide macrocyclization, *Nat. Chem.*, 2011, **3**, 509–524.
- 193 K. Graumann and A. Premstaller, Manufacturing of recombinant therapeutic proteins in microbial systems, *Biotechnol. J.*, 2006, **1**, 164–186.
- 194 F. W. Studier and B. A. Moffatt, Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes, *J. Mol. Biol.*, 1986, **189**, 113–130.
- 195 F. Frottin, A. Martinez, P. Peynot, S. Mitra, R. C. Holz, C. Giglione and T. Meinnel, The proteomics of N-terminal methionine cleavage, *Mol. Cell. Proteomics*, 2006, **5**, 2336–2349.
- 196 K. F. Geoghegan, H. B. F. Dixon, P. J. Rosner, L. R. Hoth, A. J. Lanzetti, K. A. Borzilleri, E. S. Marr, L. H. Pezzullo, L. B. Martin, P. K. Lemotte, A. S. McColl, A. V. Kamath and J. G. Stroh, Spontaneous α -N-6-phosphogluconoylation of a 'His tag' in Escherichia coli: The cause of extra mass of 258 or 178 Da in fusion proteins, *Anal. Biochem.*, 1999, **267**, 169–184.
- 197 M. W. Pantoliano, E. C. Petrella, J. D. Kwasnoski, V. S. Lobanov, J. Myslik, E. Graf, T. Carver, E. Asel, B. A. Springer, P. Lane and F. R. Salemme, High-density miniaturized thermal shift assays as a general strategy for drug discovery, *J. Biomol. Screen.*, 2001, **6**, 429–440.
- 198 F. H. Niesen, H. Berglund and M. Vedadi, The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability, *Nat. Protoc.*, 2007, **2**, 2212–2221.
- 199 P. Cimmperman, L. Baranauskienė, S. Jachimovičiūtė, J. Jachno, J. Torresan, V. Michailoviene, J. Matuliene, J. Sereikaite, V. Bumelis and D. Matulis, A quantitative model of thermal stabilization and destabilization of proteins by ligands, *Biophys. J.*, 2008, **95**, 3222–3231.
- 200 N. Tsuji, K. Fukuda, Y. Nagata, H. Okada, A. Haga, S. Hatakeyama, S. Yoshida, T. Okamoto, M. Hosaka, K. Sekine, K. Ohtaka, S. Yamamoto, M. Otaka, E. Grave and H. Itoh, The activation mechanism of the aryl hydrocarbon receptor (AhR) by molecular chaperone HSP90, *FEBS Open Bio*, 2014, **4**, 796–803.
- 201 A. Soshilov and M. S. Denison, Ligand displaces heat shock protein 90 from overlapping binding sites within the aryl hydrocarbon receptor ligandbinding domain, *J. Biol. Chem.*, 2011, **286**, 35275–35282.
- 202 S. Crosson, S. Rajagopal and K. Moffat, The LOV domain family: Photoresponsive signaling modules coupled to diverse output domains, *Biochemistry*, 2003, **42**, 2–10.
- 203 M. Jerabek-Willemsen, C. J. Wienken, D. Braun, P. Baaske and S. Duhr, Molecular interaction studies using microscale thermophoresis., *Assay Drug Dev. Technol.*, 2011, **9**, 342–353.
- 204 M. Jerabek-Willemsen, T. Andre, R. Wanner, H. M. Roth, S. Duhr, P. Baaske and D.

Bibliography

- Breitsprecher, MicroScale Thermophoresis: Interaction analysis and beyond, *J. Mol. Struct.*, 2014, **1077**, 101–113.
- 205 G. A. Weiss, C. K. Watanabe, A. Zhong, A. Goddard and S. S. Sidhu, Rapid mapping of protein functional epitopes by combinatorial alanine scanning, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 8950–8954.
- 206 J. Han and G. Han, A Procedure for Quantitative Determination of Tris(2-carboxyethyl)phosphine, an odorless Reducing Agent More Stable and Effective than Dithiothreitol, *Anal. Biochem.*, 1994, **220**, 5–10.
- 207 M. C. Alliegro, Effects of dithiothreitol on protein activity unrelated to thiol- disulfide exchange: For consideration in the analysis of protein function with cleland's reagent, *Anal. Biochem.*, 2000, **282**, 102–106.
- 208 R. S. Morgan and J. M. Mcadon, Predictor for sulfur-aromatic interactions in globular proteins, *Int. J. Pept. Protein Res.*, 1980, **15**, 177–180.
- 209 S. K. Burley and G. A. Petsko, Amino-aromatic interactions in proteins, *FEBS Lett.*, 1986, **203**, 139–143.
- 210 B. L. Bodner, L. M. Jackman and R. S. Morgan, NMR study of 1:1 complexes between divalent sulfur and aromatic compounds: A model for interactions in globular proteins, *Biochem. Biophys. Res. Commun.*, 1980, **94**, 807–813.
- 211 G. Némethy and H. A. Scheraga, Strong interaction between disulfide derivatives and aromatic groups in peptides and proteins, *Biochem. Biophys. Res. Commun.*, 1981, **98**, 482–487.
- 212 R. Diamond, Real-space refinement of the structure of hen egg-white lysozyme, *J. Mol. Biol.*, 1974, **82**, 371–391.
- 213 K. N. M. M. Daeffler, H. A. Lester and D. A. Dougherty, Functionally important aromatic-aromatic and sulfur- π interactions in the D2 dopamine receptor, *J. Am. Chem. Soc.*, 2012, **134**, 14890–14896.
- 214 J. Coste, D. Le-Nguyen and B. Castro, PyBOP®: A new peptide coupling reagent devoid of toxic by-product, *Tetrahedron Lett.*, 1990, **31**, 205–208.
- 215 L. A. Carpino, H. Imazumi, B. M. Foxman, M. J. Vela, P. Henklein, A. El-Faham, J. Klose and M. Bienert, Comparison of the effects of 5- and 6-HOAt on model peptide coupling reactions relative to the cases for the 4- and 7-isomers, *Org. Lett.*, 2000, **2**, 2253–2256.
- 216 T. Clark, M. Hennemann, J. S. Murray and P. Politzer, Halogen bonding: The σ -hole, *J. Mol. Model.*, 2007, **13**, 291–296.
- 217 X. García-Llinás, A. Bauzá, S. K. Seth and A. Frontera, Importance of R-CF₃...O Tetrel Bonding Interactions in Biological Systems, *J. Phys. Chem. A*, 2017, **121**, 5371–5376.
- 218 Z. Weng and Q. Zhao, in *Protein-Protein Interactions: Methods and Applications: Second Edition*, eds. C. L. Meyerkord and H. Fu, Humana Press, 2nd edn., 2015, pp. 341–352.
- 219 S. Aydin, A short history, principles, and types of ELISA, and our laboratory experience with peptide/protein analyses using ELISA, *Peptides*, 2015, **72**, 4–15.
- 220 R. W. Bally and T. C. J. Gribnau, Some Aspects of the Chromogen 3, 3', 5, 5'-Tetramethylbenzidine as Hydrogen Donor in a Horseradish Peroxidase Assay, *Clin. Chem. Lab. Med.*, 1989, **27**, 791–796.
- 221 D. Josephy, T. Eling and R. Mason, The Horseradish Peroxidase-catalyzed Oxidation of 3,5,3',5'- Tetramethylbenzidine, *J. Biol. Chem.*, 1982, **257**, 3669–3675.

- 222 Y. Du, in *Protein-Protein Interactions: Methods and Applications: Second Edition*, 2015, vol. 1278, pp. 529–544.
- 223 N. J. Moerke, Fluorescence Polarization (FP) Assays for Monitoring Peptide-Protein or Nucleic Acid-Protein Binding, *Curr. Protoc. Chem. Biol.*, 2009, **1**, 1–15.
- 224 Y. Z. Ng, P. A. Baldera-Aguayo and V. W. Cornish, Fluorescence Polarization Assay for Small Molecule Screening of FK506 Biosynthesized in 96-Well Microtiter Plates, *Biochemistry*, 2017, **56**, 5260–5268.
- 225 A. Stefanachi, F. Leonetti, L. Pisani, M. Catto and A. Carotti, Coumarin: A natural, privileged and versatile scaffold for bioactive compounds, *Molecules*, 2018, **23**, 1–34.
- 226 J. Wang, J. Xie and P. G. Schultz, A Genetically Encoded Fluorescent Amino Acid, *J. Am. Chem. Soc.*, 2006, **128**, 8738–8739.
- 227 J. M. Cottam Jones, P. W. R. Harris, D. B. Scanlon, B. E. Forbes, M. A. Brimble and A. D. Abell, Fluorescent IGF-II analogues for FRET-based investigations into the binding of IGF-II to the IGF-1R, *Org. Biomol. Chem.*, 2016, **14**, 2698–2705.
- 228 M. P. Brun, L. Bischoff and C. Garbay, A very short route to enantiomerically pure coumarin-bearing fluorescent amino acids, *Angew. Chemie - Int. Ed.*, 2004, **43**, 3432–3436.
- 229 R. Strachan, J. Wood and R. Hirschmann, Synthesis and Properties of 4-Methyl-2-oxo-1,2-benzopyran-7-yl 6-D-Galactoside (Galactoside of 4-Methylumbelliferone), *J. Org. Chem.*, 1961, **27**, 1074.
- 230 D. A. Chan, P. D. Sutphin, N. C. Denko and A. J. Giaccia, Role of prolyl hydroxylation in oncogenically stabilized hypoxia-inducible factor-1 α , *J. Biol. Chem.*, 2002, **277**, 40112–40117.
- 231 G. L. Wang and G. L. Semenza, Desferrioxamine induces erythropoietin gene expression and hypoxia-inducible factor 1 DNA-binding activity: Implications for models of hypoxia signal transduction, *Blood*, 1993, **82**, 3610–3615.
- 232 W. L. Bragg, The Specular Reflection of X-rays, *Nature*, 1912, **90**, 410.
- 233 W. Friedrich, P. Knipping and M. Laue, Interferenzerscheinungen bei Röntgenstrahlen, *Ann. Phys.*, 1913, **346**, 971–988.
- 234 J. Kendrew, G. Bodo, H. Dintzis, R. Parrish and H. Wyckoff, A Three-Dimensional Model of the Myoglobin Molecule Obtained by X-ray Analysis, *Nature*, 1958, **181**, 662–666.
- 235 J. C. Kendrew, R. E. Dickerson, B. E. Strandberg, R. G. Hart, D. R. Davies, D. C. Phillips and V. C. Shore, Structure of Myoglobin A three-dimensional fourier synthesis at 2 Å Resolution, *Nature*, 1960, **48**, 422–427.
- 236 E. F. Garman and T. R. Schneider, Macromolecular Cryocrystallography, *J. Appl. Crystallogr.*, 1997, **30**, 211–237.
- 237 P. F. Lindley, The use of synchrotron radiation in protein crystallography, *Radiat. Phys. Chem.*, 1995, **45**, 367–383.
- 238 H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Baht, H. Weissig, I. N. Shindyalov and P. E. Bourne, The Protein Data Bank, *Nucleic Acids Res.*, 2000, **28**, 235–242.
- 239 R. PDB, PDB Current Holdings Breakdown, <http://www.rcsb.org/pdb/statistics/holdings.do>, (accessed 10 October 2019).
- 240 A. McPherson and J. A. Gavira, Introduction to protein crystallization, *Acta Crystallogr. Sect. F Struct. Biol. Commun.*, 2014, **70**, 2–20.

Bibliography

- 241 C. X. Weichenberger, P. V. Afonine, K. Kantardjieff and B. Rupp, The solvent component of macromolecular crystals, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2015, **71**, 1023–1038.
- 242 I. Müller, Guidelines for the successful generation of protein-ligand complex crystals, *Acta Crystallogr. Sect. D Struct. Biol.*, 2017, **73**, 79–92.
- 243 H. Hou, M. Shi, S. Y. Hu, F. Ahmad, B. Zhang, Z. H. Chen and D. C. Yin, A systematic comparison of sitting and hanging-drop crystallization using traditional and cross-diffusion microbatch crystallization plates, *J. Cryst. Growth*, 2019, **521**, 1–8.
- 244 M. A. Dessau and Y. Modis, Protein crystallization for X-ray crystallography, *J. Vis. Exp.*, 2010, **9**, 1–6.
- 245 J. W. Pflugrath, Macromolecular cryocrystallography - Methods for cooling and mounting protein crystals at cryogenic temperatures, *Methods*, 2004, **34**, 415–423.
- 246 G. Rhodes, *Crystallography Made Crystal Clear*, Elsevier, Third., 2006.
- 247 V. K. Pecharsky and P. Y. Zavalij, in *Fundamentals of Powder Diffraction and Structural Characterization of Materials*, 2009, pp. 239–262.
- 248 A. Brito and M. Archer, in *Practical Approaches to Biological Inorganic Chemistry*, Elsevier, 2013, pp. 217–255.
- 249 G. Bunkóczi, N. Echols, A. J. McCoy, R. D. Oeffner, P. D. Adams and R. J. Read, Phaser.MRage: Automated molecular replacement, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2013, **69**, 2276–2286.
- 250 T. Bergfors, Seeds to crystals, *J. Struct. Biol.*, 2003, **142**, 66–76.
- 251 D. Liebschner, P. V. Afonine, M. L. Baker, G. Bunkoczi, V. B. Chen, T. I. Croll, B. Hintze, L. W. Hung, S. Jain, A. J. McCoy, N. W. Moriarty, R. D. Oeffner, B. K. Poon, M. G. Prisant, R. J. Read, J. S. Richardson, D. C. Richardson, M. D. Sammito, O. V. Sobolev, D. H. Stockwell, T. C. Terwilliger, A. G. Urzhumtsev, L. L. Videau, C. J. Williams and P. D. Adams, Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in Phenix, *Acta Crystallogr. Sect. D Struct. Biol.*, 2019, **75**, 861–877.
- 252 A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni and R. J. Read, Phaser crystallographic software, *J. Appl. Crystallogr.*, 2007, **40**, 658–674.
- 253 P. Emsley, B. Lohkamp, W. G. Scott and K. Cowtan, Features and development of Coot, *Acta Crystallogr. Sect. D*, 2010, **66**, 486–501.
- 254 D. Rajalingam, K. M. Kathir, K. Ananthamurthy, P. D. Adams and T. K. S. Kumar, A method for the prevention of thrombin-induced degradation of recombinant proteins, *Anal. Biochem.*, 2008, **375**, 361–363.
- 255 J. P. S. Makkerh, C. Dingwall and R. A. Laskey, Comparative mutagenesis of nuclear localization signals reveals the importance of neutral and acidic amino acids, *Curr. Biol.*, 1996, **6**, 1025–1027.
- 256 M. Eben Massari and C. Murre, Helix-Loop-Helix Proteins: Regulators of Transcription in Eucaryotic Organisms, *Mol. Cell Biol.*, 2000, **20**, 429–440.
- 257 J. T. Henry and S. Crosson, Ligand-Binding PAS Domains in a Genomic, Cellular, and Structural Context, *Annu. Rev. Microbiol.*, 2011, **65**, 261–286.
- 258 M. J. Percy, S. M. Mooney, M. F. McMullin, A. Flores, T. R. Lappin and F. S. Lee, A common polymorphism in the oxygen-dependent degradation (ODD) domain of hypoxia inducible factor-1 α (HIF-1 α) does not impair Pro-564 hydroxylation, *Mol. Cancer*, 2003, **2**, 31.

- 259 C.-J. Hu, A. Sataur, L. Wang, H. Chen and M. C. Simon, The N-Terminal Transactivation Domain Confers Target Gene Specificity of Hypoxia-inducible Factors HIF-1 and HIF-2 α , *Mol. Biol. Cell*, 2007, **18**, 4528–4542.
- 260 V. L. Dengler, M. Galbraith and J. M. Espinosa, Transcriptional Regulation by Hypoxia Inducible Factors, *Crit. Rev. Biochem. Mol. Biol.*, 2014, **49**, 1-15.