UNIVERSITY OF SOUTHAMPTON

FACULTY OF NATURAL AND ENVIRONMENTAL SCIENCES

Chemistry

Volume 1 of 1

Towards a small molecule inhibitor of the HIF-1α/HIF-1β protein-protein interaction

by

Charlotte Eve Lawrence

Thesis for the degree of Doctor of Philosophy

February 2015

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF NATURAL AND ENVIRONMENTAL SCIENCES

Chemistry

Doctor of Philosophy

TOWARDS A SMALL MOLECULE INHIBITOR OF THE HIF-1α/HIF-1β PROTEIN-PROTEIN INTERACTION

Charlotte Eve Lawrence

Hypoxia-inducible factor (HIF) is a heterodimeric, oxygen-dependent, transcription factor that regulates the cellular response to hypoxia by directing the expression of multiple genes, such as those involved in angiogenesis and glucose transport. HIF activation has been shown to aid the survival of cancer cells in hypoxic regions; hence it is viewed as a potentially important target for cancer therapy.

There are two predominant isoforms of HIF, HIF-1 and HIF-2, formed by heterodimerisation of HIF-1α or HIF-2α, respectively, with HIF-1β. The dimerisation of the two subunits is necessary for DNA-binding and subsequent activation of transcription. Miranda *et al.* (2013) have recently identified a six amino acid cyclic peptide inhibitor of HIF dimerisation (*cyclo*-CLLFVY); the Tat-tagged variant is called P1. This has shown activity within several cell-based assays.[1]

This project sought to identify which amino acid residues of *cyclo*-CLLFVY were critical to its activity by synthesising five alanine analogues and testing them in cell and biophysical assays. It was not possible to identify an active motif and it could be concluded that the specific conformation of the intact cyclic peptide is required for activity. The functionality of independently bacterially expressed fragments of HIF-1α and HIF-1β was also validated by an EMSA. The Tavassoli group used these proteins to establish the binding location of the inhibitor to the HIF-1α-PAS-B domain (work by A. Tavassoli and A. Male).

Table of Contents

List of tables

List of figures

DECLARATION OF AUTHORSHIP

I, Charlotte Lawrence

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.

Towards a small molecule inhibitor of the HIF-1α/HIF-1β protein-protein interaction

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. Parts of this work have been published as:

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, *Journal of the American Chemical Society*, 2013, **135**, 10418-10425.

Signed: .....

Date: .....

Acknowledgements

Firstly, I would like to thank the Gerald Kerkut Charitable Trust, Dr Niall Walker, Dr Alan Thomas and Professor Robert Walker for funding the research and their support.

Thank you to Dr Ali Tavassoli and Professor Paul Townsend for making it possible for me to carry out this research, their guidance and encouragement. I would like to extend this thanks to their research groups, for their helpful discussions and technical support.

I am also grateful to Professor Keith Fox and Scott Kimber for guidance with the EMSA studies. To Dr Julie Herniman and Dr Stuart Findlow for their expertise in mass spectrometry and NMR, respectively. Also to Dr Emilia Danilowicz-Lueber (Nanotemper technologies) for her advice with the MST studies and analysis.

I would like to thank the following for their continued support, advice, friendship, helpful discussions and proofreading: Dr Ben Ainsworth, Dr Elizabeth Willow-Hill, Mary Rook, Dick Richmond, Prune Richmond, Phil Lawrence, Hannah Yuill, Nicky Macgregor, Anne Edmondson, Nick Cherry, Dr Joanna Brigg, Maths Mark, Dr Charles Ainsworth, Dr Suzy Shuttleworth and Bethan Carden.

I would like to extend a special thank you to Colette Richmond, Dr Becky Sands and Dr Abigail Male, without whom I would never have been able to get through my PhD. In the toughest times you have been the brightest stars. Thank you.

Throughout this project, I have transitioned from a graduate chemistry student to that of a chemical biology researcher. Many different biological and cell-based techniques have been investigated during the project that have increased my skill set and led to a greater understanding into the field of chemical biology methodology.

Definitions

Hypoxia: A state where oxygen demand exceeds oxygen supply, oxygen concentrations are below those required for normal cell function.

P1: Tat-tagged *cyclo*-CLLFVY, HIF-1 dimerisation inhibitor, discovered by Miranda *et al*. (2013).[1]

Abbreviations

ACF - Acriflavine

AICAR - Aminoimidazole carboxamide ribonucleotide

APS - Ammonium persulphate

ARNT - Aryl hydrocarbon receptor nuclear translocator

ATIC - Gene encoding inosine monophosphate synthase

BCA - Bicinchoninic assay

bHLH - Beta-helix-loop-helix

BSA - Bovine serum albumin

CAIX - Carbonic anhydrase 9

CBP - CREB-binding protein

CREB - cAMP response element-binding

C-TAD - C-terminal transactivational domain

CTBP - C-terminal binding protein

DCM - Dichloromethane

DIC - *N*,*N*′-Diisopropylcarbodiimide

DIPEA - *N*,*N*- Diisopropylethylamine

DFX - Desferrioxamine

DMEM - Dulbecco's Modified Eagles Medium

DMF - Dimethyformamide

DMSO - Dimethyl sulfoxide

DNA - Deoxyribonucleic acid

dNTPs - Deoxynucleotide triphosphates

DTT - DL-Dithiothreitol

EDC - N-(3-Dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride

EDTA - Ethylenediaminetetra acetic acid

EDT - 1,2-Ethanedithiol

ELISA - Enzyme-linked immunosorbant assay

EMSA - Electrophoretic mobility shift assay

EPO - Erythropoietin

eq - Equivalents

ESI - Electrospray ionisation

FBS - Fetal bovine serum

Fmoc - Fluorenylmethyloxycarbonyl

FIH - Factor inhibiting HIF

FPLC - Fast protein liquid chromatography

GAPDH - Glyceraldehyde-3-phosphate dehydrogenase

GLUT1 - Glucose transporter 1

GST - Glutathione S-transferase

HIF - Hypoxia-inducible factors

HIS3 - Gene encoding imidazoleglycerol-phosphate

HOBt - Hydroxybenzotriazole

HPLC - High-performance liquid chromatography

HRE - Hypoxic response element

HRMS - High resolution mass spectrometry

IPTG - Isopropyl β-D-1-thiogalactopyranoside

ITC - Isothermal titration calorimetry

MCF-7 - Michigan Cancer Foundation-7

MDM2 - Mouse double minute 2 homolog

MDMX - MDM2-like p53 binding protein

MS - Mass spectrometry

MST - Microscale thermophoresis

m/z - Mass to charge ratio

NEB - New England Biolabs

Npu - *Nostocpunctiforme*

NMP - *N*-Methylpyrrolidone

NMR - Nuclear magnetic resonance

N-TAD - N-terminal transactivational domain

ODDD - Oxygen dependent degradation domain

PAS-A - Per (period circadian protein)-ARNT-Sim (single-minded protein)-A

PAS-B - Per (period circadian protein)-ARNT-Sim (single-minded protein)-B

PBS - Phosphate buffered saline

PCNA - Proliferating cell number antigen

PCR - Polymerase chain reaction

pH - Potential of hydrogen

PLA - Proximity ligation assay

PLB - Passive lysis buffer

PLP - Pyridoxal phosphate

PNK - Polynucleotide kinase

PTEN - Phosphatase and tensin homolog

PyBOP - (Benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate

RACK1 - Receptor for activated C kinase 1

RIPA - Radioimmunoprecipitation assay

rpm - Revolutions per minute

rt - Retention time

RTHS - Reverse two-hybrid system

RT-PCR – Real time Polymerase chain reaction

SDM - Site-directed mutagenesis

SDS - Sodium dodecyl sulphate

SDS-PAGE - Sodium dodecyl sulphate polyacrylamide gel electrophoresis

SICLOPPS - Split-intein circular ligation of peptides and proteins

SOC - Super optimal broth containing glucose

SPR - Surface plasmon resonance

Ssp - *Synechocystis* sp. PCC6803

Spy - Aldrithiol protecting group

SSAT1/2 - Spermidine/spermine N^1-acetyltransferase-1/2

TCEP - Tris(2-carboxyethyl)phosphine

Tat - Trans-activator of transcription

TAE - Tris-acetate-EDTA

TBE - Tris-borate-EDTA

TFA - Trifluoroacetic acid

TIS - Triisopropylsilane

T-Jump - Temperature jump

Trx - Thioredoxin

TSAP - Thermosensitive alkaline phosphate

UV - Ultra violet

VEGF - Vascular endothelial growth factor

VHL - Von Hippel-Lindau

3-AT - 3-Amino-1,2,4 triazole

Nuclear Magnetic resonance (NMR) Abbreviations

H - Hydrogen

m - Multiplet

s - Singlet

d - Doublet

dd - Doublet of doublet

br s - Broad singlet

br d - Broad doublet

Amino Acids Abbreviations

Amino Acid	3-Letter Abbreviation	1-Letter Abbreviation
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamine	Gln	Q
Glutamic acid	Glu	E
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V
Propargylglycine (non-natural)	Pra	Pr

1 Introduction

1.1 Protein-Protein interactions

Protein-protein interactions are essential for biological function within cells. They are fundamental to many biological processes, such as transcription, cell signalling and differentiation.[2] Protein-protein interactions can be classified as either stable or transient.[3] Stable interactions are associated with large complexes; while transient interactions are thought to control the majority of cellular processes. There are hundreds of thousands of protein-protein interactions occurring in each human cell.[2, 4] Targeting protein-protein interactions is complex and sometimes referred to as the 'high-hanging fruits of drug discovery' (reviewed by Wells *et al.* (2007)).[5] However, many active small molecules have been discovered that disrupt protein-protein interactions.[5-8] Using small molecules to inhibit protein-protein interactions fundamental to a specific biological pathway is an important emerging area of cancer drug design.[9] This has been extensively reviewed by Fry *et al.* (2005) and White *et al.* (2008).[9-11] An example is the hypoxia-inducible factor-1 (HIF-1), a transcription factor that promotes cancer cell survival under hypoxic conditions.[12] The activation of HIF-1 is dependent on the protein-protein interaction of its two subunits, hence providing a target for drug design.[13]

1.2 The transcription factor: hypoxia-inducible factor-1

HIF-1 is an oxygen-regulated heterodimeric transcription factor.[12, 14] It regulates the expression of many hypoxia-inducible genes, such as angiogenic growth factors and erythropoietin (EPO).[15, 16] Oxygen is fundamental to many processes critical to survival in mammalian cells (i.e. metabolism, growth and cell proliferation). Under hypoxic conditions, when the local oxygen concentration is low and inadequate for normal function (which is often considered to be 1-21% of a normoxic cell of similar nature), HIF-1 activation facilitates up-regulation of genes that aid cell survival.[17]

HIF-1 is essential to oxygen homeostasis, controlling a diverse number of physiological processes in mammalian cells (Figure 1).[18] HIF-1 activates genes that promote cell survival

either through cellular adaptation to the hypoxic environment or increased oxygen transport into the cells. It has been hypothesised that in human endothelial cells, over 2% of the genome is capable of being induced by the hypoxia response network, illustrating the vast influence of the HIF-1 transcription factor (Figure 1).[19]

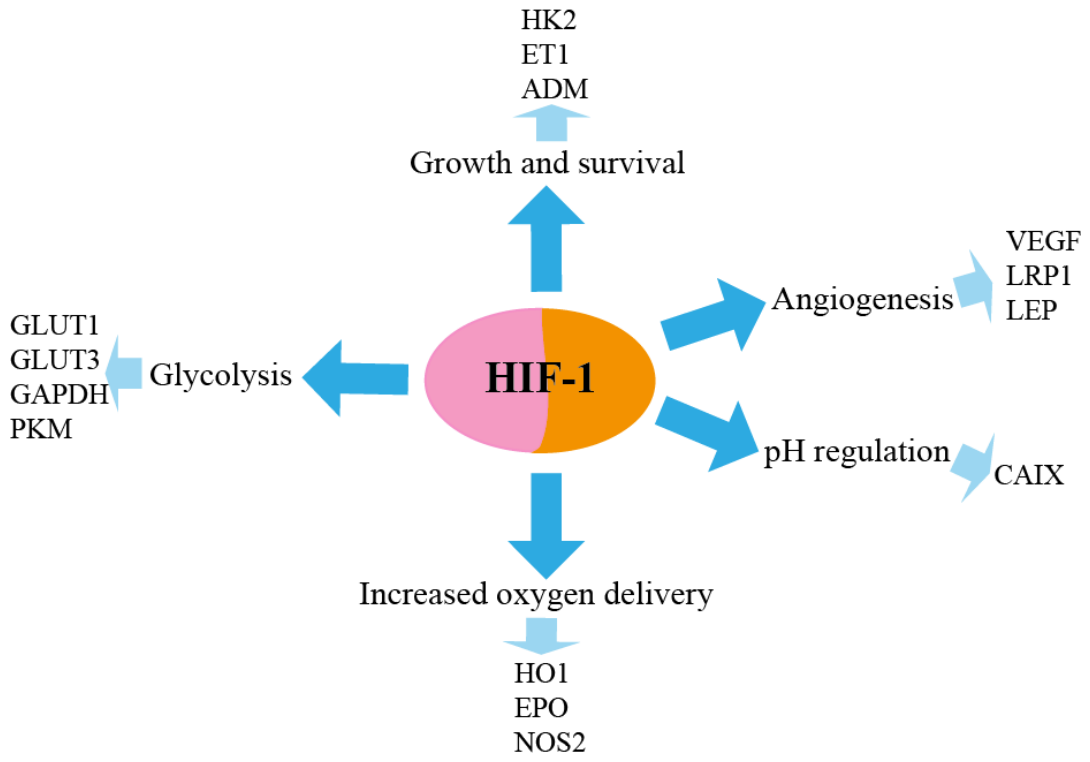


Figure 1 Examples of HIF-1 responsive genes.

HIF-1 controls the expression of many genes involved in the response to hypoxia including those involved in growth and survival, angiogenesis, pH regulation, increased oxygen delivery and glycolysis. The diagram uses the following abbreviations: ADM: adrenomedullin; CAIX: carbonic anhydrase IX; EPO: erythropoietin; ET1: endothelin-1; GAPDH: glyceraldehyde-3-phosphate-dehydrogenase; GLUT1: glucose transporter 1; GLUT3: glucose transporter 3; HK2: hexokinase 2; HO1: heme oxygenase 1; LEP: leptin; LRP1: LDL-receptor-related protein 1; PKM: pyruvate kinase M; VEGF: vascular endothelial growth factor.[17-19]

HIF-1 was first identified by Semenza *et al.* (1991), through the discovery of the hypoxic induction of the erythropoietin gene and the identification that HIF-1 binds to the hypoxia response element (HRE) upstream of the gene.[12, 15]

1.3 HIF-1 structure

HIF-1 is a heterodimeric protein comprising of two subunits HIF-1α (120-130 kDa) and HIF-1β (91-94 kDa), which dimerise to form the HIF-1 functional transcription factor.[20] HIF-1β is often referred to as the aryl hydrocarbon receptor nuclear translocator (ARNT).[21] HIF-1α and HIF-1β dimerise to form the active HIF-1 transcription factor; this then binds to specific DNA sequences upstream of its target genes (i.e. 5'-RCGTG-3', where R is A or G); known as hypoxia response elements (HRE).[13] As outlined in section 1.2, once bound to the HRE, the HIF-1 transcription factor regulates the transcription of many target genes.[18]

HIF-1α and HIF-1β both have N-terminal basic-helix-loop-helix (bHLH) domains critical for DNA-binding.[13, 22] Both proteins bind to DNA in the major groove via these domains. In addition, they also have two Per (period circadian protein)-ARNT-Sim (single-minded protein) domains (PAS-A and B) that mediate dimerisation essential for DNA-binding (Figure 2).[13, 23] Jiang *et al*. (1996) showed that amino acids 1-166 of HIF-1α, encompassing the bHLH and PAS-A domains, are sufficient for dimerisation to take place.[13] Later Yang *et al.* (2005) suggested that both PAS domains were necessary for heterodimer formation.[24]

The C-terminus of HIF-1α contains two transactivational domains: C-terminal transactivational domain (C-TAD), and an N-terminal transactivational domain (N-TAD) (Figure 2). Co-activators such as CREB-binding protein (CBP) and p300 bind to the C-TAD, leading to the activation of transcription.[25] Furthermore, an oxygen-dependent degradation domain (ODDD) resides in the C-terminus of HIF-1α. Proline residues contained in this region control degradation of the protein in the presence of oxygen (section 1.6).[26] HIF-1β also contains a C-TAD domain (Figure 2).

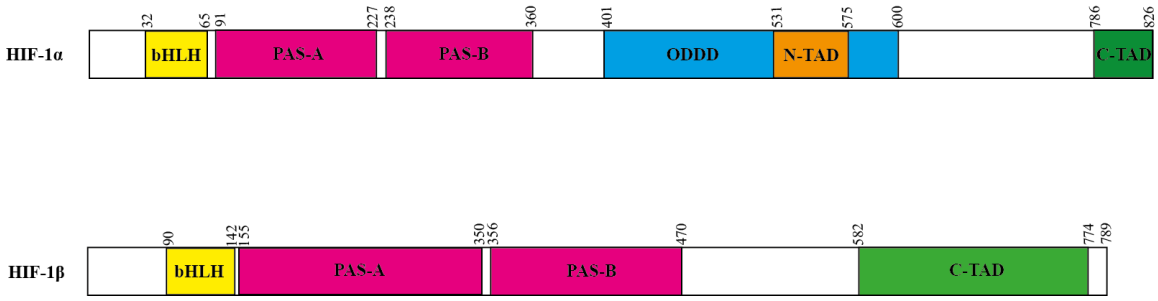


Figure 2 Schematic of HIF-1α and HIF-1β structure.[25-27]

HIF-1α contains bHLH, PAS-A, PAS-B, ODDD, N-TAD and C-TAD. HIF-1β contains bHLH, PAS-A, PAS-B and C-TAD domains.

Due to the instability of recombinant HIF-1α, no crystal structure of the full-length protein is currently reported in the literature. Structural characterisation has been impeded by homodimerisation, instability and insolubility of the protein *in vitro*.[28, 29] In contrast, owing to its increased stability *in vitro*, structural analysis of the HIF-2 isoform (section 1.4) is more readily reported.[30-32] Initially, NMR analysis of HIF-2α-PAS-B/HIF-1β-PAS-B was reported and showed that the β-sheet surface of each subunit mediates dimerisation.[23, 33] The HIF-2α-PAS-B domain adopts a α/β domain fold, characteristic of other PAS domains, signified by several α-helices and a five stranded anti-parallel β-sheet.[33, 34] Scheuermann *et al.* (2009) built upon the NMR structures and obtained a crystal structure analysis of the same domains.[30] The NMR model identified interfacial ionic interactions between the proteins. In order to study the interaction they used mutants to reverse an interfacial salt bridge (HIF-2α-PAS-B (R247E) and HIF-1β-PAS-B (E362R)).[30] This mutation increased the binding affinity of the dimer allowing further biophysical characterisation. Crystal structure analysis of the mutant complex showed an anti-parallel interaction of the domains, burying a total of 2050 $Å^2$ upon dimerisation.[30] HIF-2α-PAS-B contained a large, irregularly shaped cavity (290 $Å^3$), filled with eight ordered water molecules, that could be a target for inhibitor binding.[31,30] Later, this presence of this cavity was exploited by Scheuermann *et al.* (2013), whereby allosteric inhibitors binding within this hydrophobic cavity were identifed. These ligands were selective over the HIF-1α isoform and prevented HIF-2α/HIF-1β dimerisation and DNA-binding.[35]

The HIF-1α and HIF-2α-PAS-B domains have 74% overall identity and 91% at the HIF-1β interface.[29, 30] Zhu *et al.* (2012) utilised a crystal structure of the homologous HIF-2α-PAS-B to construct mutant forms of HIF-1α-PAS-B (R245E) and HIF-1β-PAS-B (E362R)

and generate a NMR structure of this domain (mutant expanded upon in section 2.7.1).[28] The HIF-1α-PAS-B (R245E) mutant improved the solubility and stability of the domain and opened up the possibility of structural characterisation. Later, within the same research group, Cardoso *et al.* (2012) generated a crystal structure of HIF-1α-PAS-B (R245E)/HIF-1β-PAS-B (E362R) (Figure 3).[29] Additionally, isothermal titration calorimetry (ITC) characterisation of the dimerisation, not possible with the wild-type, showed a K_D of 125 nM between these proteins. This characterisation allowed identification of key residues involved in the protein-protein interaction, and highlighted differences to the HIF-2 interaction region. Specifically, analysis showed that the β-sheet interaction was similar to that with HIF-2 and three amino acid differences were identified at the interface. These included T327, V336 and V342 in HIF-1α PAS-B, which are replaced by P329, M338, L344 in HIF-2α PAS-B. A single mutation of HIF-2α M338V led to a 10-fold decrease in binding affinity. Hence, these variations may explain the lower binding affinity of HIF-2α to HIF-1β in comparison to HIF-1α.[29] The structure also showed an internal cavity similar to that of HIF-2; however, the cavity was smaller, containing different residues and water occupancies. Specifically, different orientation of H291, and sequence differences of A275I, L317V between HIF-1α and HIF-2α, respectively.[29] This may be a defining difference in the search for inhibitors specific to each isoform.

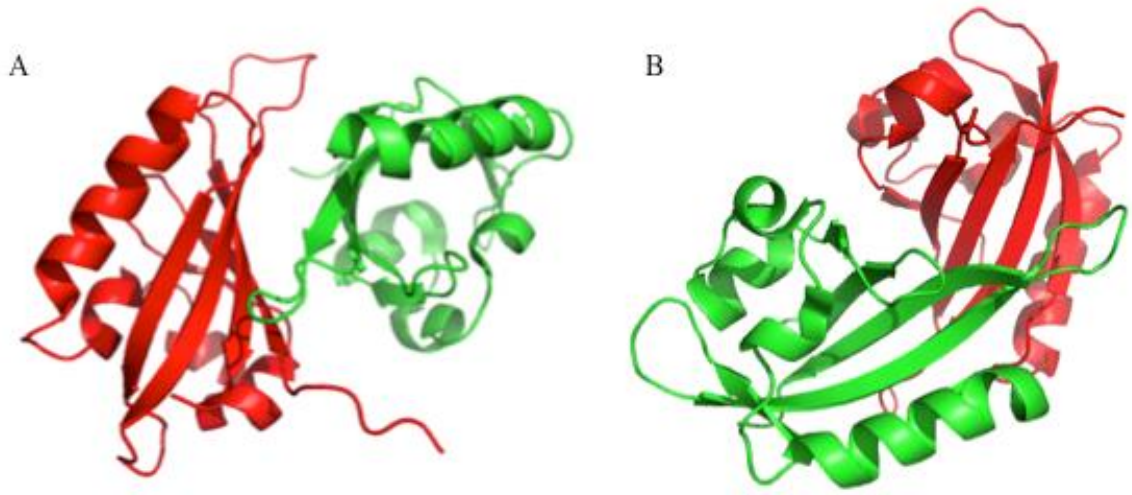


Figure 3 Crystal structure of HIF-1α-PAS-B and HIF-1β-PAS-B dimerisation interface. RCSB protein data bank structure PDB-4H6J. HIF-1α-PAS-B238-248 (R245E) mutant (red) and HIF-1β-PAS-B357-470 (E362R) mutant (green) (A) and (B) represent two different orientations. The diagram indicates the β-sheet interaction between the two proteins. Residues T327, V336 and V342 which differ from the HIF-2α isoform are located at this interface.[29]

1.4 HIF-1 isoforms

There are three HIF isoforms; HIF-1, HIF-2 and HIF-3. All three isoforms share the HIF-1β subunit, but have a different α subunit, HIF-1α, HIF-2α and HIF-3α, respectively.[36-39]

HIF-2, discovered in 1997 by several independent groups, is sometimes referred to as endothelial PAS protein (EPAS1), HIF-like factor (HLF), member of the PAS super family (MOP2) or HIF-related factor (HRF). It is also an important transcription factor, which is regulated by and functions by similar mechanisms to HIF-1.[36-38, 40] The structure of HIF-2α is similar to that of HIF-1α, containing bHLH, PAS-A and B, NTAD, CTAD and ODDD domains. The DNA-binding, PAS domains and ODDD have high sequence similarity, in comparison to the C-TAD; a sequence alignment below highlights this (Figure 4).[41] The exact sequence similarity figures and amino acid sequence identity for each domain vary between papers and reviews; this may be due to different analysis software, or different variants of the same isoform being analysed.[27, 36, 42] However, all concur that the N-terminal domains (bHLH, PAS-A and PAS-B) have a much higher

similarity between isoforms than that of the C-terminal domains (ODDD, N-TAD and C-TAD).[27, 36, 42]

Unlike HIF-1, HIF-2 is known to be present at different levels at different times in hypoxia and in certain tissue types more than others.[41] HIF-1α is expressed in nearly all cell types, however, HIF-2α expression is more restricted and more abundant in blood vessels specifically.[36, 38] HIF-1 and HIF-2 have many shared gene targets, but also have unique targets.[43] Hu *et al*. (2003) used DNA microarrays to investigate the presence of specific HIF-2 target genes.[44] They concluded that HIF-1 and HIF-2 have many overlapping targets and HIF-1 has unique targets, whereas for HIF-2 no specific targets were identified in this study.[44] The results suggested that glycolytic genes are induced only by HIF-1α in some cell lines.[44]

HIF-1 and HIF-2 also interact with other proteins in different ways and have different roles within some biological processes.[45] An example of this is their interaction with p53, hypoxia and other tumour related-stresses that activate p53. This tumour suppressor protein is crucial to cell function as it regulates the cell cycle. It is found to be frequently mutated in cancer and has a complex involvement in many pathways.[46] p53 levels are often found to be elevated in regions of hypoxia, but this is not a direct consequence of hypoxia itself.[47] HIF-1 and HIF-2 have different effects on the p53 pathway, this was reviewed by Keith *et al.* (2012) who described the complexity of the different roles and feedback systems.[43] HIF-1α is hypothesised to bind to p53 through an interaction with MDM2. This interaction leads to HIF-1α degradation, in turn, supressing HIF-1 activity.[48-50] In contrast, HIF-2α binds to p53 directly and not MDM2 and inhibits p53 through numerous pathways.[51] Furthermore, c-Myc/MAX are transcription factors that are implicated in pathways involving proliferation and apoptosis. HIF-1 and HIF-2 putatively interact with the Myc and MAX cascade in different ways.[43, 52]

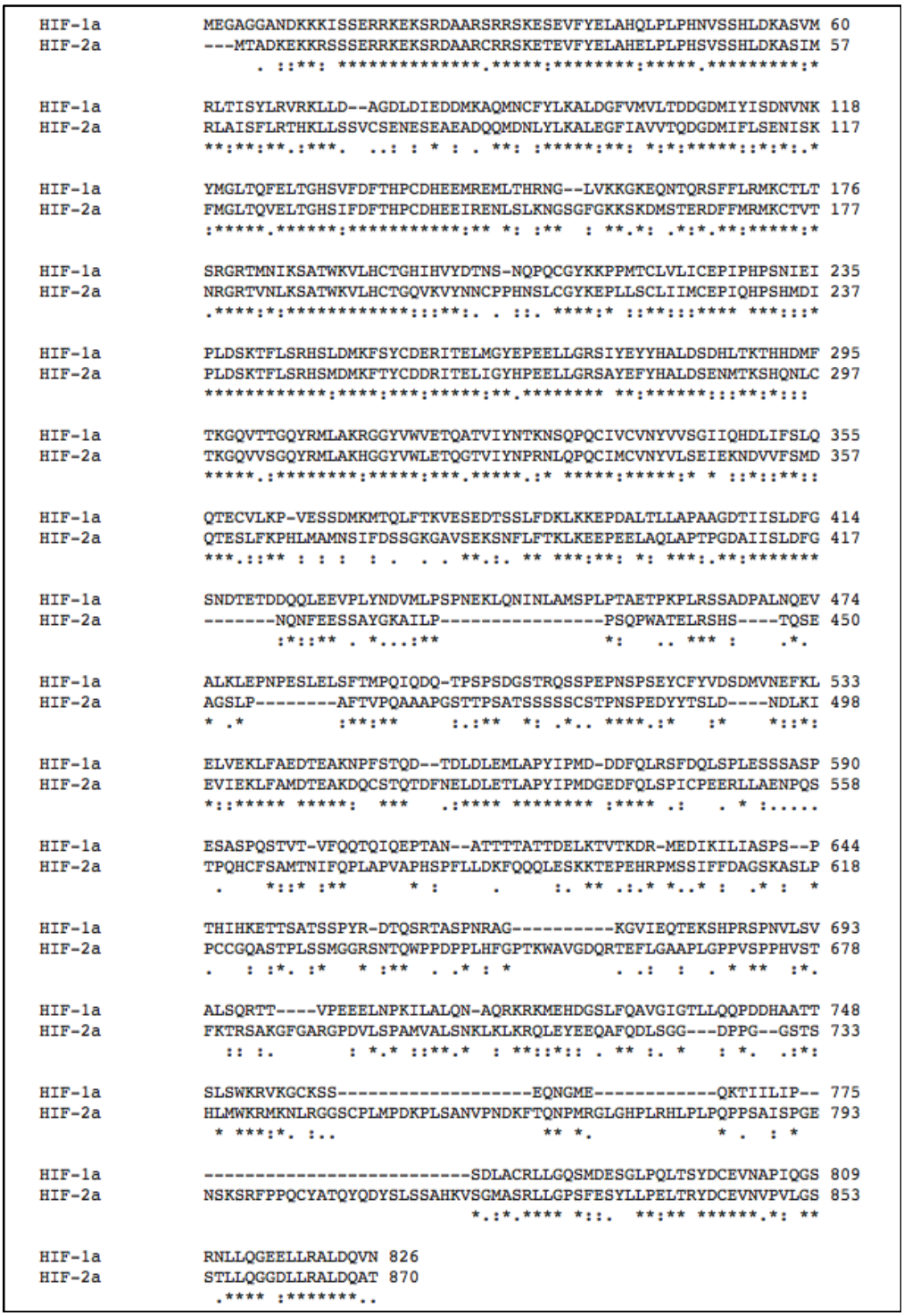


Figure 4 Clustal 2.1 multiple sequence alignment of HIF-1α and HIF-2α.[53]

(*) represents a conserved sequence, (:) represents a conservative difference, (.) represents a semi-conservative difference, ( ) respresents a non-conservative difference and (-) represents where sequences are insufficiently similar for alignment. The N-terminal domains have a greater sequence similarity than the C-terminal domains when comparing the two isoforms.

In addition, Ashcroft *et al*. (2006) investigated the effect of each isoform on the expression of target genes under hypoxia. VEGF expression in MCF-7 cells was primarily stimulated by HIF-1; whereas in renal carcinoma cells (RCC4), HIF-2 predominantly controlled VEGF expression where HIF up-regulation was a result of a loss in von Hippel-Lindau tumour suppressor protein (VHL) function.[54] A reciprocal connection between the two isoforms was observed; under hypoxia HIF-2α siRNA led to an increase in HIF-1 stimulated VEGF expression in MCF-7 cells. The body of literature indicates that HIF-1 and HIF-2 have both overlapping and distinct roles; further investigation into the implication of these variations is needed, in order to gain a clear picture of the networks.

In contrast, HIF-3 activity is less well understood, but it is hypothesised to be a negative regulator of the hypoxia response network.[39] HIF-3 was first identified in 1998 and was found to also be up-regulated at low oxygen concentrations.[55] HIF-3 is composed of HIF-3α and HIF-1β and the dimer recognises the HRE; this isoform was purified and characterised by Hara *et al.* (2001) and was found to lack the C-terminal transactivation domains of HIF-1α.[55, 56] HIF-3 was therefore suggested to be a negative regulator, supressing expression of genes under the control of HIF-1 and HIF-2.[56] One of the most well characterised variants of HIF-3 is inhibitory Per/ARNT/Sim domain protein (IPAS).[57] IPAS has sequence homology to HIF-1α in the bHLH domain (75%), PAS-A (34%) and PAS-B domains (40%).[57] This IPAS variant lacks the N-TAD and C-TAD domains required for transcriptional activation.[57] Makino *et al*. (2001) demonstrated the ability of IPAS to act as a negative regulator; here IPAS had an inhibitory effect on VEGF gene expression in the cornea.[39] Later, Augstein *et al.* (2010) described that HIF-3 has a more restricted expression pattern when compared to the other two isoforms.[58] It is restricted to certain organs; these authors interestingly report different roles of variants of HIF-3, showing only the HIF-3α2 variant expression to be inducible by hypoxia and not HIF-3α1 and HIF-3α3 in HUVEC cells.[58]

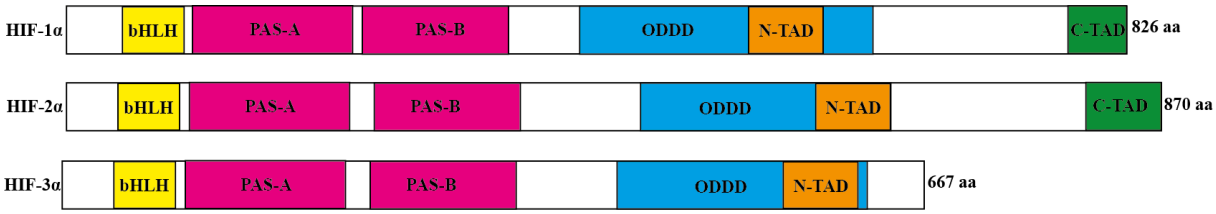


Figure 5 Illustration of HIF isoforms: HIF-1, HIF-2 and HIF-3.[42, 59]

The HIF-3α1 variant is illustrated; other splice variants exist. [42]

1.5 HIF-1 function

As previously described, HIF-1α and HIF-1β heterodimerise prior to binding to the HRE of genes under hypoxic control. Dimerisation of the two subunits can occur in the absence of DNA and is a necessary precursor to DNA-binding (Figure 6).[13, 60, 61] The active complex then recruits cofactors, such as p300/CBP, which initiate transcription of target HIF responsive genes (Figure 6).[62, 63] The C-terminal domain of HIF-1α interacts with the cysteine and histidine rich 1 (CH1) domain of p300; the interaction is regulated by a leucine rich hydrophobic interface and HIF-1α Leu795, Cys800, Leu818 and Leu822 are critical.[64] p300 and CBP are co-activators, which increase the transcription of target genes; they do this by relaxing the chromatin by histone deacetylase activity and recruiting transcription machinery.[62, 65] These co-activators bind to the C-terminal transactivation domains of HIF-1.[62]

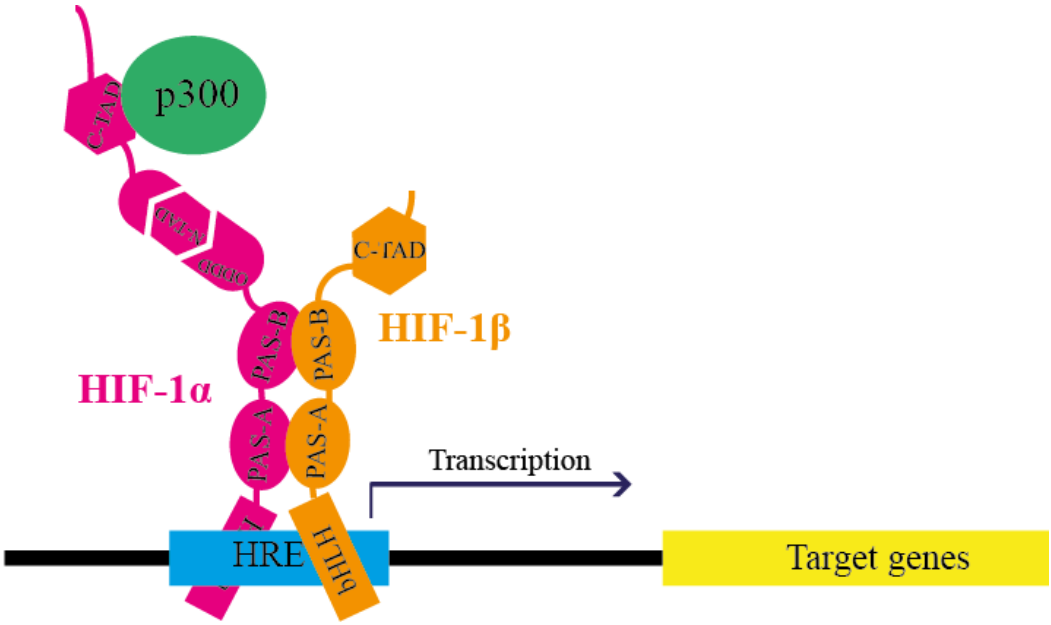


Figure 6 Illustration of HIF-1 function.

The HIF-1α/HIF-1β dimer binds to the HRE upstream of target genes. Co-activators such as p300 are recruited and initiate the transcription of target genes.

Other transcription co-activators can also be recruited by HIF-1α; for example, steroid receptor co-activator-1 (SRC-1) and transcription intermediary factor 2 (TIF2), both interact with HIF-1, enhancing transcription.[66]

1.6 HIF-1 regulation

HIF-1β is constituently expressed and present in cells at similar levels regardless of cellular oxygen concentrations. HIF-1α is regulated pre- and post-translationally and present at very low levels under normoxia. HIF-1α is regulated by both oxygen-dependent and oxygen-independent pathways. The regulation is complex with many feedback loops involved.[67] These include the ubiquitination and receptor of activated kinase C (RACK1) binding to HIF-1α. Both will be discussed later in sections 1.6.1 to 1.6.4.

1.6.1 HIF-1α ubiquitination by oxygen regulated prolyl hydroxylation

The most well cited regulation pathway is that of hydroxylation of HIF-1α (oxygen-dependent regulation) by prolyl-hydroxylases.[26] During normoxia, HIF-1α undergoes proteasomal degradation via an ubiquitin-dependent pathway.[41] This occurs by oxygen-dependent post-translational hydroxylation of proline residues (P564 and P402, located in the ODDD) within HIF-1α by HIF-prolyl-hydroxylases (PHD).[26, 68] The PHD's are part of a group of 2-oxoglutarate-dependent dioxygenases and require oxygen and iron to function. The PHD therefore function as the oxygen sensors in the pathway.[26] Hydroxylation of HIF-1α allows binding to the VHL, which leads to ubiquitination and, consequently, proteasomal degradation (Figure 6).[26] Continuous degradation of HIF-1α by this mechanism prevents transcription of HIF-responsive genes under normoxia. In contrast, under hypoxic conditions this oxygen-dependent pathway is inhibited, and hence HIF-1α is stabilised, allowing dimerisation with HIF-1β and the formation of an active HIF-1 complex (Figure 6).

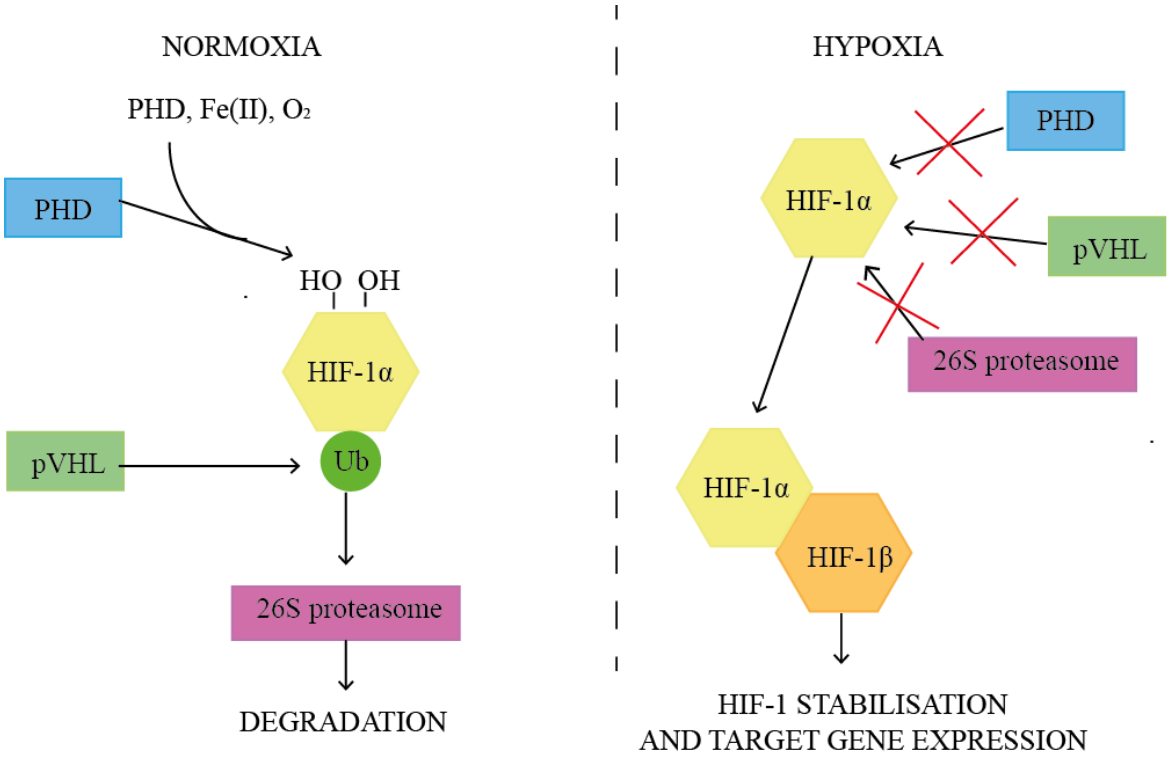


Figure 7 Schematic showing hypoxic regulation control of HIF-1α by VHL E3 ubiquitin ligase.

Under normoxic conditions HIF-1α undergoes hydroxylation by PHD and then binds to VHL proteins, which recruit an E3 ubiquitin ligase complex. This targets HIF-1α for proteasomal degradation. During hypoxia a lack of dioxygen means that hydroxylation does not occur, and therefore HIF-1α is stabilised and can dimerise with HIF-1β.[41]

1.6.2 Factor inhibiting HIF and oxygen-dependent asparagine hydroxylation regulates the hypoxic response

Mahon *et al*. (2001) identified factor inhibiting HIF (FIH), a protein that binds to HIF-1, inhibiting transactivational function and down-regulating target gene expression.[69] They found FIH to bind at the C-terminal of HIF-1α (at residues 757-826) and up-regulation of FIH inhibited reporter gene expression. They also described the interaction of FIH with VHL, demonstrating an underlying mechanism for the alternative action in HIF-1α stabilisation.[69]

Another oxygen-dependent mechanism of HIF-1α regulation is asparagine hydroxylation, identified in 2002.[70] Under normoxia, an asparaginyl hydroxylase (member of 2-oxoglutarate-dependent family) hydroxylates Asn803 in the C-TAD, which prevents its interaction with p300(CH1 domain)/CBP, in turn, preventing transcription.[70] Under hypoxia, the lack of oxygen substrate hinders the asparaginyl hydroxylase's function. Later, Lando *et al.* (2002) reported that FIH hydroxylated Asn803, identifying it to be a Fe(II)-dependent hydroxylase that utilises molecular O_2 to modify the asparagine residue.[71] Furthermore, they demonstrated that FIH also inhibits the interaction between HIF-2α and p300. Mutation of HIF-2α Asn851 to alanine diminishes the effects of FIH-1 and does not block the protein-protein interaction with p300.[71] Hydroxylation of Asn851 residue was confirmed by MS.[71]

1.6.3 Oxygen-independent Hsp90 inhibitor induced degradation of HIF-1α

Liu *et al.* (2007) identified an oxygen-independent HIF-1α regulation pathway.[72] They reported that the RACK1 bound to HIF-1α resulting in reduced HIF-1α protein levels and transcriptional activity. The mechanism by which this works is that RACK1 competes for HIF-1α binding with Hsp90. RACK1 binding destabilises HIF-1α, leading to oxygen-independent protosomal degradation. RACK1 dimerises and recruits elongin C, constructing an E3 ligase complex, which leads to ubiquitination and protosomal degradation.[72] Later, Baek *et al.* (2007), reported that spermidine/spermine N^1-acetyltransferase-1 (SSAT1) plays a role in this degradation pathway.[73] SSAT1 and SSAT2 have contrasting roles within the stabilisation and degradation pathways. SSAT1 interacts with both HIF-1α and RACK1 and is required for this pathway.[73]

1.6.4 HIF regulation by small ubiquitin-like modifier-1 (SUMO)

Initially, SUMOylation of HIF-1α affects the levels of protein; specifically increased expression of SUMO-1 has been found to increase HIF-1α stability. HIF-1α is covalently modified at lysine residues, putatively resulting in its stabilisation, in turn, leading to increased HIF-1 transcriptional activity.[74, 75] In contrast, later evidence suggested that SUMOylation of HIF-1α may in fact lead to increased degradation of HIF-1α, via a hydroxylation-independent VHL pathway.[76, 77]

There are multiple regulation pathways of HIF, and the interaction and competition between pathways is complex and sensitive (Figure 8). The pathways detailed above include VHL HIF-1α degradation by prolyl hydroxylation,[26, 78], VHL HIF-1α degradation by asparaginyl hydroxylation,[69, 71], VHL HIF-α degradation by ADP-ribosylation factor domain protein 1 (ARD1) acetylation of lysine, [79] p53-dependent ubiquitination,[49] SUMOylation of HIF-1α,[74, 75] and RACK1/SSAT1/SSAT2 mediated pathways,[72, 73] Although this is not an exhaustive list, it highlights the complexity of the network, which is further convoluted by the interaction between these pathways and feedback loops and variation between cell types and hypoxic environment.

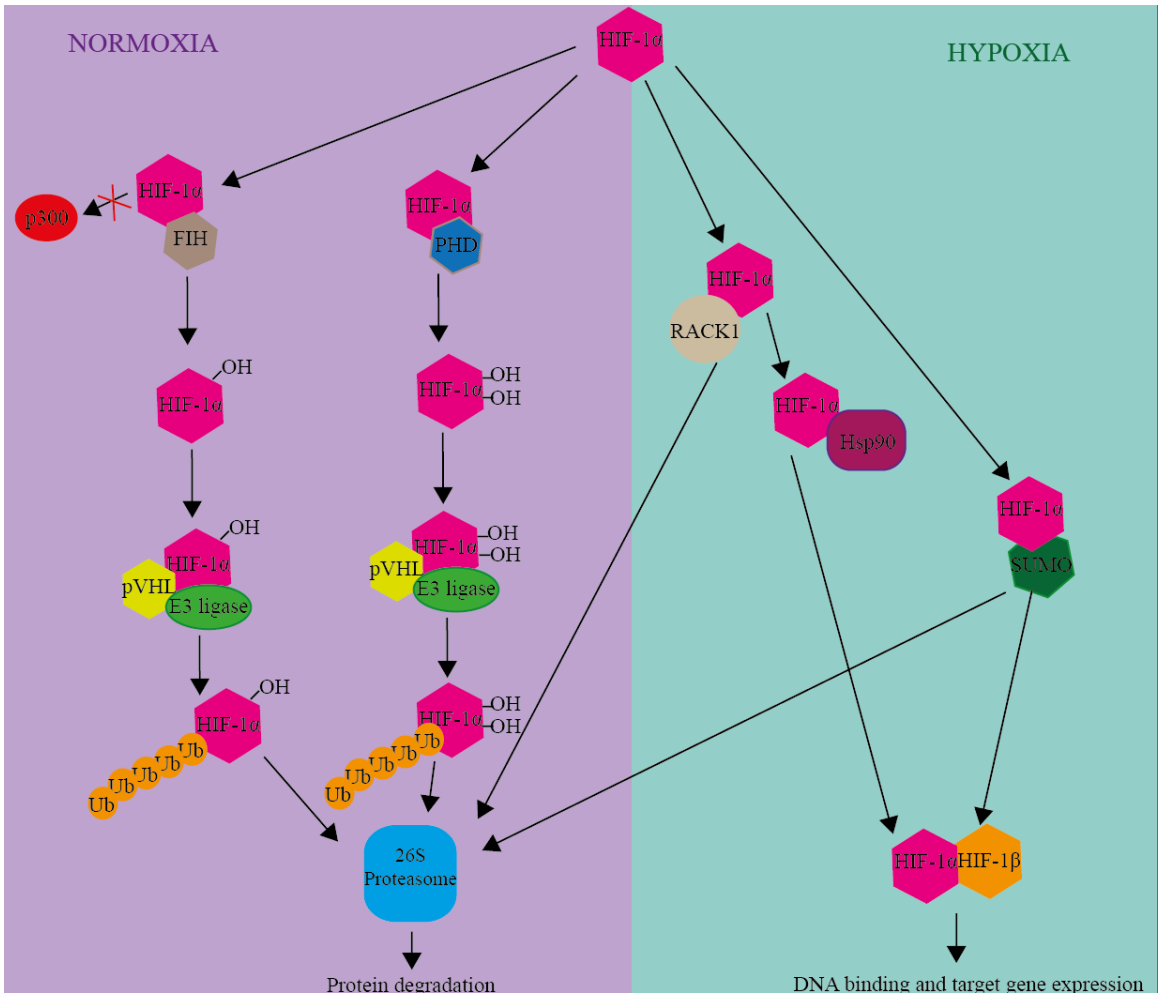


Figure 8 Illustration showing the different pathways of HIF-1α degradation and stabilisation. Under normoxia HIF-1α is degraded by both PHD and FIH mediated pathways. Hydroxyation of HIF-1α leads to ubiquitination and proteasomal degradation. SUMOylation leads to both degradation and stabilisation of HIF-1α. RACK1 and Hsp90 compete for HIF-1α binding leading to degradation and stabilisation respectively. Stabilised HIF-1α dimerises with HIF-1β resulting in DNA binding and target gene expression.

1.7 Targeting HIF-1 in cancer therapy

Having shown that HIF-1 is pivotal to cell adaptation to hypoxia, the transcription factor poses a promising drug target for cancer treatment.[10, 15, 16, 80, 81] Tumour cells divide rapidly and hence quickly outstrip their local oxygen concentration.[18, 82] HIF-1 aids tumour cell survival by increasing oxygen transport and allowing hypoxic survival.[63, 82, 83] There are many reports of HIF-1 overexpression in primary cancer cells, this overexpression has a link to poor patient outcome and increased mortality; this has been extensively reviewed by Semenza (2003 and 2010).[80, 84] There is an increasing interest in targeting HIF-1 as a cancer intervention and inhibitor development is on-going.[85]

Aside from intratumour hypoxia, genetic alterations in cancer cells can result in HIF-1 up-regulation. Loss of function of several tumour suppressor genes (for example: VHL and phosphatase and tensin homolog (PTEN)) leads to elevated HIF-1α protein. Mutation and loss of function of tumour suppressor protein VHL leads to decreased HIF-1α degradation. Loss of function of PTEN, leads to increased synthesis and stabilisation of HIF-1α as a result of regulating Akt activation of HIF-1.[86] Again inhibition of HIF-1 is sought to combat the effect of these genetic alternations within a cancer environment.

Directing therapeutic efforts towards HIF-1 within a tumour environment is based on the principle that normal tissues do not experience the same decrease in oxygen levels as a tumour environment.[78] Hence the hypoxic pathway would not be interrupted in normal tissues, and limited effect on normal tissues should be observed by HIF-1 inhibition.[81] There are many examples of HIF-1 up-regulation in tumour environments; a few of these are described below.[87, 88]

1.7.1 Evidence of HIF-1 overexpression in cancer cells

HIF-1 overexpression is reported in many cancers.[87-89] Birner *et al*. (2000) investigated HIF-1 as a marker in invasive cervical cancer.[87] They show by immunohistochemistry that expression of HIF-1α was observed in 81.3% of early-stage invasive cervical cancers at various levels but absent in normal cervical epithelia, demonstrating its potential use as a marker for tumour growth. In addition, Zhong *et al.* (1999) utilised immunohistochemistry

to identify HIF-1α overexpression in renal, breast, skin, ovarian, prostate, gastric and colon carcinomas. This correlated with cell proliferation.[88] The effect that this up-regulation has on cancer environments may differ to that of HIF-2. Hence, investigation into isoform-specific overexpression could lead to a greater understanding of cancer progression.[90]

1.7.2 Evidence of a link between poor patient prognosis and elevated HIF-1 levels

The reported overexpression of HIF-1 in tumours has been linked to tumour progression, malignancy and poor patient prognosis.[80] A preliminary study, of 953 patients, highlighted a potential relationship between HIF-1α overexpression and prognosis in patients with hepatocellular carcinomas.[91] However, further studies are required to validate these findings, due to the complex nature of cancer progression pathways. Furthermore, Rasheed *et al.* (2009) studied 90 rectal cancer patients and again established a correlation between HIF-1α expression and cancer specific mortality, as well as cancer reoccurrence. Notably, this study identified that HIF-2α overexpression had no specific correlation to cancer reoccurrence and patient prognosis.[92] Similarly, a correlation between breast cancer aggressiveness and invasiveness against HIF-1α overexpression has also been reported.[93]

These examples represent many studies in the literature demonstrating the adverse effect of HIF-1 overexpression on prognosis of a variety of cancers. This compounds the notion that HIF-1 is an important target for therapeutic agents and research into cancer pathways.

1.8 Other roles of HIF-1

HIF-1 is clearly an important cancer target with much potential and investigation into its inhibition is vast.[94-96] However, it should also be noted that in contrast to the inhibition of HIF-1 in this project, activation of HIF-1 could also demonstrate a therapeutic advantage. Heart attacks are ischemic disease caused by areas of hypoxia, ischemic preconditioning protects the heart against injury caused by prolonged ischemia.[95] HIF-1 is thought to play a role here, and there is evidence that HIF-1 activity is required for cardioprotection. Strokes are also an example of ischemic disease where HIF-1 activation may play an additional therapeutic role.[96]

1.9 HIF-1 Inhibitors

The hypoxia response network is vast, HIF-1 has complex regulation pathways which leads to many potential targets for cancer treatment.[85, 97] The targets for inhibitor development are extensive (reviewed by Onnis *et al.* (2009)), however, the most researched points for inhibition could be sub-divided as follows:[85]

1.) HIF-1α degradation pathway
2.) HIF-1 transcriptional activity
3.) HIF-1α mRNA expression and protein translation
4.) HIF-1 DNA-binding
5.) HIF-1α/HIF-1β dimerisation

An insight into each of these targets will be discussed and examples probed in the following sections.

1.9.1 Inhibiting the HIF-1α degradation pathway

As previously explored (section 1.6) the HIF-1 degradation pathway is extensive and therefore there are many opportunities to disrupt the pathway. By promoting the degradation of HIF-1α or conversely, by disrupting the stabilisation of HIF-1α, the action of HIF-1 can be decreased under hypoxia.

Previously described was the Hsp90 stabilisation of HIF-1α (section 1.6.3), Isaac *et al*. (2002) reported that the Hsp90 inhibitor (Figure 8), geldanamycin, increased the degradation of HIF-1α.[98] Hsp90 has many roles including the stabilisation of proteins against heat stress and degradation. It is also essential for tumour growth and is often a therapeutic target for cancer drugs.[99] The natural anasmycin has been reported to demonstrate anti-tumour activity as a result of binding to Hsp90 and disruption of its activity.[100, 101] The increased degradation of HIF-1α was independent of the VHL protein, as degradation was increased in the absence of active VHL. This suggested the possibility of an alternative degradation pathway. By increasing the degradation of HIF-1α, geldanamycin, in turn, affects HIF-1 transcriptional activity. Treatment decreased VEGF expression under hypoxia, indicating the effect on HIF-1 regulated transcription.[98, 100] A geldanamycin analogue, 17-allylamino-17-demethoxygeldanamycin is in clinical trials against several cancer and tumour types.[102, 103] Another Hsp90 inhibitor, Ganetespib has been found to also block HIF-1 activity *in vivo.*[104]

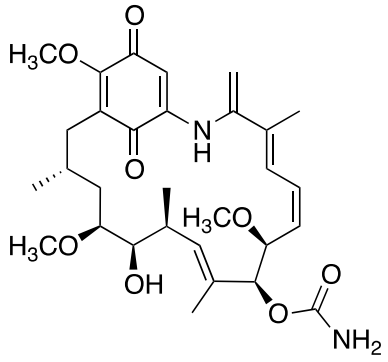


Figure 9 Structure of Hsp90 inhibitor, geldanamycin.

One potential limitation with specifically targeting a point in the degradation pathway is that there are many feedback loops within the network; thus mis-regulation caused by inhibition may be accounted for by an alternative mechanism.

1.9.2 Inhibiting HIF-1 transcriptional activity

The action of HIF-1 could be disrupted by modulating its transcriptional activity.[105-108] Specifically, inhibiting the interaction with co-activators leads to decreased recruitment of transcription machinery. Henchey *et al.* (2010) adopt this strategy when investigating the interaction between p300 and HIF-1α[63] By mimicking the alpha-helical conformation of the critical residues $_{799}$DCEVNA$_{804}$, they rationally designed a hydrogen bond surrogate alpha-helix inhibitor that bound specifically to p300 at sub-nanomolar concentrations. In cell-based assays the compound downregulated VEGF and GLUT1 transcription. This demonstrates that inhibiton of the interactions between HIF-1α and its co-activiators provides a method of modulating HIF-1 activity. Similarly, Burslem *et al.* (2014) also target interaction between HIF-1α and p300.[106] With the knowledge that the HIF-1α CTAD takes an α-helical form, wrapping itself around the CH1 domain of p300, they systematically identified residues that would be good targets for inhibition with rationally designed inhibitors to mimic the α-helix.[106, 109] Initially, two oligoamides were synthesised that mimicked the sequence of the α-helices. Further SAR, using fluorescence anisotropy, clarified the structural importance of the isobutyl group in the primary and secondary oligoamide side chain positions, determining the most potent compound to be Burslem 3 (Figure 10).[106]

Kung *et al.* (2004) identified Chemotin inhibits the same interaction, and disrupted the interaction in cell as well as *in vitro* (Figure 10). Critically, toxicity has been observed.[105]

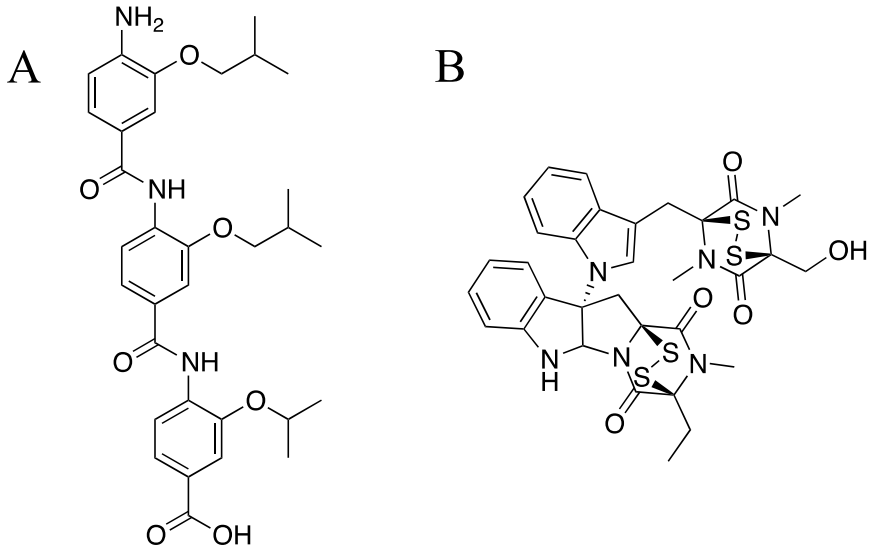


Figure 10 (*A)* Structure of Burslem 3,[106] (B) Structure of Chemotin.[105]

1.9.3 Inhibiting HIF-1α mRNA expression and protein translation

Although levels of HIF-1α are regulated through post-translational mechanisms, modulating the transcription or translation of HIF-1α would also serve to inhibit the action of the transcription factor by lowering cellular levels. The HIF-1α promoter contains a HRE and has a self-regulatory mechanism; under hypoxia HIF-1α mRNA levels are found to be elevated.[110] HIF-1α transcription leads to HIF-1α protein accumulation and increased activation of HIF-1 target genes. EZN-2968 is a HIF-1α mRNA antagonist, it is a locked oligonucleotide (5'-TGGcaagcatccTGT-3', upper case and lower case represent DNA and locked nucleic acid residues, respectively). The EZN-2968 sequence is complementary to HIF-1α mRNA coding sequence; it binds and degrades HIF-1α mRNA. The inhibitor led to reduced levels of HIF-1α mRNA in 15PC3 (prostate carcinoma), the effect was dose-dependent and almost complete inhibition was evident at 5 nM. There was a degree of selectivity towards the HIF-2 isoform: 25 nM only led to a 20% decrease in HIF-2α mRNA expression. The compound also led to a decrease in VEGF and matrix metalloproteinase-2 (both under HIF-1 regulation) secretion in U373 (glioblastoma) cells, this effect was more pronounced under normoxia than hypoxia. After successful *in vivo* testing in mouse models, phase I clinical trials were started, these are on-going. A pilot

trial showed the reduction of HIF-1α mRNA levels in several tumour biopsies, tentatively proving the concept for the use of this agent.[111]

Alternatively, levels of HIF-1α could be controlled by modulating the protein translation. There are many reports of small molecules, which inhibit the rate of HIF-1α translation, although many may not by specific.[112-115] Zhang *et al.* (2008) described that digoxin and some cardiac glycosides inhibited the translation of HIF-1 (Figure 11).[112] The exact mechanism of action here still remains to be elucidated and further clinical trials into the effect digoxin has on intratumour HIF-1α expression is on-going.

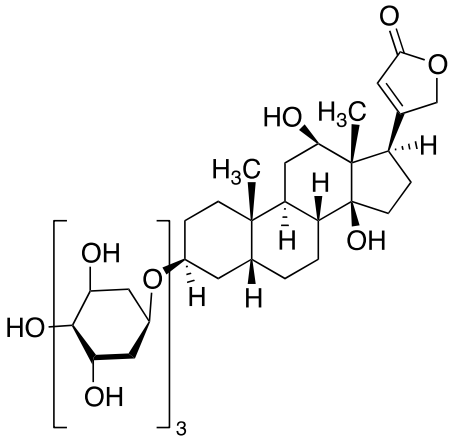


Figure 11 Structure of digoxin.[112]

1.9.4 Inhibiting HIF-1 DNA-binding

Inhibiting HIF-1 binding to the HRE prevents the transcription of genes under HIF-1 control.[116] Inhibition at this point completely bisects the hypoxia response network as DNA-binding is critical to transcriptional activation. Many of the above targets also play a part in other biological networks, therefore, inhibition of these could be categorised as non-specific. Whereas inhibition of this last step of the hypoxia response network specifically prevents HIF-1 dependent gene expression.[117] One potential drawback to the inhibitor is specificity; if the inhibitor binds to DNA, there may be off target effects by non-specific DNA-binding.

Kong *et al.* (2005) discovered that echinomycin prevented HIF-1/HRE binding through an HIF-1/DNA-binding enzyme-linked immunosorbent assay (ELISA) screen (Figure 12).[117]

The compound specifically prevented HIF-1/HRE binding relative to binding of two control proteins to their partner DNA strand. Binding was inhibited in a dose-dependent manner with near complete inhibition evident at 320 nM. When tested in a HRE-luciferase reporter assay, the EC_{50} was approximately 1.2 nM. Treatment of hypoxic U251 (human glioblastoma astrocytoma) cells with echinomycin caused a dose-dependent decrease in VEGF mRNA expression.[117]

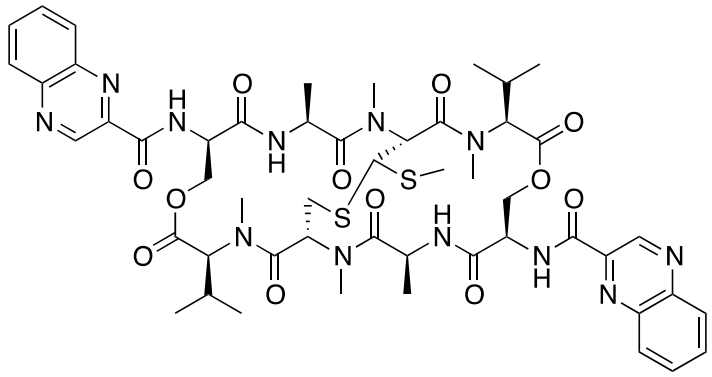


Figure 12 Structure of HIF-1 DNA-binding inhibitor, echinomycin.[117]

1.9.5 Inhibiting HIF-1α/HIF-1β dimerisation

Looking at the potential targets along the hypoxia response network, the disruption of HIF-1α/HIF-1β dimerisation is a key step. Inhibition here completely bisects the network. Without dimerisation of the two subunits, DNA-binding and transcription of target genes cannot take place, regardless of any changes within the regulatory system. Inhibition here also leads to the possibility of specifically inhibiting one isoform over the other. This protein-protein interaction could be the most efficient and specific place to inhibit and is the primary focus of this project. Critically, bisecting the hypoxia response network at this crucial protein-protein interaction disrupts the action of HIF-1. This inhibition would function independently to the reason for HIF-1α up-regulation, whether it be due to hypoxia or tumour suppressor genetic alterations (e.g. VHL mutation). As previously described, there are distinct differences between the interaction between HIF-1α-PAS-B/HIF-1β-PAS-B and HIF-2α-PAS-B/HIF-1β-PAS-B (section 1.4). These differences allow for the development of a specific inhibitor. To date there are four reported inhibitors

of HIF-1 dimerisation, three of which will be described below and the final inhibitor, P1, will be discussed later (section 1.10).

Semenza *et al*. (2003) identified Acriflavine (ACF) to be a HIF-1 inhibitor; the antibacterial drug was identified through the screening of 200 compounds using a split synthetic renilla protein fragment complementation assay.[118] The assay identified ACF, a mixture of proflavine and trypaflavine, to be the lead compound, which was then investigated further using several assays to establish more about its activity. For example, a co-immunoprecipitation (co-IP) assay in HEK293 showed a decrease in the interaction between endogenous HIF-1α and HIF-1β when cells were treated with ACF and exposed to 1% oxygen for 24 hours. Immunoprecipitation (IP) with anti-HIF-1α antibodies showed HIF-1β to be absent in the presence of ACF under hypoxia. *In vitro* GST pull-down assays also confirmed this effect; His-HIF-1α (12-396) was 'pulled-down' by GST-HIF-1β (11-510), and this interaction was inhibited by ACF in a dose-dependent manner, with the IC_{50} estimated to be 1 μM.[118]

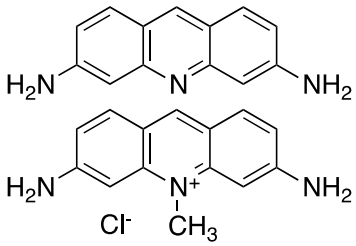


Figure 13 Structure of HIF-1α/HIF-1β dimerisation inhibitor, ACF.

ACF is a mixture of proflavine and trypaflavine.[118]

Furthermore, *in vivo* under hypoxic conditions, ACF inhibited the expression of VEGF in mammalian cells and decreased tumour-induced vascularisation. Lee *et al*. (2009), with ACF, provided a proof of principle that small molecules can inhibit HIF-1 dimerisation.[118]

Park *et al*. (2006) reported the use of an ELISA to screen over 2000 compounds from chemical libraries generated at the National Cancer Institute, to identify an inhibitor of the interaction between HIF-1α-PAS-A and HIF-1β-PAS-A.[119] Firstly, FLAG-HIF-1β-PAS-A was immobilised on a plate, then HIF-1α-PAS-A-His in the presence of test compounds was added and the binding activity measured using an anti-His-HRP antibody and

appropriate substrates (Figure 10). A decrease in signal indicated an inhibition of dimerisation, allowing inhibitors of the interaction to be identified.[119]

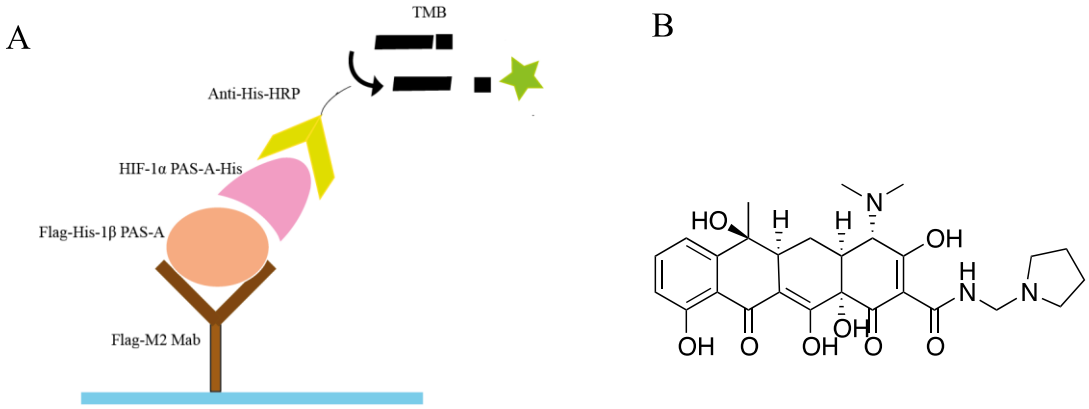


Figure 14 Schematic of the HIF-1α-PAS-A and HIF-1β-PAS-A based ELISA and the structure of rolitetracycline.

(A) ELISA: The FLAG-HIF-1β-PAS-A recombinant protein was immobilised on a plate coated with an anti-FLAG antibody. HIF-1α-PAS-A-His was added in the presence of test compounds. The binding activity was measured using anti-His-HRP and 3,3′,5,5′-tetramethylbenzidine substrates, this enabled visualisation of bound His-HIF-1α. *(B)* The structure of rolitetracycline.[119]

From this screening, rolitetracycline, NSC-50352, was found to inhibit HIF-1 dimerisation. The compound showed an IC_{50} of 1.4 μM from the ELISA, however, failed to inhibit HIF-1 dependent luciferase activity in cell-based assays. This may be explained by poor solubility or membrane permeability.[119] This example shows that HIF-1-PAS-A is a valid target for small molecule inhibitors. However, it demonstrates that turning an inhibitor found via an *in vitro* screen such as this into a viable inhibitor active in cells can be challenging.

As previously described, Cardoso *et al.* (2012) presented a crystal structure of the HIF-1α-PAS-B/HIF-1β-PAS-B interaction.[29] At the protein-protein interface, a cavity that was different to that identified in HIF-2α was identified in the HIF-1α-PAS-B subunit. Cardoso *et al.* (2012) reasoned that this cavity was a suitable target for a small molecule inhibitor that could demonstrate specificity over HIF-2α. Three cysteine residues were identified on

the interface, two of which resided within the cavity. A library of electrophilic compounds was screened to explore the covalent interactions with these cysteine residues. Irreversible binding was identified by mass spectrometry (MS). Compound 5 (as named by Cardoso *et al*. (2012)) was observed to covalently bind to Cys255, NMR analysis of the complex showed local conformational changes upon inhibitor binding (Figure 15).[29] The binding of compound 5 resulted in a ~10-fold decrease in HIF-1α-PAS-B/HIF-1β-PAS-B dimerisation. Further work is needed to establish if compound 5 will have selectivity or modulate HIF-1 driven transcription in cells.

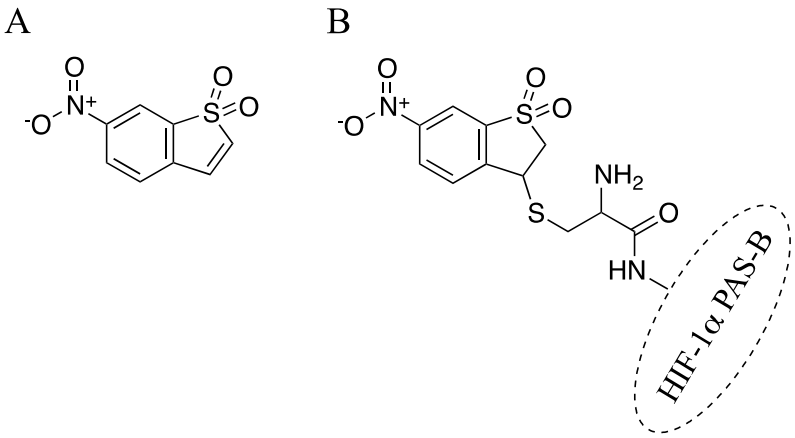


Figure 15 Structure of HIF-1α-PAS-B/HIF-1β-PAS-B dimerisation inhibitor compound 5.
(A) Structure of compound 5 (as named by Cardoso *et al*. (2012)).[29] *(B)* Illustration of compound 5 covalently bound to Cys255 within HIF-1α-PAS-B.

These examples represent an overview of the research efforts to develop a novel HIF-1 inhibitor as cancer therapeutic. This area of research is pursued by many academic groups and research institutions.[29, 118, 120, 121] Most recently Miranda *et al*. (2013) have published a HIF-1 specific inhibitor.[1]

1.10 A Cyclic peptide inhibitor of the hypoxia response network, P1

Miranda *et al.* (2013) developed a cyclic peptide inhibitor of HIF-1α/HIF-β that is specific towards HIF-1 and inhibits hypoxic signalling in cancer cells.[1] The inhibitor, P1, is a cyclic peptide (*cyclo*-CLLFVY) with a Tat-tag (arginine rich tag which aids membrane permeability and solubility, expanded upon in section 2.2) (Figure 16).

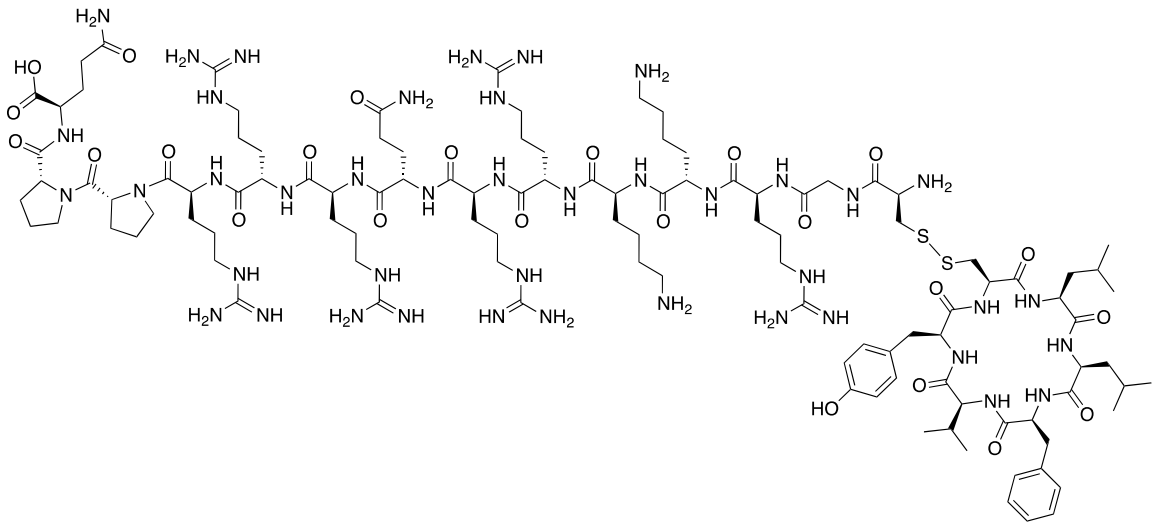


Figure 16 Structure of HIF-1α/HIF-1β inhibitor, P1

P1 was identified using a genetically encoded high-throughput screening system, which combined a bacterial reverse two-hybrid system (RTHS) with split-intein circular ligation of proteins and peptides (SICLOPPS) methodology, described in detail in section 2.12.1. This technique screened a library of 3.2 million hexapeptides as inhibitors of HIF-1α/HIF-1β dimerisation. After identification of the cyclic peptide, verification of the peptides action was investigated. P1 exhibited the ability to reduce the expression of HIF-1 target genes. Treatment of both MCF-7 cells and U2OS (bone osteosarcoma) cells led to a dose-dependent decrease in both VEGF mRNA and protein levels (Figure 17).[1]

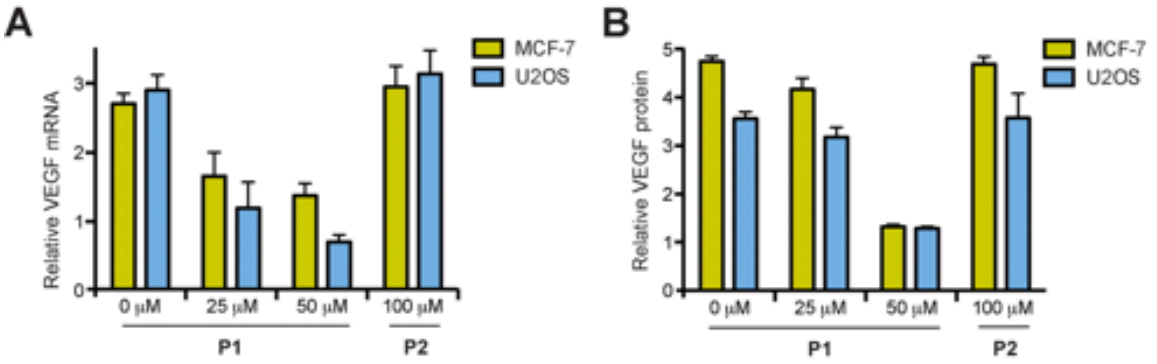


Figure 17 Inhibition of HIF-1 target gene (VEGF) by P1.[1]

Work by Dr E. Miranda. *(A)* qPCR analysis of VEGF mRNA in MCF-7 and U2OS cells after treatment with P1. Normoxia =1. P2 is a control peptide, sequence (*cyclo*-CRLMVL). *(B)* Analysis of VEGF protein levels in MCF-7 and U2OS cells after treatment with P1, normoxia =1.[1]

P1 was shown, using a proximity ligation assay (PLA), to have an effect on HIF-1α/HIF-β dimerisation in MCF-7 cells.[122] A PLA detects protein-protein interactions *in situ*; HIF-1α and HIF-1β antibodies with short DNA strands attached identified and bound to the proteins. When the antibodies were in close proximity they bind, addition of additional oligonucleotides leads to rolling circle polymerisation, amplifying the DNA. A complementary fluorescently labelled probe allows visualisation of the dimer. PLA indicated that HIF-1α/HIF-1β dimerisation was inhibited by P1 (Figure 18).[1]

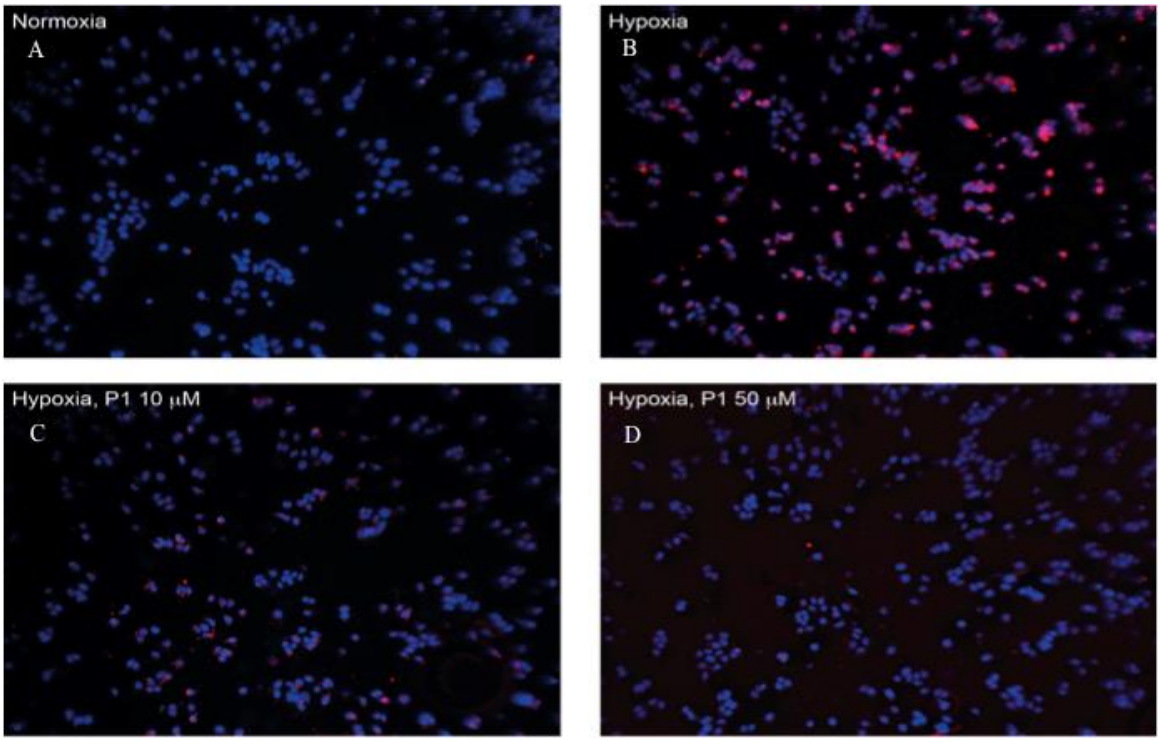


Figure 18 Inhibition of HIF-1 dimerisation by P1 visualised by PLA.

Work by Dr E. Miranda. Blue dots represent the DAPI (4',6-diamidino-2-phenylindol) stained nuclei of the MCF-7 cells. Red dots represent the PLA signal and HIF-1α/HIF-1β dimerisation. *(A)* Normoxia; *(B)* Hypoxia; *(C)* Hypoxia, P1 10 μM; *(D)* Hypoxia, P1 50 μM.[1]

P1 demonstrates selectivity for HIF-1 over HIF-2, P1 had no effect on VEGF mRNA levels in 786-0 (renal carcinoma) cells, which have minimal HIF-1α levels, but significant levels of HIF-2α. This indicated that P1 does not act on HIF-2. A HIF-2 dependent luciferase assay also confirmed this specificity.

This project aimed to develop the inhibitor and to investigate the binding location of the P1 (section 1.11).

1.11 Project aims

Previously, in the Tavassoli research group, a promising selective cyclic peptide inhibitor of HIF-1 dimerisation was identified, Tat-*cyclo*-CLLFVY (P1).[1] The main objective of this project was to identify the active motif of the parent peptide *cyclo*-CLLFVY. Aims of the project included:

- Identification of which amino acid residues of P1 (Tat-*cyclo*-CLLFVY) are critical to the activity of the compound.
- To utilise alanine scanning, where each amino acid was substituted for alanine, to ascertain if the substitution has an effect on inhibitor activity.
- To establish methods for the recombinant production of functional HIF-1α and HIF-1β.
- To set up an *in vitro* assay to test the potency of HIF inhibitors but to also confirm that recombinant proteins generated were functional and could participate in DNA-binding and dimerisation.

In order to achieve these aims a small library of alanine analogues were synthesised using Fmoc solid-phase peptide synthesis and tested in a selection of *in vitro* and cell-based assays. A secondary aim of the project was to produce functional recombinant HIF-1α and HIF-1β that could be used in *in vitro* assays and to determine the binding location of the inhibitor. The functional protein needed to dimerise and bind to DNA.

2 Results

As described previously (section 1.10), P1 is a novel selective inhibitor of HIF-1 dimerisation. However, due to the peptide's hydrophobic nature, its use as a therapeutic agent is hindered. The hydrophobicity of *cyclo*-CLLFVY leads to poor yields during synthesis, due to water insolubility and poor translocation across cell membranes. Consequently, Miranda *et al.* (2013) tagged the cyclic peptide with an arginine rich 14 amino acids tag (Tat) to improve solubility and membrane permeability. The tagged peptide (P1) disobeys Lipinski's Rule of Five, due to its high molecular weight (greater than 500 Da) and more than five hydrogen bond donors, suggesting it may have poor pharmacokinetic properties.[123] This large compound is also complicated and expensive to synthesise.

As a result, a second-generation inhibitor with improved potency, pharmacological properties and a simpler synthetic route is required. Firstly, in order to generate a library of such second-generation compounds the active motif of P1 must be identified, allowing the critical amino acids to be used as a scaffold. The initial library screen used the 20 canonical amino acids. The second-generation library could subsequently probe chemical spaces unavailable in the initial SICLOPPS screen (section 2.1). This library could then be screened in *in vitro* and *in vivo* assays to identify novel small molecule inhibitors with increased potency.

The aim of this research was to identify the active motif of P1 using alanine scanning. Each amino acid was sequentially substituted for alanine and the activity of the peptide compared to the parent compound. As cysteine was present in all peptides initially screened and was utilised for solubility tagging, the cysteine residue remained unchanged in the analogues. If substitution of an amino acid with alanine caused a loss in activity, then the residue may be deemed critical to activity. This technique was used by Spurr *et al.* (2012) to successfully develop a cyclic peptide inhibitor of aminoimidazole carboxamide ribonucleotide transformylase/inosine monophosphate cyclohydralase (ATIC) heterodimerisation into a second-generation potent dipeptide.[124] ATIC is a heterodimeric protein that catalyses the last two steps of the *de novo* purine biosynthesis, whereby dimerisation of the subunits is essential for activity. Initially, Tavassoli *et al.* (2005)

identified *cyclo*-CRYFNV (k_i = 17±3 μM) that inhibited the dimerisation.[125] Subsequently, Spurr *et al.* (2012) identified that the active motif was RY using alanine scanning. Development of this motif led to the identification of compound 14 (RY(NO_2)), which exhibited 25-fold improved inhibitory activity compared to the parent cyclic peptide.[124] Based on the successful use of alanine scanning to identify the active motif of a cyclic peptide identified by SICLOPPS screening, a series of alanine analogues of *cyclo*-CLLFVY were subsequently screened in a selection of assays.

2.1 Utilising the HIF-1 RTHS and SICLOPPS

Cyclo-CLLFVY was identified using a combination of a bacterial RTHS and the SICLOPPS screening platform. Consequently, the HIF-1 RTHS and SICLOPPS technology were used to determine the amino acid residues required for activity.[1]

The RTHS links protein-protein interactions to the expression of reporter genes critical to cell survival; this system is based on a bacteriophage regulatory circuit.[126, 127] The reporter genes include Kan^R (aminoglycoside 3'-phosphotransferase conferring kanamycin resistance) and HIS3 (imidazoleglycerol-phosphate dehydratase; this enzyme catalyses the sixth step of histidine biosynthesis).[128] These genes lead to the conditional selectivity of the strains. The heterodimeric system relies on two recombinant proteins both incorporating a dimerisation and DNA-binding domain; one protein is composed of a N-terminal repressor 434 (DNA-binding domain) attached to dimerising protein X (dimerisation domain) and a partner protein composed of an N-terminal P22 repressor (DNA-binding domain) attached to interacting protein Y (dimerisation domain).[127, 129] In the HIF-1 RTHS, proteins X and Y are HIF-1$\alpha_{1\text{-}350}$ and HIF-1$\beta_{1\text{-}459}$. HIF-1 dimerisation leads to reconstitution of the active repressor (434-P22) protein, which in turn blocks reporter gene expression, leading to cell death (a growth defect) on selective media. If HIF-1 dimerisation is inhibited then the chimeric repressor proteins (434 and P22) are unable to bind to their corresponding operators. As a result, the transcription of the reporter genes is allowed, leading to host bacteria survival (a growth advantage) on selective media (Figure 19).

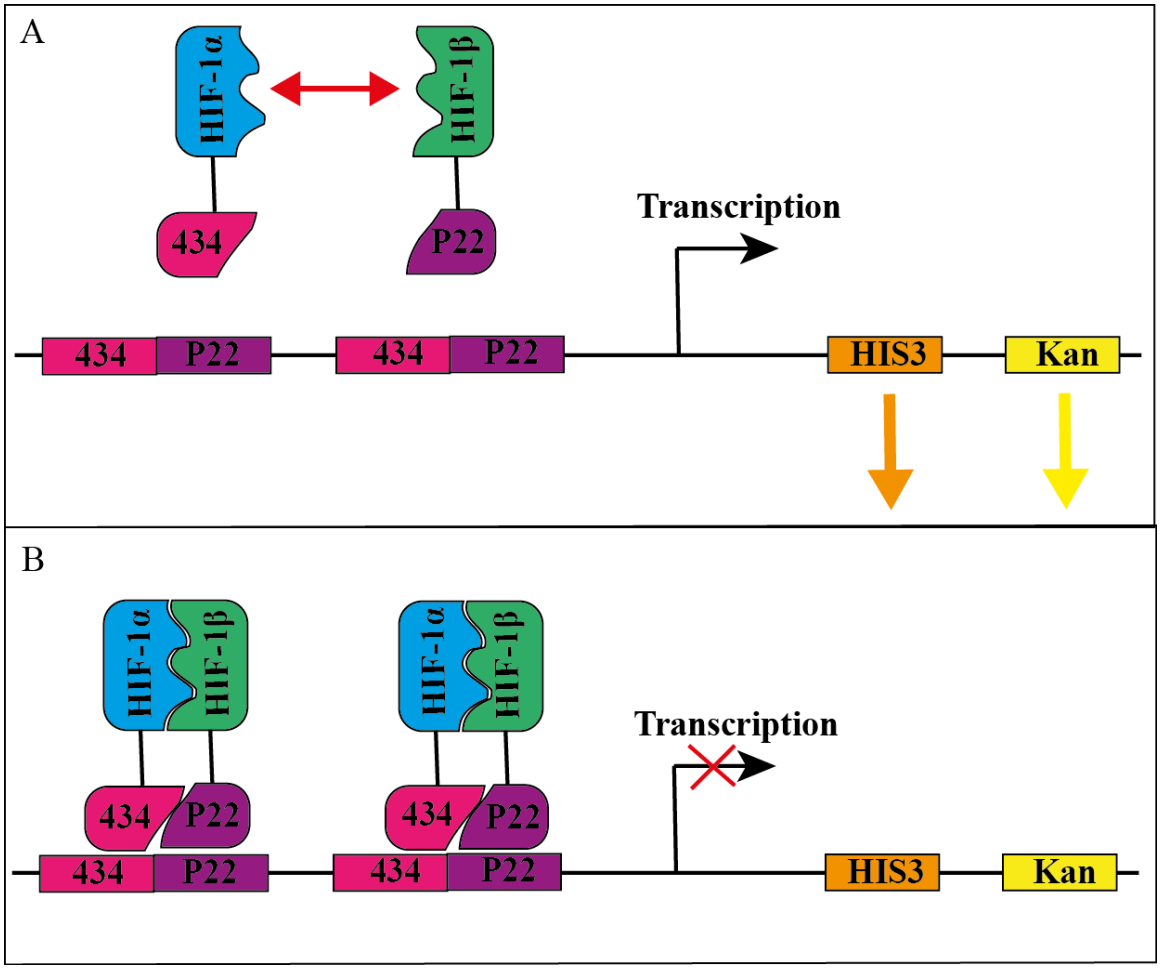


Figure 19 Illustration of the HIF-1 RTHS.

(A) Monomeric recombinant proteins cannot bind to the repressor region, therefore, downstream genes (HIS3 and Kan^R) are expressed, allowing cells to survive on selective media. *(B)* Dimerisation of HIF-1α and HIF-1β allows 434/P22 dimer to bind to the repressor region, leading to inhibition of downstream gene expression. The *LacZ* gene is also present downstream of the repressor region and can be used to quantify the levels of protein-protein interaction. This, however, does not contribute to cell survival on selective media.

An inhibitor of the protein-protein interaction can be identified using the SICLOPPS library platform.[126, 130, 131] The SICLOPPS technology uses intein (*int*ernal prot*eins*) processing to produce cyclic peptides.[130] Intein-mediated protein splicing is a self-catalysed process whereby a protein sequence, known as an intein, excises itself, leaving portions of the remaining protein to ligate. This forms a new peptide bond, which generates a linear peptide (Figure 20A). Unlike typical inteins, the SICLOPPS methodology utilises split–inteins whereby the intein elements are rearranged to produce

an active *cis*-intein (C-terminal intein: Target peptide: N-terminal intein). Intein splicing here results in cyclisation of the target peptide *in vivo* (Figure 19B and C).

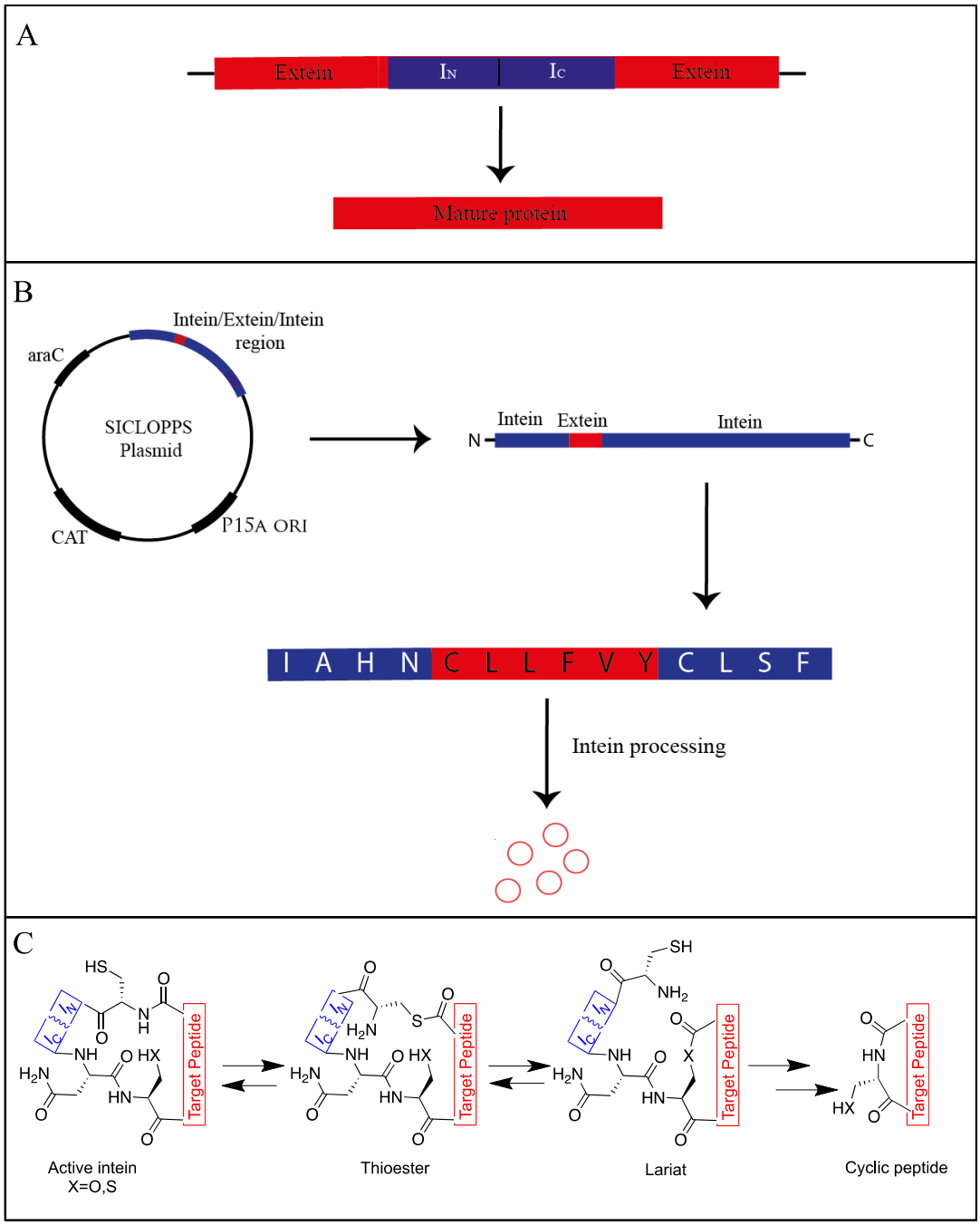


Figure 20 SICLOPPS mechanism.

(A) Illustration of standard intein splicing. *(B)* Illustration of SICLOPPS plasmid. The extein region contains the degenerate library sequence encoding CXXXXX peptides; each plasmid encodes a unique cyclic peptide that makes up a library of ~10^8 members. *(C)* Mechanism of the SICLOPPS processing: an expressed fusion protein folds to form an active intein; an N-to-S acyl shift at the intein junction produces a thioester. The thioester then undergoes transesterification, with a side chain nucleophile producing a lariat intermediate. An asparagine side chain liberates the cyclic product as a lactone and an acyl shift generates the thermodynamically favourable lactam cyclic peptide.

The combination of the SICLOPPS platform and a bacterial RTHS means that when the cyclic peptide disrupts the protein-protein interaction, the P22 and 434 proteins are unable to act as a functional repressor. Consequently, the cells survive. By constructing SICLOPPS plasmids encoding specific cyclic peptides and transforming them into the HIF-1 RTHS, it was possible to assess inhibition of HIF-1 dimerisation by visualising the effect of host cell survival on selective media. Colonies could be cultured and the plasmid extracted and sequenced, yielding the active motif.

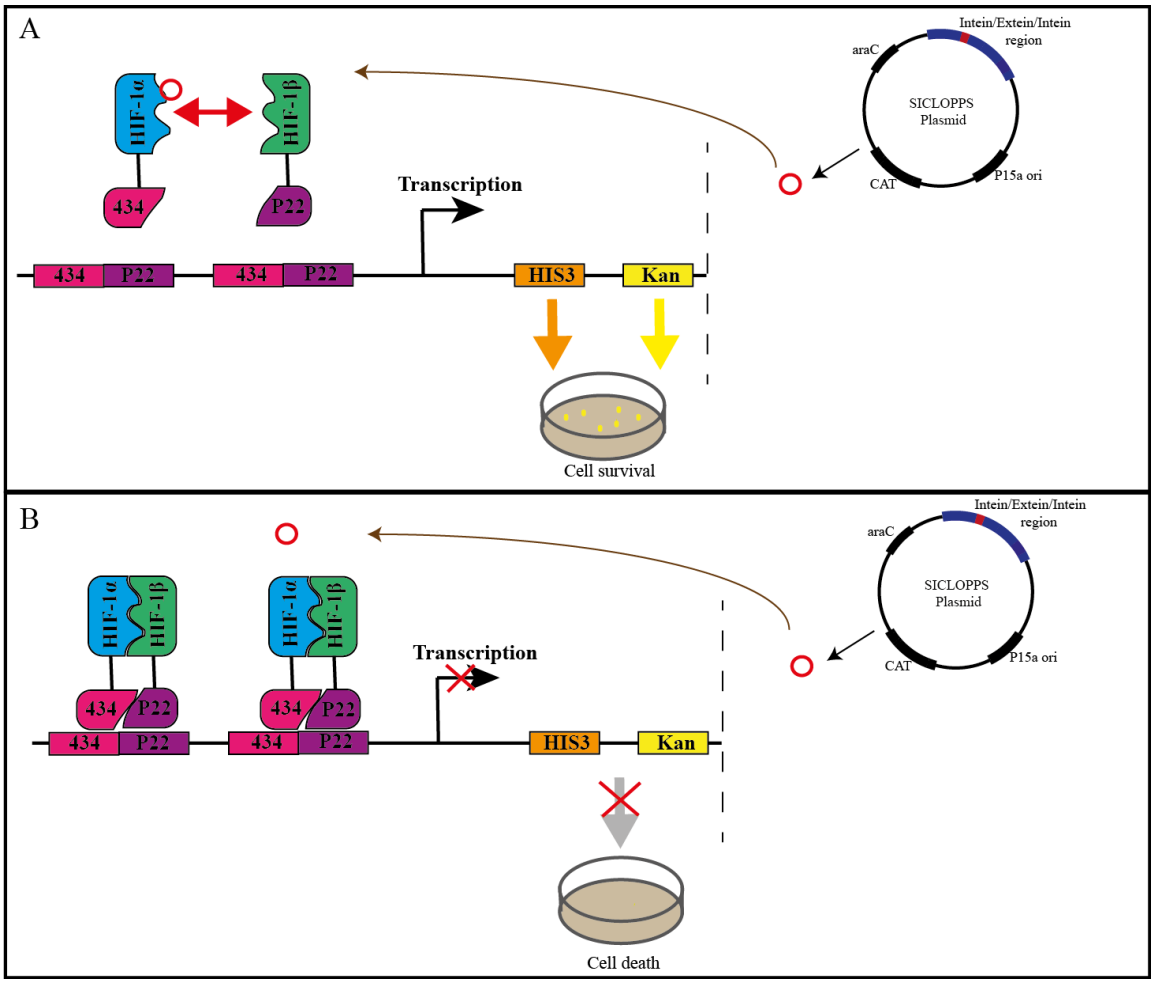


Figure 21 Genetic selection of cyclic peptide inhibitors of HIF-1α/HIF-1β dimerisation using the HIF-1 RTHS and SICLOPPS methodology.

(A) If the SICLOPPS cyclic peptide inhibits HIF-1α and HIF-1β dimerisation, monomeric recombinant proteins cannot bind to the repressor region; as a result, essential survival genes are expressed allowing cells to survive on selective media. *(B)* If the SICLOPPS cyclic peptide does not inhibit the interaction between HIF-1α and HIF-1β, HIF-1 recombinant proteins bind to the repressor region leading to inhibition of essential reporter genes expression and cell death on selective media.

In order to assess the activity of each alanine analogue in the RTHS, SICLOPPS plasmids were constructed encoding cyclic peptide alanine analogues of *cyclo*-CLLFVY, where individual residues were substituted for alanine. All residues were sequentially substituted except cysteine, as the nucleophilic residue is necessary for intein processing, which leads to the formation of the cyclic peptides *in vivo*. The alanine analogues made were as follows:

***Cyclo*-CALFVY**

***Cyclo*-CLAFVY**

***Cyclo*-CLLAVY**

***Cyclo*-CLLFAY**

***Cyclo*-CLLFVA**

Primers were designed to incorporate codons for the desired cyclic peptides (five alanine analogues and parent *cyclo*-CLLFVY). The SICLOPPS plasmid (pARCBD) was used as a template for PCR and inserts were amplified encoding each cyclic peptide intein. The inserts and pARCBD backbone were digested and ligated. Positive colonies were grown and the plasmids purified and verified by sequencing. To confirm that the inteins were capable of splicing, they were purified using the chitin-binding domain affinity tag fused to the N-intein. Firstly, the plasmids were transformed into SNS126 (blank RTHS strain), and intein expression was induced with arabinose. The cells were then lysed and inteins extracted using chitin beads; the bead solution was incubated overnight at room temperature to allow the inteins to process. A control of SNS126 cells alone was performed in parallel. The solutions were analysed by sodium dodecylsulphate-polyacrylamide gel electrophoresis (SDS-PAGE). All alanine samples showed unspliced intein bands (C-terminal intein: Target peptide: N-terminal intein, 21403 Da) as well as N-terminal intein bands (20664 Da) (Figure 22). The N-terminal bands confirmed partial splicing was occurring and the faint C-terminal bands indicated complete splicing. This suggested that all analogues were capable of splicing and could be used in the HIF-1 RTHS.

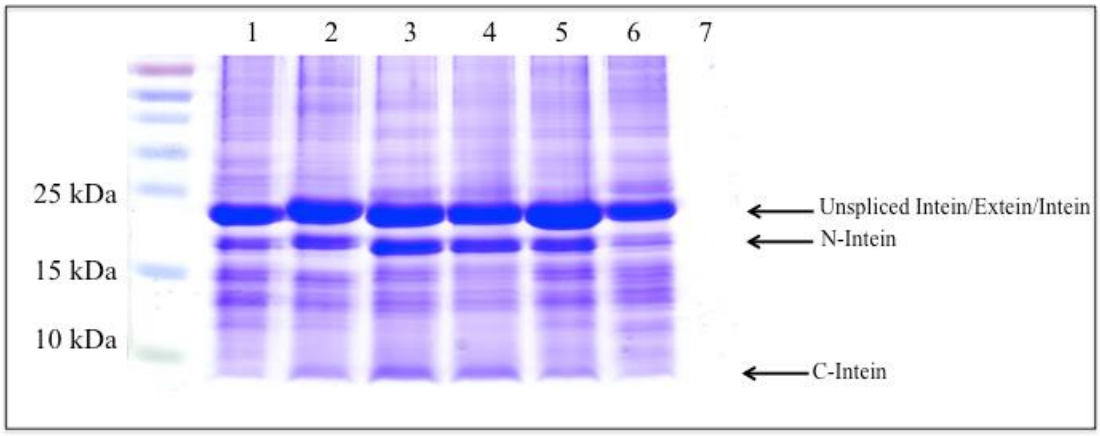


Figure 22 SDS-PAGE (15%) analysis of purified alanine analogues using *Ssp* inteins. *(Lane 1)* CLLFVY; *(Lane 2)* CALFVY; *(Lane 3)* CLAFVY; *(Lane 4)* CLLAVY; *(Lane 5)* CLLFAY; *(Lane 6)* CLLFVA; *(Lane 7)* SNS126 blank control. Bands at 21403 Da represent unspliced intein, bands at 20664 Da represent the N-intein by-product.

The HIF-1 RTHS (constructed by Dr K.Nordgren, a previous PhD student in the Tavassoli group) was grown, and integration of the HIF-1 RTHS construct was verified by chromosomal PCR.[1] In order to visualise the protein-protein interactions, the strains were drop-spotted in six, 10-fold serial dilutions onto selective media plates containing 3-amino-1,2,4-traizole (3-AT, 7.5 mM), kanamycin (25 μg/ml) and increasing concentrations of IPTG (0-50 μM) (Figure 23). 3-AT is a competitive inhibitor of HIS3 and isopropyl-β-D-1-thiogalactopyranoside (IPTG) induces the expression of the recombinant 434-HIF-1α and P22-HIF-1β proteins. As expected, the growth of the HIF-1 RTHS was reduced as IPTG concentration increased. This was because, as IPTG increases, so too does 434-HIF-1α and P22-HIF-1β recombinant protein expression, leading to the repression of reporter genes (HIS3, Kan^R) and, ultimately, cell death.

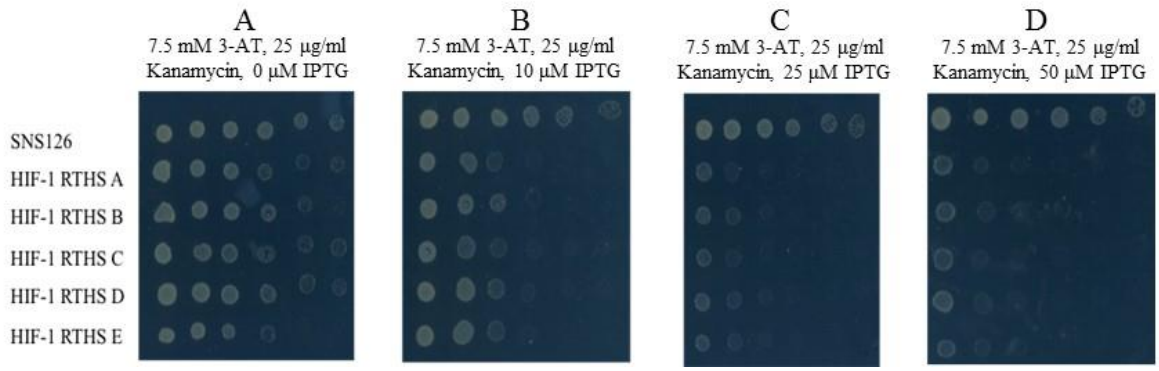


Figure 23 Drop-spotting of the HIF-1 RTHS onto minimal media containing 3-AT (7.5 mM), kanamycin (25 µg/ml) and increasing concentrations of IPTG (0 µM to 50 µM, *A* to *D*). The negative control was the blank strain SNS126. Increased repression of growth was observed for the HIF-1 RTHS strains as IPTG concentrations were increased (A-E represent different strain stocks). Spots from left to right represent 10-fold serial dilutions of the cultured strain. The first spot represents a 10-fold dilution of the culture grown overnight.

The drop-spotting analysis showed that plates supplemented with 7.5 mM 3-AT, 25 µg/ml kanamycin and 25 µM IPTG provided optimal conditions for drop-spotting of the alanine analogue SICLOPPS plasmids, as growth was sufficiently repressed.

2.1.1 Alanine scanning of SICLOPPS generated alanine analogues

The six alanine analogue SICLOPPS plasmids were transformed into chemically-competent HIF-1 RTHS. This experiment was conducted in duplicate. Resulting colonies were incubated overnight in LB media and then drop-spotted in 10-fold serial dilutions onto four minimal media plates. All plates contained 7.5 mM 3-AT and 25 µg/ml kanamycin; however, each plate had different additional supplements, as outlined below (Figure 23):

1.) **Control plate with no IPTG or arabinose**: No expression of the HIF-1 recombinant proteins or cyclic peptides was induced, therefore full growth was expected.

2.) **25 µM IPTG**: Induction of HIF-1 recombinant protein expression caused the repression of reporter gene expression, with growth shut down anticipated.

3.) **6.5 μM L-arabinose**: Cyclic peptide expression induced, therefore if the peptide was non-toxic, full growth was expected.

4.) **25 μM IPTG and 6.5 μM L-arabinose**: Both the recombinant proteins and cyclic peptide are expressed, so full growth was anticipated if the cyclic peptide inhibited HIF-1α/HIF-1β dimerisation. Reduced growth indicated an inactive peptide.

As expected, in the presence of the parent peptide *cyclo*-CLLFVY full growth was observed. In contrast, the alanine scanning drop-spotting showed cell death in the presence of IPTG and arabinose with *cyclo*-CLAFVY, *cyclo*-CLLAVY and *cyclo*-CLLFVA. This implied that amino acid residues three, four and six were critical to the inhibitor's activity. However, the plate with L-arabinose showed significant levels of cell death for all analogues other than the parent *cyclo*-CLLFVY. Toxicity of the inteins or cyclic peptide could be an explanation for the observed phenotype.

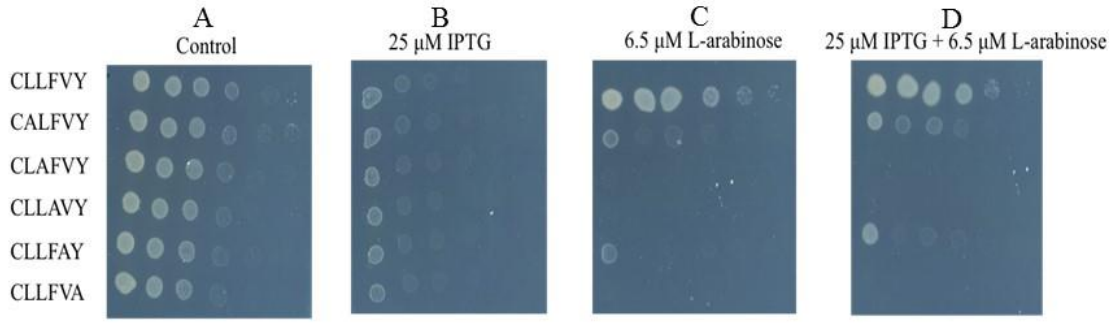


Figure 24 HIF-1 *Ssp* alanine scanning drop-spotting, all plates contained 3-AT and kanamycin.

(A) Minimal media control plate, full growth was expected as there was no IPTG to induce expression of HIF-1; (*B)* Minimal media with 25 μM IPTG, inhibition of growth was anticipated as HIF-1α and HIF-1β dimerise and block transcription of reporter genes; (*C*) Minimal media with L-arabinose, full growth was expected; (*D)* Minimal media with 25 μM IPTG and 6.5 μM L-arabinose, a restoration of growth was expected compared to 25 μM IPTG plate for any active cyclic peptides. Spots from left to right represent 10-fold serial dilutions of the cultured strain. The first spot represents a 10-fold dilution of the culture grown overnight.

Next, the reasons for the observed toxicity were investigated.

2.1.2 Investigating the toxic phenotype observed with *Ssp* SICLOPPS plasmids

To determine if toxicity was responsible for the poor growth of the alanine analogue strains, the HIF-1 RTHS strains containing the alanine analogue plasmids were drop-spotted onto rich media with IPTG and L-arabinose but no kanamycin, hence no selective pressure. All strains showed full growth, indicating no toxicity from the cyclic peptide or intein by-products when there was no selection pressure (Figure 25).

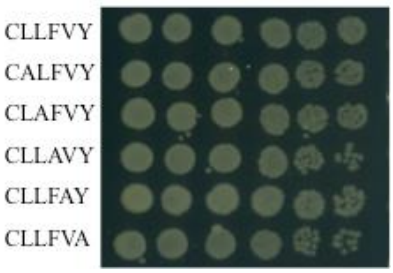


Figure 25 Alanine analogue SICLOPPS plasmids drop-spotted in the HIF-1 RTHS onto rich media with 6.5 μM L-arabinose to induce SICLOPPS induction.

Spots from left to right represent 10-fold serial dilutions of the cultured strain. The first spot represents a 10-fold dilution of the culture grown overnight.

To support the findings on rich media, the analogue strains were drop-spotted onto minimal media with no selective pressure (media lacking 3-AT and kanamycin). Here the plates were supplemented with histidine, as these strains do not have the ability to synthesise histidine without expression of the RTHS construct. These conditions should have eliminated any selective pressure from varied expression of the reporter construct, whilst maintaining the minimal media conditions consistent with the RTHS screen (Figure 26). Growth was inhibited for strains containing *cyclo*-CLAFVY, *cyclo*-CLLAVY and *cyclo*-CLLFVA encoding plasmids. This was more evident on plates containing 13 μM L-arabinose compared to those containing 6.5 μM L-arabinose, where levels of peptide expression were higher (Figure 26). This indicated that toxicity caused by these cyclic peptides or associated inteins was the reason for differential growth on the minimal media. When comparing the results on rich and minimal media, it appears that the pressure of the latter sensitises the strains.

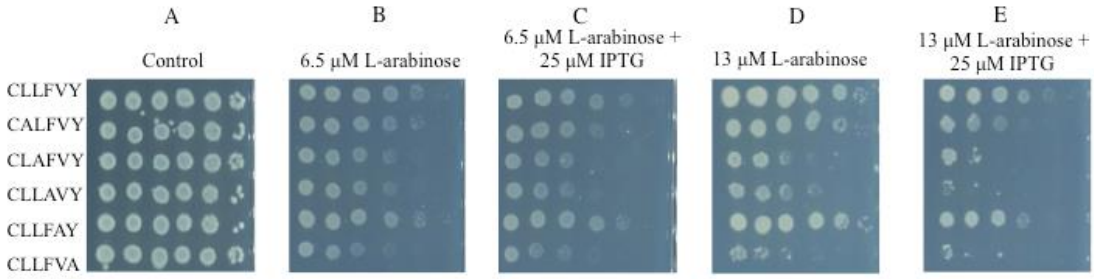


Figure 26 HIF-1 RTHS with alanine analogue SICLOPPS plasmids drop-spotted on minimal media supplemented with histidine.

(A) Minimal media control plate; (*B)* Minimal media with 6.5 μM L-arabinose; (*C)* Minimal media with 6.5 μM L-arabinose and 25 μM IPTG; (*D)* Minimal media with 13 μM L-arabinose; (*E)* Minimal media with 13 μM L-arabinose and 25 μM IPTG. Spots from left to right represent 10-fold serial dilutions of the cultured strain. The first spot represents a 10-fold dilution of the culture grown overnight.

To further verify the toxicity of the peptides, the analogue plasmids were also transformed into SNS126 cells and drop-spotted under the chosen selection conditions (Figure 27). The inherent toxicity was further confirmed by this SNS126 drop-spotting (Figure 27); here all analogues with the exception of *cyclo*-CALFVY and parent *cyclo*-CLLFVY appeared to be toxic.

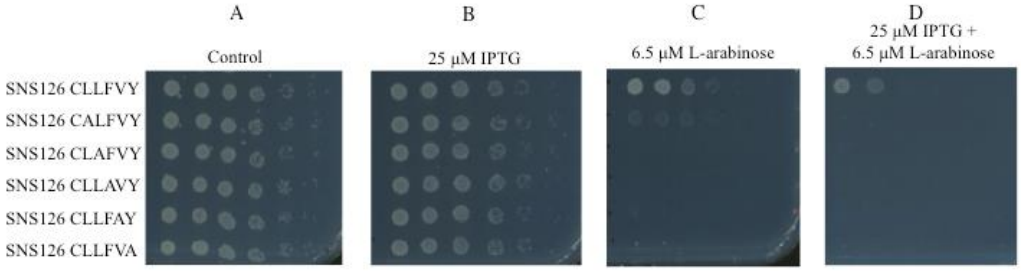


Figure 27 Drop-spotting of *cyclo*-CLLFVY alanine analogue plasmids in SNS126 on minimal media plates all containing 3-AT and kanamycin.

(A) Minimal media control plate; (*B)* Minimal media with 25 μM IPTG; (*C*) Minimal media with 6.5 μM L-arabinose; (*D)* Minimal media with 25 μM IPTG and 6.5 μM L-arabinose. Full growth was expected on all plates. Spots from left to right represent 10-fold serial dilutions of the cultured strain. The first spot represents a 10-fold dilution of the culture grown overnight.

Reasons for this toxicity could be due to different splicing rates of the cyclic peptides leading to different build up of intein by-products or inherent toxicity of the cyclic peptides. It is known that the identity of the second amino acid residue influences intein splicing when using *Synechocystic* sp. PCC6803 inteins (*Ssp,* utilised in the SICLOPPS plasmids).[132] It is, therefore, feasible that the alanine substitution within the extein has an effect on splicing rates, hence influencing the overall toxicity caused by the cyclic peptides and intein by-products. This may explain the phenotype observed. The observed toxicity means that no conclusions on alanine analogue toxicity can be drawn from these experiments.

2.1.3 Drop-spotting the alanine analogues within the HIF-2 RTHS

In-cell testing of P1, within the Tavassoli group, has shown that the inhibitor exhibits selectivity towards HIF-1 over HIF-2.[1] To establish if a similar selectivity or toxicity was observed with the alanine analogues in the HIF-2 RTHS, the alanine analogues were transformed into the HIF-2 RTHS and drop-spotted. Initially, the samples were drop-spotted onto varying concentrations of IPTG to establish the optimal conditions. As with HIF-1, cell growth decreased as IPTG was increased because the expression of 434-HIF-2α and P22-HIF-1β recombinant protein increased. This led to the repression of reporter gene expression. The plates showed 7.5 mM 3-AT, 25 μg/ml kanamycin and 25 μM IPTG were suitable conditions for drop-spotting the alanine analogue SICLOPPS plasmids (Figure 28).

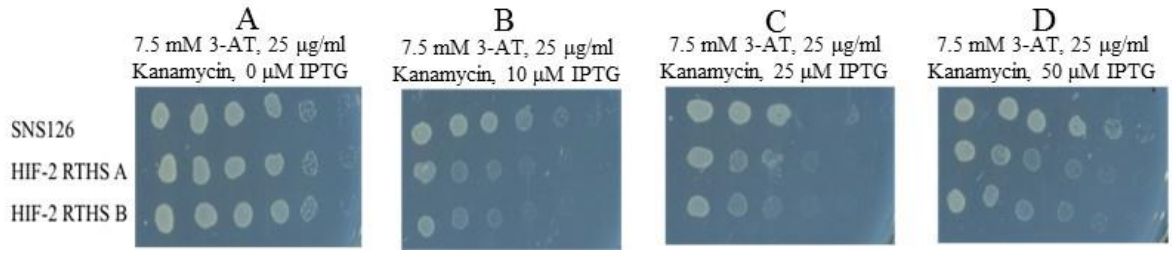


Figure 28 Drop-spotting of HIF-2 RTHS onto minimal media containing 3-AT (7.5 mM), kanamycin (25 μg/ml), IPTG was increased from 0 μM to 100 μM (*A*-*D*).

The negative control was the strain, SNS126. Repression of growth was observed for the HIF-2 RTHS strains in line with increasing IPTG. Spots from left to right represent 10-fold serial dilutions of the cultured strain. The first spot represents a 10-fold dilution of the culture grown overnight.

As with the HIF-1 RTHS assays, the alanine analogues were then drop-spotted onto the selection plates, in the presence or absence of IPTG and arabinose (Figure 29). The HIF-2 drop-spotting on minimal media showed growth comparable to that observed with the HIF-1 RTHS. Again, cell death with L-arabinose was observed for some analogues. *Cyclo*-CLLFVY and to a lesser extent *cyclo*-CALFVY and *cyclo*-CLLFAY are active inhibitors of both HIF-1 and HIF-2 dimerisation within this system. The other three SICLOPPS analogues demonstrated a toxic phenotype impeding the ability to access their relative activity.

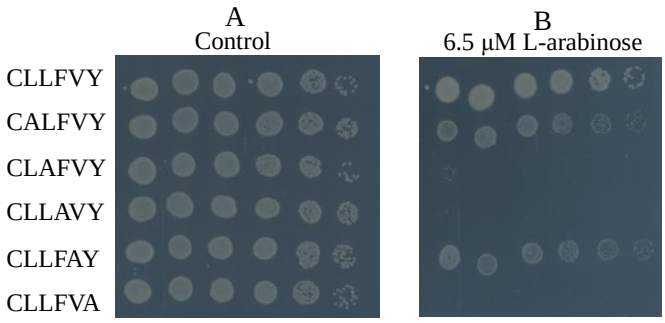


Figure 29 HIF-2 *Ssp* alanine scanning drop-spotting.

All plates contained 3-AT and kanamycin. *(A*) Minimal media control plate; full growth was expected as there was no IPTG to induce expression of 434-HIF-2α and P22-HIF-1β; *(B)* Minimal media with L-arabinose, full growth was expected. Spots from left to right represent 10-fold serial dilutions of the cultured strain. The first spot represents a 10-fold dilution of the culture grown overnight.

The *Ssp* intein toxicity hindered the ability to rank the alanine analogue activity. Next the use of a different intein was investigated.

2.1.4 Investigating the use of *Npu* inteins to mitigate intein-processing bias

The inteins used in the experiments described above originated from *Ssp.* Recently, Iwai *et al.* (2006) demonstrated that the *Nostoc punctiforme (Npu)* inteins have a trans-splicing activity superior to that of the *Ssp.*[132] The study also showed that the protein splicing ability of the *Npu* inteins was less affected by amino acid changes within the extein. With

this in mind, it seemed that using *Npu* inteins for the alanine scanning could eliminate hypothesised differences in splicing rates between the analogues. We reasoned that this change would eliminate the variable toxicity levels and potentially allowing unbiased comparison between the activities of the analogues. An *Npu* derivative of the *Ssp* pARCBD plasmid had been previously designed and constructed by J. Townend, a PhD student in the Tavassoli lab. As with the *Ssp* inteins, six alanine analogues were cloned into this plasmid in a similar manner to the *Ssp* analogues, to generate six *Npu* alanine analogue plasmids, (*cyclo*-CLLFVY and its five analogues). As described in section 2.1.1, the inteins were expressed and purified to confirm their ability to splice. Like the *Ssp* inteins, the N-terminal intein had a C-terminal chitin-binding domain enabling affinity purification with chitin beads. The inteins were expressed by induction with L-arabinose in the SNS126 strain, purified by chitin affinity chromatography and visualised by SDS-PAGE (Figure 30). The SDS-PAGE image strongly indicated that the *Npu* constructs spliced; all samples showed an unspliced band (C-terminal intein: Target peptide: N-terminal intein: 23033.34 Da) and also two bands indicating splicing (N-terminal intein 17950.32 Da and C-terminal intein 4362.09 Da). There was also an unexplained band, however, as it appeared in the control SNS126 it was discounted.

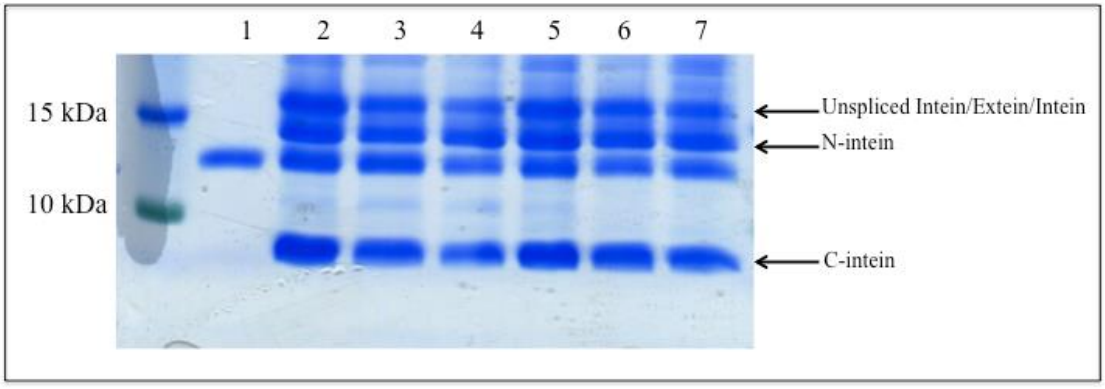


Figure 30 SDS-PAGE (15%) analysis of purified alanine analogues using *Npu* inteins. *(Lane 1)* SNS126 blank control; *(Lane 2)* CLLFVY; *(Lane 3)* CALFVY; *(Lane 4)* CLAFVY; *(Lane 5)* CLLAVY; *(Lane 6)* CLLFAY; *(Lane 7)* CLLFVA.

Once splicing had been confirmed, the *Npu* analogues were transformed into the HIF-1 RTHS. Resulting colonies were grown and 10-fold serial dilutions of the cultures were drop-spotted onto four minimal media plates. All plates contained 7.5 mM 3-AT and 25 μg/ml kanamycin, with each plate containing additional supplements, as outlined in

section 2.1.1 (Figure 31). As an additional control to observe the effects of toxicity the plasmids were transformed into blank SNS126 and drop-spotted at the same time.

The *Npu* drop-spotting indicated that induction of the *Npu* intein/extein construct was toxic. Addition of L-arabinose to both HIF-1 RTHS and SNS126 strains caused shut down of growth (Figure 31). The effect was most noticeable within the HIF-1 RTHS strain, which may be due to additional growth pressure caused by induction of the HIF-1 constructs.

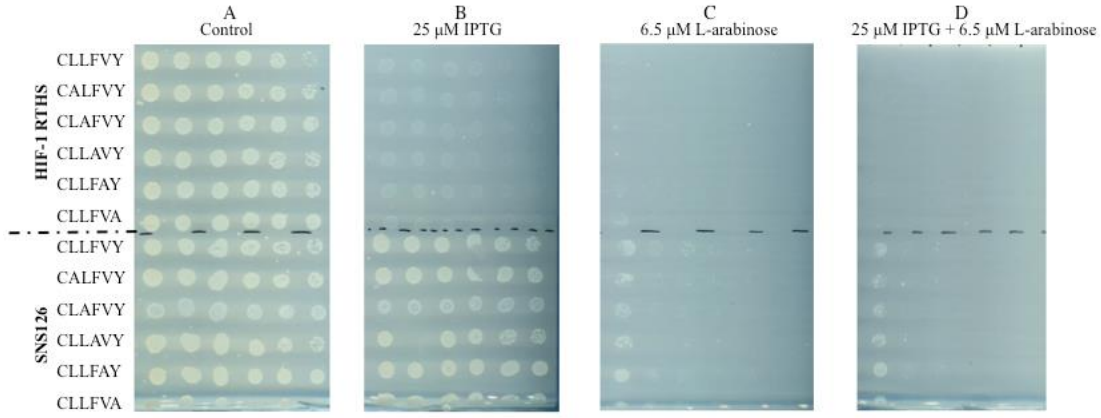


Figure 31 HIF-1 *Npu* alanine scanning drop-spotting.

All plates contained 3-AT and kanamycin. *(A)* Minimal media control plate, full growth was expected as there was no IPTG to induce expression of 434-HIF-1α and P22-HIF-1β; (*B)* Minimal media with IPTG, inhibition of growth was expected as HIF-1α and HIF-1β dimerise and block transcription of reporter genes; *(C)* Minimal media with L-arabinose, full growth was expected; (*D)* Minimal media with IPTG and L-arabinose, restoration of growth was expected compared to IPTG plate for any active cyclic peptides. Spots from left to right represent 10-fold serial dilutions of the cultured strain. The first spot represents a 10-fold dilution of the culture grown overnight.

Due to the toxic phenotype it was not possible to rank the alanine analogue activities; the toxicity observed was investigated further.

2.1.5 Investigating the toxic phenotype observed with *Npu* SICLOPPS plasmids

To confirm the toxicity effects observed with the *Npu* constructs, two additional drop-spotting experiments were carried out:

1.) *Ssp* and *Npu* P1 expressing plasmids were drop-spotted onto the same plates to visualise the differences in toxic phenotype and to confirm L-arabinose induced death with the *Npu* constructs (Figure 32A).

2.) P1 *Npu* Strains (HIF-1 RTHS and SNS126) were drop-spotted on plates with different L-arabinose concentrations, to establish if varying induction of the intein/extein construct would affect the toxic phenotype observed (Figure 32B and C).

The drop-spotting showed distinct differences between the *Ssp* and *Npu* plasmids. As previously observed (section 2.1.1), the HIF-1 RTHS *Ssp* P1 drop-spotting shut down in the presence of IPTG and displayed regrowth with the addition of L-arabinose. Whereas, L-arabinose addition (1.625-6.5 μM) caused death to all *Npu* variants.

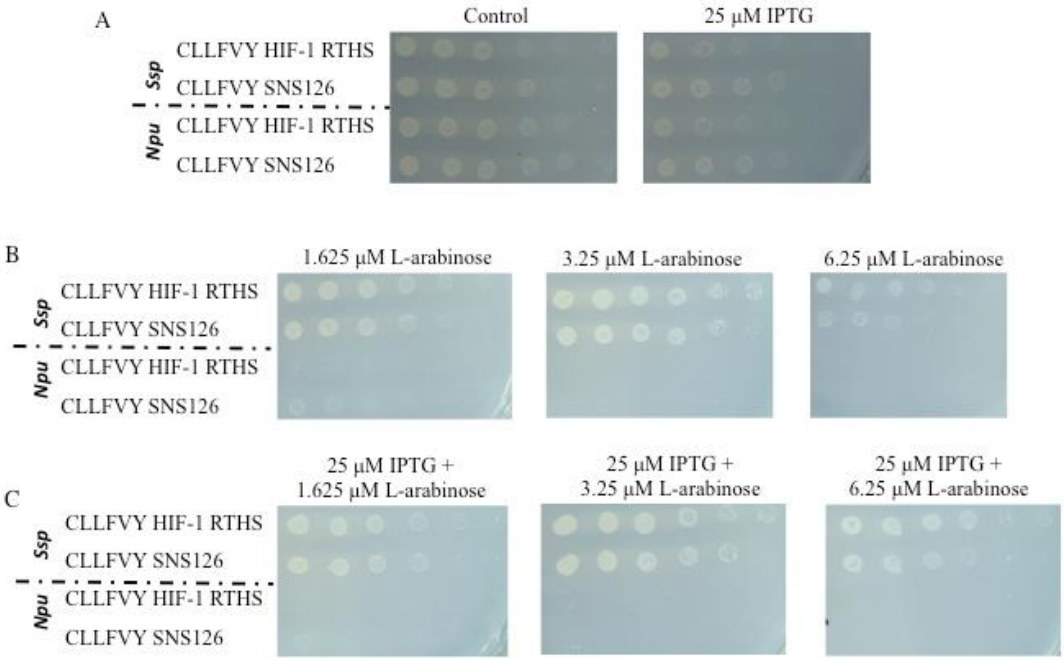


Figure 32 Drop-spotting to investigate the toxic phenotype of *Npu* intein/extein expression. *(A left)* Minimal media control plate, full growth was expected as there was no IPTG to induce expression of 434-HIF-1α and P22-HIF-1β; *(A right)* Minimal media with IPTG, inhibition of growth was expected for HIF-1 RTHS strains as HIF-1α and HIF-1β dimerise and block transcription of reporter genes; *(B left)* Minimal media with 1.625 μM L-arabinose, full growth expected; *(B middle)* Minimal media with 3.25 μM L-arabinose, full growth expected; *(B right)* Minimal media with 6.25 μM L-arabinose, full growth expected; *(C left)* Minimal media with IPTG and 1.625 μM L-arabinose, restoration of growth was expected compared to IPTG plate for any active cyclic peptides in the HIF-1 RTHS; *(C middle)* Minimal media with IPTG and 3.25 μM L-arabinose, restoration of growth was expected compared to IPTG plate for any active cyclic peptides within the HIF-1 RTHS; *(C right)* Minimal media with IPTG and 6.25 μM L-arabinose, restoration of growth was expected compared to IPTG plate for any active cyclic peptides within the HIF-1 RTHS. Spots from left to right represent 10-fold serial dilutions of the cultured strain. The first spot represents a 10-fold dilution of the culture grown overnight.

The inherent toxicity of the *Npu* inteins or cyclic peptides makes this drop-spotting technique inappropriate to rank the activity of the P1 alanine analogues. Current work by J. Townend has confirmed the observed toxicity by the *Npu* inteins. Further research into the use of the *Npu* inteins in combination with the SICLOPPS methodology is being carried out within the Tavassoli group (J. Townend, personal communication). Consequently, other methods of testing the alanine analogues were probed.

2.2 Synthesis of the P1 alanine analogues using Solid-phase peptide synthesis

As the RTHS was deemed to be unsuitable for determining the difference in the alanine analogue activity, the five alanine analogues (*cyclo*-CALFVY, CLAFVY, CLLAVY, CLLFAY and CLLFVA) were synthesised, by fluorenyloxycarbonyl (Fmoc) solid-phase peptide synthesis, so their activity could be tested *in vivo* and *in vitro* (Figure 33).

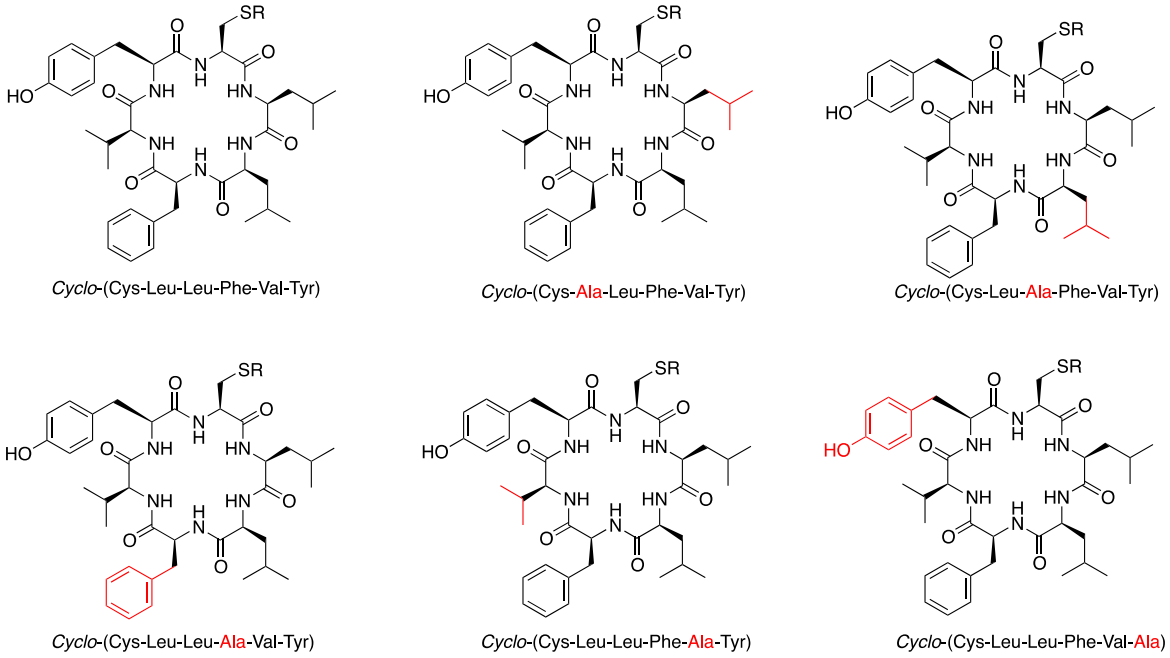


Figure 33 Structure of *cyclo*-CLLFVY and its five alanine analogues.

Cyclo-CALFVY, *cyclo*-CLAFVY, *cyclo*-CLLAVY, *cyclo*-CLLFAY, and *cyclo*-CLLFVA). R = H, spy (cysteine protecting group) or Tat (solubility tag).

2.2.1 Solid-phase peptide synthesis

Solid-phase peptide synthesis was first reported by Merrifield (1963).[133] It is a process by which small solid beads have covalently attached linkers from which amino acid chains can be built. The synthesised peptide becomes immobilised and soluble side-products and reagents from coupling reactions can be easily washed from the product containing resin. The peptides described here were synthesised using Fmoc solid-phase synthesis,[134] whereby, the carboxy-terminal for the first amino acid was attached to the resin via a Wang linker. The amino-terminus was protected with an Fmoc group, cleavable with 20% piperidine. Coupling of the next Fmoc-protected amino acid (activated by a suitable coupling reagent), could then take place, and the deprotection-coupling-deprotection-coupling cycle was continued until the desired peptide length was reached (Figure 33). At each step it was easy to wash away by-products and reagents. Once the peptide chain was completed, the peptide was then cleaved from the resin using 88-95% trifluoroacetic acid (TFA), alongside deprotection of acid-labile side chains (e.g. *tert* butyl or trityl groups).

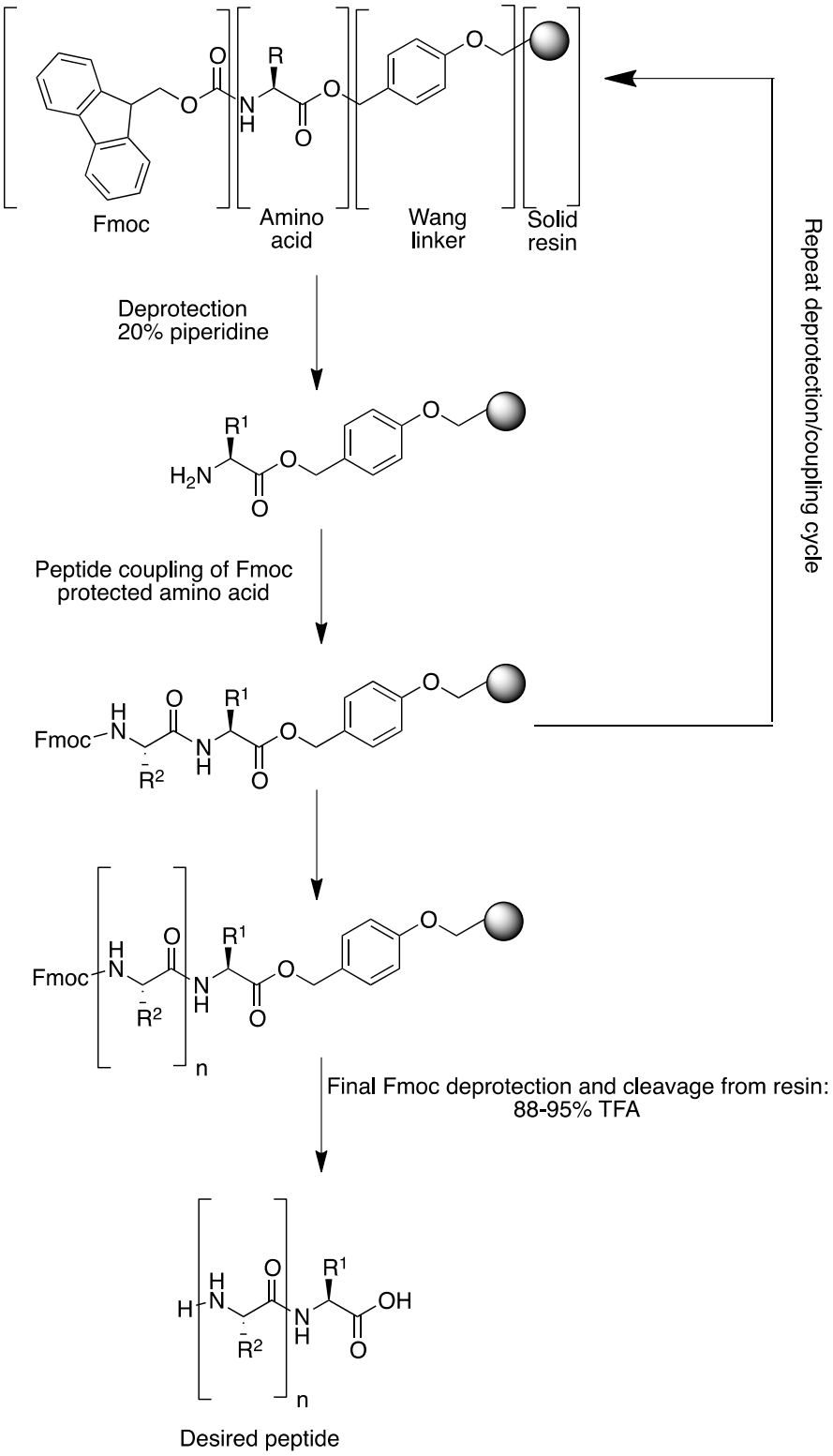


Figure 34 General schematic of Fmoc solid-phase peptide synthesis.

The first amino acid is attached to an insoluble resin via a Wang linker. The Fmoc protecting group is removed with 20% piperidine prior to coupling the subsequent Fmoc protected amino acid. The deprotection/coupling cycle is repeated until the desired peptide is complete. The final product can be cleaved from the Wang resin and simultaneously deprotected with 88-95% TFA.

All five alanine analogues of P1 were made using the same methodology as the parent compound. Based on previous optimisation of P1 synthesis, (work by Dr F. Ho Kwie, a Post Doctoral researcher in the Tavassoli lab), the cyclic peptide was split between the two leucine residues to minimise steric hindrance during cyclisation, compared to the more bulky aromatic side chains. Consequently, linear-LFVYCL was synthesised starting with an Fmoc-leucine loaded Wang resin. Diisopropylcarbodiimide (DIC) in combination with hydroxybenzotriazole (HOBt) was used to activate the carboxylic acid during coupling.[135] HOBt functions as a good leaving group, enhancing the rate of the coupling reaction and consequently, minimising racemisation.[135, 136]

In order to monitor the progress of the reaction, the ninhydrin test was used.[137] When a free amine is present, reaction with ninhydrin results in a blue colour; in the absence of primary amines, in the case of complete coupling, the solution remains a pale yellow colour. As described above, amino acid residues with reactive side chains were protected with orthogonal protecting groups. The *tert*-butyl (tyrosine) and trityl (cysteine) protecting groups were removed with TFA during the resin cleavage step. Triisopropylsilane (TIS), phenol and water were included in the cleavage cocktail to trap carbocations generated by protecting group removal, in order to avoid undesired side reactions.

For some of the analogues (LFAYCL and LFVACL) a microwave peptide synthesiser was used to synthesise the linear peptide. Unlike during manual synthesis benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBOP) and *N*,*N*-diisopropylethylamine (DIPEA) were used as the peptide coupling reagent/activator and activator base, respectively.[138]

The cysteine residue presented a synthetic problem. Firstly, if left unprotected during linear purification, disulphide bond formation between linear units may occur. Secondly, during cyclisation a thioester may have formed in place of a peptide bond and identification of this side product could be difficult. To avoid these problems, post cleavage, the thiol was protected with aldrithiol-2 and then the aldrithiol-linear peptide was purified by reverse-phase high-performance liquid chromatography (HPLC).[139] The purified linear peptides were analysed by nuclear magnetic resonance (NMR) spectroscopy (including ^{1}H NMR, total correlation spectroscopy (TOCSY) and correlation spectroscopy (COSY)), analytical HPLC and MS. Carbodiimide mediated peptide coupling using *N*-(3-dimethylaminopropyl)-*N*-ethylcarbodiimide hydrochloride (EDC) and HOBt was used to

cyclise the aldrithiol protected linear peptide. The cyclic peptide *cyclo*-C(spy)LLFVY was then purified by reverse-phase HPLC (Figure 35). The compounds were again analysed by NMR (including ^{1}H NMR, TOCSY and COSY), analytical HPLC and MS.

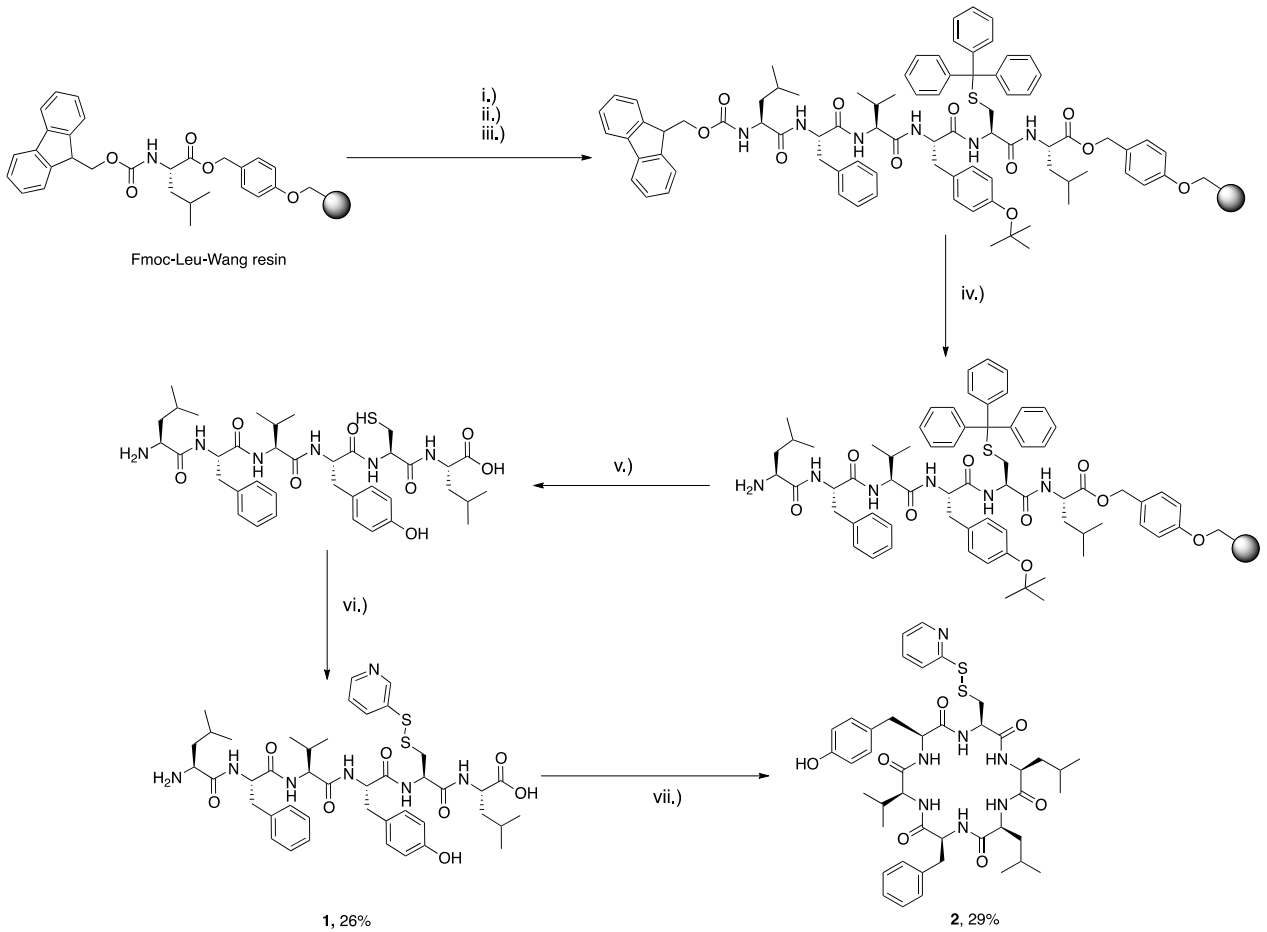


Figure 35 Scheme for the synthesis of P1.

i.) 20% piperidine/DMF (v/v); ii.) DIC (3 eq), HOBt (3 eq), Fmoc-Cys(Trt)-OH (3 eq); iii.) steps i.) and ii.) repeated for Fmoc-Tyr(tBu)-OH, Fmoc-Val-OH, Fmoc-Phe-OH, and Fmoc-Leu-OH (where the Liberty 1 peptide synthesiser was used, i.) 20% piperidine/DMF (v/v); ii.) PyBOP (4.4 eq), DIPEA (0.82 eq, in NMP 1:3, (v,v)), Fmoc-Cys(Trt)-OH (4.8 eq); iii.) steps i.) and ii.) repeated for Fmoc-Tyr(tBu)-OH, Fmoc-Val-OH, Fmoc-Phe-OH, and Fmoc-Leu-OH). iv.) 20% piperidine/DMF (v/v), v.) TFA/phenol/H_2O/TIS, 88/5/5/2, vi.) 2,2'dithiopyridine (5 eq), DMF, vii.) EDC (3 eq), HOBt (6 eq), DMF.

Analysis of the initial synthesis of the parent linear LFVYC(spy)L compound yielded a by-product with a *m/z* change of -3 units by mass spectrometry. On further analysis of the mass spectrometry, it was speculated that the impurity was generated from nucleophilic attack of the deprotected cysteine to the Wang linker (Figure 36).

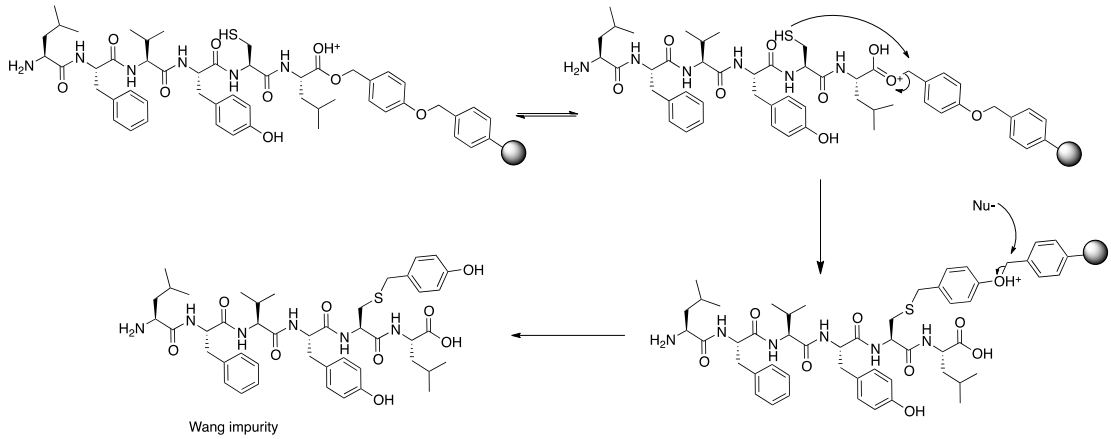


Figure 36 Proposed mechanism for synthesis of the Wang linker impurity.

This mechanism is thought to take place in the acidic cleavage cocktails; the nucleophilic sulphur of cysteine attacks the Wang linker.

Further confirmation of this impurity came from ^{1}H NMR analysis of the linear peptides. On each NMR two doublets at approximately 7.10 and 6.67 ppm were present, which could be assigned to the aromatic ring in the Wang linker impurity (Figure 37). Integration of the peaks showed the impurity to be approximately 10% in all analogues; due to its similarity to the parent compound it would prove difficult to remove by HPLC. However, the purity of the final tagged peptide should not be affected, as the Wang linker impurity will be unable to react in later tagging reactions and will subsequently be removed in the final Tat-tagged purification step.

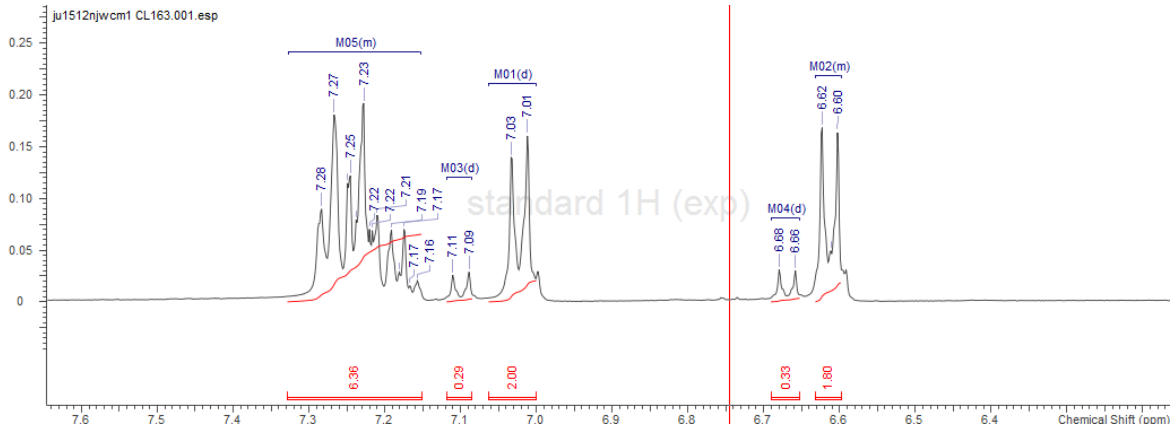


Figure 37 ^{1}H NMR trace indicating the Wang linker impurity.

Peaks at 7.10 and 6.67 ppm, d, J= 8.6 Hz. By ^{1}H NMR analysis suggests approximately 10% impurity.

Cyclo-C(spy)LLFVY and its analogues have poor aqueous solubility and require a tag to improve solubility and membrane permeability for *in vitro* and in-cell testing. A short peptide derived from the human immunodeficiency virus trans-activating-transcriptional activator (Tat) was attached to *cyclo*-CLLFVY, via a disulphide bond between the cysteine of the cyclic peptide and a cysteine placed at the beginning of the Tat sequence (CGRKKRRQRRRPPQ).[139] The Tat-tag has been used to aid the translocation of a number of peptides across the cell membrane of mammalian cells.[139-141] The Tat-tag was synthesised using a Liberty 1peptide synthesiser, starting from an Fmoc protected glutamine (trityl-protected) loaded Wang resin. Once completed, the linear peptide was cleaved from the resin and purified by reverse-phase HPLC.

A disulphide bond between the Tat-tag and cyclic peptide cysteine needed to be formed. Initially, the cysteine of the Tat-tag was protected with aldrithiol prior to HPLC purification. Tris(2-carboxyethyl phosphine) (TCEP), a reducing agent that is irreversibly oxidised during the reaction, was used to remove the aldrithiol group from the cyclic peptide prior to addition of the aldrithiol-protected Tat-tag. However, this additional protection step was not necessary and direct addition of the Tat-tag (1-2 eq), without aldrithiol protection, to the aldrithiol protected peptide was adequate for disulphide bond formation (Figure 38). Release of the aldrithiol resulted in a colour change to a yellow solution based on the release of thiopyridone. The identity and purity of the Tat-tagged compounds were confirmed by MS and liquid chromatography MS (LCMS), respectively.

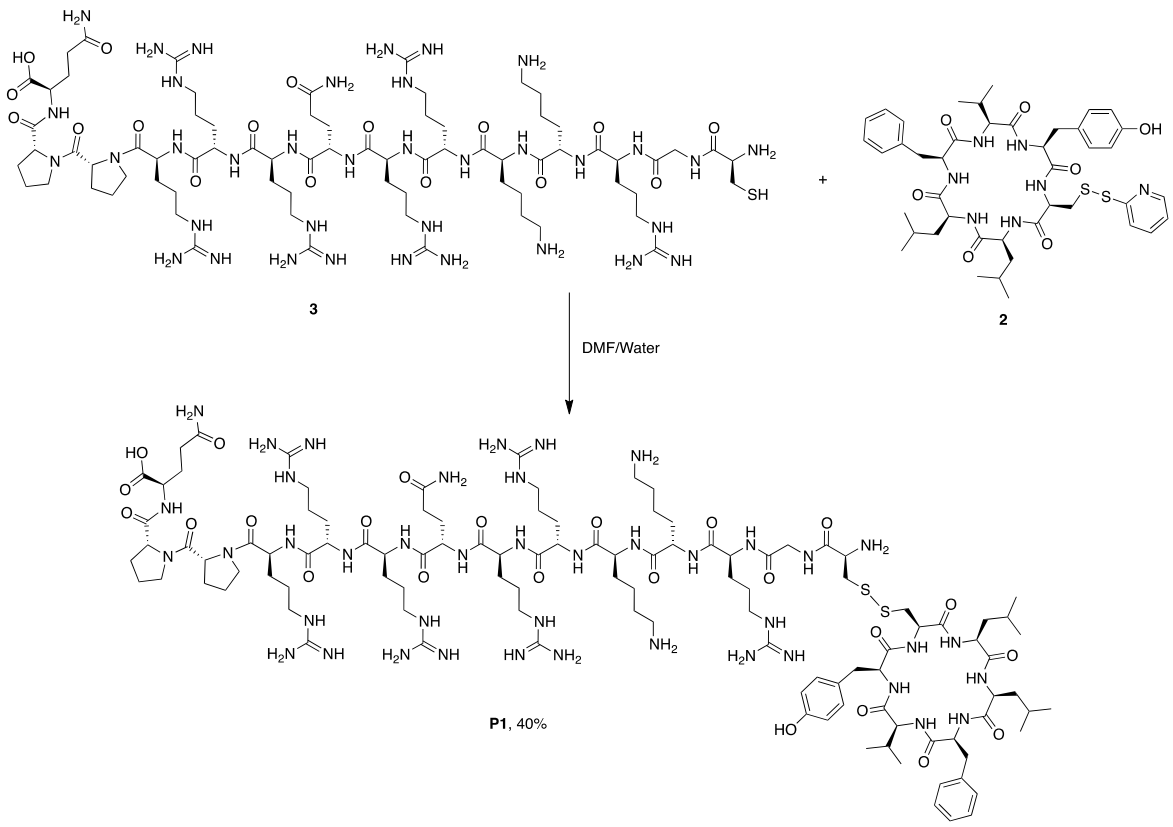


Figure 38 Structure of Tat-tag (CGRKKRRQRRPPQ).

The Tat-tag cysteine forms a disulphide bond with the cysteine in *cyclo*-C(spy)LLFVY; the aldrithiol protecting group in the cyclic peptide is displaced.

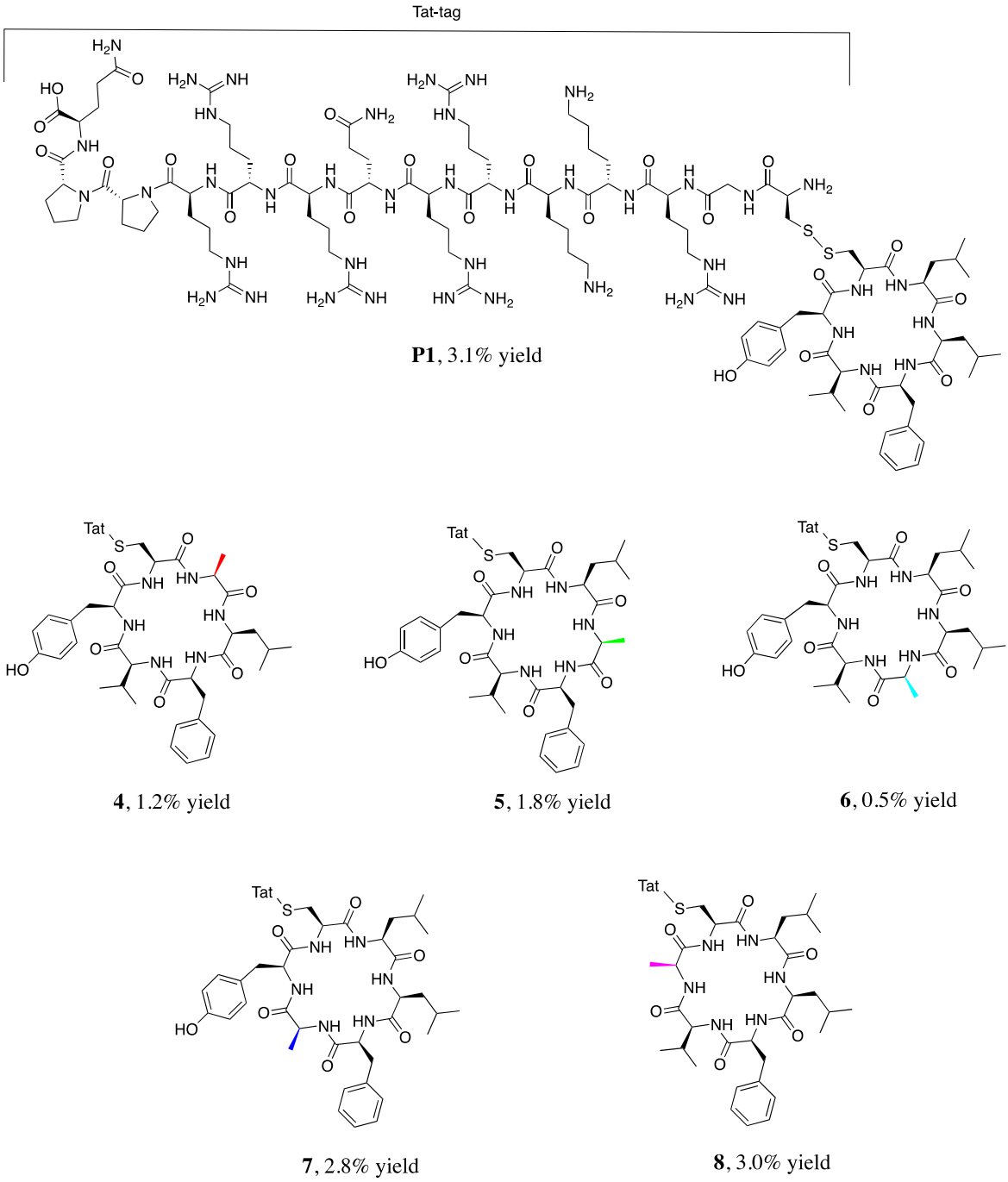


Figure 39 Structure and yields of P1 and its five Tat-tagged alanine analogues.

The successful synthesis of all five alanine analogues and the parent peptide, P1, allowed them to be tested in a selection of biological function and biophysical assays.

2.3 Biological function assay: testing of the alanine analogues: Hypoxia response element-dependent luciferase assay

The use of HRE-dependent luciferase assays to monitor HIF-1 dimerisation is widely reported, and has previously been used to assess the activity of P1.[1, 121, 142-144] These reporter assays place the transcription of the enzyme firefly (*Photinus pyralis*) luciferase gene under the transcriptional control of HIF-1 dimerisation. Firefly luciferase catalyses luciferin oxidation to form oxyluciferin, and this bioluminescent reaction releases light (Figure 40).

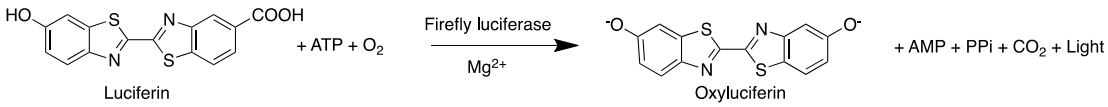


Figure 40 Bioluminescent reaction catalysed by firefly luciferase.
Luciferin is oxidised to oxyluciferin and light is released.

With the peptides in hand, a HRE-dependent luciferase assay was used to determine the relative activity of the alanine analogues to P1.

2.3.1 Development of the HRE-dependent luciferase assay in the literature

Rapisarda *et al.* (2002) first developed a HRE-dependent luciferase assay for HIF-1 dimerisation.[142] They constructed the pGL2-TK-HRE plasmid by cloning three repeats of the nitric oxide synthase promoter canonical HRE into a pGL2-TK vector (a Promega plasmid, with firefly luciferase expression controlled by a thymidine kinase (TK) promoter). The plasmid was stably transfected into U251 cells (human Neuronal Glioblastomal cells) where hypoxic conditions (1% O_2) or treatment with desferrioxamine mesylate salt (DFX, Figure 41B and C) induced a significant increase in luciferase expression. In this assay, DFX acted as a hypoxia mimic; by inhibiting the action of the PHDs via iron chelation, in turn, preventing HIF-1α degradation. Consequently, HIF-1α

accumulates as if hypoxia had occurred. Due to experimental difficulties simulating hypoxia, DFX is often used as a mimic. Concentrations of DFX used vary in the literature, up to approximately 500 μM.[121] The cells were used in a high throughput screening assay to identify inhibitors of HIF-1 transcriptional activity. Specifically, stably transfected cells were treated with library compounds under hypoxia. Seven non-toxic compounds were found to inhibit the hypoxic induction of luciferase expression.[142]

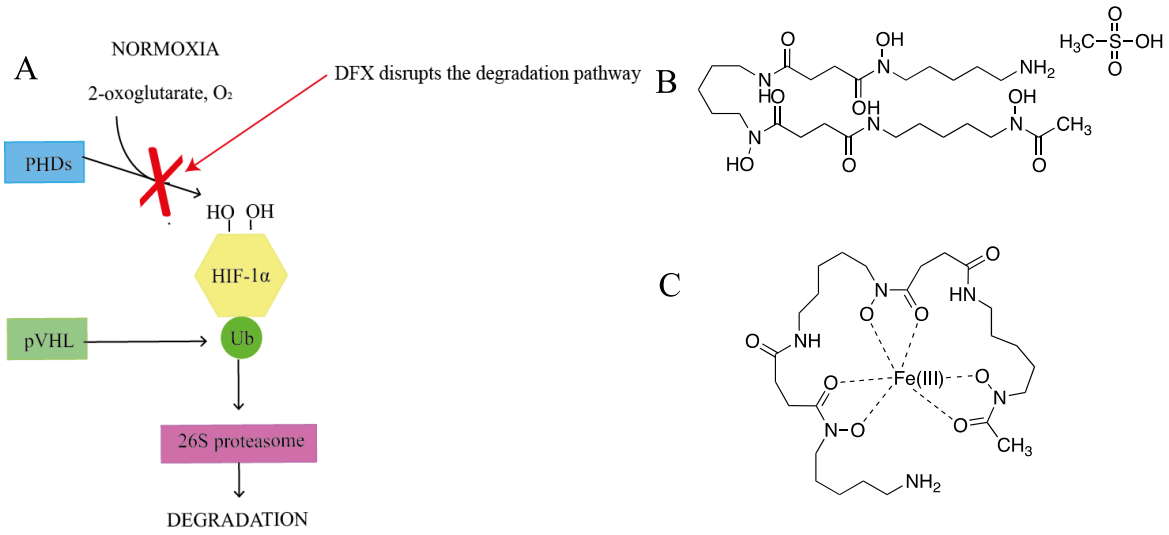


Figure 41 The action of DFX as a hypoxia mimic.

(A) DFX inhibits PHDs, disrupting the degradation of HIF-1α. *(B)* Structure of DFX. *(C)* Structure of DFX bound to iron.[145]

Similarly, Wang *et al.* (2009) transiently transfected the pGL2-TK-HRE plasmid into Michigan Cancer Foundation-7 (MCF-7) cells and utilised the HRE-dependent luciferase reporter assay to demonstrate the ability of phenethyl isothiocyanate to inhibit HIF-1 activity.[144] Again, they found upon treatment of hypoxia or DFX a significant increase in luciferase expression was observed by an increase in bioluminescence. However, at concentrations between 3.25-26 μM of the inhibitor (phenethyl isothiocyanate) this increase was mitigated. Recently, it was shown that P1 was able to disrupt HIF-1 dimerisation in such an assay, by transiently transfecting MCF-7 cells with pGL2-TK-HRE which were subsequently treated with P1 (0-50 μM) under hypoxia.[1] Under hypoxic conditions the luciferase activity increased over 10-fold, however, the addition of P1 (0-50 μM) reversed this, so luciferase activity was decreased in a dose-dependent manner (approximately 50% decrease in luciferase assay at 50 μM P1).[1]

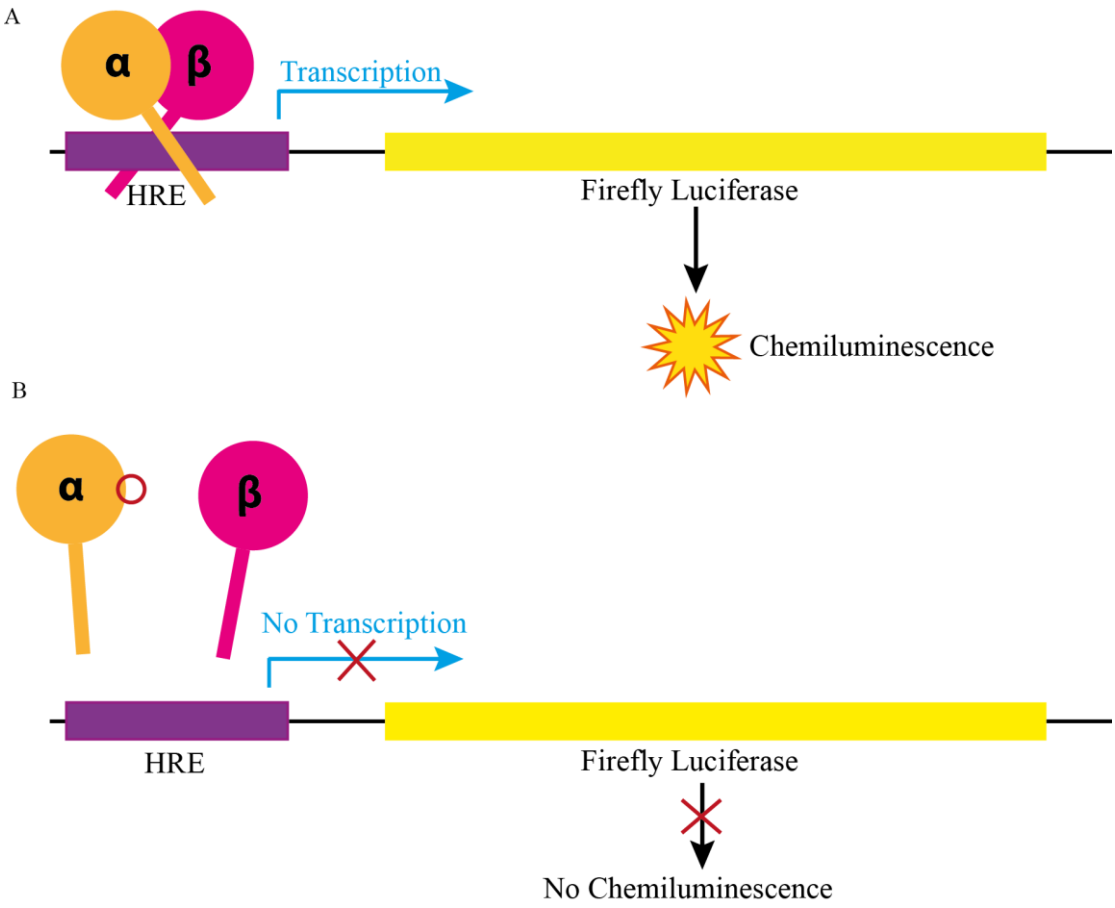


Figure 42 Illustration of the HRE-dependent luciferase assay.

(A) HIF-1α and HIF-1β dimerise and form an active transcription factor which binds to the HRE, leading to the expression of firefly luciferase. *(B)* Dimerisation of HIF-1α and HIF-1β are inhibited by a cyclic peptide, the monomeric proteins cannot bind to the HRE and expression of firefly luciferase is suppressed.

The successful application HRE luciferase assays, for the identification of HIF-1 inhibitors, sets a precedent for their use to investigate the activity of the P1 alanine analogues.

2.3.2 Optimisation of a HRE-dependent luciferase assay

A HRE-dependent luciferase assay was established for use in assessing the activity of alanine analogues. In order to compare the activity of the analogues, MCF-7 cells were co-transfected with a *Renilla* (*Renilla reniformis)* luciferase control plasmid (pRL-SV40)

alongside the pGL2-TK-HRE plasmid to allow normalisation of the firefly luciferase activity to the *Renilla* luciferase activity.[146] This enabled an accurate comparison between experiments and analogues to be made.

Before testing the inhibitors, the experimental parameters were optimised. These included the optimisation of the ratio of FuGENE 6 transfection reagent (Promega, UK) to DNA. This optimisation was carried out using pEGFP-N12.13 (a plasmid expressing green fluorescent protein (GFP)) and the transfection efficiency was observed with a fluorescent microscope. A 1:6 DNA:FuGENE ratio provided the most efficient transfection, and this ratio was used for further experiments (Figure 43). To confirm successful transfection during the HRE-dependent luciferase assays, a GFP transfection was performed in parallel to the transfection of pGL2-TK-HRE (HRE/firefly) and pRL-SV40 (*Renilla*) plasmid during each experiment enabling successful transfection to be visually confirmed.

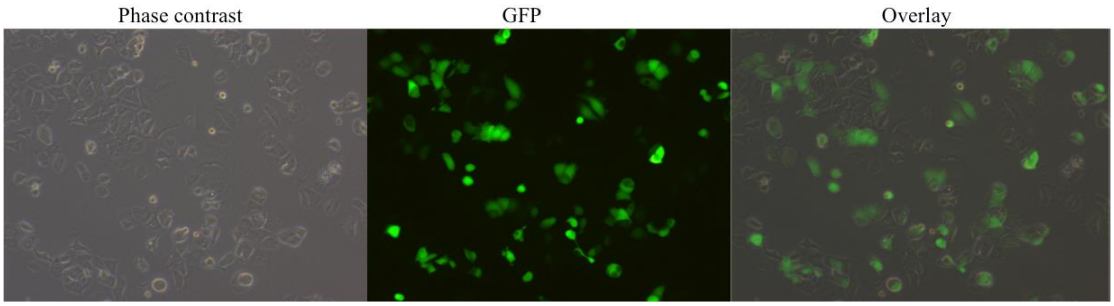


Figure 43 FuGENE transfection of pEGFP-N12.13 into MCF-7 cells.
Cells were seeded, 200,000 cells per well of a six well plate and transfected with a 1:6 ratio of DNA:FuGENE, 1 μg of plasmid was transfected into each well. At 24 hours post-transfection fluorescent and phase contrast images were captured and transfection efficiency was assessed. The images are representative of six individual experiments.

A suitable transfection ratio of the pGL2-TK-HRE (HRE/firefly) and luminescence readout was required (but it needed to be high enough to mask any residual firefly luminescence), a ratio of 1.25:98 of *Renilla* (pRL-SV40) to firefly reporter plasmid (pGL2 -TK-HRE) was decided upon. This generated a sizeable read out of both reporters and upon addition of DFX led to a Fluc/Rluc ratio of approximately 100-fold (DFX= 25 μM) which was deemed appropriate for the assay (Figure 44). A concentration of 25 μM DFX was chosen as it induced a large fold increase in luciferase activity allowing changes to be observed without saturating the system, high levels of DFX may cause off target effects or contribute to toxicity.

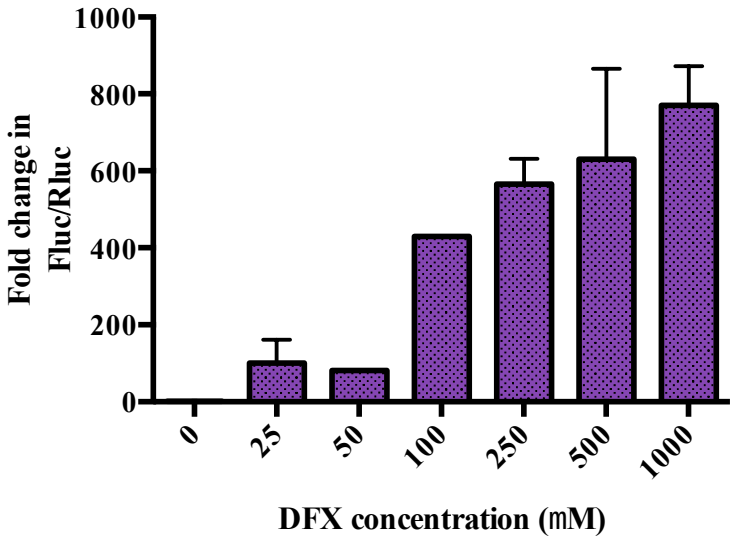


Figure 44 MCF-7 cells treated with DFX (0-1000 μM) transfected with pGL2 -TK-HRE and pRL-SV40 (*Renilla*) at a ratio of 98:1.25.

Data reported as fold change in Fluc/Rluc when treatment with water =1. All experiments done in triplicate well repeats, n=experiment repeat: 25 μM n=4, 50 μM n=1, 100 μM n=1, 250 μM n=2, 500 μM n=8, 1000 μM n=2.

The effect of DFX on HRE-dependent firefly luciferase activity was also investigated. Cells co-transfected with pGL2-HRE-luc or pGL3-luc (SV40 luciferase control plasmid) and pRL-SV40 (*Renilla*) were treated with varying concentrations of DFX (0-1 mM), using water as a vehicle. Increasing DFX concentration led to a significant increase in HRE-dependent firefly luciferase activity (as expected, *Renilla* activity did not change upon addition of DFX). As described above, 25 μM of DFX was used for testing the analogues. Addition of DFX to cells transfected with pGL3-luc and pRL-SV40 showed no significant changes in Fluc/Rluc with increasing DFX. This implied that the effect observed was due to HRE regulation not an effect on firefly luciferase activity (Figure 45).

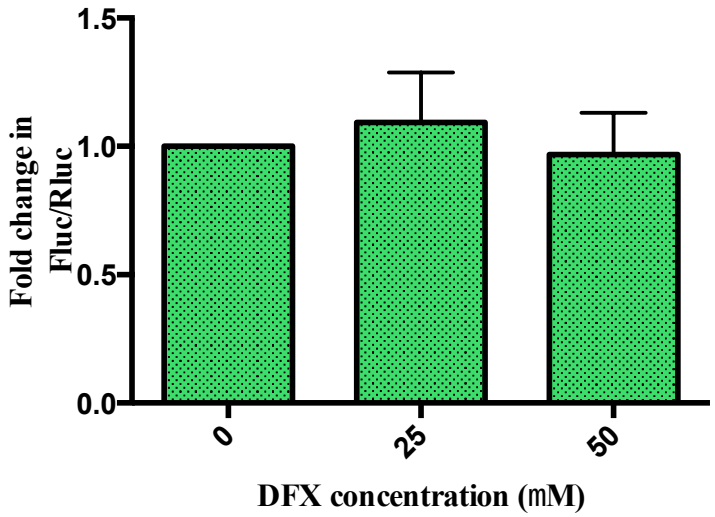


Figure 45 MCF-7 cells treated with DFX (0-500 μM) transfected with pGL3-promoter and pRL-SV40 at a ratio of 98:1.25.

Data reported as fold change in Fluc/Rluc when treatment with water =1. All experiments done in triplicate well repeats, n=experiment repeat: 25 μM n=4, 500 μM n=6.

Treatment with DFX leads to increased levels of HIF-1α within the cells, and to verify the activity of the DFX within our system, MCF-7 cells were treated with DFX (0, 250 and 500 μM) and the levels of the HIF-1α subunit were analysed by Western blot. HIF-1α was not detected when cells were treated with the vehicle (water). In contrast, in cells treated with DFX (250 and 500 μM) HIF-1α protein was evident (Figure 46). These concentrations were used to induce maximal HIF-1α levels that could be detected by Western blot. For all samples housekeeper protein glyceraldehyde 3-phosphate dehydrogenase (GAPDH) levels were similar for all treatments. The resulting Western blot confirms that in our system addition of DFX leads to the stabilisation of the HIF-1α protein (Figure 46).

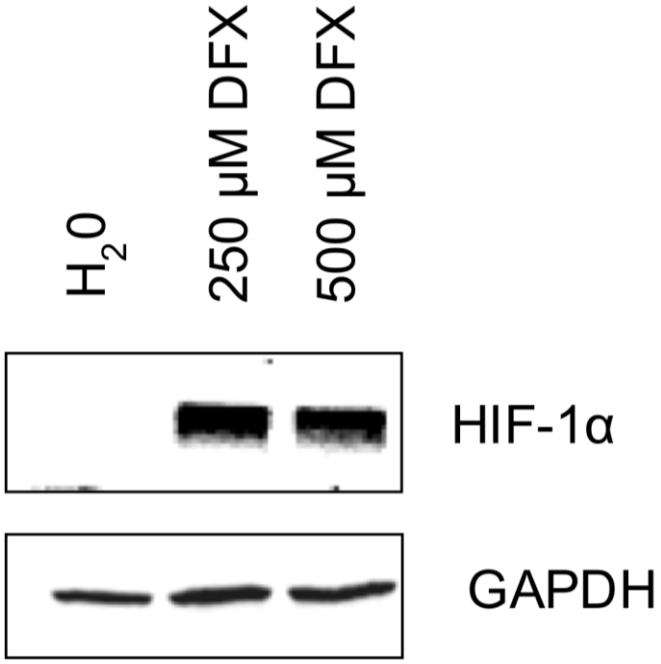


Figure 46 Western blot of HIF-1α from MCF-7 cells treated with DFX.

Anti-HIF-1α used to detect HIF-1α in cells treated with (from left to right) vehicle (water), 250 μM DFX and 500 μM DFX. GAPDH was used as a control.

2.3.3 Testing the P1 alanine analogues in the HRE-dependent luciferase assay

Once assay conditions were confirmed, the assay was used to test the activity of the alanine analogues. Firstly, solutions of the analogues were made. Arginine rich peptides are often found to be hygroscopic, and this is the case with the Tat-tagged alanine analogues. NMR analysis showed water to be present post lyophilisation and a comparison between concentrations of P1 by weight and by measurement of absorbance at 280 nm, found 38% (mass) water to be present. Therefore, peptide stocks were made by dissolving the peptides in sterile deionised water, measuring the absorbance at 280 nm and calculating the concentration using a calculated extinction coefficient and the Beer Lambert Law (as described in section 5.3.2.10).[147, 148] The tyrosine chromophore made this methodology possible for all analogues except Tat-*cyclo*-CLLFVA. Here, the concentration had to be calculated by the mass of the compound and adjusted to allow for the presence of 38% (mass) water. To do this, the peptide was weighed and dissolved simultaneously with a

peptide containing tyrosine. To ensure that the adjustment could be as accurate as possible, both peptides were lyophilised prior to being weighed.

To test the analogues, cells were transfected with pGL2-TK-HRE and pRL-SV40, and then 24 hours later plated onto opaque 96-well plates and incubated for five hours to allow adhesion to the plate surface. Alanine analogues were then added and after 30 minutes incubation, DFX (25 μM) was added. The firefly and Renilla luciferase activities were measured after 16 hours incubation and cell lysis. The HRE-dependent luciferase assay showed differences between the analogues, as previously addition of DFX (25 μM) caused a large increase in Fluc/Rluc (approximately 150-fold). A summary of the effect that each analogue had on HRE-dependent luciferase activity is described below (Figure 47).

Tat-*cyclo*-CLLFVY (P1): Fluc/Rluc decreased in a dose-dependent manner with addition of P1. This agreed with previous reports.[1] 50 and 100 μM P1 led to approximately a 50% decrease in activity when compared to the vehicle.

Tat-*cyclo*-CALFVY: No trend was observed, the peptide appeared to have little effect on HRE-dependent luciferase activity.

Tat-*cyclo*-CLAFVY: The general trend observed here was a dose-dependent increase in HRE-dependent luciferase activity; 100 μM of Tat-*cyclo*-CLAFVY led to almost double Fluc/Rluc compared to the vehicle.

Tat-*cyclo*-CLLAVY: Addition of inhibitor led to a small decrease in HRE-dependent luciferase activity; 25, 50 and 100 μM of inhibitor showed a similar effect (70-100-fold increase in Fluc/Rluc) compared to 150-fold increase of the vehicle.

Tat-*cyclo*-CLLFAY: Addition of 50 μM and 100 μM of inhibitor led to a small increase in Fluc/Rluc compared to the vehicle, a dose-dependent trend was not evident.

Tat-*cyclo*-CLLFVA: A dose-dependent decrease in Fluc/Rluc was observed with the addition of inhibitor. The inhibitor appeared to be more potent than P1. 100 μM showed a 100-fold decrease in Fluc/Rluc. It should be noted that for 100 μM both Rluc and Fluc individually decreased, which could indicate that this concentration of the inhibitor was toxic. Reporting the ratio Fluc/Rluc should account for this discrepancy, however, drawing

conclusions from the 50 μM concentration only might provide more reliable comparisons. Furthermore, cells transfected with pGL3 (control plasmid) and treated with alanine analogues showed little variation in Fluc/Rluc (Figure 48); this indicated that the effects observed were due to disruption of the hypoxia response network, not a direct effect on firefly luciferase activity.

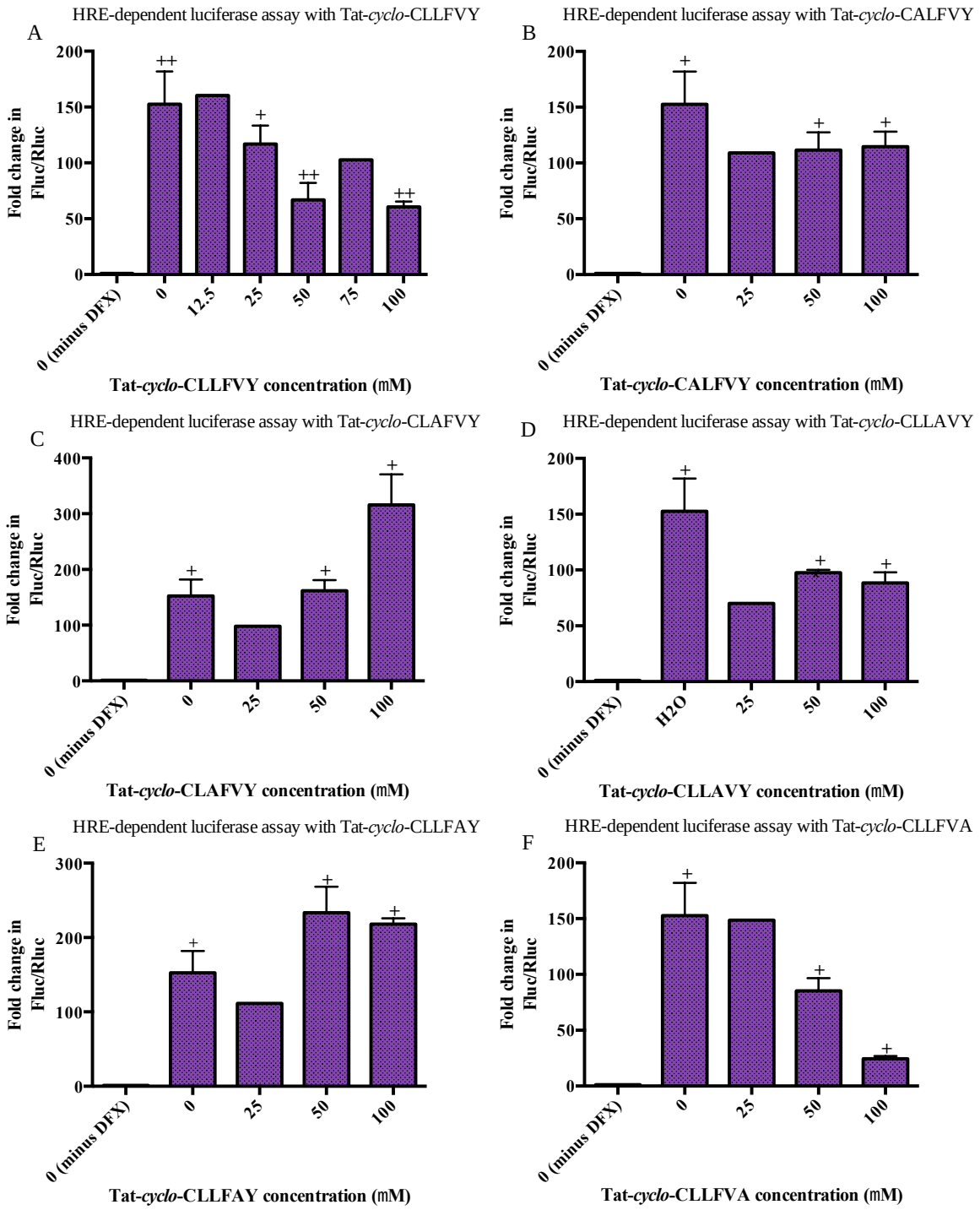


Figure 47 HRE-dependent luciferase assay alanine scanning.

MCF-7 cells transfected with pGL2-TK-HRE and pRL-SV40, were treated with all six alanine analogues (0-100 µM). Firefly and *Renilla* luciferase activity was measured and the relative firefly luciferase activity calculated (Fluc/Rluc). All experiments were performed in triplicate well repeats to ensure result consistency. Experimental repeats (n) denotes as follows: $^{+}$ n=2, $^{++}$ n=3, no cross n=1.

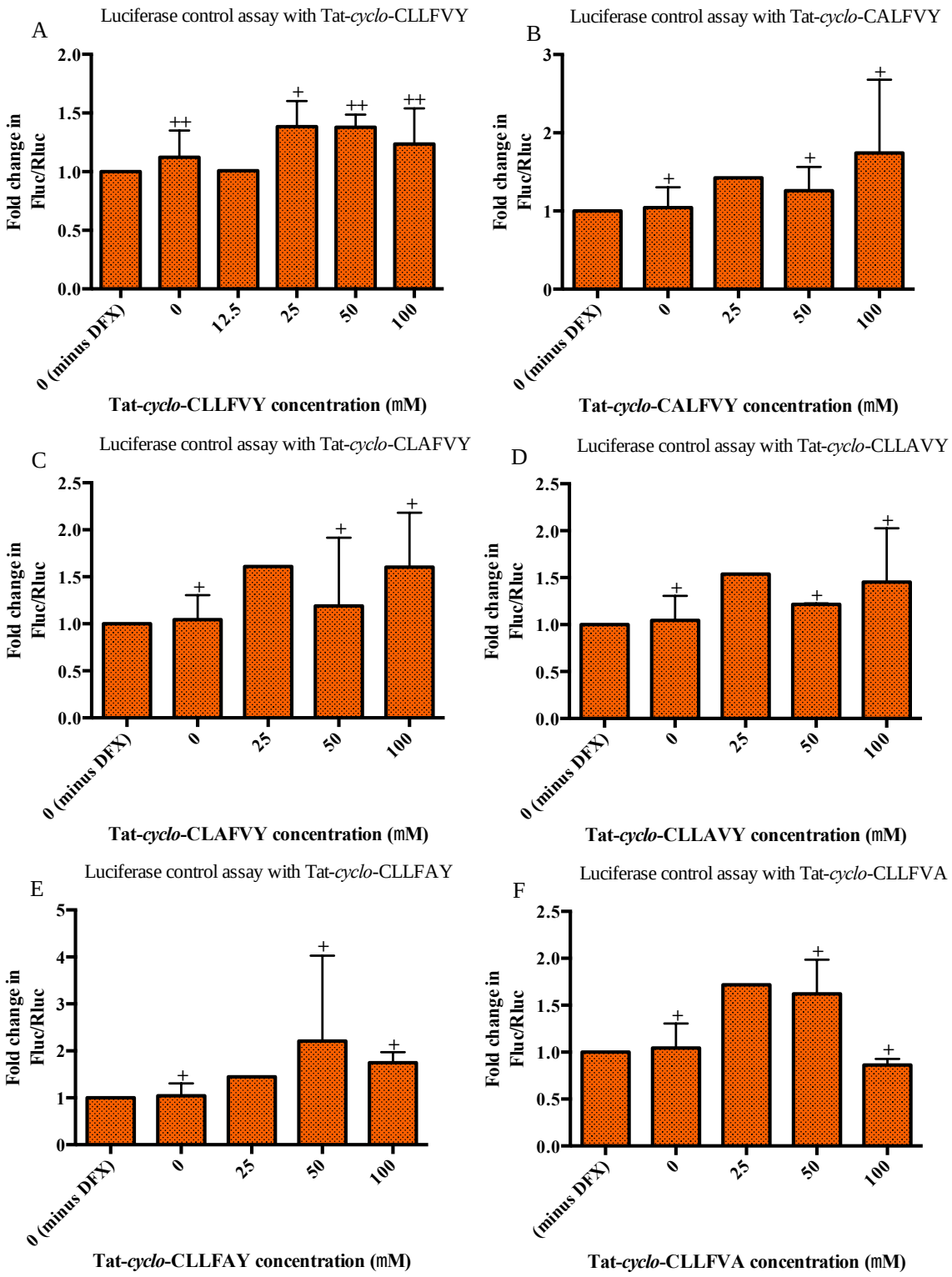


Figure 48 Luciferase control assay alanine scanning.

MCF-7 cells transfect). Firefly and *Renilla* luciferase activity was measured and the relative firefly luciferase activity calculated (Fluc/Rluc). All experiments were performed in triplicate well repeats to ensure result consistency. Experimental repeats (n) denoted as follows: $^{+}$ n=2, $^{++}$ n=3, no cross n=1.

Finally, to allow better comparison, Figure 49 shows the difference in the activity of the analogues at 50 μM. Treatment with P1 showed the expected dose-dependent decrease in luciferase activity. In addition, 50 μM treatment of Tat-*cyclo*-CLLFVA showed similar inhibitory action to the parent compound. Similarly, Tat-*cyclo*-CLLAVY showed some evidence of inhibitory action. However, Tat-*cyclo*-CALFVY and Tat-*cyclo*-CLLFAY showed no significant trends for activity. Whilst, Tat-*cyclo*-CLAFVY led to an increase in HRE–dependent luciferase activity. This was an unexplained result, however, the hypoxia response network is highly complex and this could possibly be due to off target actions. Based on the results observed here no conclusions about the involvement of the second leucine in P1 can be made. These results tentatively suggest that the tyrosine residue is not critical to the inhibitors activity, on the other hand, the first leucine residue and the valine appear to be critical to activity. The involvement of phenylalanine is unclear as some inhibitory action is indicated.

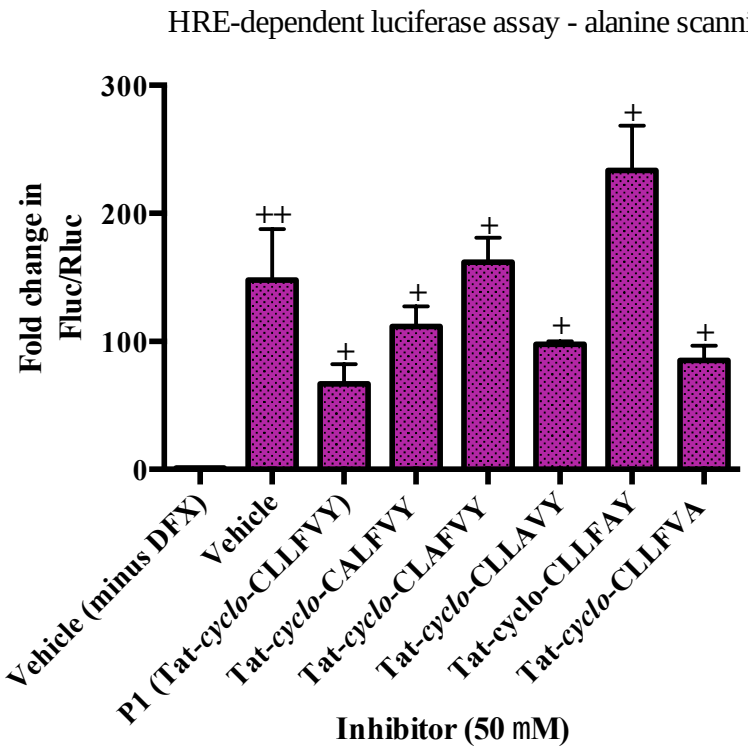


Figure 49 HRE-dependent luciferase assay alanine scanning comparison at an inhibitor concentration of 50 μM.

MCF-7 cells transfected with pGL2-TK-HRE and pRL-SV40. Firefly and *Renilla* luciferase activity were measured and the relative firefly luciferase activity calculated (Fluc/Rluc). All experiments were performed in triplicate well repeats to ensure result consistency. Experimental repeats (n) denoted as follows: $^{+}$ n=2, $^{++}$ n=3, no cross n=1.

Critically, these results and conclusions regarding the active motif of P1 require further verification. Another assay was required to test the alanine analogue activities so a comparison could be made.

2.4 Generation of an in vitro HIF-1 assay to test the alanine analogues

As the alanine scanning using the HRE-dependent luciferase showed inconclusive results, an *in vitro* assay was investigated to test the alanine analogues activity instead of using an *in vivo* approach. A protein-based assay would eliminate off target effects of the analogues and theoretically facilitate comparison of the activity of the analogues. Firstly, the expression of recombinant proteins was explored. It was also necessary to show that the recombinant proteins were functional, i.e. able to dimerise and bind DNA. There are references within the literature reporting inactive recombinant HIF, principally when recombinant HIF-1α and HIF-1β had been independently expressed.[60, 149]

2.4.1 Initial expression of recombinant HIF-1α and HIF-1β proteins

The expression of recombinant HIF-$1\alpha_{1\text{-}350}$ and HIF-$1\beta_{1\text{-}459}$ was performed.[150] These N-terminal fragments were chosen because they incorporate the bHLH, PAS-A and PAS-B domains, which are critical to DNA-binding and dimerisation and also because these fragments were used by Dr K. Nordgren in the HIF-1 RTHS. It can be assumed that P1 binds here.[1, 13] The principal isoform 1 of HIF-1α and isoform 3 of HIF-1β were used in all cases.

In order to individually and co-express HIF-$1\alpha_{1\text{-}350}$, the gene was cloned into pET32a (ampicillin resistant, Novagen, UK) and pET28a (kanamycin resistant, Novagen, UK), cloning performed by Dr K. Nordgren. Initially, HIF-$1\alpha_{1\text{-}350}$ was expressed from pET32a, which encodes a Trx-His-tagged recombinant protein (pET32a-Trx-His-HIF-$1\alpha_{1\text{-}350}$). Thioredoxin (Trx) is a 109 amino acid protein, which aids the solubility of recombinant proteins. A band visualised by SDS-PAGE at approximately 55 kDa indicated that Trx-His-HIF-$1\alpha_{1\text{-}350}$ (58.0 kDa) was expressed from pET32a-Trx-His-HIF-$1\alpha_{1\text{-}350}$ in BL21 (DE3) (Figure 50).

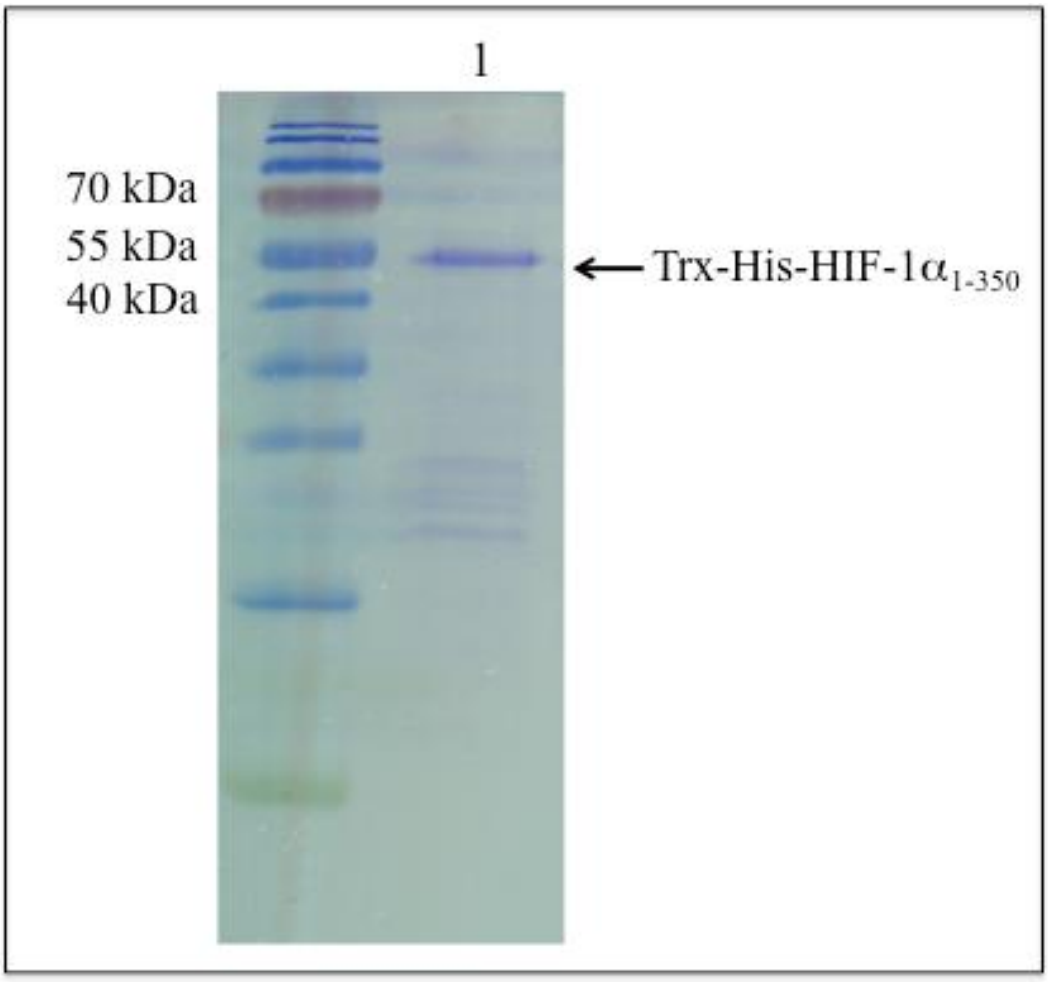


Figure 50 SDS-PAGE (12%) analysis of purified Trx-His-HIF-1α_{1-350}.

(Lane 1) Purified and desalted sample of Trx-His-HIF-1α_{1-350}.

To obtain HIF-1α_{1-350} from a kanamycin resistant plasmid, His-HIF-1α_{1-350} was expressed from pET28a. Expression of soluble His-HIF-1α_{1-350} from this vector was problematic, suggesting that the N-terminal Trx-tag expressed from pET32a could aid stabilisation of HIF-1α_{1-350}.[149] Therefore, to have Trx-His-HIF-1α_{1-350} in a kanamycin resistant plasmid, Trx-His-HIF-1α_{1-350} was cloned into pET28a; by digesting Trx-His-HIF-1α_{1-350} from pET32a-Trx-His-HIF-1α_{1-350} and ligating into pET28a to give pET28a-Trx-His-HIF-1α_{1-350} (kanamycin resistant). Analysis by SDS-PAGE indicated successful expression of Trx-His-HIF-1α_{1-350} from the plasmid in the expression strain BL21 (DE3).

HIF-1β_{1-459} had been cloned into pGEX-2TK (GE Healthcare, ampicillin resistant, cloning by Dr E. Miranda), encoding a GST-tagged recombinant protein. Analysis by SDS-PAGE showed a band at approximately 80 kDa, indicating successful expression of the desired

recombinant protein, GST-HIF-1β_{1-459} (78.1 kDa), in BL21 (DE3) (Figure 51). All plasmids used for protein expression had a lac promoter and could be induced by IPTG.

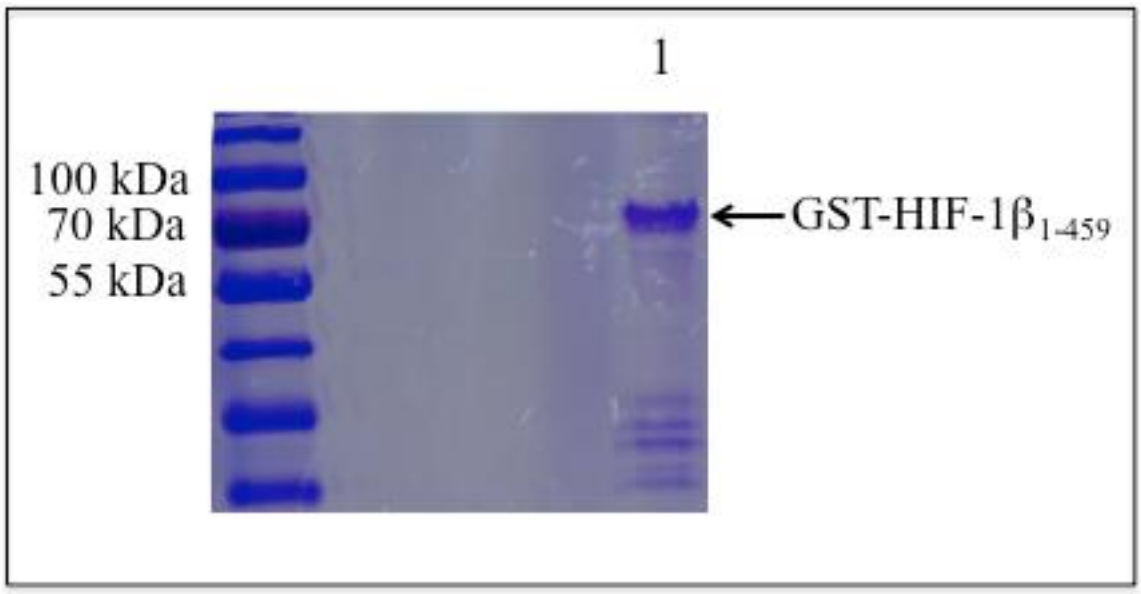


Figure 51 SDS-PAGE (12%) analysis of purified GST-HIF-1β_{1-459}.

(Lane 1) Purified and desalted sample of GST-HIF-1β_{1-459}.

As previously described, the solubility of expressed recombinant HIF-1α_{1-350} and HIF-1β_{1-459} was initially problematic; a significant amount of protein was detected in the insoluble cell fraction by SDS-PAGE.[151] Optimum expression conditions were found by induction at OD_{600} at 0.6. At this OD, most host cells have reached exponential growth, which is optimal for protein expression. Lower induction levels can lead to decreased levels of insoluble protein. Expression of Trx-His-HIF-1α_{1-350} with induction using both 0.1 mM and 0.5 mM IPTG was carried out. There appeared to be little difference in the ratio of protein in the soluble and insoluble fraction of the cell lysate, therefore, 0.1 mM IPTG was used for induction of all protein expression.[151] Furthermore, post induction the cultures were incubated at 18°C overnight. It is thought that incubating the expression culture at a lower temperature can help to achieve maximum recombinant protein solubility.[151] Lower induction levels are postulated to aid the production of soluble protein, additionally the use of auto-inducing media to express Trx-His-HIF-1α_{1-350} was tried. The media contains glucose and lactose; the glucose inhibits induction in the early stages of growth and then, when there is no glucose left, lactose slowly induces protein expression. The use of this media did not appear to improve solubility and therefore LB media was used.[152] The

identified optimal conditions allowed the production of soluble HIF-1 proteins in sufficient quantity, to enable the development of *in vitro* HIF heterodimerisation assays and biophysical binding assays.

The proteins were routinely purified using nickel and GST affinity Sepharose beads as this allowed simultaneous purification of both proteins at 4°C. This was achieved by centrifuging the expression cultures, removing the supernatant and storing the pellets at -80°C to aid lysis. The frozen pellets were then resuspended in lysis buffer (GST lysis buffer for GST-fusion proteins, and MHW buffer for His-fusion proteins).[118] Both lysis buffers were tris-based and contained Triton X-100, a surfactant that aided protein solubility, and protease inhibitors to prevent proteolysis. Furthermore, lysozyme was added to the lysates to aid degradation of the cell wall. The solutions were sonicated to further lyse the cells and then centrifuged, the soluble layer was added to Sepharose beads (GST-HIF-1$\beta_{1\text{-}459}$ to GST Sepharose and Trx-His-HIF-1$\alpha_{1\text{-}350}$ to Ni Sepharose). The beads were agitated at 5°C for an hour and then centrifuged, the supernatant containing unbound proteins removed and the Sepharose beads washed twice with appropriate wash buffers. Elution buffer (containing 500 mM imidazole to elute His-tagged recombinant proteins and 10 mM reduced glutathione to elute GST-tagged proteins) was added to the Sepharose beads, the beads were incubated at 5°C with agitation for 1 hour. The beads were centrifuged and the supernatant containing the protein retained and analysed by SDS-PAGE. To remove the imidazole or reduced glutathione, for *in vitro* assays, the proteins were desalted using PD10 size exclusion columns. The protein concentrations were calculated using a Bradford reagent assay. The proteins were flash frozen at -80°C until use. Attempts to purify the proteins by fast purification liquid chromatography (FPLC) often led to degradation, this may be because in comparison to beads purification, FPLC purification was carried out at room temperature in our lab.

Having confirmed that the proteins could be reproducibly independently expressed, Trx-His-HIF-1$\alpha_{1\text{-}350}$ and GST-HIF-1$\beta_{1\text{-}459}$ were coexpressed. This was to avoid independent expression, which in the literature has been shown to result in compromised functionality.[60, 149]

2.4.2 Inhibition of dimerisation of co-expressed Trx-His-HIF-1α_{1-350} and GST-HIF-1β_{1-459}

Firstly the proteins were co-expressed, due to reports that independently expressed HIF-1α and HIF-1β can lack functionality.[60, 149, 153] pGEX-GST-HIF-1β_{1-459} (ampicillin resistant) and pET28a-Trx-His-HIF-1α_{1-350} (kanamycin resistance) were transformed into the same BL21 (DE3) strain, sequentially. GST pull-down assays were then used to establish if *cyclo*-CLLFVY could inhibit the dimerisation of co-expressed HIF-1 *in vitro*. The assays were also carried out using the published inhibitors of HIF dimerisation, ACF and NSC-50352.[118, 119]

Firstly, an experiment was done without eluting the proteins from the GST beads during purification; after bead washing the beads were resuspended in tris lysis buffer. The bead slurries were incubated with varying concentrations of ACF (0.1 – 500 μM) overnight at 4°C. The beads were then washed thoroughly and analysed by SDS-PAGE. The gel showed no differences in the amounts of Trx-His-HIF-1α_{1-350} with increasing ACF concentration (Figure 17). This result suggests that ACF does not disrupt the interaction between HIF-1α_{1-350}/HIF-1β_{1-459} within the assay. It should, however, be noted that non-specific binding of Trx-His-HIF-1α_{1-350} to the GST Sepharose beads could not be ruled out. This experiment was repeated with inhibitor NSC-50352 (1- 100 μM) and *cyclo*-CLLFVY (1- 500 μM), again no decrease in Trx-His-HIF-1α_{1-350} was observed with increasing concentration of inhibitor, suggesting that dimerisation was not inhibited.

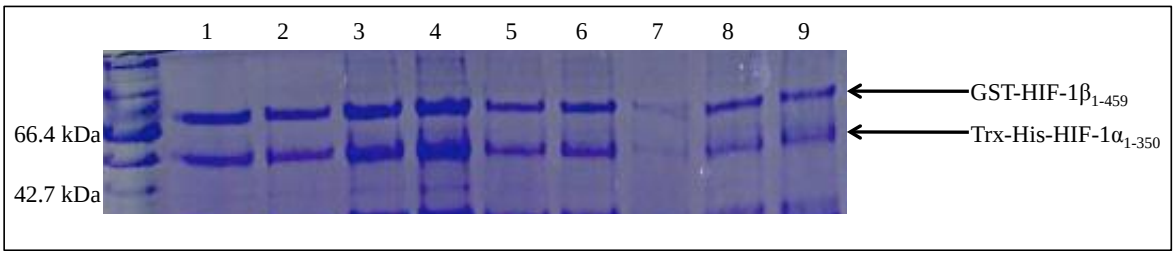


Figure 52 SDS-PAGE of GST pull-down experiment with co-expressed Trx-His-HIF-1α$_{1-350}$ and GST-HIF-β$_{1-459}$ and ACF.

(Lane 1) Trx-His-HIF-1α$_{1-350}$/GST-HIF-β$_{1-459}$; *(Lane 2)* Trx-His-HIF-1α$_{1-350}$/GST-HIF-β$_{1-459,}$ DMSO (vehicle control); *(Lane 3)* Trx-His-HIF-1α$_{1-350}$/GST-HIF-β$_{1-459,}$ 0.1 μM ACF; *(Lane 4)* Trx-His-HIF-1α$_{1-350}$/GST-HIF-β$_{1-459,}$ 0.5 μM ACF; *(Lane 5)* Trx-His-HIF-1α$_{1-350}$/GST-HIF-β$_{1-459,}$ 1 μM ACF; *(Lane 6)* Trx-His-HIF-1α$_{1-350}$/GST-HIF-β$_{1-459,}$ 5 μM ACF; *(Lane 7)* Trx-His-HIF-1α$_{1-350}$/GST-HIF-β$_{1-459,}$ 50 μM ACF; *(Lane 8)* Trx-His-HIF-1α$_{1-350}$/GST-HIF-β$_{1-459,}$ 100 μM ACF; *(Lane 9)* Trx-His-HIF-1α$_{1-350}$/GST-HIF-β$_{1-459,}$ 500 μM ACF.

The failure to observe inhibition of dimerisation in this assay may be due to the inherent nature of the inhibitors activity arising from the conditions under which it was selected. The P1-RTHS, NSC-50352 ELISA and ACF luciferase assays, used to select inhibitors, are likely to present a dynamic mixture of monomer and dimer proteins. Whereas, in this assay, the inhibitor will only encounter a preformed dimer. If the inhibitor acts by binding a monomer and preventing its dimerisation, its activity will not be observed in this assay. The following pull-down assays were carried out using independently expressed proteins to eliminate this factor.

2.4.3 Pull-down of individually expressed proteins to confirm dimerisation

Having successfully purified recombinant HIF-1α and HIF-1β proteins (section 2.4.1), a Ni pull-down was carried out to confirm dimerisation of the two recombinant HIF-1α and HIF-1β proteins and perhaps ultimately investigate if P1 and the alanine analogues could be probed using this assay. Trx-His-HIF-1α$_{1-350}$ has Ni affinity via the His-tag, therefore Ni beads will 'pull-down' the protein and any interacting proteins as depicted below (Figure 53).

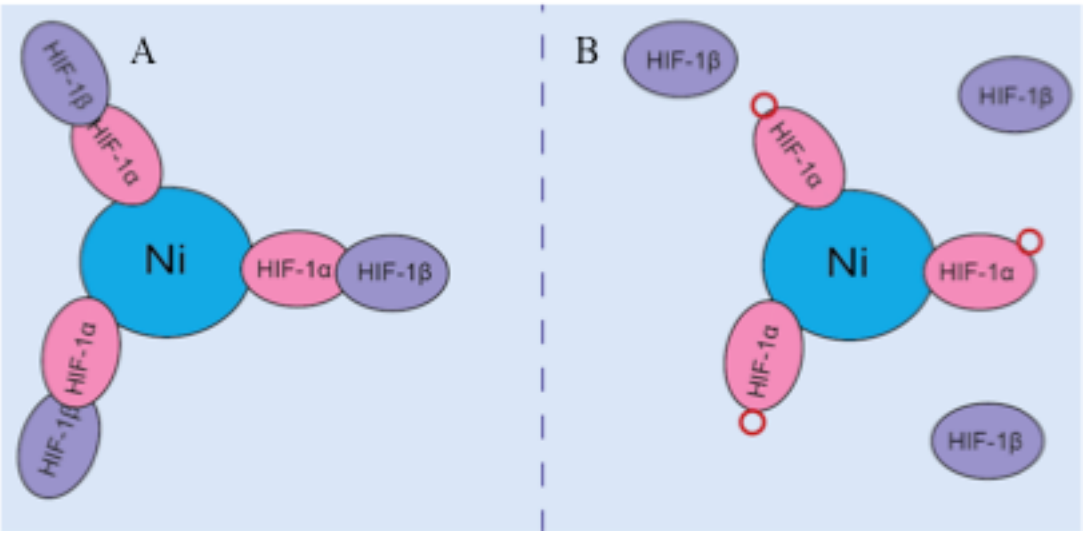


Figure 53 The principle of the Ni pull-down assay.

(A) Trx-His-HIF-1$\alpha_{1\text{-}350}$ binds to Ni Sepharose beads, if GST-HIF-1$\beta_{1\text{-}459}$ binds to HIF-1$\alpha_{1\text{-}350}$ then after bead washing both proteins would be evident by SDS-PAGE analysis; *(B)* If HIF-1$\alpha_{1\text{-}350}$ and HIF-1$\beta_{1\text{-}459}$ do not interact, or the interaction is inhibited, after extensive bead washing only Trx-His-HIF-1$\alpha_{1\text{-}350}$ will be evident by SDS-PAGE analysis. Peptide depicted by red circles in *(B)*.

An initial experiment to show dimerisation of the two proteins was performed. The purified and desalted recombinant proteins (30 μg) were added sequentially to clean Ni Sepharose beads, Trx-His-HIF-1$\alpha_{1\text{-}350}$ and GST-HIF-1$\beta_{1\text{-}459}$ were added to beads in tandem. After 1 hour and 30 minutes the beads were washed and then analysed by SDS-PAGE. On analysis, it was apparent that GST-HIF-1β_1-459 was only evident in the presence of Trx-His-HIF-1α1-350, suggesting that the two proteins interact (Figure 54). This also indicated that GST-HIF-1β1-459 does not non-specifically bind to the Ni beads. This experiment was repeated and the same results were observed.

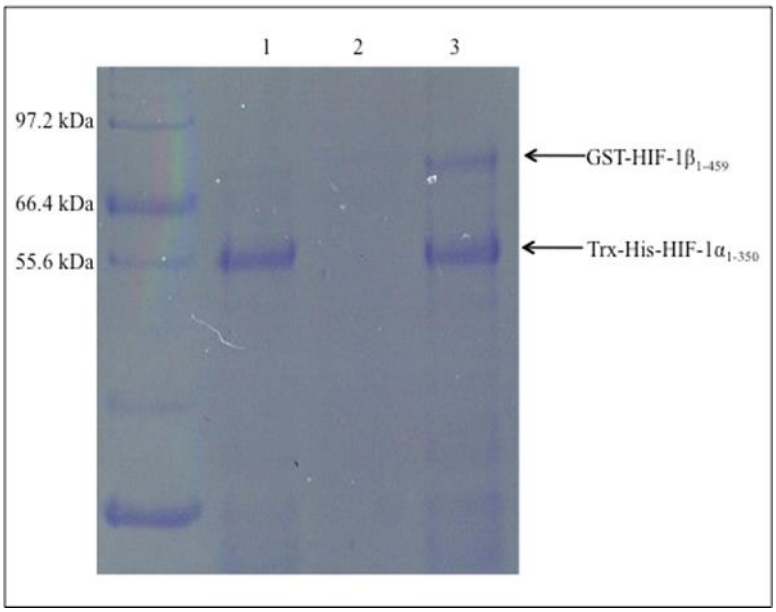


Figure 54 SDS-PAGE gel of Ni pull-down experiment.

(Lane 1) Trx-His-HIF-1$\alpha_{1\text{-}350}$ alone; *(Lane 2)* GST-HIF-1$\beta_{1\text{-}459}$ alone; *(Lane 3)* Trx-His-HIF-1$\alpha_{1\text{-}350}$ and GST-HIF-1$\beta_{1\text{-}459}$.

A GST pull-down was also attempted but non-specific binding of Trx-His-HIF-1α to GST made this assay configuration unfeasible.

2.4.4 Inhibition of dimerisation of independently expressed Trx-His-HIF-1$\alpha_{1\text{-}350}$ and GST-HIF-1$\beta_{1\text{-}459}$

Ni pull-down assays were carried out with HIF-1 dimerisation inhibitors, ACF and *cyclo*-CLLFVY.[118] ACF (1 μM-1 mM) was added to solutions of Trx-His-HIF-1$\alpha_{1\text{-}350}$ and GST-HIF-1$\beta_{1\text{-}459}$ (approximately 30 μg each). The protein solutions were incubated overnight at 4°C and then added to clean Ni Sepharose beads. The beads were washed three times and their eluents analysed by SDS-PAGE (Figure 55). The gel did not show decreasing levels of GST-HIF-1$\beta_{1\text{-}459}$ with increasing ACF concentration as expected, implying that in the assay ACF did not break apart the dimer.

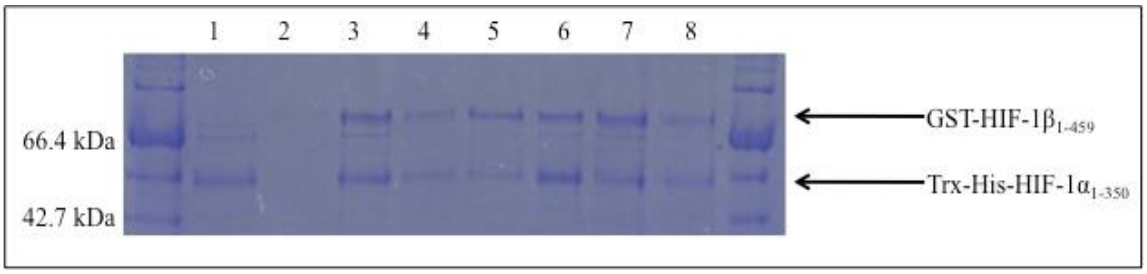


Figure 55 Ni pull-down experiment with independently expressed Trx-His-HIF-1α_{1-350} and GST-HIF-1β_{1-459} and ACF.

(Lane 1) Trx-His-HIF-1α_{1-350} alone; *(Lane 2)* GST-HIF-1β_{1-459} alone; *(Lane 3)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459}, DMSO (vehicle control); *(Lane 4)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459}, 0.1 μM ACF; *(Lane 5)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459}, 1 μM ACF; *(Lane 6)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459}, 10 μM ACF; *(Lane 7)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459}, 100 μM ACF; *(Lane 8)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459}, 1 mM ACF.

The experiment was repeated with *cyclo*-C(spy)LLFVY, and the gel appeared very similar to the above gel (Figure 55) showing no trend for decreasing GST-HIF-1β_{1-459} with increasing *cyclo*-C(spy)LLFVY. The intensity of the bands was quantified with Image J (NIH), with no significant dose-dependent inhibition evident.

Our RTHS data strongly indicated that P1 functions by disruption of HIF-1 dimerisation, these results suggest that assay optimisation may be required. One factor that may have been causing the assay failure was protein concentration. If the protein concentration used was such that the equilibrium strongly favoured dimerisation, then very high concentrations of inhibitor may be required to inhibit the interaction. Whereas, at a lower protein concentration dimerisation of Trx-His-HIF-1α_{1-350} and GST-HIF-β_{1-459} would be less favourable and the effect of the inhibitor may be observed more clearly. Consequently, a lower concentration of each recombinant protein (4 μg) was used in the following pull-down assays. At this concentration Western blot analysis was required to visualise the result, this was preferential to the previously used SDS-PAGE because the specificity of the antibodies would confirm the identity of the bands. The proteins were purified and their concentrations calculated using a Bradford reagent assay. In order to establish if the addition order of proteins and inhibitor is important, *cyclo*-C(spy)LLFVY (no tag) (5 μM-500 μM) or vehicle (DMSO) was added to purified His-HIF-1α_{1-350} (4 μg). The solutions were incubated for 1 hour at 4°C and then GST-HIF-1β_{1-459} (4 μg) was added. The

solutions were incubated for a further hour; the reverse experiment was also carried out by binding GST-HIF-1β_{1-459} to the beads first. The solutions were then added to Ni Sepharose beads, the beads were washed several times and then analysed by Western blot. The primary antibodies used were mouse-anti-GST and mouse-anti-His. A secondary anti-mouse red fluorescent antibody was used to allow visualisation. Unfortunately, the Western blot was of poor quality and this could be due to many reasons such as antibody concentration, blocking or incubation times. The clearest Western blot was where Trx-His-HIF-1α_{1-350} was added first, however, it did not show decreasing quantities of GST-HIF-1β_{1-459} levels with increasing *cyclo*-C(spy)LLFVY, as determined when the bands were quantified using Image J (NIH). This confirmed that there was no correlation between *cyclo*-C(spy)LLFVY concentration and levels of GST-HIF-1β_{1-459} (Figure 56).

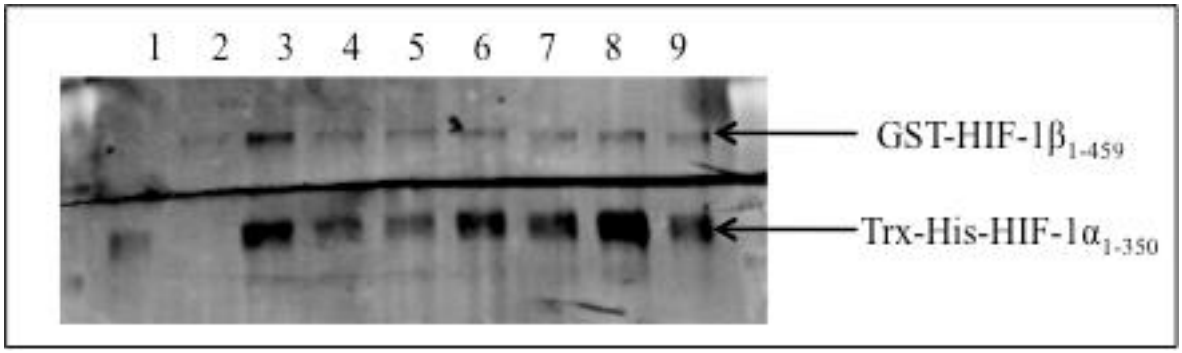


Figure 56 Western blot of Ni pull-down of Trx-His-HIF-1α_{1-350} and GST-HIF-1β_{1-459} with inhibitor *cyclo*-C(spy)LLFVY.

(Lane 1) Trx-His-HIF-1α_{1-350} alone; *(Lane 2)* GST-HIF-1β_{1-459}; *(Lane 3)* Trx-His-HIF-1α_{1-350}/GST-HIF-1β_{1-459}; *(Lane 4)* Trx-His-HIF-1α_{1-350}/GST-HIF-1$\beta_{1-459,}$ DMSO (vehicle control); *(Lane 5)* Trx-His-HIF-1α_{1-350}/GST-HIF-1β_{1-459}, 5 μM *cyclo*-C(spy)LLFVY; (Lane 6) Trx-His-HIF-1α_{1-350}/GST-HIF-1$\beta_{1-459,}$ 50 μM *cyclo*-C(spy)LLFVY; *(Lane 7)* Trx-His-HIF-1α_{1-350}/GST-HIF-1$\beta_{1-459,}$ 100 μM *cyclo*-C(spy)LLFVY; *(Lane 8)* Trx-His-HIF-1α_{1-350}/GST-HIF-1$\beta_{1-459,}$ 250 μM *cyclo*-C(spy)LLFVY; *(Lane 9)* Trx-His-HIF-1α_{1-350}/GST-HIF-1β_{1-459}, 500 μM *cyclo*-C(spy)LLFVY.

This may have been due to the insolubility of the cyclic peptide, but a small test experiment with P1 (Tat-*cyclo*-CLLFVY) also showed no inhibition. Pull-down assays are not very sensitive and also difficult to quantify because a relatively large amount of protein is required. After several attempts it was decided to halt the investigation into inhibition of dimerisation by pull-down. The pull-down experiment had tentatively indicated that the proteins were functionally able to dimerise.

To further confirm the functionality of the recombinant proteins, to enable their use in further biophysical assays and to attempt to construct a robust assay an EMSA was investigated. An EMSA requires low amounts of protein, in comparison to previously described pull-down assays and hence may enable inhibition of dimerisation to be visualised.

2.5 Confirmation of the DNA-binding ability of Trx-His-HIF-1$\alpha_{1\text{-}350}$ and GST-HIF-1$\beta_{1\text{-}459}$ by EMSA

Having shown that the recombinant Trx-His-HIF-1$\alpha_{1\text{-}350}$ and GST-HIF-1$\beta_{1\text{-}459}$ dimerise using a pull-down assay, the ability of this dimer to bind to DNA was investigated. In addition, this confirms the functionality of the proteins as dimerisation is a critical precursor to DNA-binding. One technique of achieving this is an electrophoretic shift assay (EMSA). An EMSA requires a low concentration of the proteins as detection is sensitive, in comparison to the previous pull-down experiments.

Specifically, the Trx-His-HIF-1$\alpha_{1\text{-}350}$ and GST-HIF-1$\beta_{1\text{-}459}$ recombinant proteins were used in an EMSA assay for two reasons: firstly to show that the HIF-1$\alpha_{1\text{-}350}$ and HIF-1$\beta_{1\text{-}459}$ produced are fully functional and can bind to DNA. Several sources indicate that independently expressed HIF-1α and HIF-1β have a compromised DNA-binding ability.[60, 149] Consequently, if this is the case the proteins may not be correctly folded and hence are unsuitable for use within the *in vitro* assays. Secondly, P1 could be added to the proteins within an EMSA, and inhibition of dimerisation and DNA-binding could be observed. The technique could then be used to compare the alanine analogues.

An EMSA is generally used to detect protein binding to DNA. The principle of the assay is that when protein is bound to DNA, it will have less mobility through a polyacrylamide gel matrix than unbound DNA and will therefore migrate more slowly (Figure 57).[154]

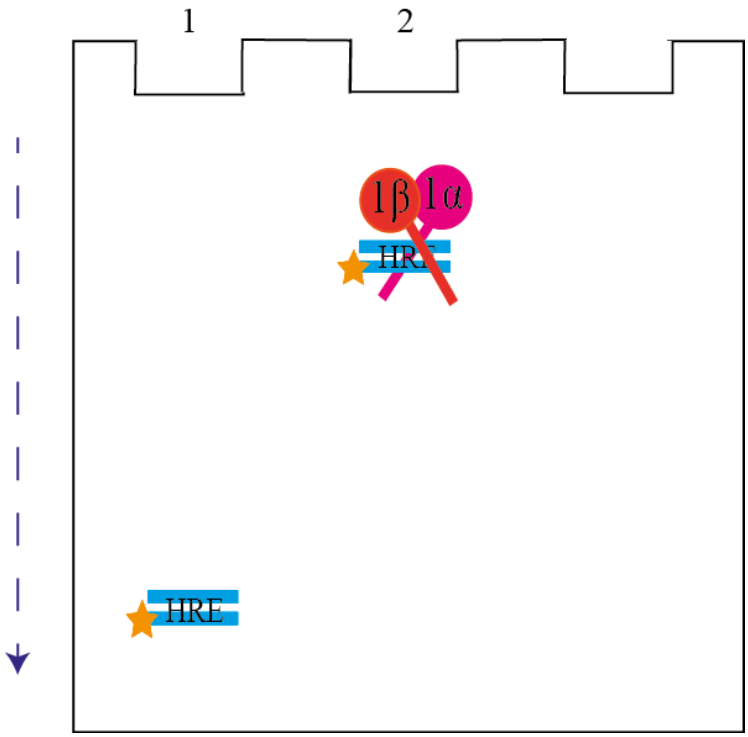


Figure 57 Principle of an EMSA assay.

Unbound DNA probe moves rapidly through an acrylamide gel (*Lane 1*) whereas, complexation with protein retards its mobility (*Lane 2*).

HIF-1 binds to specific DNA sequences (HRE's). The DNA probe chosen for the EMSA here was a 30 bp section of the erythropoietin (EPO) promoter (5'-GGCTGGGCCTACGTGCTGTCTCACACAGCC -3'), that incorporates a HRE (underlined) that is bound by HIF-1.[12] Typically, the DNA probe is firstly radiolabelled with [^{32}P] at the 5' end by phosphorylation using T4 polynucleotide kinase (PNK) with [^{32}P]-γATP. Only the forward strand of the HRE probe was 5' phosphorylated with [^{32}P]γATP by PNK at 37°C within one hour. The probe was then gel purified and detected by exposure of the gel to X-ray film. The DNA was then extracted from the gel in TE buffer overnight and ethanol precipitated; the DNA pellet was redissolved in TE buffer to a concentration of 10 Ci/s/µl. The labelled strand was annealed to its unlabelled reverse partner to leave the double stranded [^{32}P] HRE probe. Presence of the doubled stranded species was confirmed by running the duplex against the single strand on a gel and observing the different mobility (Figure 58).

All radiation work was carried out in designated radiation areas behind appropriate screens, radiation training, local rules and risk assessements were followed at all times. All used and disposed of radioactive material were recorded appropriately.

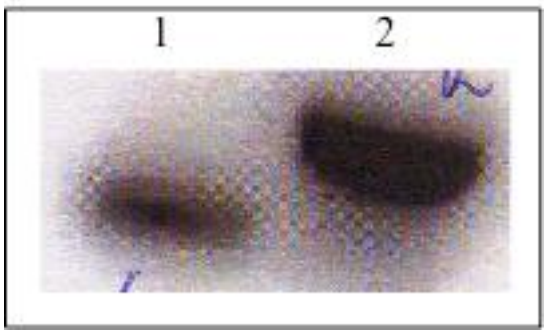


Figure 58 Agarose gel analysis of single stranded and duplex ^{32}P-HRE probe.
(Lane 1) Single stranded ^{32}P-HRE probe; *(Lane 2)* Double stranded ^{32}P-HRE probe.

As measurement of probe concentration was difficult due to its radioactivity, the concentration of the double stranded probe was estimated to be 0.3 μM (this was based on initial starting concentration and estimated losses).

2.5.1 Verification of Trx-His-HIF-1$\alpha_{1\text{-}350}$ and GST-HIF-1$\beta_{1\text{-}459}$ functionality

Firstly, an initial EMSA needed to be run to try and establish if the binding of Trx-His-HIF-1$\alpha_{1\text{-}350}$ and GST-HIF-1$\beta_{1\text{-}459}$ to the DNA probe could be visualised. The proteins were expressed, and purified and buffer exchanged into EMSA buffer (tris-based buffer).[149] EMSA samples (10 μl) contained poly(dI-dC) (a non-specific binding competitor), bovine serum albumin (BSA, to prevent non specific binding to DNA) and the [^{32}P] HRE probe.

On each EMSA gel control Lanes were present containing:

1.) [^{32}P] HRE probe only

2.) Trx-His-HIF-1$\alpha_{1\text{-}350}$ only

3.) GST-HIF-1$\beta_{1\text{-}459}$ only

These were run to establish where just the probe and monomeric proteins migrated to on a gel. Mixtures of Trx-His-HIF-1α_{1-350} and GST-HIF-1β_{1-459} and the DNA probe were also run, with increasing protein DNA ratios (1:3 to 1:50). The protein samples were incubated for 30 minutes prior to the addition of labelled probe (Figure 59).

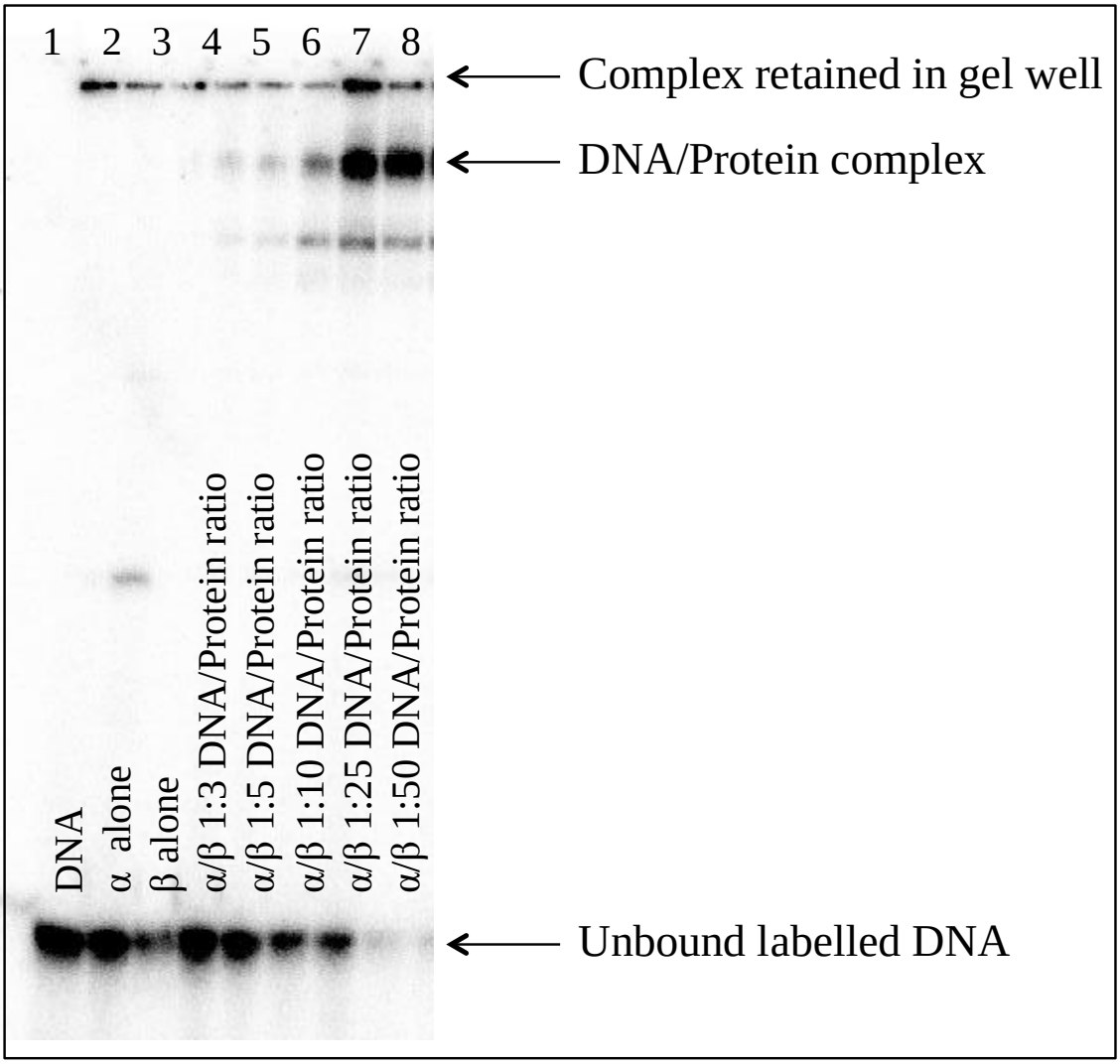


Figure 59 EMSA 1 (preliminary EMSA): Independently expressed Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex binding to [^{32}P] HRE probe.

(*Lane 1)* DNA probe alone; (*Lane 2)* Trx-His-HIF-1α_{1-350} alone; (*Lane 3)* GST-HIF-1β_{1-459} alone; (*Lane 4)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex 1:3 DNA/protein ratio; (*Lane 5)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex 1:5 DNA/protein ratio; (*Lane 6)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex 1:10 DNA/protein ratio; (*Lane 7)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex 1:25 DNA/protein ratio; (*Lane 8)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex 1:50 DNA/protein ratio.

This initial EMSA (EMSA 1) confirmed that the monomeric proteins would not bind to the HRE probe, however, the Trx-His-HIF-1$\alpha_{1\text{-}350}$/GST-HIF-1$\beta_{1\text{-}459}$ complex did appear to bind. The ratio of DNA/Protein binding did not appear to be 1:1; the unbound probe band (bottom of the gel) disappeared at a 1:25 ratio. This may have been due to the binding affinity of the HIF-1 complex or a proportion of inactive protein. The gel also showed a certain amount of DNA remaining in the gel wells across all samples and also a secondary band, which was unexplained. Further optimisation eliminated these bands, for example, the initial EMSA used loading buffer containing EDTA, using a buffer lacking the chelator eliminated this second band, which may have been due to EDTA reacting with the protein. Running the gel at a slower voltage, i.e. run at approximately 100 V for the first 30 minutes and then sped up to approximately 200 V, minimised the amount of DNA retarded in the gel well. To confirm HIF-1 was binding to the HRE probe a competition experiment was carried out. An unlabelled probe was titrated into the protein/labelled probe (25:1) mixtures (EMSA 2, Figure 60). On addition of the unlabelled HRE probe (EMSA 2, Figure 60), the DNA/protein complex band decreased to nothing and at the same time the free DNA probe band increased. This indicated that the HIF-1 complex was indeed binding to the labelled HRE probe and that there was a dynamic equilibrium.

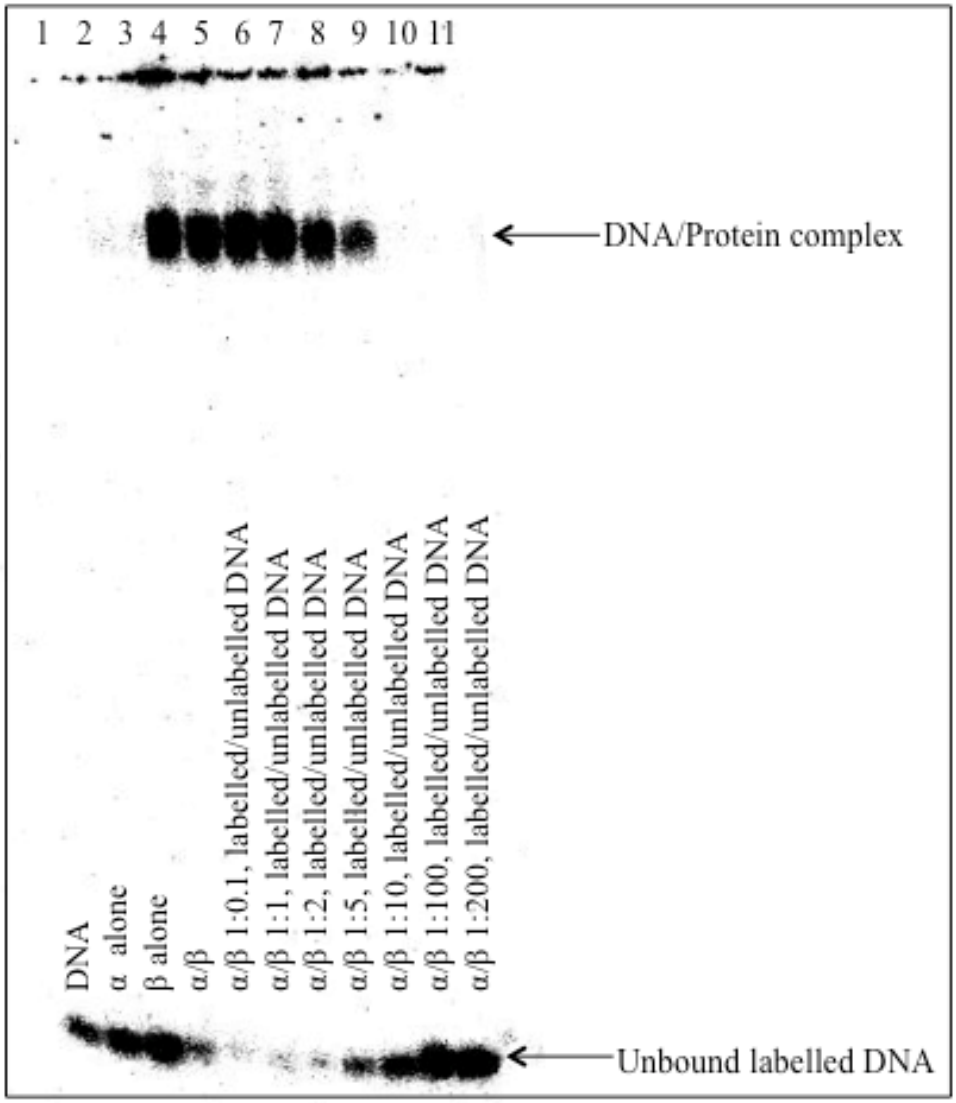


Figure 60 EMSA 2: Independently expressed Trx-His-HIF-1$\alpha_{1\text{-}350}$/ GST-HIF-1$\beta_{1\text{-}459}$ complex binding to [^{32}P] HRE probe in competition to unlabelled probe.

1:25, labelled probe/ protein ratio. (*Lane 1)* DNA probe alone; (*Lane 2)* Trx-His-HIF-1$\alpha_{1\text{-}350}$ alone; (*Lane 3)* GST-HIF-1$\beta_{1\text{-}459}$ alone; (*Lane 4)* Trx-His-HIF-1$\alpha_{1\text{-}350}$/ GST-HIF-1$\beta_{1\text{-}459}$; (*Lane 5)* Trx-His-HIF-1$\alpha_{1\text{-}350}$/ GST-HIF-1$\beta_{1\text{-}459,}$ 1:0.1, labelled/unlabelled DNA probe; (*Lane 6)* Trx-His-HIF-1$\alpha_{1\text{-}350}$/ GST-HIF-1$\beta_{1\text{-}459,}$ 1:1, labelled/unlabelled DNA probe; (*Lane 7)* Trx-His-HIF-1$\alpha_{1\text{-}350}$/ GST-HIF-1$\beta_{1\text{-}459,}$ 1:2, labelled/unlabelled DNA probe; (*Lane 8)* Trx-His-HIF-1$\alpha_{1\text{-}350}$/ GST-HIF-1$\beta_{1\text{-}459,}$ 1:5, labelled/unlabelled DNA probe; (*Lane 9)* Trx-His-HIF-1$\alpha_{1\text{-}350}$/ GST-HIF-1$\beta_{1\text{-}459,}$ 1:10, labelled/unlabelled DNA probe; (*Lane 10)* Trx-His-HIF-1$\alpha_{1\text{-}350}$/ GST-HIF-1$\beta_{1\text{-}459,}$ 1:100, labelled/unlabelled DNA probe; (*Lane 11)* Trx-His-HIF-1$\alpha_{1\text{-}350}$/ GST-HIF-1$\beta_{1\text{-}459,}$ 1:200, labelled/unlabelled DNA probe.

As mentioned previously, the complete binding of the labelled probe did not occur when the DNA/protein ratio was 1:1, once the assay was better optimised the concentration of Trx-His-HIF-1α_{1-350}/GST-HIF-1β_{1-459} was increased from a 1:0.1 to 1:55 ratio (EMSA 3, Figure 61). Full probe binding was observed at a ratio of 1:20, quantification of the intensity of the bands allowed a binding curve to be plotted and a K_d of 0.37 μM was found for the HIF-1 complex to the labelled DNA probe to be obtained (Figure 62). This experiment was repeated for consistency and similar results were observed. Although this dissociation constant has been calculated it should taken wth caution, as it does not reflect additional equilibria that may be occurring between monomeric proteins and DNA, as well as homodimerisation. Aside from this, the concentrations of functional protein is an estimate as some may be aggregated or truncated, and the concentration of DNA is an estimate due to the practical limitations of measuring radioactive DNA concentration.

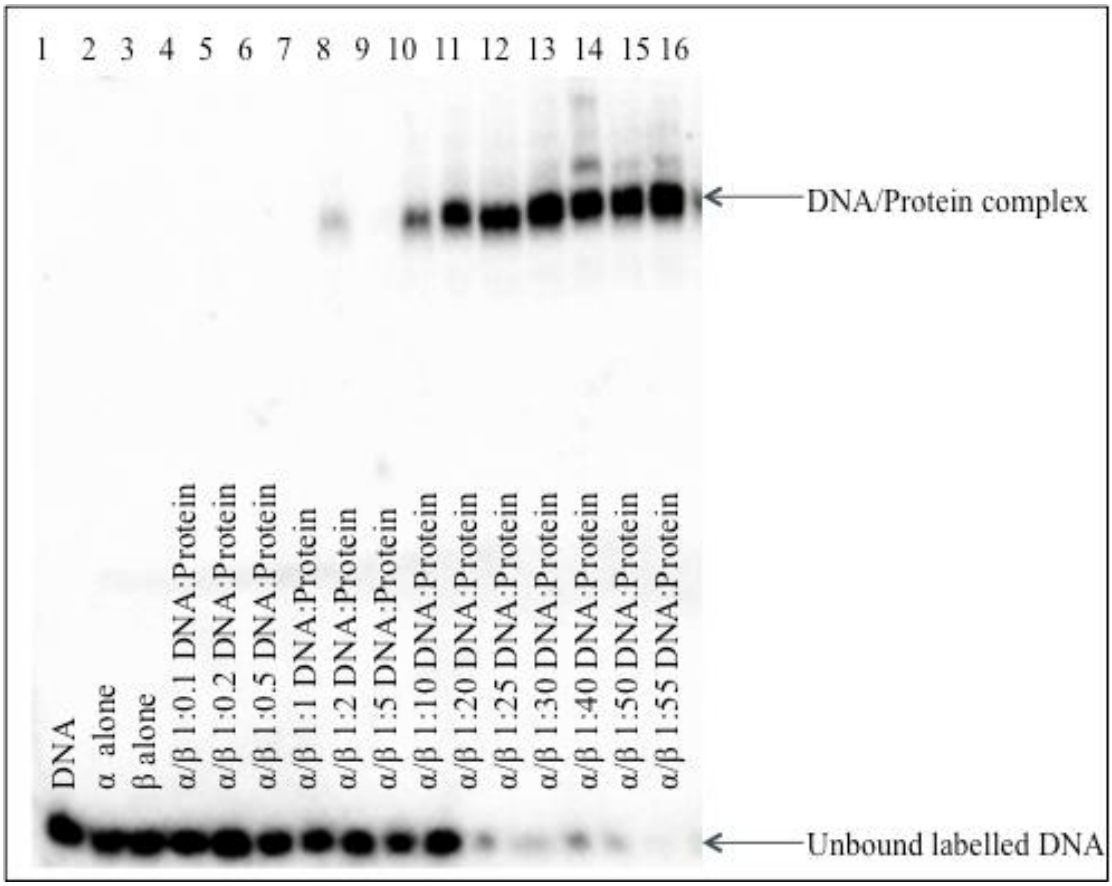


Figure 61 EMSA 3: Independently expressed Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex binding to [^{32}P] HRE probe.

(*Lane 1)* DNA probe alone; (*Lane 2)* Trx-His-HIF-1α_{1-350} alone; (*Lane 3)* GST-HIF-1β_{1-459} alone; (*Lane 4)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex 1:0.1 DNA/protein ratio; (*Lane 5)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex 1:0.2 DNA/protein ratio; (*Lane 6)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex 1:0.5 DNA/protein ratio; (*Lane 7)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex 1:1 DNA/protein ratio; (*Lane 8)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex 1:2 DNA/protein ratio; (*Lane 9)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex 1:5 DNA/protein ratio; *(Lane 10)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex 1:10 DNA/protein ratio; (*Lane 11)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex 1:20 DNA/protein ratio; (*Lane 12)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex 1:25 DNA/protein ratio; (*Lane 13)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex 1:30 DNA/protein ratio; (*Lane 14)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex 1:40 DNA/protein ratio; (*Lane 15)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex 1:50 DNA/protein ratio; (*Lane 16)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex 1:55 DNA/protein ratio.

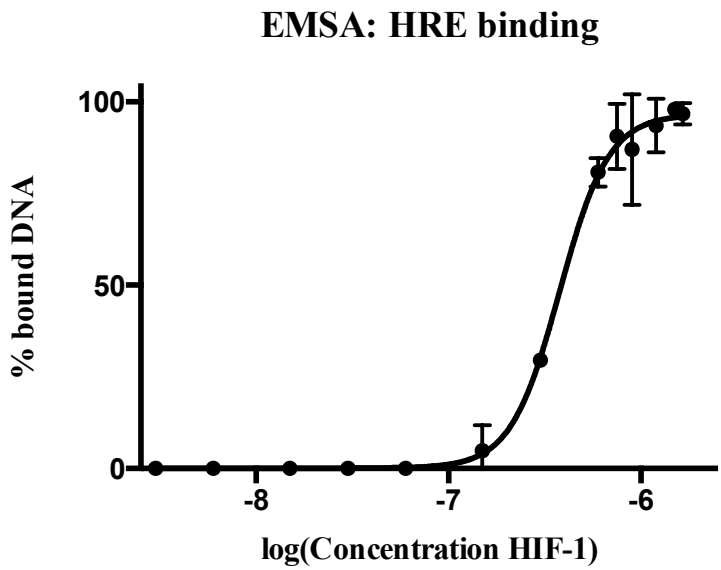


Figure 62 Binding curve for binding of independently expressed Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex [^{32}P] HRE probe.

K_d = 0.37 μM. Error bars represent standard deviation from two repeat experiments. Statistical analysis was carried out in Prism 6, with Sigmoidal 4PL fit.

Having confirmed the functionality of the proteins, these recombinant proteins were used to probe the action of P1 further (described later in section 2.5.2.2). As a control for P1 specificity HIF-2α was also investigated, Trx-His HIF-2α_{1-360} (from pET32a-Trx-His-HIF-2α_{1-360}, constructed by Dr K. Nordgren) was expressed and purified in the same manner as Trx-His-HIF-1α_{1-350} (Figure 63), section 2.4.1). An initial HIF-2 EMSA was run to show the functionality of the isoform (EMSA 4, Figure 64).

Figure 63 SDS-PAGE (12%) analysis of purified Trx-His-HIF-2α_{1-360} (59.0 kDa).

(*Lane 1)* purified and desalted sample of Trx-His-HIF-2α_{1-360}.

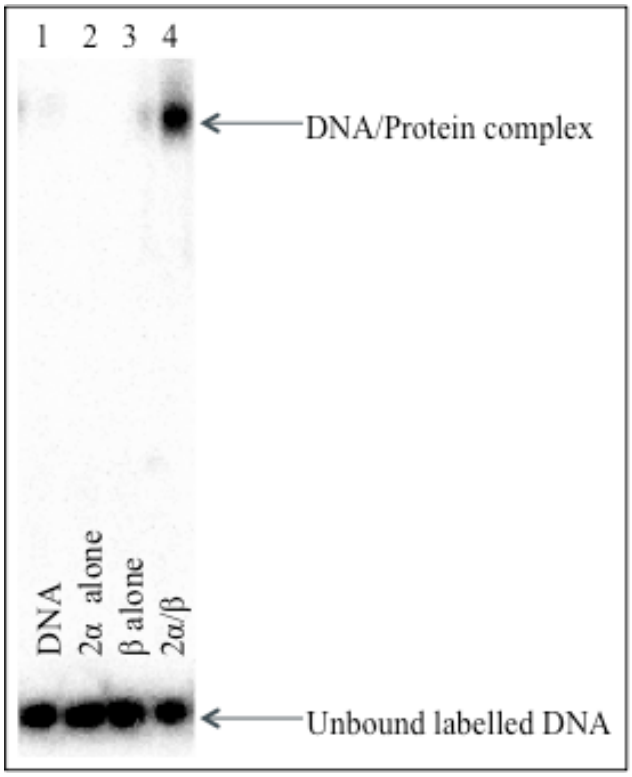


Figure 64 EMSA 4: Independently expressed Trx-His-HIF-2α_{1-360}/ GST-HIF-1β_{1-459} complex binding to [^{32}P] HRE probe.

1:5, labelled probe/ protein ratio. (*Lane 1)* DNA probe alone; (*Lane 2)* Trx-His-HIF-2α_{1-360} alone; (*Lane 3)* GST-HIF-1β_{1-459} alone; (*Lane 4)* Trx-His-HIF-2α_{1-360}/ GST-HIF-1β_{1-459}.

The HIF-2 EMSA indicated Trx-His-HIF-2$\alpha_{1\text{-}360}$ was functional, as dimerisation and DNA-binding was evident at a 1:5 DNA/protein ratio (EMSA 4, Figure 64).

2.5.2 HIF-1 EMSA to test inhibitors

The EMSA was reproducible and used very little sample, so investigation into whether the technique could be used to test HIF-1 inhibitors was carried out.

2.5.2.1 Verification of EMSA with HIF-1 inhibitor ACF

To verify the assay, ACF was used as a control, ACF has been shown to inhibit HIF-1α/HIF-1β dimerisation.[118] The concentration of protein to use was considered as it may have a large effect on the success of the inhibitor to prevent complexation, as it will be necessary to disrupt the equilibrium between both dimerisation and DNA-binding. Initially, a 1:25, DNA/protein ratio was used as this was approximately where complete DNA-binding occurred (Figure 61). ACF was titrated into the Trx-His-HIF-1$\alpha_{1\text{-}350}$/GST-HIF-1$\beta_{1\text{-}459}$ mixture (0-500 μM), the vehicle was 10% DMSO. The addition order was Trx-His-HIF-1$\alpha_{1\text{-}350}$, ACF and then GST-HIF-1$\beta_{1\text{-}459}$ in succession and then the solutions were incubated for 30 minutes prior to the addition of the labelled probe (EMSA 5, Figure 64).

The addition of ACF prevented DNA-binding, presumably by inhibiting the dimerisation of HIF-1α and HIF-1β. This was seen in a dose-dependent manner, as the protein/DNA complex band decreased and the unbound DNA band reappeared with the addition of the inhibitor (EMSA 5, Figure 65). On quantification of the bands an IC_{50} = 40.04 μM was calculated. Semenza *et al.* (2009) observed an IC_{50} of 1 μM and complete inhibition at 5 μM by both an *in vitro* pull down and HRE reporter assay.[118] The discrepency between these could be due to assay conditions, however, both are in agreement that ACF inhibits HIF-1 in a dose-dependent manner leading to complete inhibition of its activity.

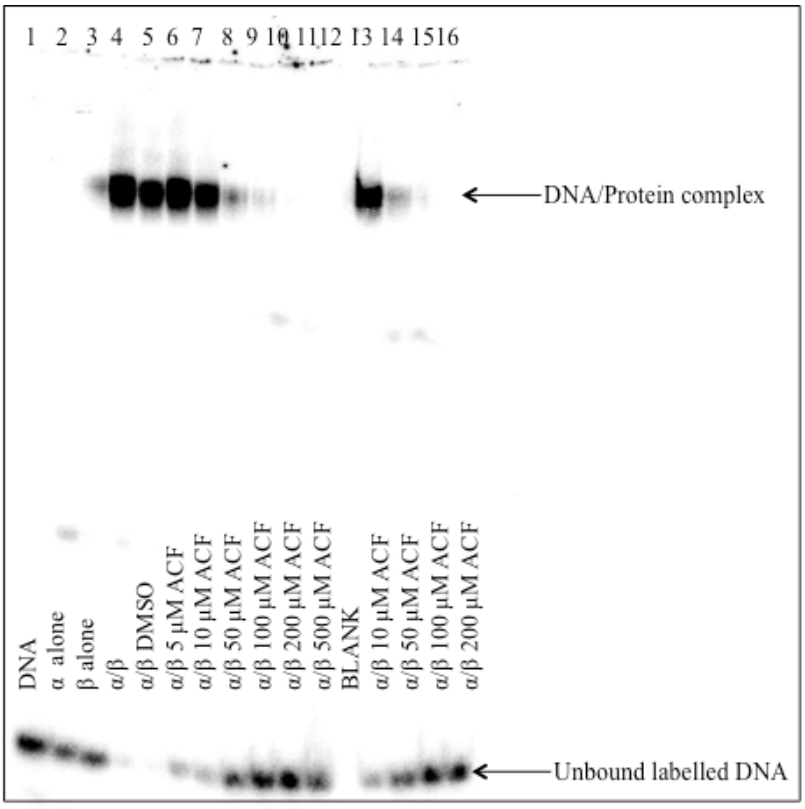


Figure 65 EMSA 5: Independently expressed Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex binding to [^{32}P] HRE probe with the addition of HIF-1 dimerisation inhibitor ACF.

A 1:25 DNA/protein ratio was used. (*Lane 1)* DNA probe alone; (*Lane 2)* Trx-His-HIF-1α_{1-350} alone; (*Lane 3)* GST-HIF-1β_{1-459} alone; (*Lane 4)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex; (*Lane 5)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex, 10% DMSO; (*Lane 6)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex, 5 μM ACF; (*Lane 7)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex, 10 μM ACF; (*Lane 8)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex, 50 μM ACF; (*Lane 9)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex, 100 μM ACF; (*Lane 10)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex, 200 μM ACF; (*Lane 11)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex, 500 μM ACF; (*Lane 12)* Blank; (*Lane 13)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex, 10 μM ACF; (*Lane 14)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex, 50 μM ACF; (*Lane 15)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex, 100 μM ACF; (*Lane 16)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex, 200 μM ACF. This result is representative of repeats.

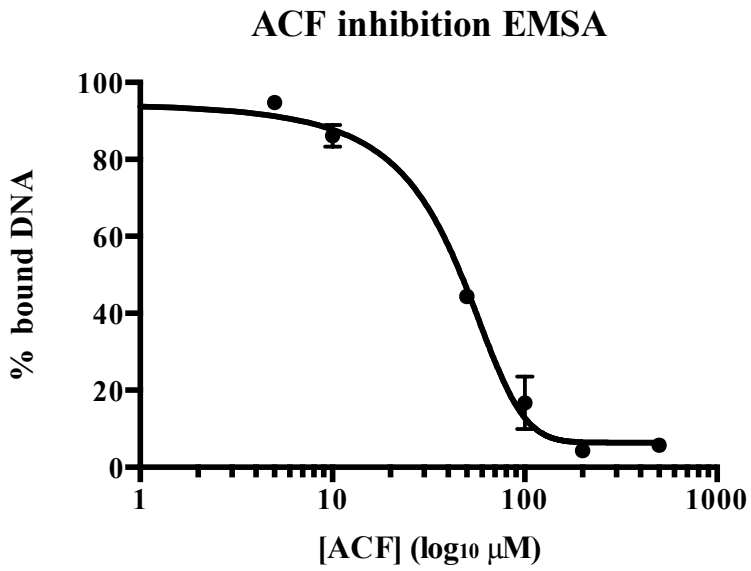


Figure 66 EMSA 5 binding curve, ACF inhibition of HIF-1/HRE binding.

ACF (0-500 μM) titrated into Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} mixture. Inhibition of HRE binding visualised by EMSA shift and band intensity calculated. Error bars represent two repeat experiments. ACF IC_{50} = 40.04 μM. Statistical analysis was carried out in Prism 6 with a Sigmoidal 4PL fit.

The ACF experiment was reproducible and indicated that this EMSA was a suitable assay in which to test HIF-1 dimerisation inhibitors. The protein/DNA ratio of 1:25 was suitable for the ACF assay; however, this may have been due to the low IC_{50} of ACF. Next the assay was used to test the activity of P1.

2.5.2.2 Investigation into using the EMSA to test the activity of P1

Cyclo-CLLFVY was found to be soluble in 10% DMSO, so this was used for the next experiments in place of the P1.The IC_{50} of *cyclo*-CLLFVY was expected to be lower than that of ACF (from previous in cell assays looking at the activity of the compound) so when testing the peptide in the EMSA the DNA/protein ratio was reduced to 1:10, where not all DNA is bound to attempt to be able to visualise the effect of the peptide.[1] Initially, the un-tagged cyclic peptide was used, which was spy deprotected using the reducing agent TCEP (Figure 67).

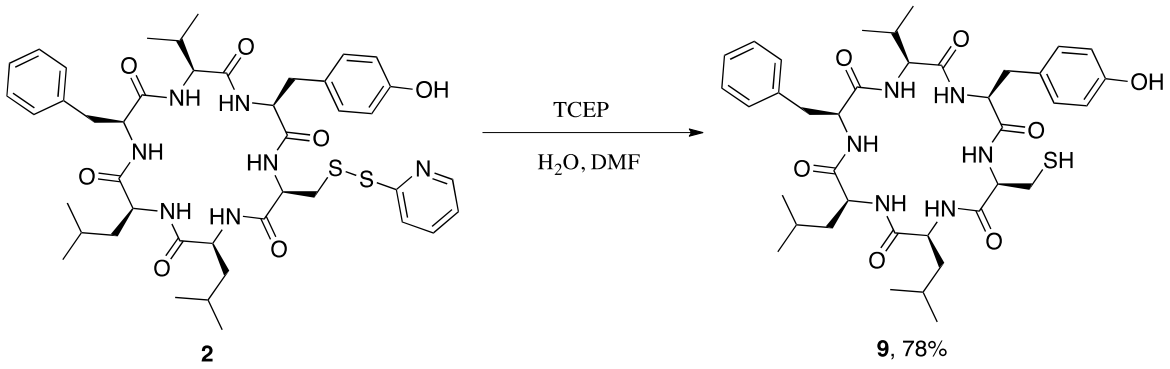


Figure 67 *Cyclo*-C(spy)LLFVY deprotection with TCEP to free the deprotected thiol.

As described for ACF, c*yclo*-CLLFVY was soluble in DMSO so this was used as the vehicle, it was added to the Trx-His-HIF-1$\alpha_{1\text{-}350}$/GST-HIF-1$\beta_{1\text{-}459}$ solutions (0-500 μM) (addition order Trx-His-HIF-1$\alpha_{1\text{-}350,}$ *cyclo*-CLLFVY, GST-HIF-1$\beta_{1\text{-}459}$ solutions) and incubated on ice for 30 minutes prior to the addition of the labelled probe. The final DMSO concentration was 10%, a control sample with just 10% DMSO was run to eliminate the effect of the vehicle. The gel was run as previously described after 1 hour incubation. EMSA 6 indicated a DNA/protein interaction. Unfortunately, it appeared that the addition of *cyclo*-CLLFVY did not affect DNA/protein binding as the size of the shift band (DNA/protein complex) did not decrease with increasing concentration of *cyclo*-CLLFVY, as would be expected if dimerisation was inhibited (EMSA 6, Figure 68).

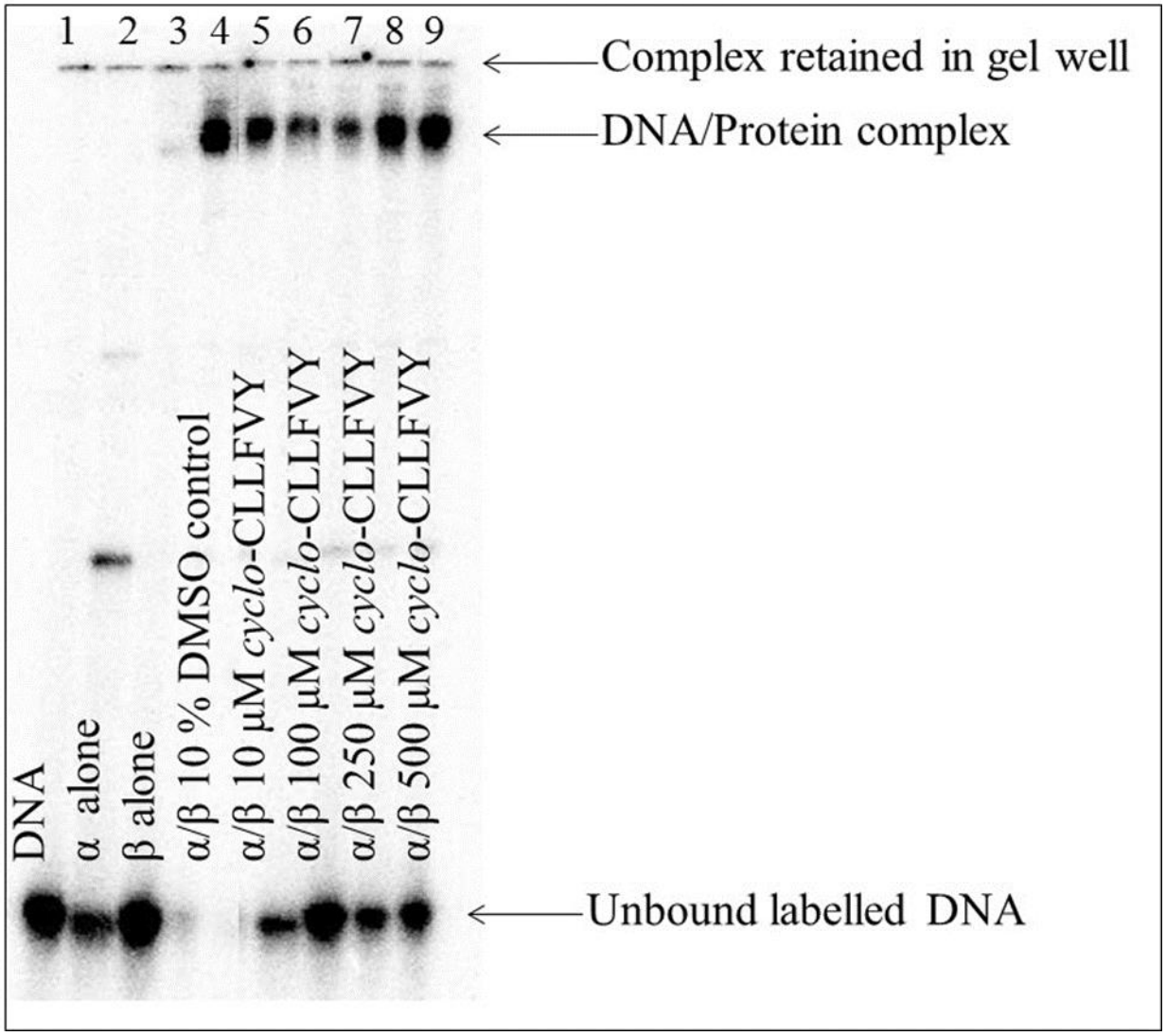


Figure 68 EMSA 6: Independently expressed Trx-His-HIF-1$\alpha_{1\text{-}350}$/ GST-HIF-1$\beta_{1\text{-}459}$ complex binding to [^{32}P] HRE probe with the addition of *cyclo*-CLLFVY.

A 1:10 DNA/protein ratio was used. (*Lane 1)* DNA probe alone; (*Lane 2)* Trx-His-HIF-1$\alpha_{1\text{-}350}$ alone; (*Lane 3)* GST-HIF-1$\beta_{1\text{-}459}$ alone; (*Lane 4)* Trx-His-HIF-1$\alpha_{1\text{-}350}$/ GST-HIF-1$\beta_{1\text{-}459}$ complex; (*Lane 5)* Trx-His-HIF-1$\alpha_{1\text{-}350}$/ GST-HIF-1$\beta_{1\text{-}459}$ complex, 10% DMSO; (*Lane 6)* Trx-His-HIF-1$\alpha_{1\text{-}350}$/ GST-HIF-1$\beta_{1\text{-}459}$ complex, 10 μM *cyclo*-CLLFVY; (*Lane 7)* Trx-His-HIF-1$\alpha_{1\text{-}350}$/ GST-HIF-1$\beta_{1\text{-}459}$ complex, 100 μM *cyclo*-CLLFVY; (*Lane 8)* Trx-His-HIF-1$\alpha_{1\text{-}350}$/ GST-HIF-1$\beta_{1\text{-}459}$ complex, 250 μM *cyclo*-CLLFVY; (*Lane 9)* Trx-His-HIF-1$\alpha_{1\text{-}350}$/ GST-HIF-1$\beta_{1\text{-}459}$ complex, 500 μM *cyclo*-CLLFVY.

Therefore, additional EMSA experiments with *cyclo*-CLLFVY were run. Several experimental parameters were changed, including preincubation of both Trx-His-HIF-1$\alpha_{1\text{-}350}$ and GST-HIF-1$\beta_{1\text{-}459}$ with *cyclo*-CLLFVY and then addition of their complementary

binding partner for a short incubation time, followed by addition of the DNA and incubation for only five minutes. This experiment still did not show a change in the band size of HIF-1/DNA complex upon addition of *cyclo*-CLLFVY.

It was speculated that the thiol of the cysteine residue in *cyclo*-CLLFVY was causing an adverse effect, and inhibitor dimerisation could be occurring. Therefore, *cyclo*-SLLFVY was synthesised by Fmoc solid-phase peptide synthesis as described previously in section 2.2.1, (Figure 69).

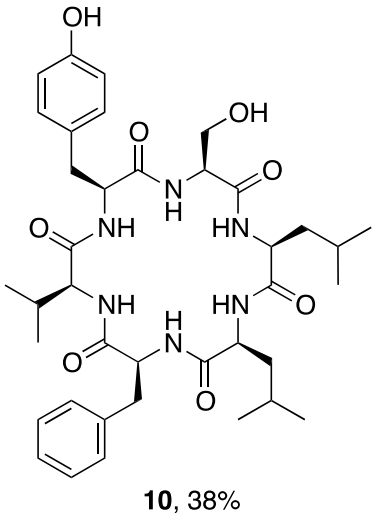


Figure 69 Structure of synthesised *cyclo*-SLLFVY.
Synthesised by solid-phase peptide synthesis.

Trx-His-HIF-1$\alpha_{1\text{-}350}$ and GST-HIF-$1_{1\text{-}459}$ were separately incubated with SLLFVY (10-500 μM) overnight before addition of their complementary subunit and further incubation for 5 minutes at room temperature before addition of labelled DNA. No experiment showed SLLFVY to inhibit HIF-1 dimerisation and DNA-binding.

One hypothesis for the inactivity of *cyclo*-CLLFVY in the EMSA was its insolubility, despite the binding mixture being in 10% DMSO, once the running buffer was added or the gel was run and solutions cooled, the peptide was likely to have precipitated and therefore allowed protein dimerisation to resume. The linear peptide (LFVYC(spy)L) may be more soluble so this was also tested (EMSA 7, Figure 70). This time a slight decrease in the band size was observed with the highest concentration of peptide (500 μM and 1 mM)

but as these were with such high concentrations of peptide compared to the protein and DNA, no conclusions could be drawn (EMSA 7, Figure 70).

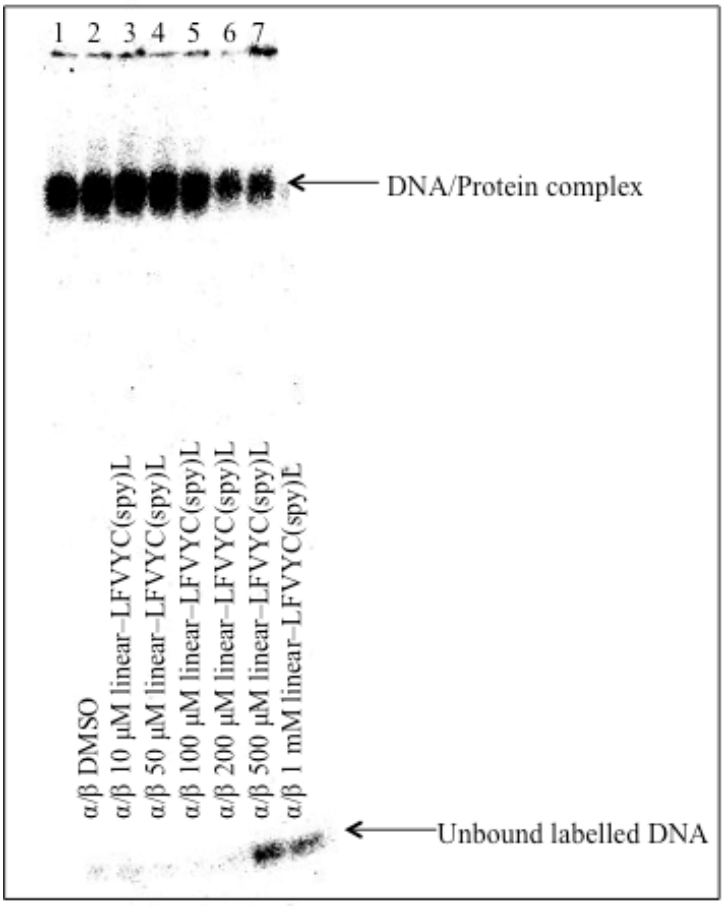


Figure 70 EMSA 7: Independently expressed Trx-His-HIF-1$\alpha_{1\text{-}350}$/ GST-HIF-1$\beta_{1\text{-}459}$ complex binding to [^{32}P] HRE probe with the addition of LFVYC(spy)L.

A 1:25 DNA/protein ratio was used. (*Lane 1)* Trx-His-HIF-1$\alpha_{1\text{-}350}$/ GST-HIF-1$\beta_{1\text{-}459}$ complex, 10% DMSO; (*Lane 2)* Trx-His-HIF-1$\alpha_{1\text{-}350}$/ GST-HIF-1$\beta_{1\text{-}459}$ complex, 10 μM LFVYC(spy)L; (*Lane 3)* Trx-His-HIF-1$\alpha_{1\text{-}350}$/ GST-HIF-1$\beta_{1\text{-}459}$ complex, 50 μM LFVYC(spy)L; (*Lane 4)* Trx-His-HIF-1$\alpha_{1\text{-}350}$/ GST-HIF-1$\beta_{1\text{-}459}$ complex, 100 μM LFVYC(spy)L; (*Lane 5)* Trx-His-HIF-1$\alpha_{1\text{-}350}$/ GST-HIF-1$\beta_{1\text{-}459}$ complex, 200 μM LFVYC(spy)L; (*Lane 6)* Trx-His-HIF-1$\alpha_{1\text{-}350}$/ GST-HIF-1$\beta_{1\text{-}459}$ complex, 500 μM LFVYC(spy)L; (*Lane 7)* Trx-His-HIF-1$\alpha_{1\text{-}350}$/ GST-HIF-1$\beta_{1\text{-}459}$ complex, 1 mM LFVYC(spy)L.

As previously mentioned, there were concerns that the reason for the lack of inhibitor activity seen in these previous assays was due to poor solubility of *cyclo*-CLLFVY, so it was decided to try the experiment using the P1 (compound with Tat-tag) This would ensue

complete solubility in the buffer and while running the gel regardless of temperature and also eliminated the need to use DMSO.

To verify experimental conditions P1 (10-500 μM) was run in parallel to ACF (10-500 μM, as an experimental control). Also two different incubation times for the proteins and inhibitors were used (30 minutes and overnight) prior to the addition of the labelled probe. The samples were run on the same gel so that a comparison could be made (EMSA 8, Figure 71). Here, P1 appeared to not migrate from the well, with the positively charged Tat-tag hindering migration. This effect was observed regardless of the incubation time (EMSA 8, Figure 71). However, as previously observed ACF was able to disrupt dimerisation and DNA-binding (EMSA 8, Figure 71).

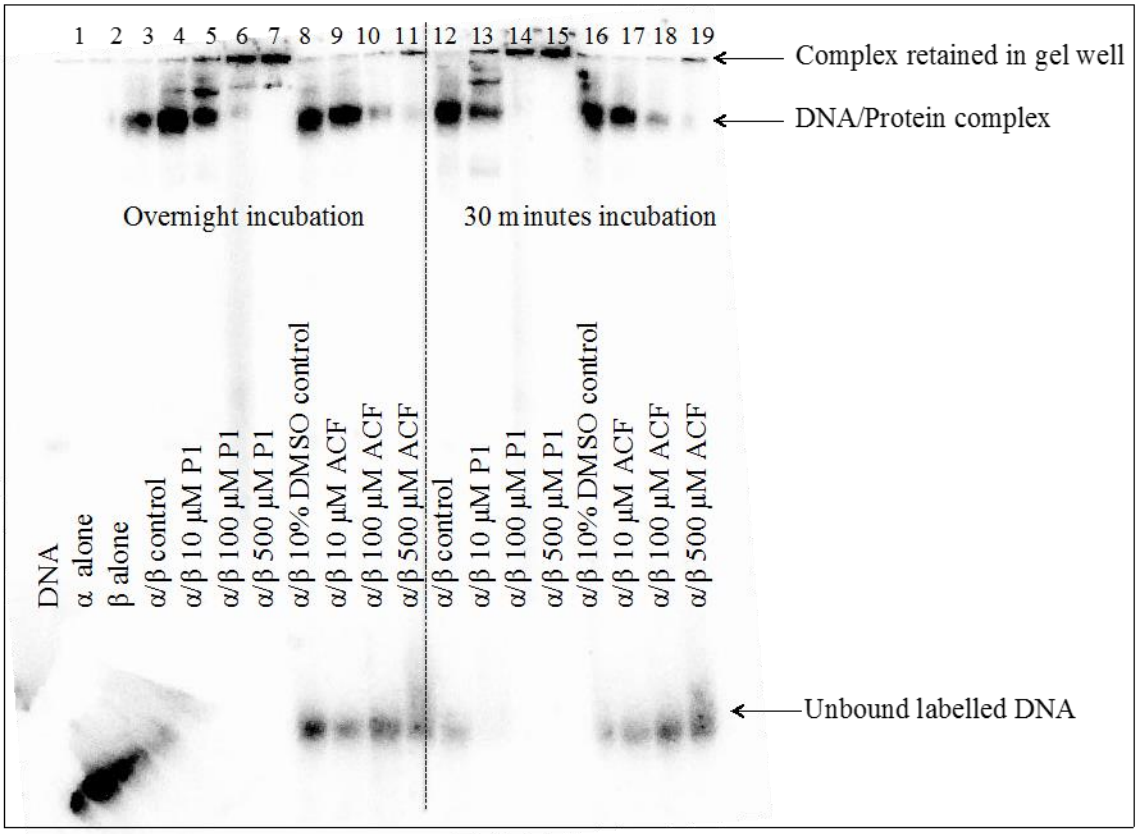


Figure 71 EMSA 8: Independently expressed Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex binding to [^{32}P] HRE probe with the addition of P1 and ACF.

A 1:25 DNA/protein ratio was used. (*Lane 1)* DNA probe alone; (*Lane 2)* Trx-His-HIF-1α_{1-350} alone; (*Lane 3)* GST-HIF-1β_{1-459} alone; (*Lane 4)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex; (*Lane 5)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex, 10 μM P1; (*Lane 6)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex, 100 μM P1; (*Lane 7)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex, 500 μM P1; (*Lane 8)*; Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex, 10% DMSO; (*Lane 9)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex, 10 μM ACF; *(Lane 10)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex, 100 μM ACF; *(Lane 11)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex, 500 μM ACF; (*Lane 12)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex; (*Lane 13)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex, 10 μM P1; (*Lane 14)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex, 100 μM P1; (*Lane 15)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex, 500 μM P1; (*Lane 16)*; Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex, 10% DMSO; (*Lane 17)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex, 10 μM ACF; *(Lane 18)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex, 100 μM ACF; *(Lane 19)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex, 500 μM ACF. The protein/inhibitor solutions for samples 1-11 were incubated overnight and samples 12-19 were incubated for 30 minutes.

This showed that the assay worked due to the ACF, but the Tat-tagged compound was not suitable for the assay. To investigate this further each individual component (protein or DNA) was incubated with the Tat-tag alone (EMSA 9, Figure 72). The EMSA showed that the phenomenon was due to the positive arginine rich Tat-tag interacting with the DNA probe because the DNA/Tat-tag samples alone were retarded in the gel wells (Figure 72). The EMSA was therefore not suitable to test the Tat-tagged alanine analogues.

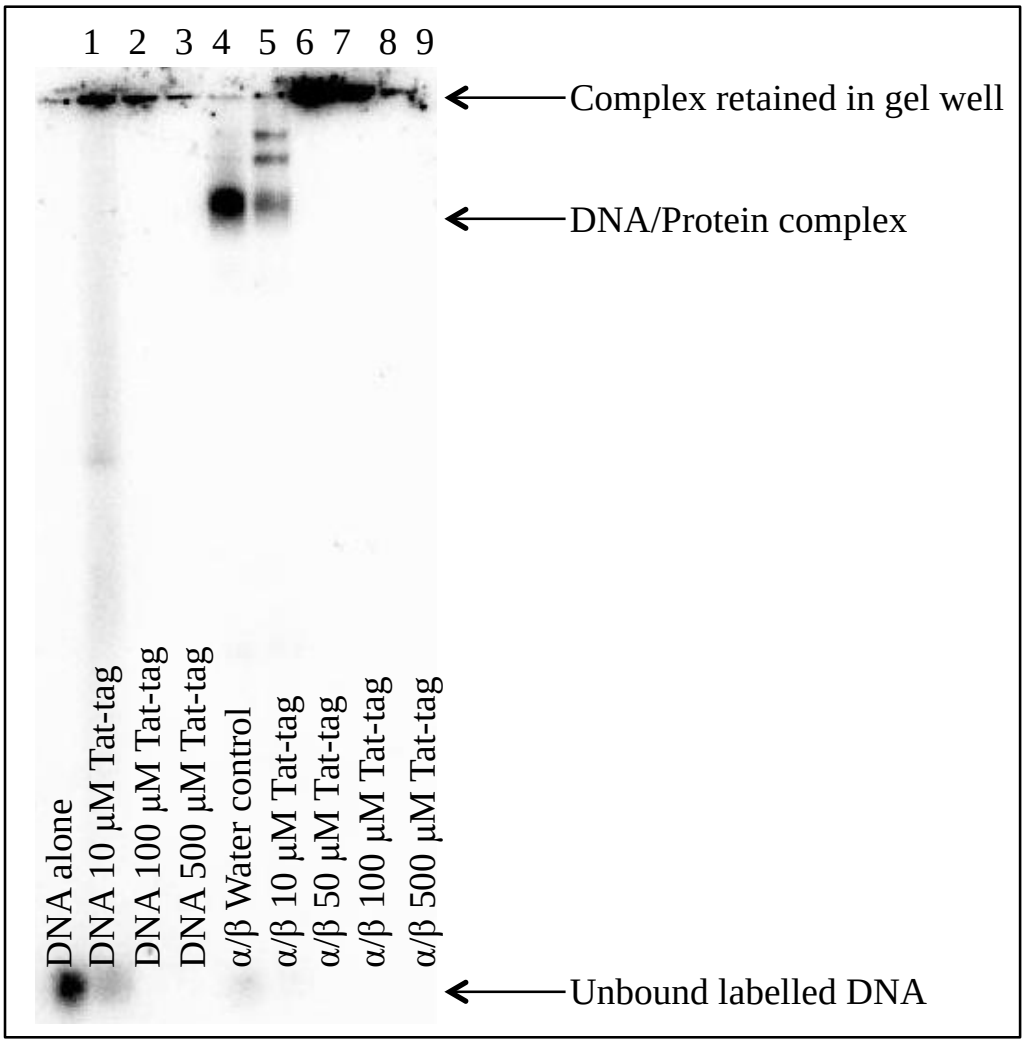


Figure 72 EMSA 9: Testing the effect of the Tat-tag on individual EMSA components. Independently expressed recombinant proteins. A 1:25 DNA/protein ratio was used. All samples contained DNA probe *Lane 1)* DNA probe alone; (*Lane 2)* 10 μM Tat-tag; (*Lane 3)* 100 μM Tat-tag; (*Lane 4)* 500 μM Tat-tag; (*Lane 5)* Trx-His-HIF-1$\alpha_{1\text{-}350}$/ GST-HIF-1$\beta_{1\text{-}459}$ complex, water control; *(Lane 6)* Trx-His-HIF-1$\alpha_{1\text{-}350}$/ GST-HIF-1$\beta_{1\text{-}459}$ complex, 10 μM Tat-tag; *(Lane 7)* Trx-His-HIF-1$\alpha_{1\text{-}350}$/ GST-HIF-1$\beta_{1\text{-}459}$ complex, 50 μM Tat-tag; *(Lane 8)* Trx-His-HIF-1$\alpha_{1\text{-}350}$/ GST-HIF-1$\beta_{1\text{-}459}$ complex, 100 μM Tat-tag; *(Lane 9)* Trx-His-HIF-1$\alpha_{1\text{-}350}$/ GST-HIF-1$\beta_{1\text{-}459}$ complex, 500 μM Tat-tag.

Due to the effect observed with the Tat-tag and the solubility issues with *cyclo*-CLLFVY alone, a different solubility tag was investigated. A polyethylene glycol (PEG) tag was investigated to improve solubility. This tag should not interact with the HRE probe in a negative manner. After initial synthetic tests, described below, work was continued by Dr F. Cuda. The cysteine residue was replaced with a propargylglycine residue then click chemistry was used to attach to a PEG-azide (Figure 73).[155, 156]

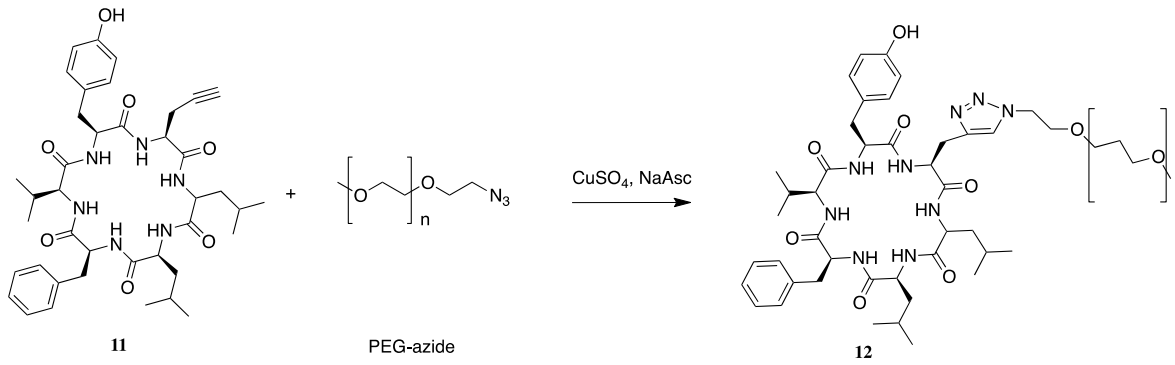


Figure 73 Schematic of the click reaction to generate PEG-*cyclo*-PrLLFVY 12.
Pr= propargylglycine. Click reaction carried out by Dr F. Cuda.

An EMSA was then run with PEG-*cyclo*-PrLLFVY, as previously described, with both a 1:5 and a 1:10 DNA/protein ratio. PEG-*cyclo*-PrLLFVY. The inhibitor was still not fully water soluble, therefore, it was dissolved in DMSO and added to the EMSA sample to leave a final DMSO concentration of 5%. No disruption of DNA-binding was observed (EMSA 10, Figure 74). The experiment was repeated with a 1:25 DNA/protein ratio with similar results. This experiment was also run with Trx-His-HIF-2$\alpha_{1\text{-}360}$ and no inhibition was observed.

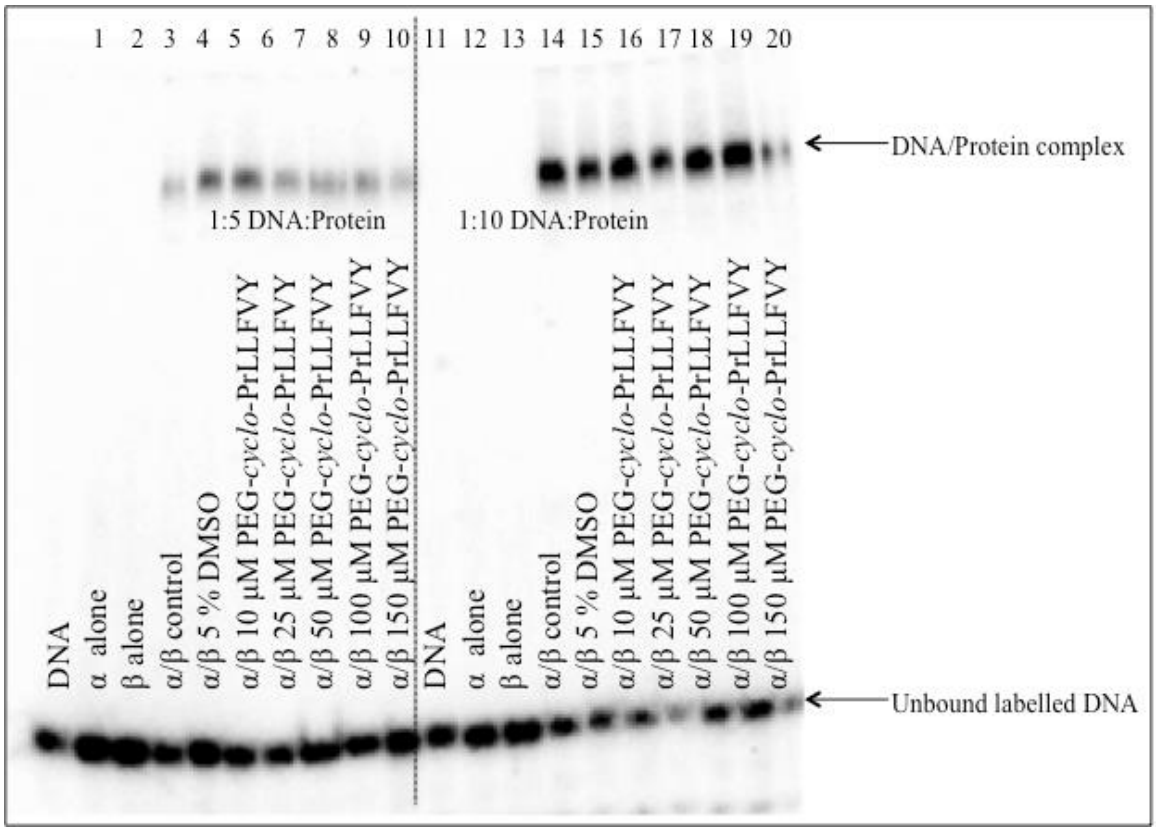


Figure 74 EMSA 10: Independently expressed Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex binding to [^{32}P] HRE probe with the addition of PEG-*cyclo*-PrLLFVY (Pr=propargylglycine).

A 1:5 DNA/protein ratio was used for *lanes 1-10*. (*Lane 1)* DNA probe alone; (*Lane 2)* Trx-His-HIF-1α_{1-350} alone; (*Lane 3)* GST-HIF-1β_{1-459} alone; (*Lane 4)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex; (*Lane 5)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex, 5% DMSO control; (*Lane 6)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex, 10 μM PEG-*cyclo*-PrLLFVY; (*Lane 7)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex, 25 μM PEG-*cyclo*-PrLLFVY; (*Lane 8)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex, 50 μM PEG-*cyclo*-PrLLFVY; (*Lane 9)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex, 100 μM PEG-*cyclo*-PrLLFVY; (*Lane 10)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex, 150 μM PEG-*cyclo*-PrLLFVY. A 1:10 DNA/protein ratio was used for *lanes 11-20*. (*Lane 11)* DNA probe alone; (*Lane 12)* Trx-His-HIF-1α_{1-350} alone; (*Lane 13)* GST-HIF-1β_{1-459} alone; (*Lane 14)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex; (*Lane 15)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex, 5% DMSO control; (*Lane 16)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex, 10 μM PEG-*cyclo*-PrLLFVY; (*Lane 17)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex, 25 μM PEG-*cyclo*-PrLLFVY; (*Lane 18)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex, 50 μM PEG-*cyclo*-PrLLFVY; (*Lane 19)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex, 100 μM PEG-*cyclo*-PrLLFVY; (*Lane 20)* Trx-His-HIF-1α_{1-350}/ GST-HIF-1β_{1-459} complex, 150 μM PEG-*cyclo*-PrLLFVY.

The EMSA has validated the functionality of the recombinant proteins, Trx-His-HIF-1α_{1-350}, Trx-His-HIF-2α_{1-360} and GST-HIF-1β_{1-459}. Consequently, these could then be used in assays (mentioned in section 2.6). Unfortunately, due to the properties of P1, it was evident that the assay was not suitable to test P1 or its analogues. This may be due to the K_d of our inhibitor being too low to prevent the equilibrium shifting to the DNA/protein complex. Attempts to use the assay for alanine scanning were not continued, however, this assay does provide a repeatable way to demonstrate the DNA-binding ability of the HIF-1 complex and to test more potent small molecule inhibitors.

2.6 Work carried out for the publication of initial findings of the ability of P1 to disrupt HIF-1α/HIF-1β dimerisation

In 2013, Miranda *et al.* published a paper describing the discovery and validation of P1 (section 1.10, Appendix C).[1] In order to publish this data, *in vitro* analysis was required. Recombinant Trx-His-HIF-1α_{1-350}, Trx-His-HIF-2α_{1-360} and GST-HIF-1β_{1-459} were expressed and purified (these were now known to be functional from the EMSA study). Furthermore, the individual domains of HIF-1α were also expressed and purified (work by myself and A. Male, a PhD student in the Tavassoli lab) (Figure 75). The recombinant proteins were used within the group (work by Dr Tavassoli) to ascertain the binding location of P1 to be the PAS-B domain of HIF-1α over the PAS-A or bHLH domain. Interaction between P1 and HIF-1α-PAS-$B_{238-349}$ was confirmed by ITC.[1] In addition, a biotinylated analogue of P1 was synthesised alongside a fluorescent derivative for use in pull-down binding assays. These compounds were synthesised by myself in collaboration with Dr F. Cuda (a Postdoctoral researcher in the Tavassoli lab). pET28a-HIF-1α-PAS-$B_{238-349}$ plasmids were constructed by A. Male and the expression vectors for bHLH and PAS-A were a kind gift from Dr G. Melillo, AstraZeneca.

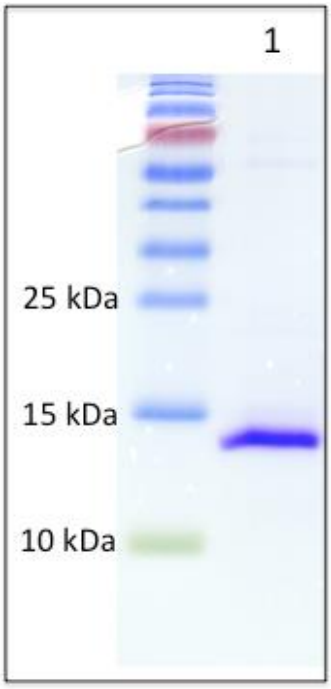


Figure 75 SDS-PAGE of purified His-HIF-1α-PAS-B$_{238-349}$.
The protein was indicated with a molecular weight of 16708.77 Da* (Lane 1). *The lab had a problem whereby the ladder ran slowly in comparison to the protein, this was the case for all protein with this batch of ladders.

The proteins and protein fragments were used to ascertain the binding affinity of P1 to HIF-1α and specifically to the PAS-B domain. Based on this information, a biophysical approach could be taken to determine the activity of each analogue.

2.7 Determining the binding affinity of the alanine analogues to the PAS-B domain of HIF-1α

As stated above, work within the Tavassoli group has established, by ITC, that P1 binds to His-HIF-1α-PAS-$B_{238-349}$ with a K_d = 124±23 nM.[1] This approach could be exploited for the alanine scanning; the direct binding of all the alanine analogues to His-HIF-1α-PAS-$B_{238-349}$ could be investigated using a biophysical technique.

One problem to overcome was the difficulty in expressing and purifying His-HIF-1α-PAS-$B_{238-349}$, which has poor solubility and precipitated during purification and dialysis. This is likely to indicate aggregation and is undesirable when trying to compare the binding affinity of the alanine analogues. In the literature the isolated HIF-1α-PAS-$B_{238-349}$ domain is well-known to have low solubility and unfavourable solution characteristics.[28] Zhu *et al.* (2012) overcame these problems by generating a more soluble and stable mutant of HIF-1α-PAS-$B_{238-349}$ (1106 mutant) that still retained the physiological properties of HIF-1α.[28] The mutants were designed based on a homology model derived from the crystal structure of the HIF-2α/ARNT-PAS-B domain complex. The generation of the quadruple mutant allowed the NMR structure of the domain to be obtained. NMR analysis additionally showed that dimer formation occurred only at concentrations above 80 μM. Zhu *et al.* (2012) constructed three mutants (Table 1), the single R245E diminished protein aggregation associated with the wild-type, however, it had much lower expression than the 1106 quadruple mutant. However, the quadruple mutant still has the propensity to homodimerise at high concentrations. (Table 1).[28]

HIF-1α-PAS-B construct name[28]	Domain boundary	Mutation	Protein yield (mg/ml)	Tm (°C)
900 Wild-type	238-349	None	0.7	49.2
947	238-349	R245E	1.5	49.0
965	238-349	E266H, R311H, S330L	7	52.3
1106	238-349	R245E, E266H, R311H, S330L	10	51.9

Table 1 HIF-1α-PAS-B mutants constructed by Zhu *et al.* (2012).

Table compares the mutations made, the yield and stability of each mutant. Mutant 1106, highlighted in pink, was used for assays reported later.[28]

In this study, the His-HIF-1α-PAS- $B_{238-349}$ 1106 mutant was used to test the relative binding of each of the alanine analogues. The HIF-1α-PAS-$B_{238-349,}$ domain used by the Tavassoli group in previous studies, was the same fragment used by Zhu *et al.* (2012) (amino acids 238-249).[1, 28] Our construct was in a pET28a vector generating a His-tagged recombinant protein, to allow for easy purification. The 1106 mutant was constructed using this vector, in combination with site-directed mutagenesis (SDM). The four mutations were sequentially cloned in and for each new amino acid codon the most abundant *Escherichia coli* codon was selected. The vector was amplified by PCR, and the methylated template vector was digested using *Dpn*I (specifically digests methylated DNA), prior to transformation of the PCR product into chemically-competent DH5α cells. Successful colonies were grown and sent for sequencing to confirm incorporation of these changes. This resulted in the successful construction of pET28a-HIF-1α-PAS-$B_{238-349}$ 1106 (Figure 76).

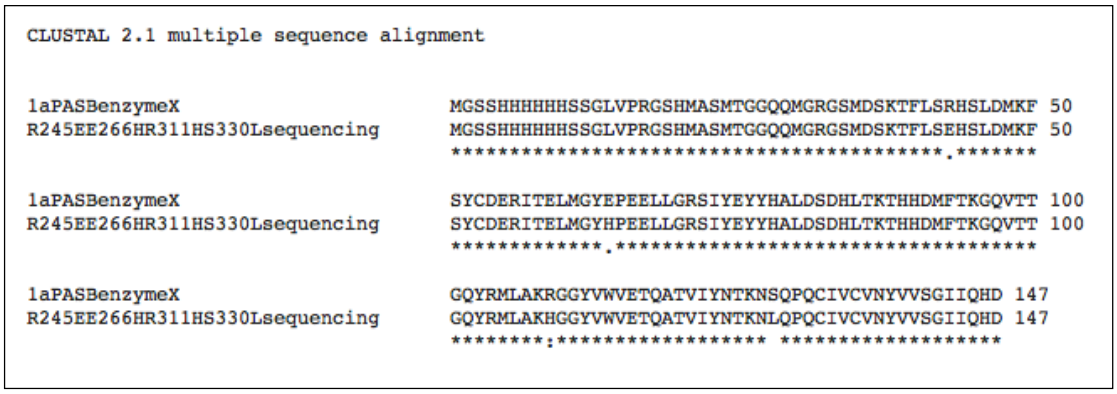


Figure 76 Clustal W alignment of protein sequences of native HIF-1α-PAS-B and constructed mutant HIF-1α-PAS-B 1106.[53]

Once the clone of the pET28-HIF-1α-PAS-$B_{238-349}$ 1106 plasmid had been verified it was transformed into BL21 (DE3) chemically-competent cells and expressed and purified following the same Ni affinity bead protocol as described for Trx-His-HIF-$1\alpha_{1-350}$ (section 2.4.1). Successful purification was confirmed by SDS-PAGE, by the occurrence of a single band at approximately 16 kDa (Figure 77).

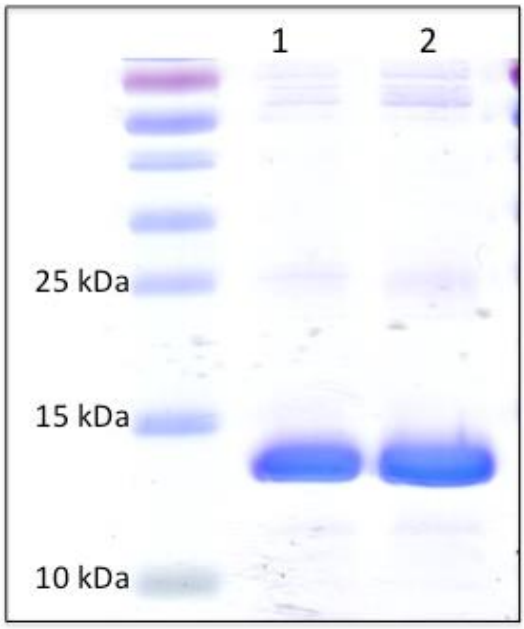


Figure 77 SDS-PAGE gel of purified His-HIF-1α-PAS-B$_{238-349}$ 1106.
The protein was indicated with a molecular weight of 16696.76 Da*. (Lane 1) and (Lane 2) represent two different expressions. * The lab had a problem whereby the ladder ran slowly in comparison to the protein, this was the case for all protein with this batch of ladders.

2.7.1 *In Vitro* alanine scanning: Microscale Thermophoresis

With the knowledge that P1 binds to His-HIF-1α-PAS-B and with the stable mutant of PAS-B (His-HIF-1α-PAS-B$_{238-349}$ 1106) in hand,[1, 28] the next step was to select a biophysical assay to allow comparison of binding affinities of the five alanine analogues to HIF-1α-PAS-B$_{238-349}$ 1106. When selecting a technique the following criteria needed to be considered:

- **Minimal inhibitor use**- synthesis and characterisation of the Tat-tagged compounds yielded finite amounts.
- **Minimal protein use**- to ensure consistency of results the same protein sample would ideally be used for testing all analogues; hence, the assay would need to use minimal amounts of protein.
- **Ability to measure a large range of K_d values**- the alanine analogues may have very different K_d values and all require measurement.

- **Cost**
- **Availability**

A selection of biophysical techniques was considered including ITC and surface plasmon resonance (SPR). MST was chosen as it satisfied the outlined criteria; the technique was accessible, cost effective, has low consumption of both protein and ligand (microliters at nanomolar concentration) and can detect a wide range of binding affinities (K_d 10^{-11}-10^{-2} M). It is also time effective, allows rapid assay optimisation and troubleshooting, and measures the interaction in free solution so immobilisation is not necessary.[157]

2.7.2 An introduction to Nanotemper MST

MST is a biophysical technique that can measure the binding of substrates to biomolecules. This is achieved by measuring changes in the mobility of molecules within microscopic temperature gradients, termed 'thermophoresis'. Thermophoresis is dependent upon changes in molecular size, charge and the hydration shell of molecules. Here, the thermophoresis of a fluorescently labelled biomolecule was measured in experiments where a non-labelled binding ligand is titrated. Binding of the ligand caused a change in at least one of these properties (molecular size, charge or hydration shell) and was then quantified by measurement of the thermophoretic change. Practically, the fluorescent-biomolecule/ligand mixture was placed in a small capillary and heated with an IR laser, causing molecules to move from the heated region to the cold exterior. The concentration of molecule in the 'hot' region eventually reached a steady state.

Specifically, when a labelled biomolecule (for example protein or DNA) was heated a fluorescence change (T-Jump) was observed, prior to thermophoresis occurring. Once heating was finished 'inverse T-jump' occurred followed by back diffusion of molecules (Figure 78).

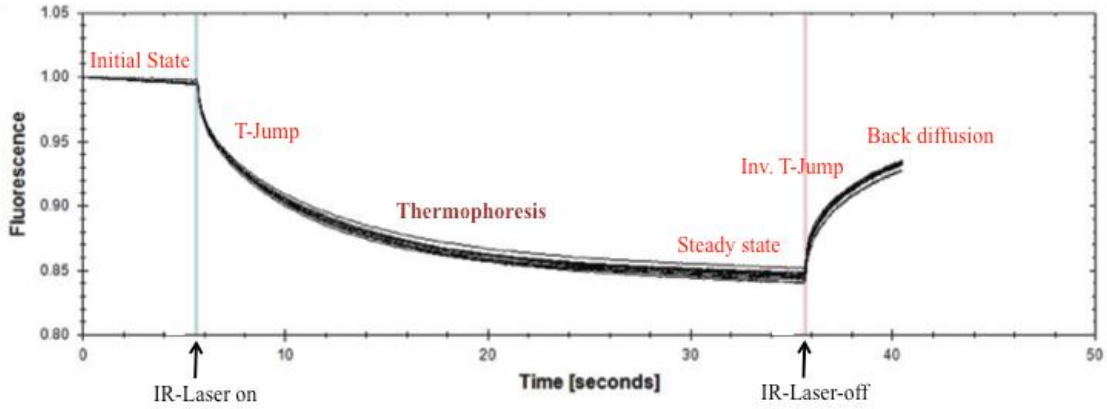


Figure 78 Typical MST signal HIF-1α-PAS-$B_{238-349}$ 1106 with P1.

Initially molecules are equally distributed, once the IR laser is turned on a fast T-jump is observed followed by thermophoresis where the fluorescence decreases. On termination of laser light an inverse T-jump occurs followed by back diffusion of molecules.[157]

Processing this signal provided information about the molecular interaction taking place within the capillary. A temperature difference (ΔT) leads to movement of molecules from the heated region, and can be quantified by the Soret coefficient S_T (Equation 1).[158]

$$C_{hot}/C_{cold} = \exp(-S_T\Delta T) \approx 1-(-S_T\Delta T) \text{ (linearisation due to relatively small } \Delta T)$$

Equation 1

The MST experiments use fluorescence to detect the changes in molecule concentration, so an expression for normalised fluorescence (F_{norm}) can be written, where F_{inital} is fluorescence prior to IR-laser heating (Equation 2).

$$F_{norm}=F_{hot}/F_{inital}$$

Equation 2

Here the MST experiments were set up as a titration series with increasing ligand concentration. The MST signal will detect binding by a change in F_{norm} (Equation 3).

$$F_{norm} = (1-X)F_{norm}(unbound) + XF_{norm}(bound)$$

Equation 3

F_{norm}(unbound) is normalised fluorescence of unbound molecules, likewise F_{norm} (bound) is the normalised fluorescence of bound molecules and X is fraction of bound molecules.[157, 159]

Differences observed in the F_{norm} of samples with and without ligand will enable the calculation of the fraction bound, and hence the dissociation constant (K_d) of the interaction. Prior to each MST experiment a 'capillary scan' was carried out; fluorescence was measured across the capillaries so that the exact position of each capillary (in μm) was known.

As a technique, MST has been used to probe the interaction between proteins and small molecules. For example Tan *et al.* (2012) identified small molecule inhibitors of proliferating cell nuclear antigen (PCNA). The anticancer target, PCNA, is required for numerous cellular processes and is required for DNA replication.[160] MST was used to establish that the inhibitor, PCNA-I1, bound selectively to PCNA trimers with a K_d of 0.41±0.71 μM.

2.7.3 Initial MST experimental set up

Prior to testing all five analogues with this technique, binding of P1 to the His-HIF-1α-PAS-B $1106_{238-349}$ mutant needed to be confirmed. Simultaneously, optimisation was also carried out. Firstly, His-HIF-1α-PAS-$B_{238-349}$ 1106 was labelled using a Monolith™ NT-647 Red NHS dye containing NHS-ester groups (excitation 650 nm and emission 670 nm), which modified primary amines within the protein (His-HIF-1α-PAS-$B_{238-349}$ 1106 contains six lysine residues). The labelled protein was isolated from the unreacted label by size exclusion chromatography. The protein concentration used for the assays needed to be calculated next. This was determined by the observed fluorescence; approximately 100 fluorescent units would be suitable for the assays. Full conversion of unlabelled protein to

labelled was assumed and a serial dilution of the protein was made (10-200 nM), loaded into capillaries and a cap scan (fluorescence scan along the capillary length) was done. 40 nM of labelled His-HIF-1α-PAS-$B_{238-349}$ 1106 gave an ideal fluorescence read out and this would be used for all subsequent experiments. Initial optimisation runs indicated that the MST parameters (MST power-40%/20% and LED power-70%) and that hydrophilic capillaries should be used (hydrophilic polymers covalently attached to the glass to prevent the adsorption of proteins). Importantly, it should be noted that above approximately 50 μM P1 aggregation was observed and no thermophoretic shift could be observed. Therefore this became the maximum peptide concentration used for the experiments.

Furthermore, prior to alanine scanning the viability of the assay had to be assessed, so preliminary runs of P1 and negative control samples were run. His-HIF-1α-PAS-$B_{238-349}$ 1106 was run and binding was evident, yielding a K_d value of 1.6±0.2 μM (this is the average from two runs) (Figure 79). The K_d value obtained from ITC with wild-type His-HIF-1α-PAS-$B_{238-349}$ was 124±23 nM. Although these values were 10-fold different, this maybe due to the different technique, protein sample or different binding of the mutant in comparison to the wild type.

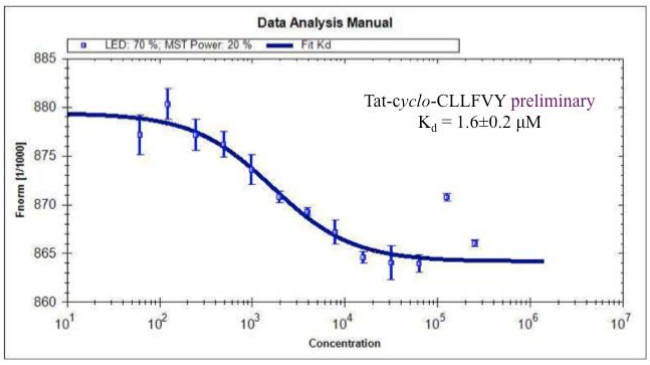


Figure 79 Preliminary MST trace of P1 with His-HIF-1α-PAS-$B_{238-349}$ 1106, n=2. NT-647 labelled HIF-1α-PAS-$B_{238-349}$ 1106 concentration kept constant (40 nM).

The concentration of P1 was varied (25 μM-6.1 nM). MST analysis performed using the Monolith NT.115. Concentrations plotted on the x-axis are plotted in nM. K_d = 1.6±0.2 μM.

Next, a Tat-tag control was run, as well as an unrelated small molecule control (pyridoxal phosphate, (PLP). Surprisingly, the traces indicated that Tat bound to HIF-1α-PAS-$B_{238-349}$ 1106 with a K_d value of 18.9±0.242 μM (Figure 80). This was 10-fold higher than the K_d

obtained for P1 so could be attributed to non-specific binding. It was decided to proceed in testing the binding of the analogues regardless of this but to attribute this K_d to no binding of the cyclic peptide. In contrast, the PLP compound showed no binding to His-HIF-1α-PAS-$B_{238-349}$ 1106 indicating that the binding observed previously was due to protein/ligand interaction not a non-specific interaction (Figure 80).

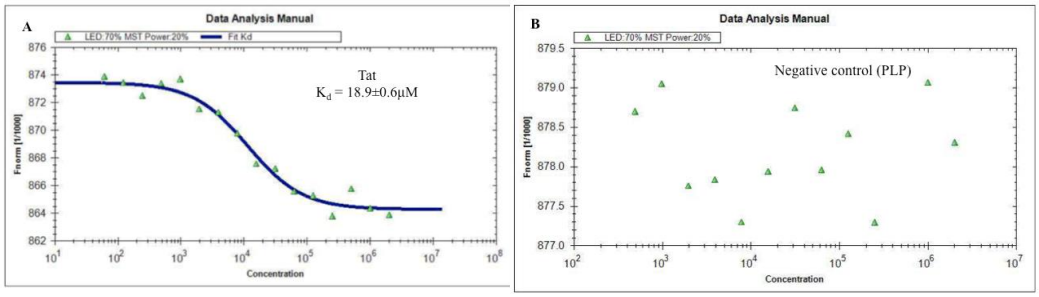


Figure 80 Negative control MST traces of Tat and PLP with His-HIF-1α-PAS-$B_{238-349}$ 1106. The concentration of compounds was varied Tat (2 mM-61 nM (by weight)) and PLP (2 mM-61 nM) MST analysis performed using the Monolith NT.115. Concentrations plotted on the x-axis are plotted in nM. *(A)* Tat, K_d = 18.9±0.242 µM, calculated K_d value 11.7 µM, but this was adjusted to account for 38% water as the peptide concentration was obtained by weight and not absorbance at 280 nm. *(B)* PLP no binding observed.

2.7.4 Microscale thermophoresis alanine scanning

For the alanine scanning, a new protein sample was labelled and again a fluorescence scan was carried out to select an appropriate concentration of labelled HIF-1α-PAS-$B_{238-349}$ 1106 to use, which was determined to be 60 nM. Prior to testing the alanine analogues the binding of P1 to the new protein sample was confirmed and compared to the results in section 2.2.2. Analysis of two independent MST runs showed P1 to bind to HIF-1α-PAS-$B_{238-349}$ 1106 with a K_d of 2.680±0.242 µM and 0.182±0.0387 µM (Figure 81). These runs confirmed binding of P1 to HIF-1α-PAS-$B_{238-349}$ 1106 and are low micromolar in concentrations, as previously observed (Figure 12), however, there is a 15-fold difference between the individual runs here.

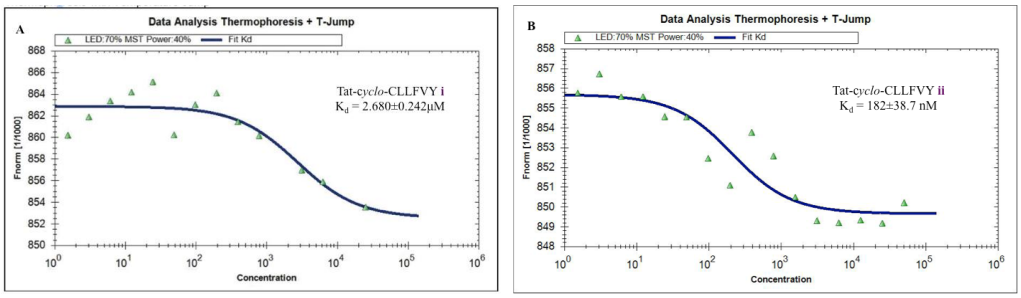


Figure 81 MST trace of P1 with His-HIF-1α-PAS-B$_{238-349}$ 1106 , two repeat experiments (*i* and *ii*) with the same protein sample.

NT-647 labelled HIF-1α-PAS-B$_{238-349}$ 1106 concentration kept constant (60 nM). The concentration of P1 was varied (50 µM-1.5 nM). MST analysis performed using the Monolith NT.115. Concentrations plotted on the x-axis are plotted in nM. *(A)* P1 *i*, K_d = 2.680±0.242 µM. *(B)* P1 *ii*, K_d = 182±38.7 nM.

Next, the binding of each alanine analogue to His-HIF-1α-PAS-B$_{238-349}$ 1106 was determined by MST. Due to limited time and supply of hydrophilic capillaries it was not possible to do repeat MST runs for all the analogues. One way of gaining duplicate runs was by moving the LED along the capillary and repeating the MST experiment so that a different part of the sample was subjected to the IR laser. This was not successful in all cases and some runs showed inconclusive results, which may have been due to laser-induced protein damage/degradation. Where successful repeat runs were obtained they have been reported.

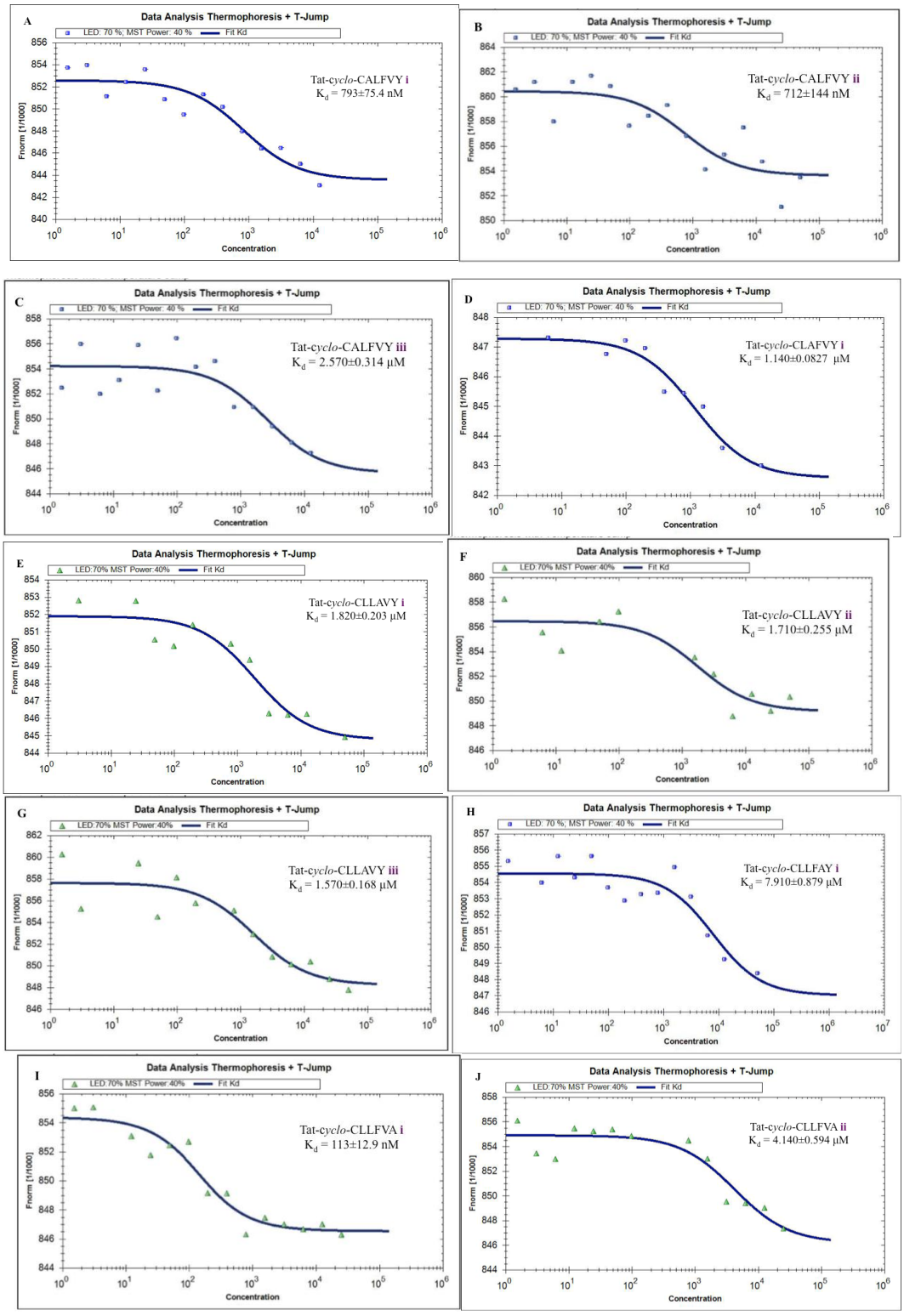


Figure 82 MST traces of alanine analogues with His-HIF-1α-PAS-B$_{238-349}$ 1106, repeat experiments with the same capillary indicated (*i*, *ii* and *iii*).

All analogues were tested with the same protein sample, NT-647 labelled His-HIF-1α-PAS-B$_{238-349}$ 1106 concentration kept constant (60 nM). The concentration of peptides was varied (50 µM-1.5 nM). MST analysis performed using the Monolith NT.115. Concentrations plotted on the x-axis are

plotted in nM. *(A)* Tat-*cyclo*-CALFVY *i*, K_d = 793±75.4 nM. *(B)* Tat-*cyclo*-CALFVY *ii*, K_d = 712±144 nM. *(C)* Tat-*cyclo*-CALFVY *iii*, K_d = 2.570±0.134 μM. *(D)* Tat-*cyclo*-CLAFVY *i*, K_d = 1.140±0.0827 μM. *(E)* Tat-*cyclo*-CLLAVY *i*, K_d = 1.820±0.203 μM. *(F)* Tat-*cyclo*-CLLAVY *ii*, K_d = 1.710±0.255 μM. *(G)* Tat-*cyclo*-CLLAVY *iii*, K_d = 1.570±0.168 μM. *(H)* Tat-*cyclo*-CLLFAY *i*, K_d = 7.910±0.879 μM. *(I)* Tat-*cyclo*-CLLFVA *i*, K_d = 113±12.9 nM. *(J)* Tat-*cyclo*-CLLFVA *ii*, K_d = 4.140±0.594 μM.

A tabulation of the K_d values of each analogue is presented in Table 2.

Alanine analogue	**K_d (repeat 1)**	**K_d (repeat 2)**	**K_d (repeat 3)**	**K_d average**
Tat-*cyclo*-CLLFVY (P1)	2.680±0.242μM	182±38.7 nM	1.6±0.2 μM (from preliminary experiments)	1.49±1.56 μM
Tat-*cyclo*-CALFVY	793±75.4 nM	715±144 nM	2.570±0.314 μM	n/a
Tat-*cyclo*-CLAFVY	1.14±0.0827 μM	n/a	n/a	n/a
Tat-*cyclo*-CLLAVY	1.82±0.203 μM	1.71±0.255 μM	1.57±0.168 μM	1.70±0.125 μM
Tat-*cyclo*-CLLFAY	7.91±0.879 μM	n/a	n/a	n/a
Tat-*cyclo*-CLLFVA	113±12.9 nM	4.14±0.594 μM	n/a	2.13±2.85 μM

Table 2 Summary of MST obtained K_d values for binding of alanine analogues to His-HIF-1α-PAS-$B_{238-349}$ 1106.

Errors values reported for averages obtained by standard deviation. Purple highlighted values represent average values or where only one repeat was possible.

2.7.4.1 Microscale thermophoresis alanine scanning summary

In summary, the results confirmed that P1 binds to His-HIF-1α-PAS-$B_{238-349}$ 1106. Repeats of MST runs showed large variation between obtained K_d values (Figure 82). However, for some a trend was evident.

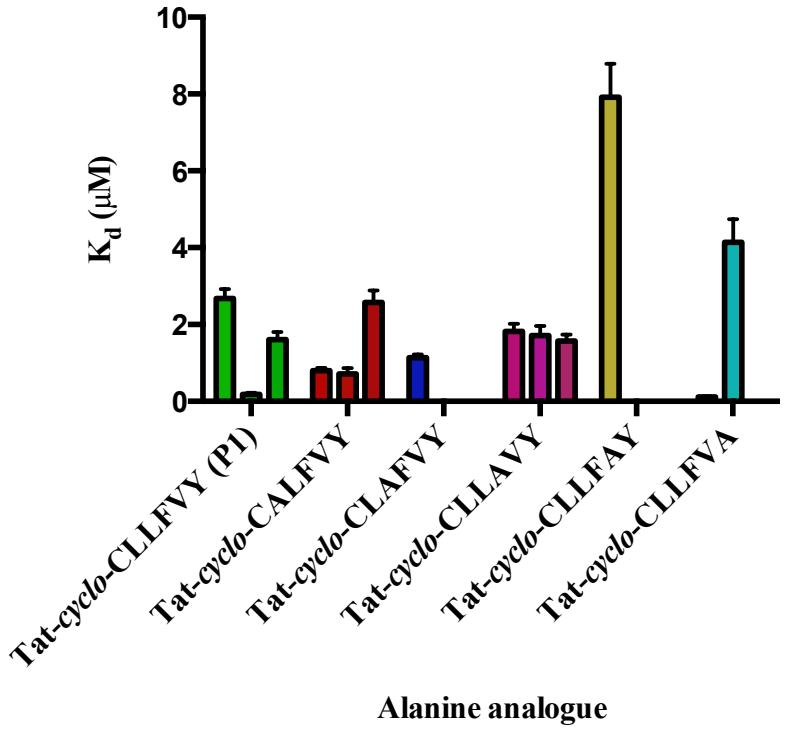


Figure 83 Bar graph summary of the K_d values obtained by MST for all five alanine analogues.

The data suggested that all of the five analogues also bind to the protein. The alanine substitution does not prevent binding in any analogue. These are the only conclusions that can be drawn from the data. Due to the large variation between runs it is not possible to distinguish between the alanine analogues and no definitive conclusions can be drawn with respect to the active motif of P1.

3 Discussion and summary

HIF-1 regulates cell adaption to hypoxia by up-regulating genes that either aid hypoxic cell survival or increase oxygen transport to a hypoxic region.[18] Tumour environments are inherently hypoxic, due to rapid cell proliferation; thus up-regulation of HIF-1 in these areas can increase the aggressiveness of cancer and lead to poor patient prognosis.[87, 91, 161] Genetic alterations in tumour suppressor genes can also result in changes in HIF-1 regulation.[162] Hence, HIF-1 is viewed as an attractive cancer therapy target.[105, 112, 117, 163] Here, we focused on the development and verification of P1, a HIF-1 inhibitor discovered by Miranda *et al.* (2013).[1] P1 inhibits HIF-1α/HIF-1β dimerisation, which represents an ideal point to disrupt the function of HIF-1 for several reasons:

1.) Disruption of HIF-1α/HIF-1β dimerisation completely bisects the hypoxia response network. In the absence of dimerisation of HIF, HIF is unable to bind and activate transcription of its downstream target genes.

2.) Targeting the interface between HIF-1α and HIF-1β facilitates selectivity over the other HIF-2 and HIF-3 isoforms. This gives the inhibitor a potential therapeutic advantage over non-specific inhibitors of HIF and may enable P1 to be used as a tool to study the different functions of each isoform in nature.

3.) HIF-1 is only up-regulated in areas of hypoxia or genetic alternation, so intervention here should minimise off-target effects; a key problem with other sites of inhibition.

The aims of this project were three-fold. Primarily, to investigate which residues within *cyclo*-CLLFVY are critical to its activity and what is the active motif for P1. Secondly, to express and purify recombinant HIF-1α and HIF-1β. Finally, to allow investigation into their functionality, and the use of HIF-1α and HIF-1β in assays to confirm the binding location of P1.

Chapter 2 firstly describes the choice of alanine scanning to establish the critical residues within *cyclo*-CLLFVY. Alanine scanning is readily reported as a technique to establish the contribution of specific amino acids to the function of proteins or peptides.[124, 164, 165] For example, Spurr *et al.* (2013) demonstrate successful alanine scanning converting a cyclic

peptide inhibitor of AICAR into a modified dipeptide, which increased the inhibitors potency and facilitated a simpler synthetic route.[124] Alanine is chosen as the amino acid substituted into the peptide being tested due to its lack of chemical functionality and relative small size; both should prevent large perturbations to peptide folding. The one drawback was the similar chemical properties between alanine, leucine and valine. This may have meant that any changes in P1 activity found following changing these residues to alanine would be minimal. Previously, valine has been used in place of alanine for scanning amino acid function.[166] However, both valine and leucine are significantly larger than alanine and cause different hydrophobic interactions so a difference would have been expected. The use of alanine scanning for protein functionality rat ATP-gated $P2X_2$ receptor has shown capacity for alanine scanning to mitigate function of the receptor when valine, isoleucine and leucine are substituted.[167]

Initially, the RTHS and SICLOPPS methodology was used to alanine scan the inhibitor; this appeared a logical place to start, as it was the method by which the inhibitor was first identified.[1] Drop-spotting conditions for the alanine scanning were verified and the protein-protein interaction was demonstrated via the growth retardation on selective media that was observed in response to increasing IPTG concentrations. This was in agreement with other protein-protein interactions investigated using a bacterial RTHS.[1, 9, 125, 139, 168] It was hypothesised that this was due to increased expression of 434-HIF-1α and P22-HIF-1β, leading to increased dimerisation and DNA-binding. This in turn, resulted in an inhibition of DNA polymerase binding and decreased transcription of reporter genes. The chosen conditions were kanamycin (25 μg/ml), 3-AT (7.5 mM) and IPTG (25 μM). These conditions resulted in adequate shut down of cell growth and were in agreement with previously reported conditions for inhibitor isolation.[1] It is worth noting that different protein-protein interactions require more or less stringent conditions in order to visualise the interaction on minimal media. For example, for the p6-UEV RTHS the conditions used were 30 μM IPTG, 25 μg/ml kanamycin and 2.5 mM 3-AT; while for the AICAR transformylase RTHS, 200 μM IPTG, 50 μg/ml kanamycin and 7.5 mM 3-AT was required.[125, 139] An explanation for these deviations could include the strength of interaction between the two proteins. If the interaction is weak and transient increased expression may be required to obtain a higher ratio of dimer. Also, mammalian proteins such as HIF-1 may not be adequately codon-optimised for *E. coli* expression and therefore, lower expression than anticipated may be occurring. In addition, the P22 and 434 domains

could also hinder the dimerisation and may lead to a decreased K_d value for the protein-protein interaction.

Next, the alanine scanning SICLOPPS plasmids encoding the desired cyclic peptides were successfully constructed and verified by DNA sequencing. They were then transformed into the HIF-1 RTHS. Initially, the dnaE intein from *Ssp* was used.[131, 169, 170] Chitin bead pull-down assays showed all alanine analogues underwent splicing at similar levels to *cyclo*-CLLFVY. Complete splicing was not observed in this bead pull-down assay, although, the ability to splice was eluded to. Previously, since incomplete splicing by SDS-PAGE analysis of chitin bead pull-down was observed, it is therefore reasonable to assume that in the context of the bead setting used some functionality of the intein is lost.[131, 171] Scott *et al.* (2001) reasoned a difference between intein 'processing' i*n vitro* and 'splicing' *in vivo*.[169] Alternatively, complete intein splicing may be hindered for these extein sequences.

Alanine scanning was then performed within the HIF-1 RTHS; the plasmids were drop-spotted onto selective minimal media with induction of IPTG and L-arabinose. For all alanine sequences (with the exception of CLLFVY) a toxic phenotype was observed following addition of L-arabinose. The strains were drop-spotted on varying supplemented plates and the toxicity was attributed to the cyclic peptide or intein by-product. Perler *et al.* (2000), reported that ligation can be influenced by the identity of the extein amino acids and those adjacent to the N- and C-intein, thus substitutions of the extein sequence could be affecting splicing rates. Hence, leading to different levels of cyclic peptide or intein by-products being produced.[172] These varying levels of product could be contributing to the observed toxicity. Moreover, a change in the I_c+2 extein residue can affect protein splicing, however, hydrophobic residues in this position should not hinder circular ligation.[169] Scott *et al.* (2001) further concluded that variation at the I_c+3 position should have little effect on circular ligation.[169] This information indicates that the CALFVY and CLAFVY should have similar splicing rates to the parent peptide and the observed toxicity should be independent of extein sequence.

High level expression of SICLOPPS plasmids can result in notable toxicity to the host cells and a reduction in toxicity maybe observed with lower induction levels.[169] Lower induction levels may have decreased toxicity, however, the difference in toxicity of analogues to CLLFVY would have biased activity comparison. The codon usage of the amino acids

may also have affected expression rates, however, the most abundant codons for all residues in question (i.e. C, L, F, V, Y and A) were used (0.34-0.57 ratio of codon use relative to all other codons for the particular amino acid for *E. coli*), so minimal expression differences should have been observed. The observed toxicity made it not possible to compare the activity of the alanine analogues. This information also indicates a bias towards a specific splicing rate and toxicity level in the original inhibitor screening using the HIF-1 RTHS and SICLOPPS methodology reported by Miranda *et al.* (2013).[1]

As it was not possible to assess the alanine analogues using the *Ssp* inteins due to this toxic phenotype described above, inteins from the dnaE gene from *Npu* were used as an alternative. Iwai *et al.* (2006) suggest that these *Npu* derived inteins have a superior trans-splicing ability compared to the previously used *Ssp* variant.[132] In addition, to the previous observations made by Scott *et al.* (2000), Iwai *et al.* (2001) showed evidence that the I_c+2 extein residue has a large effect on the ligation efficiency of the *Ssp* intein splicing.[132, 169] However, poor efficiency with leucine, alanine and valine (approximately 5-10%) and good efficiency with tyrosine and phenylalanine (approximately 80%) was observed.[132] This again indicated that there should be minimal differences in splicing rates between CLLFVY and CALFVY, however, a small variation may have caused the toxic phenotype observed.

It is not known why extein residues near the splice junctions affect splicing activity despite not being involved in the catalytic reaction themselves.[132] Iwai *et al.* (2006) observed that the *Npu* inteins are more 'promiscuous' and that the extein residues appear to have less influence on the splicing rates.[132] The *Npu* alanine analogue SICLOPPS plasmids were successfully constructed and their ability to splice was verified by chitin bead pull-down assays. In contrast to the *Ssp* versions, a higher ratio of spliced inteins were shown via SDS-PAGE. The strains were next drop-spotted onto minimal media under selection conditions. Once again, toxicity was observed upon addition of L-arabinose, with even CLLFVY also exhibiting toxicity in this instance. Lower L-arabinose levels did not alleviate the retardation of growth. It could be speculated that the increased efficiency of the *Npu* inteins may cause an increase in the levels of cyclic peptide and intein by-products, again resulting in a toxic phenotype. Work within the Tavassoli group is on-going to understand this phenomenon and circumvent this issue with the use of degradation tags and varying levels of L-arabinose (unpublished work by J. Townend).

In comparison to our work, Datta *et al.* (2010) adopted a similar RTHS/SICLOPPS (*Ssp*) alanine scanning strategy to establish an active motif of two cyclic peptide inhibitors, CIFYYV and CDLRWF.[168] These were found to be inhibitors of the protein-protein interactions between MDM2/p53 and MDM2-like p53 binding protein (MDMX)/p53, respectively. Inhibition of these interactions results in the restoration of tumour suppressor protein, p53, activity. In line with our observations, Datta *et al.* (2010) observed that the alanine analogues had different toxicity levels and stated that 'In all cases, the gene-reporter effects generally matched the mutant survival trend which are likely to be affected by reporter independent factors, such as the toxicity of the expressed peptide'.[168] This supports our findings that alanine substitution of the cyclic peptide affects host strain survival. However, the authors of this study did tentatively conclude that all analogues of CIFYYV were inert, demonstrating the need for all five residues.[168] This observation could, however, be biased towards the toxic phenotype as is supported by their *ortho*-nitrophenyl-β-galactosidase assay results included in this study. Substitution of the phenylalanine and charged residues in CDLRWF affected activity the most.[168] Taken together with our results the RTHS and SICLOPPS methodology is not an ideal system to perform alanine scanning due to the inherent effects of substituting extein residues. This may be overcome with more development of the *Npu* SICLOPPS variant.

Due to the inability to rank the alanine analogues within a bacterial RTHS, the alanine analogues were synthesised chemically and tested via other means. Due to insolubility, of the cyclic peptides each alanine analogue was Tat-tagged, using a similar methodology to that previously reported in the literature.[1, 9, 139] This methodology yielded the desired Tat-tagged peptides with a purity in excess of 90%. However, the overall yields were low, less than 3% in all instances.

The low yield could be attributed to the following factors. Firstly, *cyclo*-CLLFVY is predominantly composed of aliphatic and aromatic residues, leading to it being hydrophobic (67%). This level of hydrophobicity is not unusual for cyclic peptide inhibitors identified from a RTHS and SICLOPPS screen. Hydrophobic inhibitors of mammalian and bacterial protein-protein interactions, as well as both hetero- and homodimers have been reported.[139, 168, 173] The hydrophobic nature of these inhibitors, may be explained by the hydrophobic 'hot spots' they target.[174-176] Secondly, *cyclo*-CLLFVY and its analogues have poor solubility in most solvents with the exception of DMF, in which the cyclisation was performed. Purification of both the linear and cyclic peptides is

likely to have resulted in a loss of product due to insolubility. In total, the final tagged peptide was reverse-phase HPLC purified three times, which would have adversely affected the overall yield, accentuated by its poor solution phase behaviour.

For future synthesis of these hydrophobic cyclic peptides several avenues of optimisation could be explored. For example, by minimising or completely eliminating the reverse-phase HPLC step. To avoid multiple purification steps an alternative synthetic route must be considered. To do this, the method of cysteine protection during cyclisation will need to be altered. The use of on-resin cyclisation or orthogonal protection would achieve this. For example, a chlorotrityl resin could be used in place of a Wang resin. The peptide could be cleaved from the chlorotrityl resin with dilute TFA (1-5% in DCM), leaving the side chain protecting groups intact. After which cyclisation could occur removing the need to re-protect the cysteine residue (Figure 84).[177] Steric bulk may hinder cyclisation in some case and a suitable site for cyclisation would need to be investigated. Additionally, this may eliminate the production of the Wang impurity observed during the synthesis of the linear peptide (Figure 36). Development of such a pathway would require detailed optimisation.

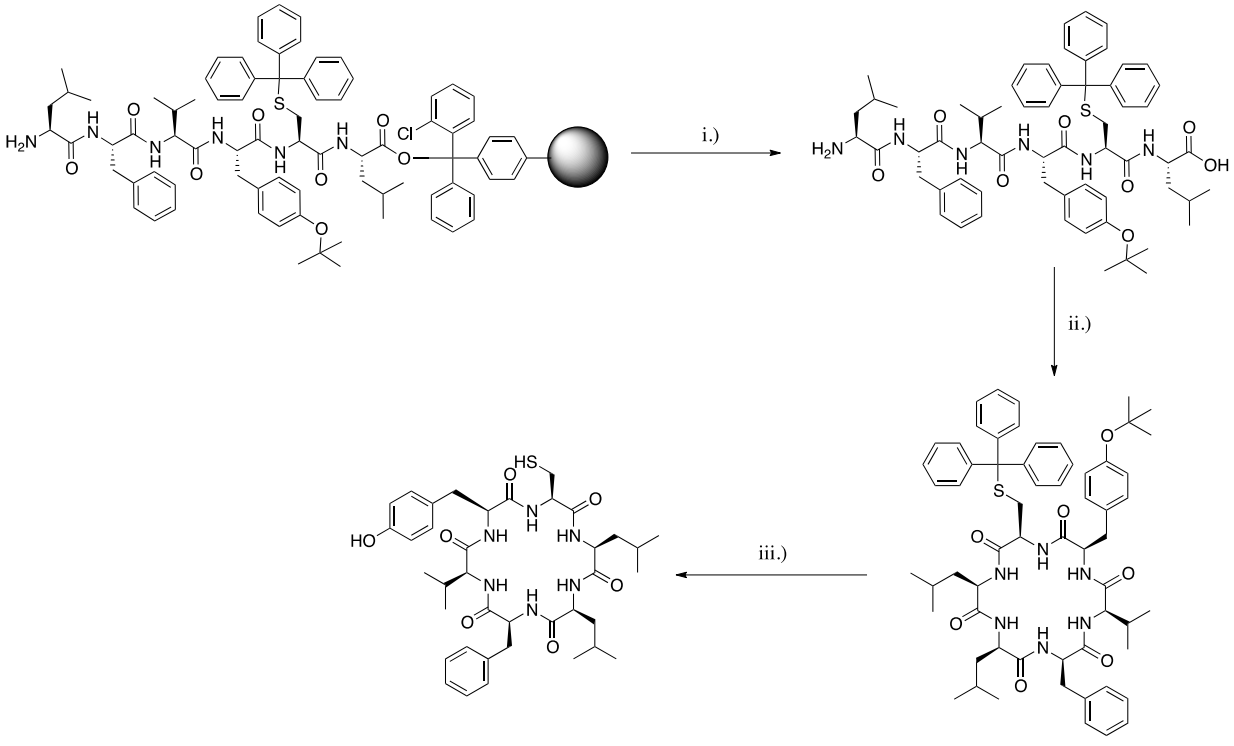


Figure 84 Proposed synthetic pathway to *cyclo*-CLLFVY using a 2-chlorotrityl chloride resin.
i.) 1-5% TFA in DCM, ii.) EDC (3 eq), HOBt (6 eq), DMF, iii.) TFA/phenol/H_2O/TIS, 88/5/5/2.

Alternatively, a C-terminally protected cysteine could be attached to a resin via its thiol moiety and then build the peptide chain and finally perform on-resin cyclisation. An orthogonal C-terminal protecting group would be required, such as an α-allyl ester. Cleavage from the resin in parallel with deprotection would yield the desired cyclic peptide. Tavassoli *et al.* (2005) and Horswill *et al.* (2004) demonstrated the cyclisation element of this concept by re-attaching a complete linear cysteine containing peptide to a modified amino PEG acrylamide copolymer resin, via disulphide bond formation. On-resin cyclisation was then carried out; yielding the desired cyclic peptide after TCEP mediated resin cleavage.[125, 126] This could be adapted to attach cysteine, orthogonally protected with an α-allyl ester protecting group, to the resin and build the linear peptide prior to on-resin cyclisation and cleavage (Figure 85).[178] Again, such a strategy would require careful optimisation and extensive verification. One such consideration is that effective cyclisation may be affected by the properties of adjacent residues.

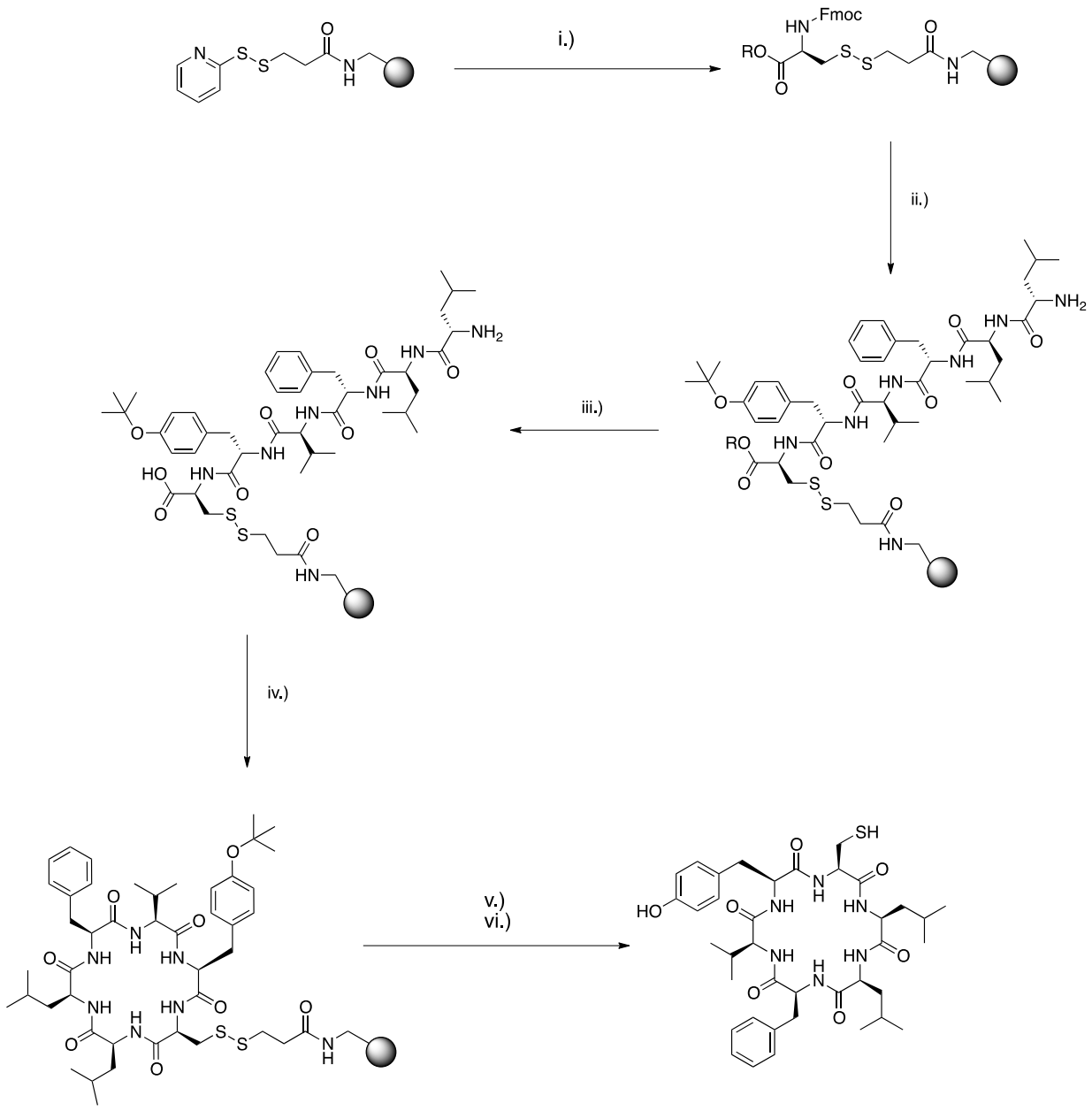


Figure 85 Proposed on-resin synthetic pathway to *cyclo*-CLLFVY using disulphide bond attachment to resin.

R= an orthogonal C-terminal protecting group such as an allyl ester. i.) Fmoc-Cys-OAll, ii.) 20% piperidine/DMF, DIC (3 eq), HOBt (3 eq), Fmoc-Tyr(tBu)-OH, deprotection/coupling repeated for Fmoc-Val-OH, Fmoc-Phe-OH, Fmoc-Leu-OH, and Fmoc Leu, iii.) $Pd(Ph_3P)_4$/ $CHCl_3$/AcOH/NMM, iv.) EDC/1-Hydroxy-7-azabenzotriazole /DMF, v.) TCEP/water/DMF, vi.) TFA/phenol/H_2O/TIS, 88/5/5/2.[126, 178]

Furthermore, changes to the purification itself could improve yield. Condron *et al.* (2008) report difficulties in purifying hydrophobic peptides derived from amyloid β-protein, they encountered <1% yield from reverse-phase HPLC.[179] The poor yield was attributed to the

peptide's poor aqueous and organic solubility. By taking advantage of the peptide insolubility that make reverse-phase HPLC difficult, they adopted an alternative protocol. The peptides were precipitated in water and scavengers removed by diethyl ether washes. This methodology produced purer peptides with better yields.[179] Such a method could be utilised for purification of hydrophobic peptides such as linear and *cyclo*-CLLFVY.[179] Moreover, changes to the mobile and solid phase of the HPLC system may be beneficial when purifying these hydrophobic peptides. Using only 25-40% water in methanol and the use of DMF within the mobile phase have been reported.[180, 181]

Once the compounds had been synthesised, optimisation of a HRE-dependent luciferase assay was started. Such an assay is readily reported in the literature for testing inhibitors of HIF-1.[1, 121, 142-144] Two ways have been described, either with transient transfection of the pGL2-TK-HRE plasmid,[1, 143, 144] or with the use of a stably transfected cell line (U2OS-HRE-luc or U251-HRE).[121, 142] Both strategies were considered, however, transient transfection was chosen as this would allow co-transfection of an additional reporter plasmid (*Renilla* luciferase), in turn, allowing normalisation for cell number. This would ideally allow a more objective comparison between experiments and alanine analogues inhibitors. DFX was chosen to mimic hypoxia in concurrence with other HRE assays.[142, 143] The use of hypoxia itself was not viable due to the lack of access to a hypoxia workstation; $CoCl_2$ was also considered, however, DFX was chosen due to its prevalence in the literature.[105, 108, 121, 142, 143, 182] Work optimising the assay made it apparent that replication of assay conditions are not always facile. Optimisation of transfection, firefly/*Renilla* plasmids and luciferase agents were considered. DFX clearly increased HRE-dependent luciferase activity in our system, notably HIF-1α levels were increased. 25 μM of DFX was used for the assay as this gave an adequate increase in luciferase activity. Addition of P1 at this 'hypoxic' state caused a dose-dependent decrease in luciferase activity in line with that observed previously.[1] This level of DFX was lower than that reported in many examples (i.e. 500 μM DFX).[121, 142, 143] However, when we used higher quantities it appeared that our system was saturated and no difference in luciferase activity was observed with P1 addition. Furthermore, the system did not appear to require more that 25 μM for a consistent and notable increase of approximately 150-fold when compared to 'normoxia'.

The active motif of *cyclo*-CLLFVY could not be elucidated from this data alone. The following conclusions were made:

1.) P1 caused a dose-dependent decrease in HRE-dependent luciferase activity, as expected.

2.) Tat-*cyclo*-CLLFVA showed a similar activity to that of the parent compound, suggesting tyrosine was not a critical residue.

3.) Tat-*cyclo*-CLAFVY and Tat-*cyclo*-CLLFAY resulted in an increase in luciferase activity, which was unexplained but could be due to off-target effects of the inhibitor.

4.) Tat-*cyclo*-CALFVY and Tat-*cyclo*-CLLAVY appeared inactive; this could be due to these residues being essential to activity or an assay off-target effect.

Further verification was required using an *in vitro* assay, this necessitated the purification of recombinant HIF-1α and HIF-1β. Typically, N-terminal fragments of each protein are used to demonstrate dimerisation *in vitro.*[118, 119] Heterologous protein expression can lead to difficulties due to codon bias, improper folding and solubility.[183, 184] The bacterial expression described here was not an exception. Trx-His-HIF-$1\alpha_{1\text{-}350}$ and GST-HIF-$\beta_{1\text{-}459}$ were successfully independently and co-expressed and purified using bead purification. The proteins precipitated upon concentration if left in elution buffer or, if left at 4ºC overnight. Protein samples were therefore flash frozen after purification. Some degradation products were observed by SDS-PAGE analysis. This difficulty in purification was anticipated, as previously reported expressions exhibit poor yields, insolubility and evidence of degradation.[28, 60, 149] Moreover, a reduction in activity has been reported for proteins (i.e. PAS-B) after SEC, storage and handling.[60]

The pull-down assay used here failed to show any inhibition of dimerisation, this could be due to the binding affinity of the PAS-B domains being in the nanomolar range (i.e. between 100-1000 nM).[29] This rendered the pull-down assay ineffective, as the concentration of proteins was high, leading to the favoured dimerisation, potentially making inhibition difficult to be visualised. In contrast, lower concentrations of protein are required in the more sensitive EMSA, putatively allowing inhibition to be observed.

Firstly, the EMSA demonstrated the functionality of independently expressed Trx-His-HIF-$1\alpha_{1\text{-}350}$ and GST-HIF-$\beta_{1\text{-}459}$ and confirmed their ability to dimerise and bind to DNA.

This had previously only been observed for co-expressed variants.[60, 149] The functionality of Trx-His-HIF-2$\alpha_{1\text{-}360}$ was also confirmed. Here, the K_d for the HIF-1 to HRE DNA interaction was 370 nM. This K_d was comparable to the K_d stated, previously, by Chapman-Smith *et al.* (2004) with co-expressed proteins containing a Trx-His and GST tag.[60] Inhibition of dimerisation using ACF was observed here, with an IC_{50} of 40.04 μM, similar to that described in Lee *et al.* (2009) of 5 μM.[118] By using ACF, the capacity of the EMSA to show disruption of the protein-protein interaction was verified. However, this assay was not suitable for testing P1, *cyclo*-CLLFVY or the analogues due to poor solubility and the interaction of the positive Tat-tag with the HRE probe.

We next turned our efforts towards assessing the direct binding of the alanine analogues to His-HIF-1α-PAS-$B_{238\text{-}349}$ 1106. As described in section 2.7.1, MST failed to show any clear trends in binding, nevertheless, all analogues and the Tat-tag showed some degree of binding. Drawing on both this and the HRE-dependent luciferase assay, it can be concluded that the parent compound is the most favourable conformation and sequence; however, further validation of this conclusion is essential. Datta *et al.* (2010) have demonstrated the requirement for the whole peptide after alanine scanning with their HDMX/p53 inhibitor.[168] Typically, ITC would be the 'gold standard' for the determining binding affinity of inhibitors to their protein partner. Miranda *et al.* (2013) used ITC to show that P1 bound to the His-HIF-1α-PAS-$B_{238\text{-}349}$ domain.[1] During the course of this project ITC was used to compare the binding affinity of the analogues, however, due to the influence of pH, buffer mismatching and the *c* value, results obtained here were variable and inconclusive. This has primarily been attributed to inconsistent experimental design and parameters. Here, the non-mutant of His-HIF-1α-PAS-$B_{238\text{-}349}$ was used; this has an inherent tendency to aggregate and misfold, leading to variation in active protein concentration.[28] This could explain why initial experiments described by Miranda *et al.* (2013) were difficult to replicate.[1] Specifically, Wiseman *et al.* (1989) state that to obtain a good binding curve and accurate thermodynamic data, knowledge of the protein concentration and K_d value must accurately be known.[185, 186] This was not possible here due to the unknown concentrations of inactive protein. Furthermore, buffer mismatch could be observed due to small variations in pH, caused by the inherent properties of the peptides. Further optimisation of this technique could be future work but was not possible here, due to financial constraints.

This study has conclusively shown that independently expressed recombinant HIF-1α and HIF-1β have the capacity to dimerise and bind to DNA at specific HRE sites. This work has enabled a parallel successful study into the binding location of P1. In addition, the EMSA assay described here represents a repeatable and robust assay for determining protein functionality and inhibitor validation. Although no active motif was identified, the construction and use of the more stable HIF-1α-PAS-$B_{238-349}$ 1106 has laid the foundation for its in future assays.

4 Conclusions and future work

In conclusion, it was not possible to identify an active motif or critical residues for *cyclo*-CLLFVY. The MST data showed that all the analogues bind to HIF-1α-PAS-B, however, only Tat-*cyclo*-CLLFVA displayed activity similar to P1 within the HRE-dependent luciferase assay. It may be that the whole cyclic peptide sequence is required for its cell-based activity, as has been observed for the inhibitor of MDM2/p53.[168] The conformation adopted by this sequence may be essential.

The EMSA clarified that our Trx-His-HIF-1α_{1-350} and GST-HIF-1β_{1-459} were functional and had the ability to dimerise and bind to DNA. This meant that the proteins could be used in the group to establish the binding location of P1. This EMSA would also serve as an assay to test the effect of inhibitors with a lower K_d value. The assay could be suitable to test a more potent second generation P1 lacking the positively charged Tat-tag. This assay provides a method to test the functionality of recombinant proteins.

In terms of future work, as well as the work described here, initial experiments were carried out to set up a real-time PCR (RT-PCR) assay (Appendix A). This would complement the data obtained for the HRE-luciferase assay. The concept was to test the alanine analogues against the expression of reporter gene mRNA levels for VEGF, GLUT1 or CAIX. This would enable a comparison between analogues on their effect on HIF-1 mediated transcription avoiding the necessity for plasmid transfection. Recently, the Tavassoli lab purchased a hypoxia workstation allowing the RT-PCR to be carried out under hypoxia; unlike the HRE-dependent luciferase assay, which utilised the hypoxia mimic, DFX. This would enable a comparison between hypoxia and DFX to be made. Future work includes the optimisation of the RT-PCR assay and its utilisation to test the alanine analogues. Ultimately, a comparison against the results obtained by the HRE-luciferase assay could be made. One problem with this assay that may be inhibitive is the cost of the large quantity of probes required to make a full comparison between analogues.

For future development of the inhibitor, knowledge about which peptide residues come into contact with the protein would be valuable. Now that it has been established that P1 binds to PAS-B and that we have obtained a more stable mutant, saturation-transfer difference (STD)-NMR or crystal structure characterisation could be used to establish which residues are responsible for binding to HIF-1α-PAS-B.[187] Initial trials to generate a sample of the HIF-1α-PAS-$B_{238-349}$ 1106 mutant to be used for STD-NMR were carried out, however, due to the sensitivity of the technique, higher concentrations of protein are necessary and methods of overcoming protein precipitation are required. Also buffer constraints are rigid so lengthy optimisation of conditions would be necessary to generate a good solution structure.

Another area where further research would be beneficial is the action that P1 has upon HIF-3. Miranda *et al*. (2013) demonstrate that P1 has specificity towards HIF-1 over HIF-2; however, the effect on HIF-3 was not probed.[1] As HIF-3 is known to be a negative regulator of the hypoxia response network, if HIF-3 is also inhibited then off target or contrary effects could occur in some cells lines or at some stages of hypoxia.

It is clear that HIF-1 is an unstable protein and replication of assays and protein purification is difficult.[149] The balancing act of buffers and protein concentration is critical to avoid HIF-1α homodimerisation or non-functional protein. Verbal communication with other research groups has highlighted the difficulty in expressing and purifying recombinant HIF-1 from bacteria. Further optimisation of protein purification should be extensively investigated.

5 Methodology

5.1 Molecular biology

5.1.1 Molecular biology materials and equipment

Primers were synthesised by Eurofins MWG Operon (Germany). PCR was carried out using a Bio-Rad MyCycler Thermal Cycler. Sequencing was carried out by Source Bioscience (Oxford, UK) or Eurofins MWG Operon (Germany). DNA concentrations were measured using a NanoDrop ND-1000 Spectrometer (NanoDrop Technologies, UK). Optical Density of cell cultures was measured using a Cary 100 Bio UV-Visible spectrometer (Agilent Technologies, UK). Enzyme X and Serial Cloner were used for the selection of restriction sites and primer design. pDraw was used to calculate annealing temperatures. SDS-PAGE was run using a mini-PROTEAN Tetra electrophoresis system (Bio-Rad, USA). Agarose gels were run using Bio-Rad sub-cell systems and power pack and were imaged using a Bio-Rad Gel Doc and Quantity-One software (Bio-Rad, USA). Cultures were spun in a Heraeus Bioruge primo R centrifuge (Thermo Scientific, UK). For imaging EMSA gels, a Typhoon FLA 7000 phosphorimager (GE Healthcare, UK) was used. All radioactive work was performed behind an appropriate screen in a dedicated room. DNA purification was carried out using QIAprep spin mini kits (QIAGEN, UK) and DNA extraction form agarose gel was carried out using QIAquick gel extraction kits (QIAGEN, UK). Media was sterilised using a Touchclave-R (LTE Scientific Ltd, UK). Buffer pH was measured using a Hydrus 300 pH meter (Fisher Scientific, UK) (the meter was routinely calibrated using appropriate controls). All restriction endonucleases and buffers were purchased from New England Biolabs (NEB, UK). All other molecular biology reagents were purchased from Fisher Scientific (UK), NEB (UK) or Promega (UK), unless otherwise stated.

5.1.2 Strains and plasmids

Strain	Chromosomal genotype
DH5α	*huA2 lac(del)U169 phoA glnV44 Φ80' lacZ(del)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR1*
BL21 (DE3)	*fhuA2 [lon] ompT gal (λ DE3) [dcm]* Δ*hsdS λ DE3 = λ sBamHIo* Δ*EcoRI-B int::(lacI::PlacUV5::T7 gene1) i21* Δ*nin5*
SNS126	F-, *Δ(araD-araB)567, ΔlacZ4787*(::rrnB-3), *&lambda⁻, Δ(araH-araF)570(::FRT), ΔaraEp-532::FRT, φP_{cp13}araE534, Δ(rhaD-rhaB)568, hsdR514 Δ(hisB), Kan^R, His3, LacZ*

Table 3 List of bacterial strains used and their genotype.

BL21 (DE3) was kindly given by Professor P. Roach, University of Southampton. DH5α was from the *E. coli.* Genetic Stock Center (Yale, UK). SNS126 and pARCBD-SICLOPPS (*Ssp*) were kindly given by Professor S. Benkovic, Pennsylvania State University. pET32a-HIF-1α_{1-350}, pET28a-HIF-1α_{1-350} and pGEX-2TK-HIF-1β_{1-459} were constructed by Dr K. Nordgren, University of Southampton. pARCBD-SICLOPPS (*Npu*) was constructed by J. Townend, University of Southampton. pET28a-HIF-1α-bHLH and pET28a-HIF-α-PAS-A were a kind gift from Dr G. Melillo.

5.1.3 Molecular biology general procedures

Antibiotic concentration used in all cases: ampicillin (10 μg/ml), kanamycin (50 μg/ml), spectinomycin (25 μg/ml (concentration for resistance gene integrated on the chromosome, this was the case for SNS126)), chloramphenicol (35 μg/ml).

5.1.3.1 Preparation of Luria Broth (LB) media

10 g LB powder was used per 250 ml water. The LB powder was suspended in deionised water and autoclaved at 121°C for 20 minutes, the media was left to cool to room temperature prior to use.

5.1.3.2 Preparation of Luria Broth (LB) agar plates

9.25 g of LB agar powder was used per 250 ml water. The LB agar powder was suspended in deionised water and autoclaved at 121°C for 20 minutes, the solution was left to cool to approximately 60°C prior to the addition of antibiotics and pouring into sterile petri dishes.

5.1.3.3 Preparation of SOC media

1 M $MgCl_2$ (500 µl), 2 M $MgSO_4$ (500 µl) and 20% glucose solution (w/v) (500 µl) (all solutions were made under sterile conditions with autoclaved deionised water) were added to 48.5 ml LB media. The solution was mixed, filtered through a sterile filter (Millipore) and stored in aliquots at 4°C.

5.1.3.4 Preparation of E. coli strain cultures and stocks

The required *E. coli* strain from an agar plate or a frozen stock was grown overnight in LB media (10 ml) with the appropriate antibiotic at 37°C with agitation. After overnight incubation, 1 ml of the culture was taken and placed in a sterile tube and DMSO (110 µl) was added. The stock was stored at -80°C.

5.1.3.5 Preparation and transformation into chemically-competent cells

The *E.coli* strain was grown from an agar plate or a frozen stock in LB media (10 ml) with the appropriate antibiotic at 37°C with agitation. After overnight incubation, fresh LB media (25 ml) was inoculated with the overnight culture (250 µl) with no antibiotics. The culture was incubated at 37°C with shaking. When the OD_{600} had reached approximately

0.6 the cultures were centrifuged (3100 rpm, 15 minutes, 4°C). The supernatant was discarded and the cell pellet was resuspended in ice cold TBF I buffer (5 ml) (Table 4). The solution was again centrifuged (3100 rpm, 15 minutes, 4°C) and the supernatant was discarded. The cell pellet was then resuspended in ice-cold TBF II buffer (1 ml) (Table 5). The cell solution was aliquoted (100 µl) into PCR tubes on dry ice and then stored at -80ºC.

Compound	Final concentration (mM)
Potassium acetate	30
Rubidium chloride	100
Calcium chloride	10
Manganese chloride	50
Glycerol	15% (v/v)

Table 4 Composition of TBF I buffer.

All components dissolved in sterile deionised water and adjusted to pH 5.8 with 1% acetic acid (v/v). 15 ml aliquots were stored at -80°C.

Compound	Final concentration (mM)
MOPS	10
Rubidium chloride	10
Calcium chloride	75
Glycerol	15% (v/v)

Table 5 Composition of TBF II buffer.

All components were dissolved in sterile deionised water. 1.5 ml aliquots were stored at -80°C.

When needed for transformation the competent cell aliquots (100 µl) were thawed on ice. Plasmid (5 µl) or ligation mixture (5 µl) was added and the cells were maintained on ice for 30 minutes. The cells were then heat shocked for 30 seconds at 42°C. They were then returned to ice for approximately 2 minutes before being added to pre-warmed SOC (900 µl, section 5.1.3.3). The cultures were incubated at 37°C with shaking for 1 hour prior to being plated on LB agar containing the appropriate antibiotics. The plates were then incubated overnight at 37°C.

5.1.3.6 Preparation of agarose gels and analysis of DNA samples using gel electrophoresis

1% agarose gel (w/v) was prepared by suspending agarose powder (2 g) in TAE buffer (200 ml) (Table 6). Three drops of ethidium bromide solution (0.625 mg/ml) were added to the solution prior to being poured into a mould and left to set.

Component	**Concentration (M) (50 X)**
Tris base	2
Glacial acetic acid	1
EDTA	0.050

Table 6 Composition of TAE buffer (50 X stock solution).

All components were dissolved in deionised water and pH adjusted to pH 8.0 with NaOH solution.

The DNA solutions were mixed with 5 X loading buffer (Green GoTaq reaction buffer, Promega), loaded onto the gel and run in 1 X TAE buffer at 100 V for 30 minutes, or until the dye front reached the bottom of the gel (Table 6). The gels were imaged under UV light.

5.1.3.7 Plasmid purification

The required strain was grown in LB media (10 ml) with the appropriate antibiotic at 37°C, after overnight incubation the culture was centrifuged (3100 rpm, 15 minutes, 4°C). The supernatant was discarded and the plasmids extracted from the cell pellet using a QIAGEN miniprep kit, as per manufacturer's instructions.

5.1.3.8 Polymerase chain reaction

PCR were carried out using GoTaq polymerase (Promega, UK) unless otherwise stated. The following reaction mixture was made (Table 7) and run on the GoTaq polymerase program based on the manufacturer's instructions (Table 8).

Component	Volume (μl)
Water	31.75
GoTaq colourless reaction buffer	10
dNTPs	5
Forward primer (10 pmol/μl)	1
Reverse primer (10 pmol/μl)	1
Template (100-300 ng/μl)	1
GoTaq DNA polymerase	0.25

Table 7 Composition of GoTaq polymerase PCR reaction mixtures.

Step	Temperature	Time	Number of cycles
Initial denaturation	95°C	2 min	1
Denaturation	95°C	30 s	30
Annealing	42-65°C (dependent on melting temperature of primers)	30 s	
Extension	72°C	1 min/kb	
Final extension	72°C	5 min	1
Cool down	4°C	∞	1

Table 8 GoTaq PCR program.

The PCR product was column purified using a gel purification column (section 5.1.3.12). In some cases the PCR product was gel purified using a QIAGEN gel purification kit, as per manufacturer's instructions (section 5.1.3.11).

5.1.3.9 Restriction digestion

Restriction digestions of PCR products and plasmids were carried out by incubating DNA with the restriction endonucleases (NEB, UK) and appropriate buffers determined by the

enzyme manufacturer's instructions. The following digestion mixtures were made and BSA was added if required (per manufacturer's instructions) (Table 9 and Table 10).

Component	Volume (µl)
Plasmid or PCR product (approximately 100 ng/µl)	43
Manufacturer's 10 X buffer	5
Restriction endonuclease 1	1
Restriction endonuclease 2	1

Table 9 Composition of restriction digestion mixture.

Component	Volume (µl)
Plasmid or PCR product (approximately 100 ng/µl)	38
Manufacturer's 10 X buffer	5
BSA (1 mg/ml)	5
Restriction endonuclease 1	1
Restriction endonuclease 2	1

Table 10 Composition of restriction digestion mixture with BSA.

The digestion mixtures were incubated at 37°C for either five hours, or for some restriction enzymes overnight (manufacturer's instructions followed). The restriction endonuclease was then heat inactivated at 65°C for 20 minutes, if required. 10 µl loading buffer (5 X green GoTaq reaction buffer, Promega, UK) was added to the mixtures and they were visualised on 1% agarose by gel electrophoresis (section 5.1.3.6) and imaged under UV.

5.1.3.10 Dephosphorylation of digested plasmids using thermosensitive alkaline phosphatase

Digested plasmids were dephosphorylated using thermosensitive alkaline phosphatase (TSAP, Promega, UK). If the restriction endonucleases had been heat inactivated, TSAP (2 µl) was added directly to the restriction digestion mixture and incubated at 37°C for 1 hour. The TSAP was then inactivated at 74°C for 15 minutes. Finally, the fragment was then analysed on a 1% agarose gel by electrophoresis and gel purified. However, if the

restriction endonuclease was unsuitable for heat inactivation, the plasmid was gel purified (section 5.1.3.11) and then dephosphorylated in the presence of buffer 4 (NEB, UK) (Table 11).

Component	**Volume (µl)**
Gel purified vector	43
Buffer 4 (NEB restriction enzyme buffer)	5
TSAP	2

Table 11 Composition of TSAP mixture.

The TSAP mixture was incubated for 1 hour in a 37°C water bath and then heat inactivated for 15 minutes at 74°C. This solution could be used directly in a ligation.

5.1.3.11 Gel purification of digested DNA fragments

The completed restriction digest mixture was run on a 1% agarose gel (section 5.1.3.6) and the band of interest was excised from the gel. The DNA was extracted from the gel using a QIAquick gel extraction kit (QIAGEN, UK), as per manufacturer's instructions.

5.1.3.12 Column purification of PCR products and digestion products

Purification of PCR products and digestion products was performed using a QIAGEN (UK) purification kit, as per manufacturer's instruction.

5.1.3.13 Ligation

The amount of digested plasmid and insert used for ligations were calculated assuming 30 x 10^6 nmol of plasmid was used for each ligation (10 µl), which was then converted into ng using Equation 4.

$$\text{Amount of vector (ng)} = \text{Size of vector (bp)} \times 660 \text{ (g/mol)} \times 30 \times 10^{6} \text{ (nmol)}$$

Equation 4

Equation 5 was then used to calculate the amount of insert used:

$$\text{Amount of insert (ng)} = \frac{\text{Amount of vector (ng)} \times \text{Size of insert (bp)}}{\text{Size of vector (bp)}} \times \text{Molar ratio insert vector}$$

Equation 5

Molar vector:insert ratios used varied from 1:1 to 1:6. Using the amounts calculated above the ligation mixtures were prepared. The T4 DNA ligase (Promega, UK) was the final component to be added to the mixture (Table 12).

Component	**Volume (μl)**
Digested plasmid	Volume calculated depending on concentration of solution
Insert	Volume calculated depending on concentration of solution
Ligase buffer (Promega)	1
T4 DNA ligase	1
Water	Up to 10 μl

Table 12 Composition of ligation mixture using T4 DNA ligase.

The ligation mixture was incubated overnight at 4°C and then either used directly for transformation into *E.coli*, or heat inactivated for 10 minutes at 70°C and stored at -20°C for transformation at a later time.

5.1.3.14 Colony polymerase chain reaction

GoTaq polymerase (Promega, UK) was used for all colony PCR and set up as described in Table 13.

Component	**Volume (µl)**
Water	6.4
GoTaq green reaction buffer	2
dNTPs	1
Forward primer (10 pmol/µl)	0.25
Reverse primer (10 pmol/µl)	0.25
GoTaq	0.1

Table 13 Composition of GoTaq colony PCR mixture.

A small amount of each colony to be analysed was picked from an LB agar plate and placed into the PCR reaction mixture. This PCR was run using the following program, which followed the manufacturer's instructions (Table 14).

Step	**Temperature**	**Time**	**Number of cycles**
Initial denaturation	95°C	2 min	1
Denaturation	95°C	30 s	30
Annealing	42-65°C (dependent on melting temperature of primers)	30 s	
Extension	72°C	1 min/kb	
Final extension	72°C	5 min	1
Cool down	4°C	∞	1

Table 14 GoTaq colony PCR program.

The mixtures were then analysed on a 1% agarose gel by electrophoresis and visualised under UV (section 5.1.3.6).

5.1.3.15 Drop-spotting of RTHS strains

Minimal media plates were prepared by adding minimal media salts (Table 15), glycerol, $MgSO_4$, spectinomycin and sterile deionised water to agar (Table 16). The media was then supplemented as necessary with kanamycin, 3-AT, IPTG and L-arabinose.

Reagents	**Amounts**
$(NH_4)_2SO_4$	5.0 g
KH_2PO_4	22.5 g
K_2HPO_4	52.5 g
Sodium citrate	2.5 g
Sterile deionised water	up to 500 ml

Table 15 Composition of minimal media salts.

Reagent	**Volume**	**Final concentration**
Agar (3.75 g in 200 ml water)	200 ml	N/A
Minimal media (10 X)	25 ml	1 X
50% glycerol	10 ml	12.5%
$MgSO_4$	0.25 ml (1 M stock)	1 mM
Spectinomycin	0.625 ml (stock)	25 μg/ml
Sterile deionised water	Up to 250 ml	N/A

Table 16 Composition of minimal media agar (250 ml) used for drop-spotting.

Overnight cultures (10 ml LB media) were grown and a 10-fold dilution series was prepared in 10% glycerol. 2.5 μl of each serial dilution was drop-spotted onto the minimal media plates. The plates were then incubated at 37°C for 48-72 hours.

5.1.3.16 Chitin bead purification of inteins

100 ml LB media was inoculated with 1% overnight culture of the alanine analogue strains (pARCBD alanine analogues in SNS126). The cultures were incubated at 37°C until an OD_{600} of 0.6 was reached. The cultures were then induced with 0.5% L-arabinose and incubated for a further 2 hours at 37°C. Following centrifugation (3100 rpm, 15 min), the supernatant was removed and the pellets were flash frozen and stored at -80°C until required. The pellets were thawed and resuspended in chitin buffer (4 ml/g) (Table 17).[131]

Component	Final Concentration (mM)
Tris-HCl	20
NaCl	5
TCEP	1

Table 17 Composition of Chitin buffer.[131]

The cells were lysed by sonication and the insoluble inclusion body was then pelleted by centrifugation (8000 rpm, 45 minutes, 4°C). The soluble fraction was filtered and added to 200 µl of chitin beads (NEB) (pre-equilibrated in chitin buffer). The beads were incubated at 4°C with gentle agitation for 1 hour, and then washed three times with chitin buffer (2500 rpm, 5 minutes, acceleration 1). After washing, the bead pellet was incubated at room temperature for 16 hours to allow the inteins to splice. 20 µl of the bead slurry was taken and analysed by SDS-PAGE using a 15% acrylamide SDS gel.

5.1.3.17 Preparation of SDS-PAGE gel and analysis of protein samples by SDS-PAGE

SDS-PAGE (8 or 15%) gels were made (Table 18, Table 19 and Table 20). 10 µl of loading buffer (Table 21) was added to 10 µl protein sample and the solutions were denatured by heating to 100°C for 10 minutes. The samples were loaded onto the gel and run in SDS running buffer at 150 V for 40-50 minutes (Table 22).

Component	Volume (ml)
Water	11.5
ProtoGel, 30% acrylamide (v/v) (National Diagnostics)	6.7
1.5 M Tris (pH 8.8)	6.3
10% SDS	0.25
10% Ammonium persulphate (APS)	0.25
Tetramethylethylenediamine (TEMED)	0.015

Table 18 Composition of 8% resolving gel for SDS-PAGE.

Component	Volume (ml)
Water	7.2
ProtoGel, 30% acrylamide (v/v) (National Diagnostics)	15
1.5 M Tris (pH 8.8)	7.5
10% SDS	0.3
10% APS	0.15
TEMED	0.02

Table 19 Composition of 15% resolving gel for SDS-PAGE

Component	Volume (ml)
Water	6.8
ProtoGel, 30% acrylamide (v/v) (National Diagnostics, UK)	1.7
1 M Tris (pH 6.8)	1.25
10% SDS	0.1
10% APS	0.1
TEMED	0.01

Table 20 Composition of stacking gel for SDS-PAGE.

Component	Concentration (mM)
Tris-HCl (pH 6.8)	100
SDS	4%
Glycerol	20%
Dithiothreitol (DTT)	50
Bromophenol blue	Trace

Table 21 Composition of SDS-PAGE loading buffer, the components were added to sterile deionised water.

Component	Concentration (M)
Tris base	0.125
Glycine	1.25
SDS	0.5%

Table 22 Composition of 5 X SDS running buffer.

The SDS-PAGE gel was then stained for eight minutes using coomassie blue stain (Table 23), which non-specifically binds to all the proteins aiding protein visualisation. Prior to being incubated in destain (Table 24) until bands could be visualised.

Composition	Amount
Coomassie blue	2 g dissolved in 250 ml water
Acetic acid	75 ml
Methanol	500 ml
Water	Make up to 1 l

Table 23 Composition of coomassie blue stain.

Composition	Amount
Acetic acid	200 ml
Methanol	100 ml
Water	800 ml

Table 24 Composition of SDS-PAGE destain.

5.1.3.18 LongAmp Taq polymerase chain reaction to verify plasmid integration onto chromosome

To confirm the integration of pAH68 plasmids, a PCR was carried out using integration verification primer 1 and integration verification primer 4 which anneal to the host SNS126 chromosome (Table 25).[188]

Primer	**Sequence**
Integration verification primer 1	GGA ATC AAT GCC TGA GTG
Integration verification primer 4	GGC ATC AAC AGC ACA TTC

Table 25 Sequences of P1 and P4 primers used for LongAmp PCR.[188]

The following reaction mixtures were made (Table 26) and run on the LongAmp *Taq* PCR program based (Table 27) on the manufacturer's instructions.

Component	**Volume (μl)**
5 X LongAmp buffer (NEB, UK)	5
dNTPs	0.75
Integration verification primer 1 (100 pmol/μl)	1
Integration verification primer 4 (100 pmol/μl)	1
Strain culture in LB media	2
LongAmp *Taq* DNA polymerase (NEB, UK)	1
Water	14.75

Table 26 Composition of LongAmp *Taq* DNA polymerase PCR reaction mixtures.

Step	Temperature	Time	Number of cycles
Initial denaturation	94°C	30 s	1
Denaturation	94°C	30 s	30
Annealing	51°C	30 s	
Extension	65°C	6 min	
Final extension	65°C	10 min	1
Cool down	4°C	∞	1

Table 27 LongAmp *Taq* PCR program

The PCR products were then analysed by agarose gel electrophoresis (section 5.1.3.6).

5.1.3.19 Expression of His and GST recombinant proteins

BL21 (DE3) competent cells were transformed with an expression vector and plated on LB agar. The strain was grown from the LB agar plate overnight in LB media (10 ml) with the appropriate antibiotic at 37°C with shaking. A 1% culture in LB media (200-1000 ml) was made from the overnight culture and incubated at 37°C with shaking. OD_{600} was monitored and when optical density reached approximately 0.6 the culture was induced with 0.1 mM IPTG. The culture was then incubated at 18°C overnight with shaking. After overnight incubation the culture was spun down (3100 rpm, 10 min), the supernatant was removed and the pellets were stored at -80°C until required.

5.1.3.20 Purification of His-tagged HIF-1α recombinant proteins (Trx-His-HIF-1α_{1-350}, Trx-His-HIF-2α_{1-360}, His-HIF-1αPAS-$B_{238-349}$, His-HIF-1αPAS-$B_{238-349}$ 1106, His-HIF-1α-PAS-A, His-HIF-1α bHLH

The cell pellets from a 300 ml expression culture generated during protein expression were thawed on ice and resuspended in MHW buffer (Table 28) (approximately 4 ml buffer per 1 g cell pellet).[118] Lysozyme (100 µg/ml) was added to the lysate and the solution was incubated with shaking for 20 minutes at 5°C. The lysate was then sonicated twice (6

cycles of 10 seconds on then 10 seconds off) and the solution was then spun down (8000 rpm, 45 minutes, 4°C). The soluble and insoluble layers were then separated; the soluble layer was filtered through a sterile filter (Millipore). During this time Ni Sepharose 6 fast flow beads (300 μl, GE Healthcare, UK) were prepared by washing with deionised water twice. The filtered soluble layer was added to the beads and incubated with shaking for 1 hour at 5°C. The beads were then spun down (2500 rpm, 5 minutes, 4°C, 1 acceleration), the supernatant containing unbound protein was removed and retained for analysis by SDS-PAGE. Ni wash buffer 1 (Table 29) (2 ml) was then added to the beads and they were incubated for 20 minutes with shaking, the beads were again spun down (2500 rpm, 5 min, 4°C, 1 acceleration) and the supernatant removed. The wash was repeated, after the wash buffer had been removed, Ni elution buffer (Table 30) (3 ml) was added to the beads and they were incubated with shaking for 1 hour 30 minutes at 5°C. Finally, the beads were spun down (2500 rpm, 5 minutes, 4°C, 1 acceleration) and supernatant containing the purified protein removed.

Component	**Concentration (mM)**
Tris	20
Imidazole	10
Sodium chloride	400
Sodium orthovanadate	1
Sodium fluoride	10
DTT	1
Protease inhibitor cocktail (Complete, EDTA-free protease inhibitor cocktail tablets (Roche))	25 X solution made from tablets and added to buffers
Glycerol	20% (v/v)
Triton X-100	1% (v/v)

Table 28 Composition of MHW lysis buffer, all components added to sterile deionised water and adjusted to pH 7.4 with dilute HCl solution.[118]

Component	Concentration (mM)
Tris	20
Imidazole	20
Sodium chloride	500
Glycerol	10% (v/v)
Triton X-100	1% (v/v)

Table 29 Composition of Ni wash buffer 1, all components added to sterile deionised water and adjusted to pH 7.4 with dilute HCl solution.[118]

Component	Concentration (mM)
Tris	20
Imidazole	500
Sodium chloride	200
Glycerol	5% (v/v)
Triton X-100	0.1% (v/v)

Table 30 Composition of Ni elution buffer, all components added to sterile deionised water and made to pH 7.4 with dilute HCl solution.

5.1.3.21 Purification of GST-tagged recombinant proteins (GST-HIF-β_{1-459} and pET28-Trx-His-HIF-1α_{1-350}/ GST-HIF-β_{1-459} co-expression strain)

The cell pellets from a 300 ml expression culture generated during protein expression were thawed on ice and resuspended in GST lysis buffer (Table 31, approximately 4 ml buffer per 1 g cell pellet). Lysozyme (100 μg/ml) was added to the lysate and the solution was incubated with shaking for 20 minutes at 5°C. The lysate was then sonicated twice (6 cycles of 10 seconds on then 10 seconds off), the solution was then spun down (8000 rpm, 45 min, 4°C). The soluble and insoluble layers were separated and the soluble layer was filtered through a sterile filter (Millipore). During this time GST Sepharose 4B beads (GE Healthcare) (300 μl) had been washed twice with tris buffer (Table 22). The filtered soluble layer was added to the beads, the beads were incubated with shaking for 1 hour at 5°C. The beads were then spun down (2500 rpm, 5 minutes, 4°C, 1 acceleration), the

supernatant containing unbound protein was removed and retained for analysis by SDS-PAGE. GST lysis buffer (2 ml) was added to the beads and they were incubated for 20 minutes with shaking, the beads were then spun down (2500 rpm, 5 minutes, 4°C, 1 acceleration) and the supernatant removed. The wash was repeated. GST elution buffer (Table 32) (3 ml) was then added to the beads, which were then incubated with shaking for 1 hour 30 minutes at 5°C. The beads were spun down (2500 rpm, 5 minutes, 4°C, 1 acceleration) and supernatant containing the purified protein isolated. The beads and elution solution were analysed by SDS-PAGE.

Component	**Concentration (mM)**
Tris	50
Sodium chloride	150
Sodium orthovanadate	1
Sodium fluoride	10
DTT	1
Protease inhibitor cocktail (Complete, EDTA-free protease inhibitor cocktail tablets (Roche))	25 X solution made from tablets and added to buffers
Triton X-100	0.1% (v/v)

Table 31 Composition of GST lysis buffer, all components added to sterile deionised water and adjusted to pH 7.5 with dilute HCl solution.[118]

Component	**Concentration (mM)**
Tris	50
Reduced glutathione	10

Table 32 Composition of GST elution buffer, tris added to sterile deionised water and adjusted to pH 8.0 with dilute HCl, reduced glutathione then added.[118]

5.1.3.22 Desalting protein solution using PD 10 desalting size exclusion column

To desalt protein solutions, PD10 columns (GE Healthcare, UK) were used, the manufacturer's instructions were followed. The gravity method was used in all cases.

5.1.3.23 Calculating protein concentration using a Bradford reagent assay

A colorimetric Bradford reagent assay was used to calculate protein concentration. BSA control protein samples were analysed each time to allow the most accurate estimation of concentration (Table 33). Protein samples (50 µl) were added to cuvettes as shown below and Bradford reagent (Sigma Aldrich, UK) (1.5 ml) was added to the protein samples and mixed.

Protein sample	**Measurement**
Water	Zero at A_{595}
BSA (0.25 mg/ml)	Measure absorbance at A_{595}
BSA (0.5 mg/ml)	
BSA (1.0 mg/ml)	
BSA (1.4 mg/ml)	
Protein buffer	Zero at A_{595}
Protein sample of interest	Measure absorbance at A_{595}

Table 33 Protein samples measured for Bradford assay.

The samples were then incubated for 10-30 minutes at room temperature before being measured at 595 nm. The protein concentration was then calculated by plotting a standard curve of known BSA concentrations.

5.1.3.24 Electrophoretic mobility shift assay

Proteins were purified (sections 5.1.3.20 and 5.1.3.21), desalted using PD10 columns (section 0) and concentration calculated using a Bradford reagent assay (section 5.1.3.23). The EMSA samples were then made up (Table 34 and Table 35).

Sample	**Poly dI-dC**	**BSA**	**Protein 1 (Trx-His-HIF-1α_{1-350})**	**Protein 2 (GST-HIF-1β_{1-459})**	**Labelled probe [^{32}P HRE] (0.03 µm)**	**EMSA buffer**
X	0.5 µl (50 ng/µl stock)	1 µl (5 mg/ml stock)	Desired concentration	Desired concentration	1 µl (0.3 µl stock solution	Make each sample up to 10 µl

Table 34 Example of the composition of EMSA samples.

Component	**Concentration (mM)**
Tris	10
Potassium chloride	50
Sodium chloride	50
DTT	5
Magnesium chloride	1
EDTA	1
Glycerol	10% (v/v)

Table 35 Composition of EMSA buffer, components added to sterile deionised water and adjusted to pH 7.4 with dilute HCl.

The samples were made and the labelled probe added last. Incubation times between protein addition and DNA addition to the samples varied between experiments. Once the DNA was added the samples were incubated at room temperature for between 5 minutes to 90 minutes, this was varied between experiments. EDTA free Ficoll (GE Healthcare, UK)

loading buffer (5 μl) was then added to each sample. The samples (8 μl) were then loaded into an 8% acrylamide non-denaturing gel in a cold room (5°C) (Table 36).

Component	Volume (ml)
AccuGel (40% acrylamide, National diagnostics, UK)	10
5 X TBE buffer	10
10% APS	0.4
TEMED	0.04
Deionised water	Make up to 50 ml

Table 36 Composition of 8% non-denaturing acrylamide gel used for EMSA.

The gel was run at 300 V, in TBE buffer, at 5°C for approximately 2 hours 30 minutes until the dye front was close to the bottom of the gel (Table 37). The gel was then submerged in 10% acetic acid/water (v/v) on one glass plate for 10 minutes, in order to fix the DNA in the gel. The gel was then transferred to a piece of filter paper, wrapped in saran wrap and dried *in vacuo*. The gel was exposed to a phosphor screen overnight; the screen was then imaged using a phosphorimager.

Component	Concentration (mM)
Tris base	89
Boric acid	89
EDTA	2

Table 37 Components of TBE buffer. A 5 X stock solution was made, in deionised water, the stock solution was diluted to 1 X in deionised water.

5.1.3.25 Site directed mutagenesis

SDM mixtures were made as outlined in Table 38, *Pfu* DNA polymerase was used (Promega, UK) and the PCR run on the SDM program based on manufacturer's instructions (Table 39).

Component	Volume (μl)
Pfu buffer (Promega)	5
dNTPs	4
Forward primer (100 pmol/μl)	0.2
Reverse primer (100 pmol/μl)	0.2
Template plasmid (100 μg/μl)	1
Pfu DNA polymerase	2
Water	37.6

Table 38 composition of *Pfu* DNA polymerase PCR mixture for SDM.

Step	Temperature	Time	Number of cycles
Initial denaturation	95°C	1 min	1
Denaturation	95°C	30 s	18
Annealing	56-63°C (dependent on primer sequence)	30 s	
Extension	72°C	13 min	
Final extension	72°C	5 min	1
Cool down	4°C	∞	1

Table 39 *Pfu* DNA polymerase PCR program.

Once the PCR was complete, a small sample was run on an agarose gel to confirm the amplification was successful (section 5.1.3.6). *Dpn*I (2 μl) (NEB) was added to the PCR solution and incubated for 2 hours at 37°C. The *Dpn*I treated mixture was then column purified (section 5.1.3.12). The eluted plasmid was transformed into DH5α chemically-competent cells (following the procedure outlined in section 5.1.3.5 with the exception that only 450 μl SOC was added to the cells after heat shock). Colonies were grown (section 5.1.3.4) and the plasmids purified (section 5.1.3.7) and sent for sequencing to confirm successful SDM.

5.1.4 Molecular biology-specific experimental procedures

5.1.4.1 Construction of the Ssp pARCBD SICLOPPS plasmids

Primers were designed to incorporate codons for the desired cyclic peptides (Table 40). The SICLOPPS plasmid (pARCBD *Ssp*), was used as a template for PCR and inserts were amplified encoding each cyclic peptide and inteins. GoTaq polymerase (Promega, UK) was used with an annealing temperature of 55°C and an extension time of 35 seconds. The inserts and pARCBD *Ssp* backbone were digested with *Hin*dIII and *Bgl*I, and the fragments were then purified and ligated using T4 DNA ligase with a vector:insert ratio of 1:6. The ligation mixtures were transformed into DH5α chemically-competent cells and plated onto selective LB agar. Colonies were grown and the plasmids purified and verified by sequencing.

Cyclic peptide	Forward primer (*Bgl*I restriction site)	Reverse primer (*Hin*dIII restriction site)
CLLFVY	GGAATTCGCCAATGGGGCGATCGCCCACAATT GCCTGCTGTTTGTGTATTGCTTAAGTTTTGGC	GGAATTCAAGCTTTCATTGAAGCTG CCACAAGG
CALFVY	GGAATTCGCCAATGGGGCGATCGCCCACAATT GCGCGCTGTTTGTGTATTGCTTAAGTTTTGGC	GGAATTCAAGCTTTCATTGAAGCTG CCACAAGG
CLAFVY	GGAATTCGCCAATGGGGCGATCGCCCACAATT GCCTGGCGTTTGTGTATTGCTTAAGTTTTGGC	GGAATTCAAGCTTTCATTGAAGC TGCCACAAGG
CLLAVY	GGAATTCGCCAATGGGGCGATCGCCCACAATT GCCTGCTGGCGGTGTATTGCTTAAGTTTTGGC	GGAATTCAAGCTTTCATTGAAGCTG CCACAAGG
CLLFAY	GGAATTCGCCAATGGGGCGATCGCCCACAATT GCCTGCTGTTTGCGTATTGCTTAAGTTTTGGC	GGAATTCAAGCTTTCATTGAAGCTG CCACAAGG
CLLFVA	GGAATTCGCCAATGGGGCGATCGCCCACAATT GCCTGCTGTTTGTGGCGTGCTTAAGTTTTGGC	GGAATTCAAGCTTTCATTGAAGCTG CCACAAGG

Table 40 Primers used to construct alanine scanning *Ssp* pARCBD SICLOPPS plasmids.
The forward primer encodes a *Bgl*I restriction site. The reverse primer encodes a *Hin*dIII restriction site. The cyclic peptide sequences are denoted in red and the restrictions sites underlined.

5.1.4.2 Construction of the Npu pARCBD SICLOPPS plasmids

Primers were designed to incorporate codons for the desired cyclic peptides (Table 41). The SICLOPPS plasmid (pARCBD *Npu*), was used as a template for PCR and inserts were amplified encoding each cyclic peptide and inteins. GoTaq polymerase used with an annealing temperature of 55°C and an extension time of 40 seconds.

Cyclic peptide	Forward primer (*Nhe*I restriction site)	Reverse primer (*Hin*dIII restriction site)
CLLFVY	CATTGCTAGCAACTGCCTGCTGTTTG TGTATT	GGAATTCAAGCTTTCATTGAAGCTGCCACA G
CALFVY	CATTGCTAGCAACTGCGCGCTGTTTG TGTATT	GGAATTCAAGCTTTCATTGAAGCTGCCAC AAGG
CLAFVY	CATTGCTAGCAACTGCCTGGCGTTTG TGTATT	GGAATTCAAGCTTTCATTGAAGCTGCC ACAAGG
CLLAVY	CATTGCTAGCAACTGCCTGCTGGCG GTGTATT	GGAATTCAAGCTTTCATTGAAGCTGCCAC AAGG
CLLFAY	CATTGCTAGCAACTGCCTGCTGTTTG CGTATT	GGAATTCAAGCTTTCATTGAAGCTGCCAC AAGG
CLLFVA	CATTGCTAGCAACTGCCTGCTGTTTG TGGCGT	GGAATTCAAGCTTTCATTGAAGCTGCCAC AAGG

Table 41 Primers used to construct alanine scanning *Npu* pARCBD SICLOPPS plasmids. The forward primer encodes an *Nhe*I restriction site. The reverse primer encodes a *Hin*dIII restriction site. The cyclic peptides sequences are denoted in red and the restrictions sites underlined.

The inserts and pARCBD *Npu* backbone were digested with *Hin*dIII and *Nhe*I, and the fragments were purified and ligated using T4 DNA ligase with a vector:insert ratio of 1:6. The ligation mixtures were transformed into DH5α chemically-competent cells and plated onto selective LB agar. Colonies were grown and the plasmids purified and verified by sequencing.

5.1.4.3 Construction of pET28a-Trx-His-HIF-1α$_{1-350}$

pET32a-Trx-His-HIF-1α$_{1-350}$ and pET28a were digested with *Xba*I and *EagI* (NEB, buffer 4 and BSA) overnight at 37°C (section 5.1.3.9). The enzymes were heat inactivated and the

digested vector was dephosphorylated with TSAP (section 5.1.3.10). The Trx-His-HIF-1$\alpha_{1\text{-}350}$ insert and pET28 vector were gel purified and ligated overnight at 4°C with a vector: insert ratio of 1:2 (sections 5.1.3.11 and 5.1.3.13). The ligation mixtures were transformed in chemically-competent DH5α and grown overnight on LB agar containing kanamycin (section 5.1.3.5). Colony PCR confirmed all colonies contained the insert; sequencing confirmed the desired sequence was present (section 5.1.3.14).

5.1.4.4 Co-expression of Trx-His-HIF-1$\alpha_{1\text{-}350}$ and GST-HIF-1$\beta_{1\text{-}459}$

To construct coexpressed strains, one plasmid was transformed into BL21 (DE3) and then grown (section 5.1.3.5 and 5.1.3.4). Chemically-competent cells were then generated from this strain and the transformation was repeated for the second plasmid, these cells were plated on LB agar plates containing both ampicillin and kanamycin.

5.1.4.5 His and GST Western blot for recombinant protein/inhibition pull down

Samples were prepared as for an SDS-PAGE in SDS loading buffer (section 5.1.3.6), the Western blot protocol was then followed from the heat inactivation step (section 5.2.3.3). Prior to incubation in primary antibody the membrane was cut just above the 58 kDa band of the protein ladder. The bottom half of the membrane where Trx-His-HIF-1$\alpha_{1\text{-}350}$ should be, was exposed to an anti-His primary antibody and the top of the membrane where GST-HIF-1$\beta_{1\text{-}459}$ should be, was exposed to an anti-GST primary antibody. The antibodies and concentrations used are described below (Table 42).

Antibody and target	Concentration	Manufacturers details
Anti-His, mouse monoclonal (*Primary)*	1:2500	Genscript, USA: A00186
Anti-GST, mouse monoclonal (*Primary)*	1:1000	Abcam, USA: Ab18183
Green goat anti-mouse *(Secondary)*	1:20,000	IRDye antibodies, LI-COR, USA: 926-32210

Table 42 Antibodies used in Western blot experiment, concentrations used and manufacturer's details stated.

5.1.4.6 5' labelling HRE probe using [γ-32P] ATP for EMSA

PNK (NEB, UK) was used to 5'label the forward strand (Table 43) of the HRE probe. The following PNK mixture was made (the PNK was added to the mixture last) (Table 44).

HRE probe strand	Sequence
Forward	5'-GGCTGGGCCTACGTGCTGTCTCACACAGCC-3'
Reverse	3'-GGCTGTGTGAGACAGCACGTAGGCCCAGCC-5'

Table 43 Sequence of forward and reverse strand of HRE probe.

HRE binding site is underlined.

Component	Volume (µl)
Water	15
30 nucleotide HRE forward DNA strand (60 µM)	1
PNK buffer	1
[γ-^{32}P] ATP	1
PNK (NEB, UK)	2

Table 44 PNK reaction mixture

The mixture was incubated for 1 hour 30 minutes and then 10 µl loading buffer was added and the mixture was run on a denaturing acrylamide gel in order to purify the insert (Table 45). The gel was run in 1 X TBE at 400 V for approximately 1 hour 30 minutes.

Components	**Volume (ml)**
Ureagel (Aqueous acrylamide/ urea solution, National diagnostics, UK)	25
5 X TBE	10 ml
20% APS	0.2
TEMED	0.04
Water	Make up to 50 ml

Table 45 Composition of denaturing polyacrylamide gel.

The gel was exposed to X-ray film and the band containing labelled probe was excised from the gel and gel purified. The DNA was extracted from the gel in TE buffer (10 mM tris, 1 mM EDTA) overnight. The solution was then filtered through glass wool to remove remaining gel, 50 µl of 3 M sodium acetate was added to the solution followed by 1 ml ethanol. The DNA was then precipitated from the ethanol on dry ice for 3 hours. The solution was spun down (15 minutes, 13,000 rpm), the supernatant was discarded and the pellet washed with 50 µl 80% ethanol. The DNA pellet was resuspended in TE buffer (10 mM tris, 0.1 mM EDTA) to make a 1 ci/s/µl solution of 5' [^{32}P] labelled single stranded HRE probe. The labelled probe was annealed to its complementary reverse strand, by addition of 1 µl of 60 µM reverse strand and heating to 95°C for 15-30 minutes. The solution finally then cooled slowly to room temperature overnight before being stored at -20°C and used for the EMSA.

5.1.4.7 Construction of pET28a-HIF-1α-PAS-B$_{238-349}$ 1106 mutant plasmid by SDM

Four SDM reactions were performed to pET28-HIF-1α-PAS-B$_{238-349}$ (R245E, E266H, R311H, S330L) (Table 46). Each SDM was carried out after successful sequencing of the previous one (section 5.1.3.25). The annealing temperatures for each mutation for each PCR are indicated in the table (Table 46).

Mutation	**Primer**	**Sequence**	**Annealing temperature (ºC)**
R245E	F	GCAAGACTTTCCTCAGTGAACACAGCCTGGATATG	63
	R	CATATCCAGGCTGTGTTCACTGAGGAAAGTCTTGC	
E266H	F	CCGAATTGATGGGATATCATCCAGAAGAACTTTTAGGCCGC	56
	R	GCGGCCTAAAAGTTCTTCTGGATGATATCCCATCAATTCGG	
R311H	F	GGATGCTTGCCAAACATGGTGGATATGTCTGGG	57
	R	CCCAGACATATCCACCATGTTTGGCAAGCATCC	
S330L	F	CACCAAGAATCTGCAACCACAGTGC	57
	R	GCACTGTGGTTGCAGATTCTTGGTG	

Table 46 Primer sequences for SDM of pET28a-HIF-1α-PAS-B$_{238-349}$ to construct HIF-1α-PAS-B$_{238-349}$ 1106.

Codon changes to each mutation are underlined. Annealing temperatures for PCR are also indicated.

5.1.4.8 MST: Protein labelling with Monolith™ NT-647 Red-NHS (amine reactive)

His-HIF-1α-PAS-B$_{238-349}$ 1106 was labelled with Monolith™ NT-647 Red-NHS (amine reactive) (NanoTemper Technologies, Germany). 20 μM (100 μl) of protein was labelled, as per the manufacturer's instructions. The labelled protein was purified using gravity flow

and spin columns following manufacturer's instructions. The protein was eluted in MST optimised buffer (Table 47).

Component	Concentration (mM)
Tris (pH 7.6)	50
NaCl	150
$MgCl_2$	10
TWEEN-20	0.05%

Table 47 Composition of MST optimised buffer

5.1.4.9 MST experiments

MST experiments were run on a Monolith™ NT.115 (NanoTemper technologies, Germany). 16 serial dilutions of the non-labelled peptides (five alanine analogues) were made in MST optimised buffer (Table 47) (100 μM-3 nM). 20 μl of these dilutions was mixed with 20 μl of the NT-647 labelled His-HIF-1α-PAS-$B_{238-349}$ 1106. The solutions were mixed and left at room temperature for 5-10 minutes. The samples were then loaded into Monolith hydrophilic capillaries (NanoTemper Technologies) and the scan run.

Monolith™ NT.115 parameters:

Laser on time: 30 seconds

Laser off time: 5 seconds

LED power: 70%

MST power: 40%

5.2 Mammalian cell culture

5.2.1 Mammalian cell-materials and equipment

SDS-PAGE gels were run using a mini-PROTEAN Tetra electrophoresis system (Bio-Rad, USA). Western blot gel to membrane transfers were carried out using a mini-trans-blot cell and power pack (Bio-Rad, USA). Western blots were imaged using an ODESSEY LI-COR imaging system and analysed using ODESSEY software (LI-COR, USA). Cell cultures were spun down using a Sorvall® Legend® centrifuge (Thermo scientific, UK). Samples prepared for Western blot were spun down using a Heraeus Fresco 17 centrifuge (Thermo scientific, UK).

Dulbecco's modified eagles media (DMEM, GlutaMAX), Trypsin-EDTA (0.5%) and PBS solutions were purchased from Life Technologies, UK. DFX was purchased from Sigma Aldrich, UK. Transfections were done using FuGENE (Promega, UK). Firefly and *Renilla* luciferase activity was measured using the Dual-luciferase reporter assay system (Promega, UK) and the luminescence and absorbance was measured using a VarioScan Flash plate reader (Thermo Scientific, UK).

5.2.2 Cell lines and plasmids

MCF-7 cells and pEGFP-N12.13 plasmid were kindly donated by Dr Sarah Bailey, University of Southampton. pGL3-control vector and pRL-SV40 plasmid were purchased from Promega, UK. pGL2-TK-HRE was kindly donated by Professor Graham Packham, University of Southampton.[142]

5.2.3 Mammalian cell culture-general experimental procedures

5.2.3.1 MCF-7 cell maintenance

MCF-7 cells were grown in DMEM media supplemented with FBS (10%, Life technologies, UK). The cells were incubated in 75 cm^2 vented flasks (Corning®, USA) in 5% CO_2/95% humidity at 37ºC. Cells were sub-cultured by firstly aspirating the media and then washing the cells with PBS (10 ml). They were then covered with trypsin-EDTA (3 ml), which was then aspirated immediately. The cells were then incubated at 37ºC until they had all detached (2-5 minutes). Finally, FBS supplemented DMEM (10 ml) was added to the flask to deactivate the trypsin. The cells were diluted to the required concentration and placed in a new culture flask (for routine passaging a 1 in 5 dilution was done).

5.2.3.2 Preparation of MCF-7 stocks and thawing cells from stocks

Cells were trypsinised and suspended in DMEM/FBS as previously described (section 5.2.3.1) and transferred to a centrifuge tube. The cells were centrifuged (300 g, 5 minutes). The supernatant was removed and the cell pellet was resuspended in FBS/DMSO (10%) to a concentration of approximately 2 x 10^6 cells/ml. The cells solution was aliquoted (1 ml) into cryovials and cooled in a Coolcell (biocision, USA) at a rate of 1°C/minute until -80°C was reached. The stocks were then stored at -80°C until required.

To regrow the cells from the stocks, warm DMEM/FBS (10%) was added to the frozen cells. Once thawed the cells were transferred to a centrifuge tube and DMEM/FBS (10%) was added up to 15 ml. The cells were centrifuged (300 g, 5 minutes). The supernatant was removed and the cell pellet was resuspended in FBS/DMSO (10%) and transferred to 75 cm^2 vented flasks (Corning®, USA) in approximately 13 ml media, then incubated at 37°C.

When cells required counting, a cell counter was used (Z™ series Beckman Coulter®, UK) following manufacturer's instructions.

5.2.3.3 Immunoblotting: Western blot

Cells were trypsinised, as previously described (section 5.2.3.1), transferred to an eppendorf tube and centrifuged (6000 rpm, 1 minute). The supernatant was removed and the cells washed once with PBS (1 ml). The pellet was resuspended in radioimmunoprecipitation assay (RIPA) buffer (300 μl) and incubated on ice for 30 minutes. After 15 minutes the solution was pipetted up and down to aid lysis (Table 48).

Component	**Concentration (mM)**
NaCl	750
IgePalCA^{650}	5%
Sodium deoxychlorate	2.5%
SDS	0.5%
Tris base (pH 8.0)	250

Table 48 Composition of 5 X RIPA buffer.

The samples were sonicated for 30 seconds using an autotune series high intensity ultra-sonic processor, 500 watt model Vibra-Cell™ (Sonics, USA) at 20% amplitude. The samples were then centrifuged to pellet cell debris (13,000 rpm, 5 minutes). The protein concentration of samples was quantified using a bicinchoninic acid assay kit (BCA) (Pierce™, Thermo Scientific), the manufacturer's instructions were followed and BSA samples (0-2,000 μg/ml) were used as standards. The assay was performed in a 96-well plate and the absorbance was measured at 563 nm.

20 μl samples were made by diluting 30 μg protein with water, 3X sample buffer containing DTT (Cell signalling technology, USA, 10 μl) was added. The sample was then heated to 100ºC for 10 minutes. The samples were run on a Mini-PROTEAN® TGX™ Any kD™ precast SDS-PAGE gel (Bio-Rad, USA), in SDS-PAGE Western running buffer at 200 V for approximately 35 minutes (Table 49).

Component	Concentration (mM)
Tris	250
Glycine	1920
SDS	10%

Table 49 Composition of 10 X SDS PAGE Western running buffer, the components were dissolved in sterile deionised water, pH was adjusted to pH 8.3 with dilute HCl.

The immobilised proteins were then transferred from the SDS-PAGE gel to a PVDF membrane. The membrane was soaked in methanol for 2 minutes and both the membrane and gel were soaked in transfer buffer, the membrane and SDS-PAGE gel were then sandwiched between filter paper and the protein transferred to the membrane by electrophoresis for 1 hour 30 minutes at 75 V in transfer buffer (Table 50).

Component	Composition (%)
10 X SDS-PAGE Western running buffer	10
Methanol	20

Table 50 Composition of transfer buffer, components dissolved in sterile deionised water.

On completion of transfer the membrane was removed from the cassette. The membrane was then submerged in blocking buffer and incubated with agitation for 1 hour (Table 51).

Component	Composition (%)
Non-fat dried milk powder	5
TWEEN-20	0.05

Table 51 Composition of blocking buffer, components dissolved in PBS.

The membrane was then incubated in primary antibody (the antibodies were diluted in blocking buffer (3 ml)) overnight at 4°C. The membrane was then washed three times with PBS/TWEEN-20 (0.05%) and then exposed to secondary IRDye® antibodies (LI-COR, USA). The secondary antibody was diluted in blocking buffer (25 ml) with 0.02% SDS.

The membrane and secondary antibody solution were agitated in the dark for 1 hour at room temperature. Again, the membrane was then washed three times with PBS/TWEEN (0.05%) and once with PBS before being imaged using a LI-COR Odyssey imager (LI-COR, USA).

5.2.3.4 MCF-7 transfection using Fugene-6

MCF-7 cells were transfected using FuGENE 6 (Promega) following manufacturer's instructions. Cells were plated one day prior to transfection, confluency at the time of transfection was to be 50-70%. To transfect one well of a six well-plate, FuGENE 6 was added to 200 µl of warm serum free DMEM and mixed. The DNA was then added, the solution was again gently mixed and incubated at room temperature for 15 minutes. The transfection mixture was subsequently added drop-wise to the cells and the plate was gently swirled to ensure even distribution of transfection complexes. The cells were then incubated until required. Successful transfection was monitored by transfection of 1 µg of pEGFP-N12.13 (GFP plasmid) and visualised using a fluorescence microscope.

5.2.4 Mammalian cell culture-specific experimental procedures

5.2.4.1 HRE-dependent luciferase assay

200,000 MCF-7 cells were plated per well of a six well plate and incubated for 24 hours to achieve approximately 50%-70% confluency. Each well was then transfected with 1 µg of DNA (pGL2-TK-HRE (HRE dependent firefly plasmid) or pGL3 control (firefly control plasmid):PRL-SV40 (*Renilla*) at a ratio of 1.25:98) using FuGENE 6 (Promega, UK) as the transfection reagent (section 5.2.3.4). After overnight incubation cells were trypsinised, counted and plated at 2,000 cells per well of an opaque 96-well plate in 90 µl of DMEM media. The plates were incubated for 5 hours and inhibitors were then added (20 X final concentration, 5 µl, dissolved in water). The cells were incubated for a further 30 minutes prior to addition of DFX (20 X final concentration, 5µl, dissolved in water) to a final

concentration of 25 μM. After a further 16 hour incubation the media was removed from the plates, and the cells were washed with PBS. The firefly and *Renilla* luciferase activity was measured in the plate reader (settings 0.2 second lag time and 10 seconds read time) using a Dual-luciferase® reporter assay system (Promega, UK). Using the plate reader's automated injectors 100 μl LARII was added and firefly luciferase activity measured. Stop and Glo reagent was then added (100 μl) and the *Renilla* luciferase activity was measured.

Each condition was repeated in triplicate wells to ensure consistency of results.

5.2.4.2 HIF-1α Western blot of MCF-7 cells treated with DFX

Cells were trypsinised and suspended in DMEM/FBS, as described previously (section 5.2.3.1), and transferred to a centrifuge tube. 200,000 cells were plated per well of a six well plate and incubated overnight. Three wells were treated:

1.) Vehicle control (H_2O)

2.) DFX 250 μM

3.) DFX 500 μM

After a further 20 hours incubation, the cells were trypsinised (section 5.2.3.1) and a Western blot was carried out, as previously described (section 5.2.3.3). After imaging, the membrane was incubated with anti-GAPDH antibody (as a control) for 1 hour at room temperature, the membrane was then washed three times with PBS/TWEEN-20 (0.05%) and once with PBS before being imaged using a LI-COR Odyssey imager (LI-COR). The antibodies and concentrations used are described below (Table 52).

Antibody and target	Concentration	Manufacturers details
Anti-HIF-1α, mouse monoclonal (*Primary)*	1:167	BD Biosciences, UK: BD: 610959
Green goat anti-mouse *(Secondary)*	1:20,000	IRDye antibodies, LI-COR, USA: 926-32210
Anti-GAPDH Mouse mab clone 6C5 (*Loading control*)	1:5000	Abcam, USA: ab8245

Table 52 Antibodies used in Western blot experiment, concentrations used and manufacturer's details stated.

5.3 Peptide synthesis

5.3.1 Peptide synthesis-materials and equipment

All amino acids and Wang resins were purchased from Novabiochem, UK or Matrix Innovations, Canada. All DMF and acetonitrile used was HPLC grade and purchased from Fisher Scientific, UK. A CEM Liberty 1 Peptide synthesiser (CEM, UK) was used to synthesise linear peptides. Reverse phase HPLC was carried out on a Waters HPLC system (Waters, UK). Samples were manually injected into a Waters flex inject system into the HPLC system containing a Waters 1525 binary pump. Elution was monitored by a Waters 2998 photo diode array detector at 220 nm (amide backbone) and 280 nm (aromatic amino acids). Solvent was removed *in vacuo* on a Buchi rotary evaporator.

ESI-MS samples were analysed using a ZMD (Waters, Manchester, UK) mass spectrometer equipped with a single quadrupole analyser. Samples were introduced to the mass spectrometer via flow injection using a Waters 600 pump (flow rate 0.1 mL / min MeCN) and Waters 2700 autosampler. Low resolution mass spectra were recorded using positive/negative ion electrospray ionisation. ESI-LC-MSand HR-MS 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 HPLC pump. Ultra performance liquid chromatography was undertaken *via* a Waters UPLC BEH C18 (50 mm x 2.1mm 1.7µm) using a gradient of 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. High resolution mass spectra were recorded using positive/negative ion electrospray ionisation.

Tat-Tag and Tat-tagged cyclic peptides were analysed by HPLC-ESI-MS, this analysis was performed because compounds with MW >1000 Da resulted in a complex array of isotopes making analysis of HRMS too complicated.

NMR was carried out using a 400 Bruker AVII400 FT-NMR Spectrometer (400/1) or a Varian Inova 600 MHz spectrometer with a cold probe.

All other chemical reagents were purchased from Fisher Scientific, unless otherwise stated.

5.3.2 Peptide synthesis-general experimental procedures

5.3.2.1 Manual solid phase peptide synthesis-Fmoc deprotection

To remove the N-terminal Fmoc groups during solid-phase peptide synthesis, 20% piperidine/DMF (v/v) (20 ml) was added to the protected peptidyl resin in a sinter funnel. The mixture was bubbled with argon for 30 minutes. The solution was then filtered and washed with DMF (3 x 10 ml) followed by alternating dichloromethane (DCM) (3 x 10 ml) and diethyl ether (3 x 10 ml). Complete deprotection was detected using the ninhydrin test. If the ninhydrin test showed incomplete Fmoc deprotection, the process was repeated.

5.3.2.2 Manual solid-phase peptide synthesis-amino acid coupling

Fmoc protected amino acid (1.5 mmol), HOBt (203 mg, 1.5 mmol) and then DIC (232 μl, 1.5 mmol) were added to DMF (3 ml). The solution was stirred for 15 minutes and then added to Fmoc deprotected peptidyl resin (0.5 mmol) in a sinter funnel. The solution was bubbled with argon for 2 hours. The resin was filtered and washed with DMF (3 x 10 ml) and alternating DCM (3 x 10 ml) and diethyl ether (3 x 10 ml). Complete coupling was detected using the ninhydrin test, if free amines were detected the process was repeated.

5.3.2.3 Automated solid phase peptide synthesis

Some of the linear peptides were synthesised using a Liberty 1 peptide synthesiser (CEM, UK). The Fmoc amino acid Wang resin was loaded into the reaction chamber. The required amount of Fmoc protected amino acids (4.8 eq) were dissolved in DMF and loaded onto the peptide synthesiser manifold. Activator base, DIPEA in NMP (1/3 (v/v), 0.82 eq) and activator, PyBOP (4.4 eq in DMF) were also loaded into the machine. For

specific quantities used for each peptide see sections 5.3.3.5, 6, 14 and 21. The microwave peptide-coupling program was run, all amino acids were coupled at 75°C apart from cysteine and histidine which were coupled at 50°C. All amino acids were coupled with microwave assistance with the exception of arginine, which was doubled coupled in the absence microwave assistance (this was to prevent intra-arginine cyclisation). Each coupling was ten minutes, at a microwave power of 25 Watts. Deprotections were carried at 75°C in 3-4 ml of 20% piperidine for three minutes, with a microwave power of 45 Watts. After each deprotection and coupling four DMF washes, with 2 ml DMF, were carried out.

5.3.2.4 Ninhydrin test: Test for free amines

To a sample of resin solution A (100 μl) and solution B (25 μl) were added and then heated to 145°C for 1-2 minutes. A resin colour change from yellow/colourless to blue/purple indicated the presence of free amines.[137]

Solution A: KCN (32 mg, 0.49 mmol) was dissolved in water (50 ml), and (1 ml) added to pyridine (49 ml). This solution was then added to a solution of phenol (20 g, 0.21 mol) in ethanol (5 ml).

Solution B: Ninhydrin (1.25 g, 7.02 mmol) was dissolved in ethanol (25 ml).

5.3.2.5 Peptide cleavage from resin and side chain deprotection

A cleavage cocktail of TFA (17.6 ml), phenol (1ml), water (1 ml) and TIS (0.4 ml) was added to the Fmoc deprotected peptidyl resin. The mixture was stirred for 2 hours under argon. The resin was then filtered and the solvent removed from the filtrate liquor *in vacuo*. In order to remove traces of TFA, ethyl acetate (20 ml) was added to the residue, and removed *in vacuo*, this was repeated. The residue was finally dissolved in minimal DMF (3-4 ml) and added to cold diethyl ether (80 ml) drop-wise. The solution was incubated at -80°C for 20 minutes and the precipitated peptide was collected by

centrifugation (4000 rpm, 20 minutes). The supernatant was removed and the pellets dried *in vacuo* overnight, to leave the crude peptide.

5.3.2.6 Cysteine aldrithiol (spy) protection

After cleavage from the resin the free thiol of cysteine was protected by a pyridinesulfenyl group (spy) using 2,2'dipyridyl sulphide (aldrithiol-2, 10 eq). The protected crude peptide was dissolved in DMF (30 ml) and aldrithiol-2 was added to the mixture and stirred for 16 hours. Solvent was removed *in vacuo* and the solution filtered. The solution was then purified using HPLC program 1 (Table 53) and lyophilised, yielding a colourless solid.

5.3.2.7 Linear peptide cyclisation

Linear spy protected peptide was dissolved in DMF (1 ml/mg), EDC (3 eq) and HOBt (6 eq). The reaction mixture was stirred for 24 hours under argon; after which the solvent was removed *in vacuo*. The remaining solution was filtered and purified by HPLC using HPLC program 1 (Table 53), the solvents were removed by lyophilisation.

5.3.2.8 HPLC purification and analysis of peptides

Linear and cyclic peptides were purified using a prep reverse-phase HPLC with a Waters Atlantis prep, T3, amide capped C18, 5 μm, 19 x 100 mm column (Waters, UK) and the solvent system HPLC program 1 described below (Table 53).

Time (minutes)	Flow (ml/min)	% A (H_2O, 0.1% TFA)	% B (MeCN, 0.1% TFA)
0	17	95	5
1.0	17	95	5
10.0	17	40	60
15.0	17	40	60
15.5	17	95	5
20.0	17	95	5
20.1	0	95	5

Table 53 HPLC program 1.

HPLC solvent program used to purify linear and cyclic peptides.

The Tat-tag and Tat-tagged peptide were purified using a prep reverse-phase HPLC with a Waters Atlantis prep, T3, amide capped C18, 5 μm, 19 x100 mm column (Waters, UK) and the solvent system described below (Table 54).

Time (minutes)	Flow (ml/min)	% A (H_2O, 0.1% TFA)	% B (MeCN, 0.1% TFA)
0	17	95	5
1.0	17	95	5
10.0	17	95	5
15.0	17	40	60
18.0	17	40	60
20.0	17	95	5
22.0	17	95	5
23.0	17	95	5
23.1	0	95	5

Table 54 HPLC program 2.

HPLC solvent program used to purify the Tat-tag and Tat-tagged peptides.

Peptide purity was analysed by running samples of the compound on a Waters Atlantis T3, amide capped C18, 5 μm, 6 x100 mm column using the solvent system HPLC program 3 described below (Table 55).

Time (minutes)	**Flow (ml/min)**	**% A (H_2O, 0.1% TFA)**	**% B (MeCN, 0.1% TFA)**
0	1	95	5
10.0	1	95	5
20.0	1	40	60
30.0	1	40	60
31.0	1	95	5
34.5	1	95	5
35.0	0	95	5

Table 55 HPLC program 3.

Solvent system used to analyse the purity of synthesised peptides.

5.3.2.9 Removal of aldrtihiol (spy) protection from cyclo-CLLFVY

Cyclo-CLLFVY(spy) (1 eq) was dissolved in DMF (100 μl/mg), and water (10 μl/mg) was added followed by TCEP (1 eq). After approximately 1 hour MeCN (0.5 ml) and water (0.5-1 ml) were added to the reaction mixture, which was filtered and purified using HPLC program 1 (Table 53) and solvents were removed by lyophilisation yielding a colourless solid.

5.3.2.10 Calculation of peptide concentrations for use in assays

With the exception of the analogues *cyclo*-CLLFVA, which lacks a required tyrosine chromophore, concentrations of peptide solutions were calculated by measurement of absorbance at 280 nm (using a nanodrop) and then calculated by the Beer Lambert law (Equation 6). Peptide solutions were made up in deionised water.

$$A = \varepsilon[C]l$$

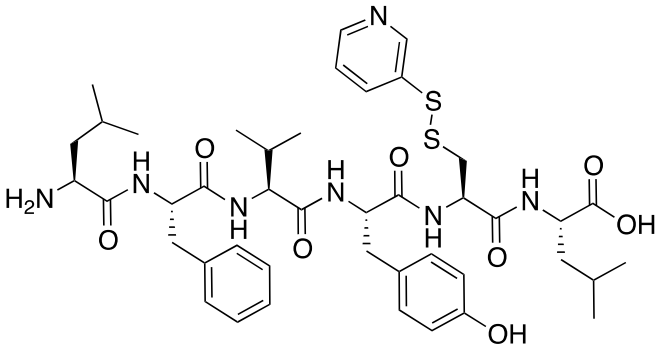
Where, A= Absorbance at 280 nm, [C]= Concentration of peptide (M), l = path length (cm) and ε = extinction coefficient at 280 nm measured in water ($M^{-1}cm^{-1}$), peptide extinction coefficients were calculated using ExPASY Bioinformatic ProtParam tool.[147]

Equation 6 The Beer Lambert Law. [148]

For CLLFVA the concentration was calculated by weighing the peptide and adjusted to 62% of this value to account for the hygroscopic nature of the peptide (this was calculated by weighing an exact amount of P1 and then dissolving it in a known amount of water and then measuring the concentration by A280).

5.3.3 Peptide synthesis-synthesised compounds and characterisation

*5.3.3.1 Synthesis of LFVYC(spy)L **1***



1

Figure 86 Structure of LFVYC(spy)L 1.

Reagent	**Equivalents**	**Quantity**
Fmoc-Leu-Wang resin	1 (1mM)	0.3-0.8 mmol/g resin loading
Fmoc-Cys(Trt)-OH	3	1.76 g
Fmoc-Tyr(tBu)-OH	3	1.38 g
Fmoc-Val-OH	3	1.02 g
Fmoc-Phe-OH	3	1.16 g
Fmoc-Leu-OH	3	1.06 g

Table 56 Amino acids used in the synthesis of LFVYCL.

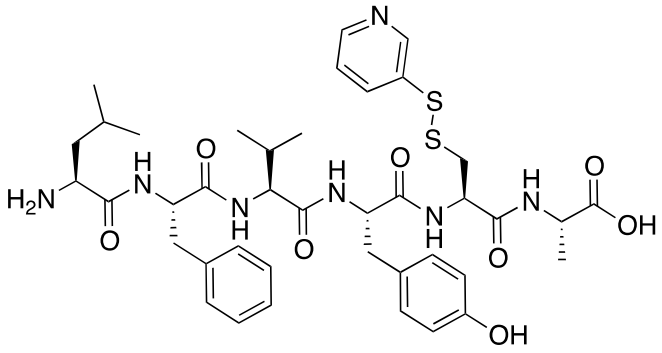
Masses of protected Fmoc amino acids used in the manual solid phase peptide synthesis.

On completion of peptide coupling (Table 56) (section 5.3.2.2) the Fmoc group was deprotected (section 5.3.2.1) and the linear peptide was cleaved from the resin and the side chain groups were deprotected (section 5.3.2.5). The peptide was spy protected (section 5.3.2.6) to leave LFVYC(spy)L **1** as a colourless solid (226 mg, 26.1%). **Analytical HPLC (280 nm)** HPLC program 3 *rt*= 20.813 minutes; **Prep HPLC (280 nm)** HPLC program 1 *rt* = 9.172 minutes; **^{1}H NMR**** δ ppm (600 MHz, DMSO-d_6) 9.14 (br s, 1H, Tyr OH), 8.65 (d, J=7.9 Hz, 1H, Cys/Phe/Tyr A NH), 8.43-8.44 (m, 1H, Cys(spy) ArH), 8.38 (d, J=8.2 Hz, 1H, Cys/Phe/Tyr B NH), 8.29 (d, J=7.9 Hz, 1H, Leu A NH),

8.03-8.06 (m, 1H, Val NH), 7.95-8.00 (m, 3H, Cys/Phe/Tyr C NH, Leu B NH_2), 7.77-7.83 (m, 2H, Cys(spy) ArH), 7.16-7.27 (m, 6H, Cys(spy) ArH, Phe ArH), 7.02 (d, J=8.5 Hz, 2H, Tyr *meta*-ArH), 6.61 (d, J=8.5 Hz, 2H, Tyr *ortho*-ArH), 4.67-4.71 (m, 1H, Cys/Phe/Tyr A αCH), 4.56-4.59 (m, 1H, Cys/Phe/Tyr B αCH), 4.50-4.54 (m, 1H, Cys/Phe/Tyr C αCH), 4.19-4.23 (m, 2H, Val αCH, Leu A αCH), 3.69-3.71 (m, 1H, Leu B αCH), 3.17 (dd, J=13.1, 5.2 Hz, 1H, Cys/Phe/Tyr B βCH_2), 2.89-3.02 (m, 3H, Cys/Phe/Tyr A βCH_2, Cys/Phe/Tyr B βCH_2, Cys/Phe/Tyr C βCH_2), 2.77-2.81 (m, 1H, Cys/Phe/Tyr A βCH_2), 2.69-2.73 (m, 1H, Cys/Phe/Tyr C βCH_2), 1.91-1.97 (m, 1H, Val βCH), 1.58-1.66 (m, 2H, Leu A γCH, Leu B γCH), 1.46-1.56 (m, 4H, Leu A βCH_2, Leu B βCH_2), 0.83-0.90 (m, 12H, Leu A CH_3, Leu B CH_3), 0.75-0.79 (m, 6H, Val CH_3);**m/z** (ESI^+, MeOH) 888.4 ($[M+Na]^+$, 100%), 910.4 ($[M+2Na-H]^+$, 54.18%), **HRMS** (ESI^+, MeOH) $[M+H]^+$ found 866.3924, $C_{43}H_{60}N_7O_8S_2$ required 866.3939 (err 1.8 ppm).

**Differentiation between the cysteine, phenylalanine and tyrosine proton environments NH, α and β protons cannot be made with the availiable spectra. Therefore A, B, C represent one of the three environments, with A having the highest NH chemical shift, Leucine A and leucine B refer to the two different leucine environments, with leucine A having the highest NH chemical shift.

*5.3.3.2 Synthesis of LFVYC(spy)A **13***



13

Figure 87 Structure of LFVYC(spy)A 13.

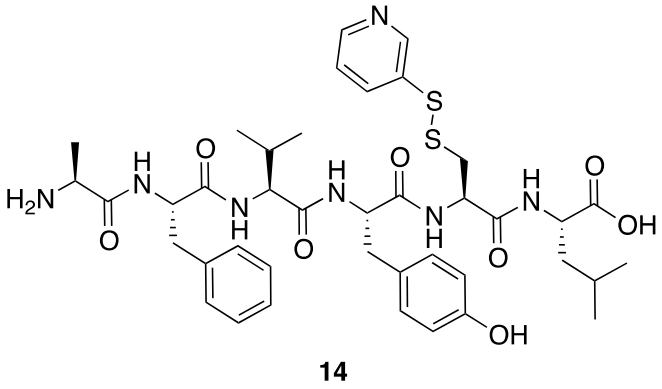
Reagent	**Equivalents**	**Quantity**
Fmoc-Ala-Wang resin	1 (0.5 mM)	0.3-0.8 mmol/g resin loading
Fmoc-Cys(Trt)-OH	3	0.88 g
Fmoc-Tyr(tBu)-OH	3	0.69 g
Fmoc-Val-OH	3	0.51 g
Fmoc-Phe-OH	3	0.51 g
Fmoc-Leu-OH	3	0.53 g

Table 57 Amino acids used in the synthesis of LFVYCA.

Masses of protected Fmoc amino acids used in the manual solid phase peptide synthesis.

On completion of peptide coupling (Table 57) (section 5.3.2.2) the Fmoc group was deprotected (section 5.3.2.1) and the linear peptide was cleaved from the resin and the side chain groups were deprotected (section 5.3.2.5). The peptide was spy protected (section 5.3.2.6) to leave LFVYC(spy)A **13** as a colourless solid (122.5 mg, 29.6%). **Analytical HPLC (280 nm)** HPLC program 3 *rt*=19.930 minutes; **Prep HPLC (280 nm)** HPLC program 1 *rt*= 7.874-8.217 minutes; **m/z** (ESI^+, MeOH) 868.3 ($[M+2Na-H]^+$, 100%), 846.3 ($[M+Na]^+$, 69.26%); **HRMS** (ESI^+, MeOH) $[M+H]^+$ found 824.3468, $C_{40}H_{54}N_7O_8S_2$ required 824.3470 (err 0.3 ppm).

*5.3.3.3 Synthesis of AFVYC(spy)L **14***



14

Figure 88 Structure of AFVYC(spy)L 14.

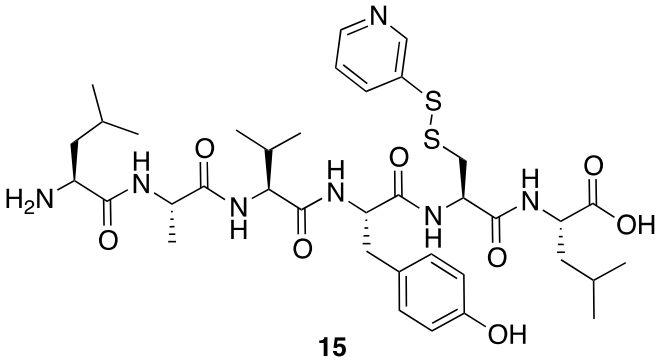
Reagent	**Equivalents**	**Quantity**
Fmoc-Leu-Wang resin	1 (0.5 mM)	0.3-0.8 mmol/g resin loading
Fmoc-Cys(Trt)-OH	3	0.88 g
Fmoc-Tyr(tBu)-OH	3	0.69 g
Fmoc-Val-OH	3	0.51 g
Fmoc-Phe –OH	3	0.58 g
Fmoc-Ala-OH	3	0.47 g

Table 58 Amino acids used in the synthesis of AFVYCL.

Masses of protected Fmoc amino acids used in the manual solid phase peptide synthesis.

On completion of peptide coupling (Table 58) (section 5.3.2.2) the Fmoc group was deprotected (section 5.3.2.1) and the linear peptide was cleaved from the resin and the side chain groups were deprotected (section 5.3.2.5). The peptide was spy protected (section 5.3.2.6) to leave AFVYC(spy)L **14** as a colourless solid (100.3 mg, 24.4%). **Analytical HPLC (280 nm)** HPLC program 3 *rt*= 21.202 minutes; **Prep HPLC (280 nm)** HPLC program 1 *rt* = 8.276 minutes; **m/z** (ESI$^+$, MeOH) 868.4 ([M+2Na-H]$^+$, 100%), 846.4 ([M+Na]$^+$, 94.57%); **HRMS** (ESI$^+$, MeOH) [M+H]$^+$ found 824.3456 $C_{40}H_{54}N_7O_8S_2$ requires 824.3470, [M+2H]$^{2+}$ found 412.6773 $C_{40}H_{55}N_7O_8S_2$ requires 412.6772 (err 1.7 ppm).

*5.3.3.4 Synthesis of LAVYC(spy)L **15***



15

Figure 89 Structure of LAVYC(spy)L 15.

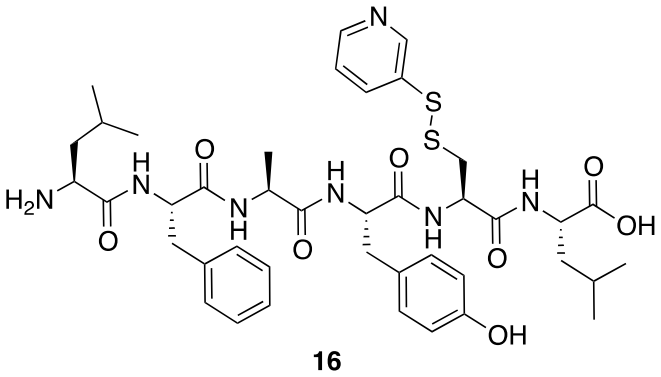
Reagent	**Equivalents**	**Quantity**
Fmoc-Leu-Wang resin	1 (0.5 mM)	0.3-0.8 mmol/g resin loading
Fmoc-Cys(Trt)-OH	3	0.88 g
Fmoc-Tyr(tBu)-OH	3	0.69 g
Fmoc-Val-OH	3	0.51 g
Fmoc-Ala-OH	3	0.47 g
Fmoc-Leu-OH	3	0.53 g

Table 59 Amino acids used in the synthesis of LAVYCL.

Masses of protected Fmoc amino acids used in the manual solid phase peptide synthesis.

On completion of peptide coupling (Table 59) (section 5.3.2.2) the Fmoc group was deprotected (section 5.3.2.1) and the linear peptide was cleaved from the resin and the side chain groups were deprotected (section 5.3.2.5). The peptide was spy protected (section 5.3.2.6) to leave LAVYC(spy)L **15** as a colourless solid (82.1 mg, 20.8%). **Analytical HPLC (280 nm)** HPLC program 3 *rt*=20.897 minutes; **Prep HPLC (280 nm)** HPLC program 1 *rt***=** 8.143 minutes; **m/z** (ESI^+, MeOH) 812.3 ($[M+Na]^+$, 100%), 834.3 ($[M+2Na-H]^+$, 41%), 790.3 ($[M+H]^+$, 14.40%); **HRMS** (ESI^+, MeOH) $[M+H]^+$ found 790.3620, $C_{37}H_{55}N_7O_8S_2$ required 790.3626 (err 0.8 ppm).

*5.3.3.5 Synthesis of LFAYC(spy)L **16***



16

Figure 90 Structure of LFAYC(spy)L 16.

Reagent	**Equivalents**	**Concentration**	**Quantity loaded into peptide synthesiser**
Fmoc-Leu-Wang resin	1 (0.5 mmol)	N/A	0.3-0.8 mmol/g resin loading
Fmoc-Phe-OH	4.8	0.2 M	0.93 g
Fmoc-Ala-OH	4.8	0.2 M	0.75 g
Fmoc-Tyr(tBu)-OH	4.8	0.2 M	1.10 g
Fmoc-Cys(Trt)-OH	4.8	0.2 M	1.41 g
Fmoc-Leu-OH	4.8	0.2 M	0.85 g
20% Piperidine/DMF (v/v)	Large excess	20%	121 ml
PyBOP	4.4	0.5 M	22 ml
DIPEA/NMP (1:3, v/v)	0.82	2 M	11 ml

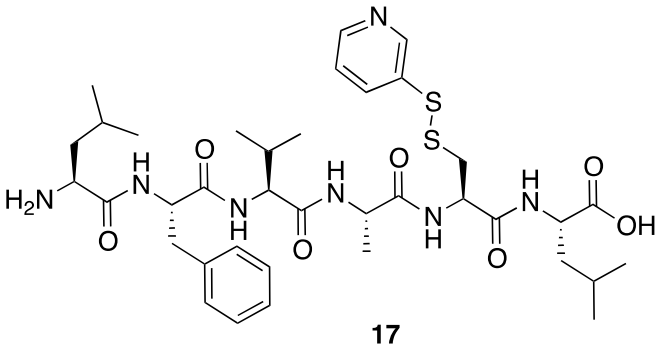
Table 60 Quantities of reagents used to make LFAYCL using a Liberty 1peptide synthesiser.

Reagents dissolved in DMF and loaded into the synthesiser.

The linear peptide was made using a Liberty 1 peptide synthesiser (CEM, UK) (section b5.3.2.3). Fmoc-Leu-Wang resin (0.5 mM) was placed in the reaction vessel. Quantities of amino acids used are shown in table above (Table 60). On completion of peptide coupling the peptidyl resin was removed from the synthesiser reaction vessel and washed with DMF (3 x10 ml) and alternating DCM (3 x 10 ml) and diethyl ether (3 x 10 ml). The Fmoc group

was deprotected manually (section 5.3.2.1) and the linear peptide was cleaved form the resin and the side chain groups were deprotected (section 5.3.2.5). The peptide was spy protected (section 5.3.2.6) to leave LFAYC(spy)L **16** as a colourless solid (83.3 mg, 19.9%). **Analytical HPLC (280 nm)** HPLC program 3 *rt*=20.603 minutes; **Prep HPLC (280 nm)** HPLC program 1 *rt*= 8.551 minutes; **m/z** (ESI^+, MeOH) 860.4 ($[M+Na]^+$, 100%), 838.4 ($[M+H]^+$, 21.79%); **HRMS** (ESI^+, MeOH) $[M+H]^+$ found 838.3622 $C_{41}H_{56}N_7O_8S_2$ required 838.3626 (err 0.6 ppm).

*5.3.3.6 Synthesis of LFVAC(spy)L **17***



17

Figure 91 Structure of LFVAC(spy)L 17.

Reagent	**Equivalents**	**Concentration**	**Quantity loaded into peptide synthesiser**
Fmoc-Leu-Wang resin	1 (0.5 mmol)	N/A	0.3-0.8 mmol/g resin loading
Fmoc-Phe-OH	4.8	0.2 M	0.93
Fmoc-Val-OH	4.8	0.2 M	0.81
Fmoc-Ala-OH	4.8	0.2 M	0.75
Fmoc-Cys(Trt)-OH	4.8	0.2 M	1.41
Fmoc-Leu-OH	4.8	0.2 M	0.85
20% Piperidine/DMF (v/v)	Large excess	20%	105 ml
PyBOP	4.4	0.5 M	22 ml
DIPEA/NMP (1:3, v/v)	0.82	2 M	11 ml

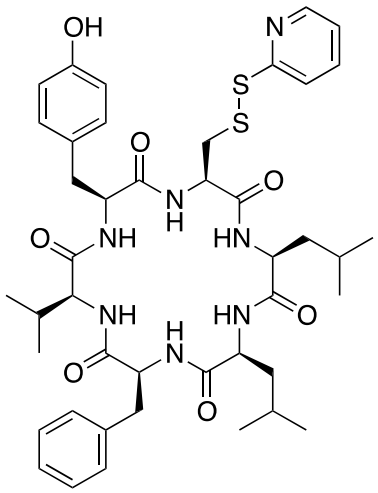
Table 61 Quantities of reagents used to make LFVACL using a Liberty 1peptide synthesiser.

Reagents dissolved in DMF and loaded into the synthesiser.

The linear peptide was made using a Liberty 1 peptide synthesiser (CEM, UK) (section 5.3.2.3). Fmoc-Leu-wang resin (0.5 mM) was placed in the reaction vessel. Quantities of amino acids used are shown in table above (Table 61). On completion of peptide coupling the peptidyl resin was removed from the synthesiser reaction vessel and washed with DMF (3 x 10 ml) and alternating DCM (3 x 10 ml) and diethyl ether (3 x 10

ml). The Fmoc group was deprotected manually (section 5.3.2.1) and the linear peptide was cleaved form the resin and the side chain groups were deprotected (section 5.3.2.5). The peptide was spy protected (section 5.3.2.6) to leave LFVAC(spy)L **17** as a colourless solid (90 mg, 23.7%). **Analytical HPLC (220 nm)** HPLC program 3 *rt*= 21.326 minutes, **Prep HPLC (280 nm)** HPLC program 1 *rt***=** 8.986 minutes; **m/z** (ESI^+, MeOH) 774.3 ($[M+H]^+$, 100%), 388 ($[M+2H]^{2+}$, 44.85%), 796.3 ($[M+Na]^+$, 15.86%); **HRMS** (ESI^+, MeOH) found 774.3663 $C_{37}H_{56}N_7O_7S_2$ required 774.3677 (err 1.8 ppm).

*5.3.3.7 Synthesis of cyclo-C(spy)LLFVY **2***



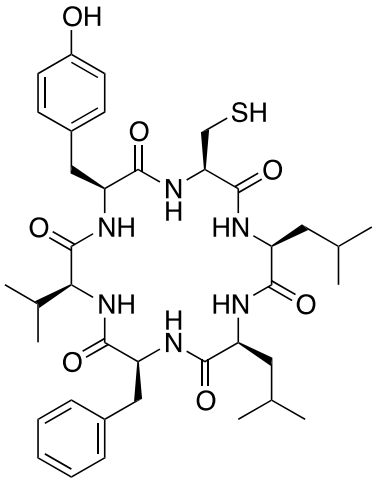
2

Figure 92 Structure of *cyclo*-C(spy)LLFVY 2

Linear LFVYC(spy)L (116 mg, 1.3 x 10^{-4} mol) was cyclised (section 5.3.2.7) to leave *cyclo*-C(spy)LLFVY **2** as a colourless solid (33.3 mg, 29.3%). **Analytical HPLC (280 nm)** HPLC program 3 *rt*= 26.704 minutes **Prep HPLC (280 nm)** HPLC program 1 *rt*= 13.124 minutes; **^{1}H NMR**** δ ppm (600 MHz, DMSO-d_6) 9.17 (br s, 1H, Tyr OH), 8.44-8.45 (m, 1H, Cys(spy) Ar), 8.24-8.27 (m, 1H, Cys/Phe/Tyr A NH), 8.12 (br d, J=7.94 Hz, 1H, Leu A NH), 7.94 (br d, J=8.24 Hz, 1H, Cys/Phe/Tyr B NH), 7.73-7.82 (m, 3H, Cys(spy) Ar, Cys/Phe/Tyr C NH, Leu B NH), 7.47-7.53 (m, 1H Val NH), 7.18-7.33 (m, 7H, Cys(spy) Ar, Phe Ar), 7.01 (d, J=8.24, 2H Tyr *meta*-ArH), 6.66 (d, J=8.24 Hz, 2H, Tyr *ortho*-ArH), 4.67-4.71 (m, 1H, Cys/Phe/Tyr C αCH), 4.39-4.43 (m, 1H, Cys/Phe/Tyr B αCH), 4.15-4.19 (m, 2H, Leu A αCH, Cys/Phe/Tyr A αCH), 4.03-4.06 (m, 1H, Val αCH), 3.94-3.99 (m, 1H, Leu B αCH), 3.21-3.30 (m, 1H, Cys/Phe/Tyr A βCH_2), 3.14-3.17 (m, 1H, Cys/Phe/Tyr C βCH_2), 3.04-3.11 (m, 1H, Cys/Phe/Tyr B βCH_2), 2.93-3.00 (m, 2H, Cys/Phe/Tyr A βCH_2, Cys/Phe/Tyr C βCH_2), 2.87 (dd, J=13.28, 10.53 Hz, Cys/Phe/Tyr B βCH_2), 1.96-2.01 (m, 1H, Val βCH), 1.51-1.77 (m, 2H, Leu A γCH, Leu B γCH), 1.43-1.51 (m, 2H, Leu A βCH_2, Leu B βCH_2), 0.69-0.89 (m, 18H, Leu A CH_3, Leu B CH_3, Val CH_3); **m/z** (ESI^+, MeOH) 870.5 ($[M+Na]^+$, 100%); **HRMS** (ESI^+, MeOH) found 870.3636 $C_{43}H_{58}N_7O_7S_2$ requires 870.3653 (err 1.9 ppm).

**Differentiation between the cysteine, phenylalanine and tyrosine proton environments. NH, α and β protons cannot be made with the availiable spectra. Therefore A, B, C represent one of the three environments, with A having the highest NH chemical shift, Leucine A and leucine B refer to the two different leucine environments, with leucine A having the highest NH chemical shift.

*5.3.3.8 Synthesis of cyclo-CLLFVY **9***



9

Figure 93 Structure of *cyclo*-CLLFVY 9.

Cyclo-C(spy)LLFVY **2** (22 mg, 2.6 x10^{-5} mol) was spy deprotected (section 5.3.2.9) to leave *cyclo*-CLLFVY **9** as a colourless solid (14.9 mg, 78%). **Analytical HPLC (280 nm)** HPLC program 3 *rt*= 24.431 minutes; **Prep HPLC (280 nm)** HPLC program 1 *rt*= 12.028 minutes; **m/z** (ESI^+, MeOH) 761.4 ($[M+Na]^+$, 100%); **HRMS** (ESI^+, MeOH) found 761.3677 $C_{38}H_{54}N_6NaO_7S$ requires 761.3667 (err 1.3 ppm).

*5.3.3.9 Synthesis of cyclo-C(spy)ALFVY **18***

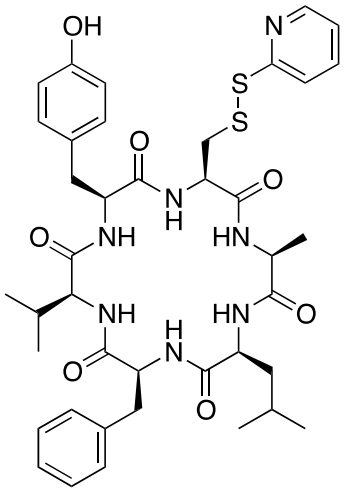


Figure 94 Structure of *cyclo*-C(spy)ALFVY 18.

LFVYC(spy)A (100 mg, 1.21 x 10^{-4} mol) was cyclised (section 5.3.2.7) to leave *cyclo*-C(spy)ALFVY **18** as a colourless solid (11.7 mg, 12.0%). **Analytical HPLC (280 nm)** HPLC program 3 *rt*= 23.437 minutes, **Prep HPLC (280 nm)** HPLC program 1 *rt*= 10.920 minutes; **m/z** (ESI^+, MeOH) 828.2 ($[M+Na]^+$, 100%); **HRMS** (ESI^+, MeOH) found 806.3346 $C_{40}H_{52}N_7O_7S_2$ requires 806.3364 (err 2.3 ppm).

*5.3.3.11 Synthesis of cyclo-C(spy)LAFVY **19***

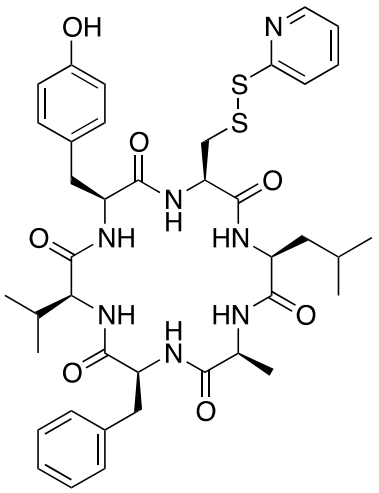
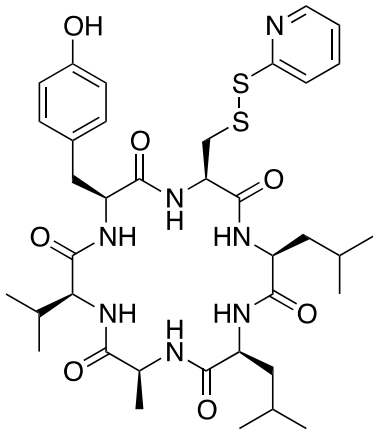


Figure 95 Structure of *cyclo*-C(spy)LAFVY 19.

AFVYC(spy)L (91 mg, 1.1 x 10^{-4} mol) was cyclised (section 5.3.2.7) to leave cyclo-C(spy)LAFVY **19** as a colourless solid (27.1 mg, 30.5%). **Analytical HPLC (280 nm)** HPLC program 3 *rt*= 22.904 minutes, **Prep HPLC (280 nm)** HPLC program 1 *rt*= 11.158 minutes; **m/z** (ESI^+, MeOH) 828.4 ($[M+Na]^+$, 100%), 806.4 ($[M+H]^+$, 100%); **HRMS** (ESI^+, MeOH) found 806.3351 $C_{40}H_{52}N_7O_2S_2$ requires 806.3364 (err 1.6 ppm).

*5.3.3.12 Synthesis of cyclo-C(spy)LLAVY **20***



20

Figure 96 Structure of *cyclo*-C(spy)LLAVY 20.

LAVYC(spy)L (70 mg, 8.87 x 10^{-5}) was cyclised (section 5.3.2.7) to leave c*yclo*-C(spy)LLAVY **20** as a colourless solid (5.6 mg, 8.2%). **Analytical HPLC (280 nm)** HPLC program 3 *rt*= 24.043 minutes, **Prep HPLC (280 nm)** HPLC program 1 *rt*= 10.944 minutes; **m/z** (ESI^+, MeOH) 794.4 ($[M+Na]^+$, 100%); **HRMS** (ESI^+, MeOH) 772.3511 found $C_{37}H_{54}N_7O_7S_2$ required 772.3521 (err 1.3 ppm), 794.3328 found $C_{37}H_{53}N_7NaO_7S_2$ required 794.3340 (err 1.6 ppm).

*5.3.3.14 Synthesis of cyclo-C(spy)LLFAY **21***

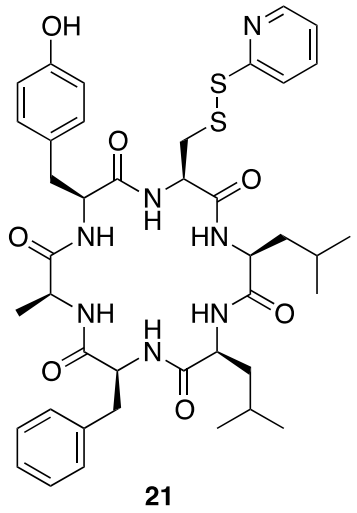
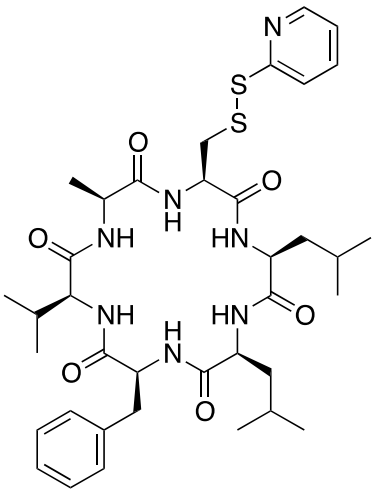


Figure 97 Structure of *cyclo*-C(spy)LLFAY 21.

LFAYC(spy)L (70 mg, 8.4 x 10^{-5} mol) was cyclised (section 5.3.2.7) to leave c*yclo*-C(spy)LLFAY **21** as a colourless solid (17.4 mg, 25.2%). **Analytical HPLC (280 nm)** HPLC program 3 *rt*= 22.586 minutes; **Prep HPLC (280 nm)** HPLC program 1 *rt***=** 11.796 minutes; **m/z** (ESI^+, MeOH) 842.4 ($[M+Na]^+$, 100%); **HRMS** (ESI^+, MeOH) 820.3511 found $C_{41}H_{54}N_7O_7S_2$ required 820.3521.

*5.3.3.16 Synthesis of cyclo-C(spy)LLFVA **22***



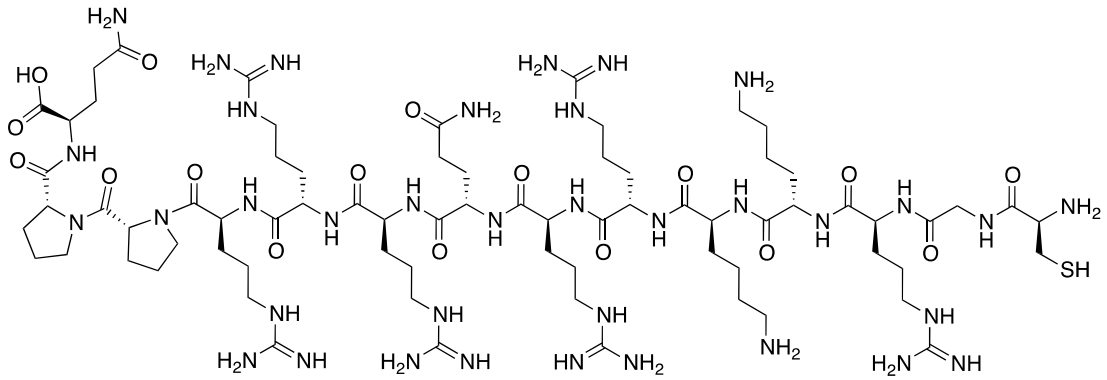
22

Figure 98 Structure *cyclo*-C(spy)LLFVA 22.

LFVAC(spy)L (80 mg, 1.03 x 10^{-4} mol) was cyclised (section 5.3.2.7) to leave c*yclo*-C(spy)LLFVA **22** as a colourless solid (15.5 mg, 19.9%). **Analytical HPLC (280 nm)** HPLC program 3 *rt*= 25.740 minutes*;* **Prep HPLC (280 nm)** HPLC program 1 *rt*= 12.601-12.867 minutes; **m/z** (ESI^+, MeOH) 778.4 ($[M+Na]^+$, 100%); **HRMS** (ESI^+, MeOH) found 756.3552 $C_{37}H_{54}N_7O_6S_2$ requires 756.3572 (err 7.6 ppm).

** Peptide is hygroscopic and the inconsistency in the integration is likely to be due to the presence of water.

*5.3.3.18 Synthesis of Tat-tag (CGRKKRRQRRRPPQ) **3***



3

Figure 99 Structure of the Tat-tag (CGRKKRRQRRRPPQ).

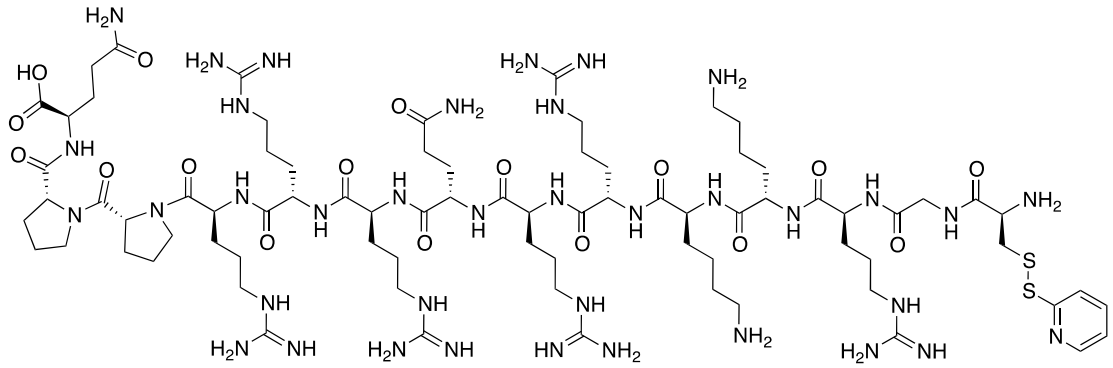
Reagent	Equivalents	Concentration	Quantity loaded into peptide synthesiser
Fmoc-Gln(Trt)-Wang resin	1 (0.5 mmol)	N/A	0.3-0.8 mmol/g resin loading
Fmoc-Pro-OH	4.8	0.2 M	1.55 g
Fmoc-Arg(Pbf)-OH	4.8	0.2 M	17.91 g
Fmoc-Lys(Boc)-OH	4.8	0.2 M	2.16 g
Fmoc-Gln(Trt)-OH	4.8	0.2 M	1.47 g
Fmoc-Gly-OH	4.8	0.2 M	0.71 g
Fmoc-Cys(Trt)-OH	4.8	0.2 M	2.69 g
20% Piperidine/DMF (v/v)	Large excess	20%	253 ml
PyBOP	4.4	0.5 M	89 ml
DIPEA	0.82	2 M	45 ml

Table 62 Amounts of amino acids and coupling reagents loaded into the microwave Liberty 1 peptide synthesiser to make Tat-tag CGRKKRRQRRRPPQ 3.

Fmoc protected amino acids were dissolved in DMF and loaded into the peptide synthesiser (section 5.3.2.3) along with Fmoc-Gln(Trt)-Wang resin (Table 62). The first resin bound Fmoc-Glu-(trt) and final proline were deprotected twice to ensure complete deprotection. All Fmoc-Arg(Pbf)-OH and Fmoc-Cys(Trt)-OH were double coupled to ensure complete coupling. On completion of peptide coupling the peptidyl resin was removed from the synthesiser reaction vessel and washed with DMF (3 x 10 ml) and alternating DCM (3 x 10 ml) and diethyl ether (3 x 10 ml). The Fmoc group was deprotected (section 5.3.2.1) and the peptide was cleaved from the resin at the same time as side chain deprotection, by stirring the resin in a solution of TFA (18.8 ml), water (0.5 ml), EDT (0.5 ml) and TIS (0.2 ml) for 3 hours. The solution was filtered and the solvent removed *in vacuo*. Ethyl acetate (20 ml) was added to the residue and then removed *in vacuo*, this was repeated again to remove all remaining TFA. The residue was dissolved in

DMF (2 ml) and added to cold ether (-80ºC, 80 ml). The solution was incubated at -80ºC for 20 minutes. The peptide was precipitated by centrifugation (4000 rpm, 20 minutes), the supernatant was removed and the pellets dried *in vacuo*. The product was dissolved in DMF (approx. 6 ml) and water (approx. 4 ml), and was filtered before purification by reverse-phase HPLC (HPLC program 2). The product containing fractions were combined, the acetonitrile was removed *in vacuo* and the water lyophilised to leave the product **3** as a colourless solid (289.7 mg, 32%). **Analytical HPLC (220 nm)** HPLC program 3 *rt* =13.660 minutes; **Prep HPLC (220 nm)** HPLC program 2 *rt***=** 15.657 minutes; **m/z** (ESI^+, MeOH) 456.6 **(**$[M+4H]^{4+}$, 72.62%) 365.5 ($[M+5H]^{5+}$, 21.70%), 911.7 ($[M+2H]^{2+}$, 27.10%), 629.9 ($[M+3Na]^{3+}$, 25.02%), 608.5 $[M+3H]^{3+}$, 19.66%), 1823.2 ($[M+H]^+$, 2.23%); **HPLC-ESI-MS** (ESI^+, MeOH, 220 nm) 0.5 minutes peak showed: 365.2 ($[M+5H]^{5+}$), 456.3 ($[M+4H]^{4+}$).

5.3.3.18.1 Synthesis of Tat(spy)-tag (C(spy)GRKKRRQRRRPPQ) **23**



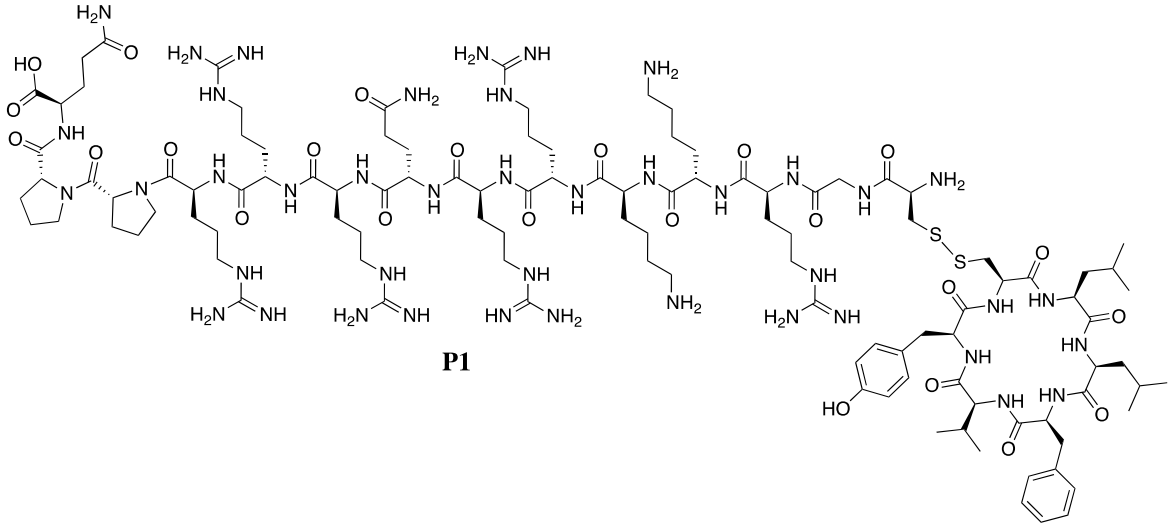
23

Figure 100 Structure of spy protected Tat-tag (C(spy)GRKKRRQRRRPPQ) 23.

In some cases the Tat-tag was spy protected (section 5.3.2.6) prior to HPLC purification. This yielded a colourless solid **23** (134.2 mg, 13.9%). **Analytical HPLC (280 nm)** HPLC program 3 *rt* = 16.657 minutes; **Prep HPLC (280 nm)** HPLC program 2 *rt***=** 13.393 minutes; **m/z** (ESI^+, MeOH/formic acid) 483.9 ($[M+4H]^{4+}$, 100%), 387.4 ($[M+5H]^{5+}$,49.01%), 644.8 ($[M+3H]^{3+}$, 47.01%), 967.1 ($[M+2H]^{2+}$, 7.56%); **HPLC-ESI-**

MS (ESI$^+$, MeOH, 220 nm) 0.2 minute peak showed: 387.2 ([M+5H]$^{5+}$), 483.8 ([M+4H]$^{4+}$), 644.2 ([M+3H]$^{3+}$).

*5.3.3.19 Synthesis of Tat-cyclo-CLLFVY **P1***

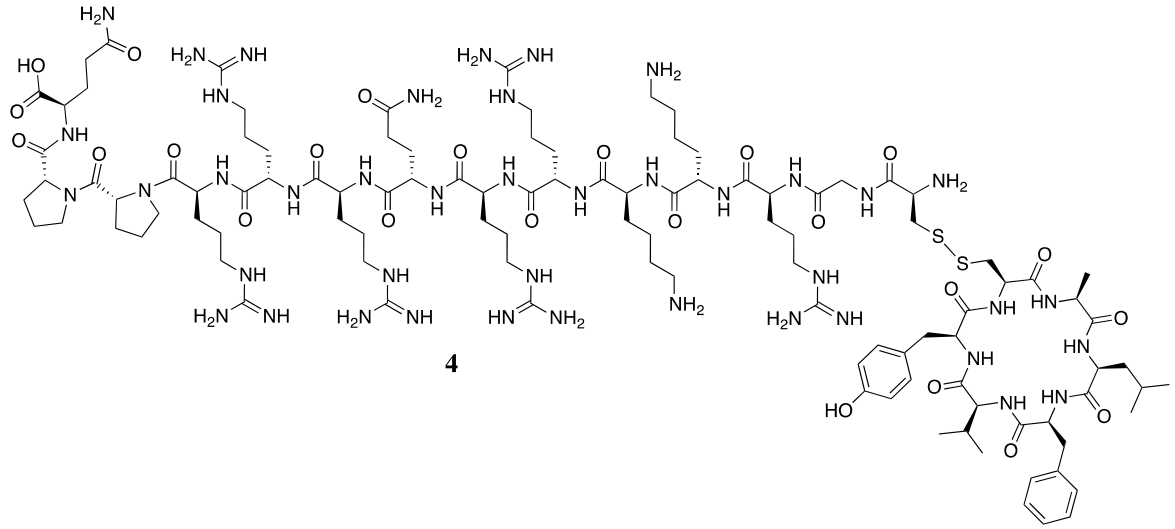


P1

Figure 101 Structure of Tat-*cyclo*-CLLFVY P1.

Cyclo-C(spy)LLFVY (10 mg, 1.18 x 10^{-5} mol) was dissolved in DMF (1 ml), water (67 µl) and Tat-tag (43 mg, 2.36 x 10^{-5} mol, 2 eq) was added and the mixture was stirred for approximately 16 hours. Water (1.5 ml) was added to the reaction mixture, and the solution was filtered and purified by reverse HPLC (HPLC program 2). Product containing fractions were combined and acetonitrile was removed *in vacuo*, the water lyophilised to leave Tat-*cyclo*-CLLFVY **P1** as a colourless solid. (12.1 mg, 40.1%). **Analytical HPLC (280 nm)** HPLC program 3 *rt*= 14.807 minutes; **Prep HPLC (280 nm)** HPLC program 2 *rt*= 20.311 minutes; **m/z** (ESI$^+$, MeOH/formic acid) 640.8 ([M+4H]$^{4+}$, 100%), 512.8 ([M+5H]$^{5+}$, 62.52%), 854.9 ([M+3H]$^{3+}$, 52.30%); **HPLC-ESI-MS** (ESI$^+$, MeOH/formic acid, 220 nm) 3.1-3.2 minute peak showed: 512.7 ([M+5H]$^{5+}$), 640.6 ([M+4H]$^{4+}$), 853.8 ([M+3H]$^{3+}$), 2560 ([M+H]$^+$).

*5.3.3.20 Synthesis of Tat-cyclo-CALFVY **4***

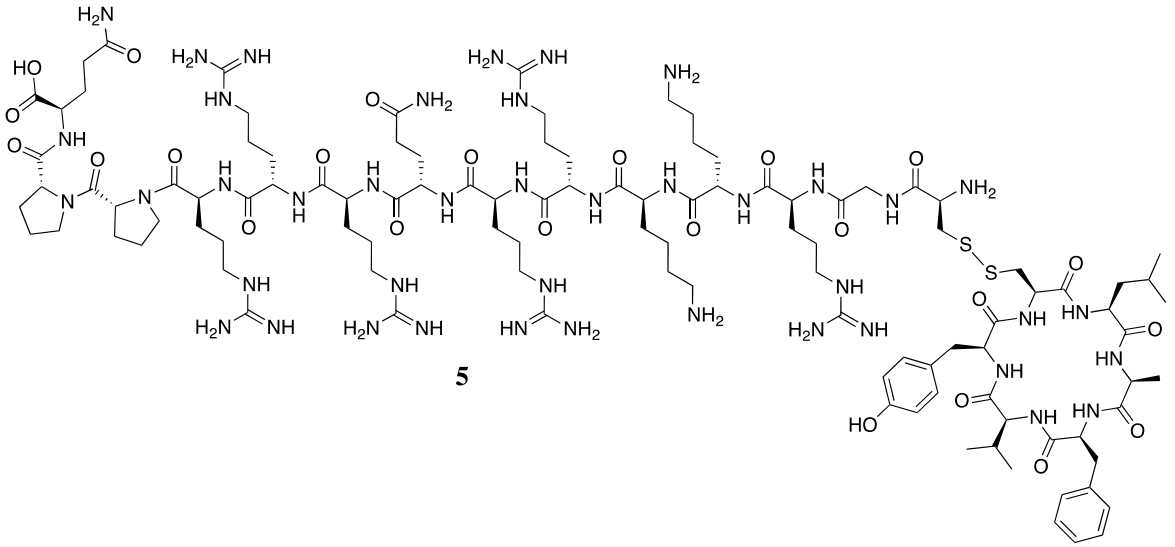


4

Figure 102 Structure of Tat-c*yclo*-CALFVY 4.

Cyclo-C(spy)ALFVY (9 mg, 1.12 x 10^{-5}), was dissolved in DMF (900 µl), water (60 µl) and Tat-tag (41 mg, 2.235 x 10^{-5}, 2 eq) was added and the reaction mixture was stirred for approximately 16 hours. Water (1.5 ml) was added to the reaction mixture, the solution was filtered and the product was purified by reverse-phase HPLC (HPLC program 2). Product containing fractions were combined and acetonitrile was removed *in vacuo*, the water lyophilised to leave Tat-*cyclo*-CALFVY **4** as a colourless solid. (9.2 mg, 32.6%). **Analytical HPLC (280 nm)** HPLC program 3 *rt*= 19.399 minutes; **Prep HPLC (280 nm)** HPLC program 2 *rt***=** 14.475 minutes; **m/z** (ESI^+, MeOH/formic acid) 630.2 ($[M+4H]^{4+}$, 100%), 840 ($[M+3H]^{3+}$, 78.28%), 504.3 ($[M+5H]^{5+}$, 35.38%), 1259.9 ($[M+2H]^{2+}$, 11.64%); **HPLC-ESI-MS** (ESI^+, MeOH/formic acid, 220 nm) 2.8-2.9 minute peak showed: 420.4 ($[M+6H]^{6+}$), 504.3 ($[M+5H]^{5+}$), 630.1 ($[M+4H]^{4+}$), 839.8 ($[M+3H]^{3+}$).

*5.3.3.21 Synthesis of Tat-cyclo-CLAFVY **5***

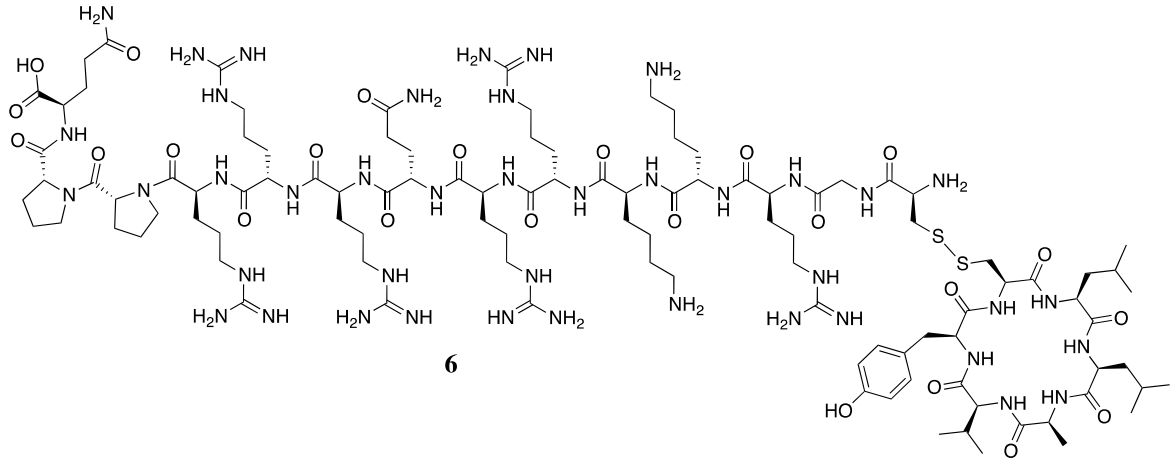


5

Figure 103 Structure of Tat-*cyclo*-CLAFVY 5.

Spy-deprotection of cyclic peptide: *cyclo*-C(spy)LAFVY (14 mg, 1.74 x 10^{-5} mol) was dissolved in DMF (1.4 ml), water (0.1 ml) and TCEP (14 mg, 1.74 x 10^{-5} mol, 1 eq) were added and the solution was stirred for 1 hour under argon. Tat(spy) (67 mg, 3.48 x 10^{-5} mol, 2 eq) was added and the reaction was stirred for 16-18 hours. DMF (2 ml) and water (2 ml) were added to the reaction mixture and the product was purified by reverse-phase HPLC (HPLC program 2). Product containing fractions were combined and acetonitrile was removed *in vacuo*, the water lyophilised to leave Tat-c*yclo*-CLAFVY **5** as a colourless solid (10.4 mg, 23.8%). **Analytical HPLC (280 nm)** HPLC program 3 *rt*= 19.078 minutes; **Prep HPLC (280 nm)** HPLC program 2 *rt*=14.377 minutes; **m/z** (ESI^+, MeOH/formic acid) 630.4 ($[M+4H]^{4+}$, 100%), 504.5 ($[M+5H]^{5+}$, 95.78%), 839.8 ($[M+3H]^{3+}$, 47.42%), 1259.9 ($[M+2H]^{2+}$, 6.19%), 360 ($[M+7H]^{7+}$, 8.655); **HPLC-ESI-MS** (ESI^+, MeOH/formic acid, 220 nm) 2.9-3.0 minute peak showed: 420.4 ($[M+6H]^{6+}$), 504.3 ($[M+5H]^{5+}$), 630.1 ($[M+4H]^{4+}$), 839.8 ($[M+3H]^{3+}$).

*5.3.3.22 Synthesis of Tat-cyclo-CLLAVY **6***

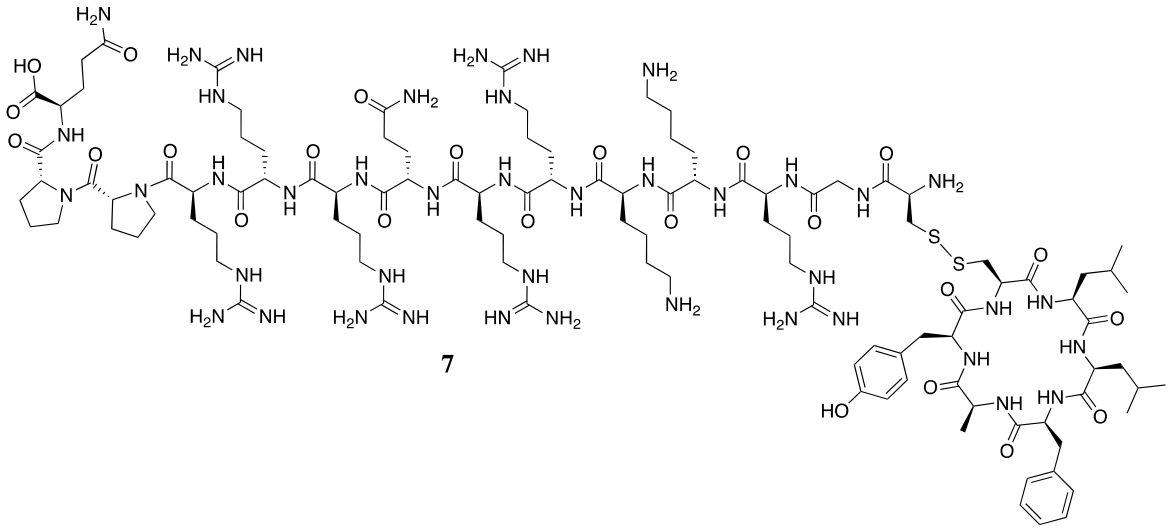


6

Figure 104 Structure of Tat-*cyclo*-CLLAVY 6.

Cyclo-CLLAVY (4 mg, 5.2 x 10^{-6} mol, 1 eq) was dissolved in DMF (400 µl), water (27 µl) and Tat-tag (19 mg, 1.0 x 10^{-5}, 2 eq) were added and the reaction mixture was stirred for approximately 16 hours. Water (1.5 ml) was added to the reaction mixture, the solution was filtered and the product was purified by reverse-phase HPLC (HPLC program 2). Product containing fractions were combined and acetonitrile was removed *in vacuo*, the water lyophilised to leave Tat-*cyclo*-CLLAVY **6** as a colourless solid (4.1 mg, 31.8%). **Analytical HPLC (280 nm)** HPLC program 3 *rt*= 19.834 minutes; **Prep HPLC (280 nm)** HPLC program 2 *rt***=** 14.310 minutes; **m/z** (ESI^+, MeOH/formic acid) 497.5 ($[M+5H]^{5+}$, 100%), 621.7 ($(M+4H)^{4+}$, 34.21%), 414.5 ($(M+6H)^{6+}$, 27.99%), 828.6 ($(M+3H)^{3+}$, 19.46%); **HPLC-ESI-MS** (ESI^+, MeOH/formic acid, 220 nm) 2.8-3.9 minute peak showed: 497.5 ($[M+5H]^{5+}$), 621.6 ($[M+4H]^{4+}$), 828.5 ($[M+3H]^{3+}$), 2483.4 ($[M+H]^+$).

*5.3.3.23 Synthesis of Tat-cyclo-CLLFAY **7***

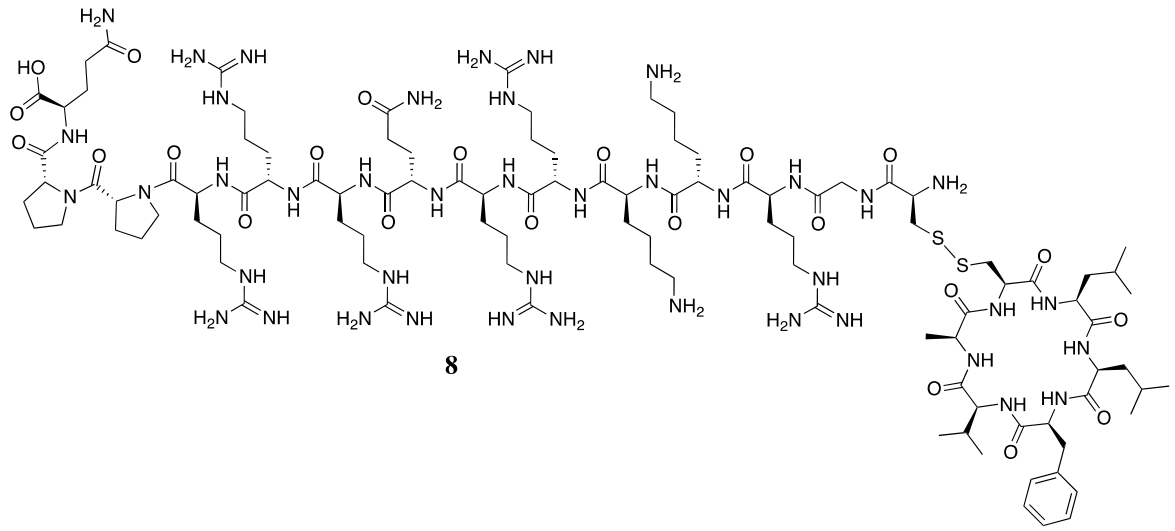


7

Figure 105 Structure of Tat-*cyclo*-CLLFAY 7.

Cyclo-C(spy)LLFAY (15 mg, 1.83 x10^{-5} mol, 1 eq) was dissolved in DMF (1.5 ml), water (0.1 ml) and TCEP (5.2 mg, 1.83 x 10^{-5}, 1 eq) were added and the solution was stirred for 1 hour under argon. Tat(spy) (71 mg, 3.66 x 10^{-5} mol, 2 eq) was added and the reaction was stirred for 16-18 hours. Water (1.5 ml) was added to the reaction mixture and the product was purified by reverse-phase HPLC (HPLC program 2). Product containing fractions were combined and acetonitrile was removed *in vacuo*, the water was lyophilised to leave Tat-c*yclo*-CLLFAY **7** as a colourless solid (23.6 mg, 56.8%). A**nalytical HPLC (280 nm)** HPLC program 3 *rt*= 20.064 minutes; **Prep HPLC (280 nm)** HPLC program 2 *rt*=14.491 minutes; **m/z** (ESI^+, MeOH/formic acid); 633.9 ($[M+4H]^{4+}$, 100%), 844.8 ($[M+3H]^{3+}$, 66.27%), 507.3 ($[M+5H]^{5+}$, 22.73%); **HPLC-ESI-MS** (ESI^+, MeOH/formic acid, 220 nm) 2.9 minute peak showed: 507.1 ($[M+5H]^{5+}$), 633.6 ($[M+4H]^{4+}$), 422.6 ($[M+6H]^{6+}$), 844.5 ($[M+3H]^{3+}$).

*5.3.3.24 Synthesis of Tat-Cyclo-CLLFVA **8***

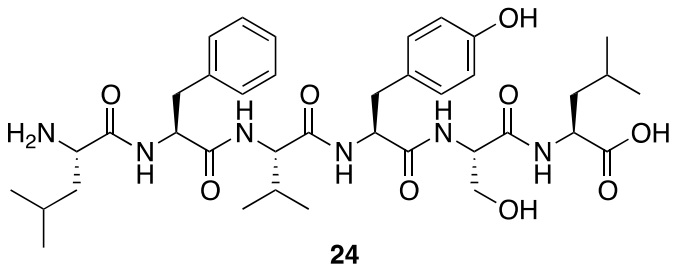


8

Figure 106 Structure of Tat-*cyclo*-CLLFVA 8.

Cyclo-C(spy)LLFVA (8 mg, 1.05 x 10^{-5} mol) was dissolved in DMF (800 µl), water (50 µl) and Tat-tag (39 mg, 2.12 x 10^{-5}) was added and the mixture was stirred for approximately 16 hours. Water (1. 5 ml) was added to the reaction mixture and the product was purified by reverse-phase HPLC (HPLC program 2). Product containing fractions were combined and acetonitrile was removed *in vacuo,* the water lyophilised to leave c*yclo*-CLLFVA **8** as a colourless product (16. 4 mg, 63%). **Analytical HPLC (220 nm)** HPLC program 3 *rt*=19.987 minutes; **Prep HPLC (220 nm)** HPLC program 2 *rt***=** 14.496 minutes; **m/z** (ESI^+, MeOH/formic acid) 494.4 ($[M+5H]^{5+}$, 72.93%), 617.7 ($[M+4H]^{4+}$, 58.84%), 823.6 ($[M+3H]^{3+}$, 25.53%), 1234.6 ($[M+2H]^{2+}$, 7.59%); **HPLC-ESI-MS** (ESI^+, MeOH/formic acid, 220 nm) 3.0-3.2 peak showed: 2467.4 ($[M+H]^+$), 494.3 ($[M+5H]^{5+}$), 617.6 ($[M+4H]^{4+}$), 823.1 ($[M+3H]^{3+}$).

*5.3.3.25 Synthesis of LFVYSL **24***



24

Figure 107 Structure of LFVYSL 24.

The linear peptide was made using a Liberty 1 peptide synthesiser. The Fmoc protected amino acids were dissolved in DMF and loaded onto the peptide synthesiser manifold, Fmoc-Leu-Wang resin (0.5 mM) was placed in the reaction vessel (Figure 108).

Reagent	**Equivalents**	**Concentration**	**Quantity loaded into peptide synthesiser**
Fmoc-Leu-Wang resin	1 (0.5 mmol)	N/A	0.3-0.8 mmol/g resin loading
Fmoc-Phe-OH	4.8	0.2 M	0.46 g
Fmoc-Val-OH	4.8	0.2 M	0.41 g
Fmoc-Tyr(tBu)-OH	4.8	0.2 M	0.55 g
Fmoc-Ser(tBu)-OH	4.8	0.2 M	0.46 g
Fmoc-Leu-OH	4.8	0.2 M	0.42 g
20% Piperidine/DMF (v/v)	Large excess	20%	77 ml
PyBOP	4.4	0.5 M	22
DIPEA	0.82	0.2 M	11

Figure 108 Quantities of reagents used to make LFVYSL using a Liberty 1peptide synthesiser.

Reagents dissolved in DMF and loaded into the synthesiser.

The linear peptide was made using a Liberty 1 peptide synthesiser (CEM, UK) (section 5.3.2.3). Fmoc-Leu-Wang resin (0.5 mM) was placed in the reaction vessel, quantities of amino acids used shown in table above (Table 61). On completion of peptide

coupling the peptidyl resin was removed from the synthesiser reaction vessel and washed with DMF (3 x 10 ml) and alternating DCM (3 x 10 ml) and diethyl ether (3 x 10 ml). The Fmoc group was deprotected manually (section 5.3.2.1) and the linear peptide was cleaved from the resin as follows: a mixture of 94% TFA, 2.5% water, 2.5% EDT and 1% TIS (10 ml) was added, the reaction mixture was stirred under argon for 3 hours at room temperature (all equipment that had come into contact with EDT was soaked in bleach). The resin was removed by filtration and washed with TFA (5 ml), the solvent was removed for the filtrate *in vacuo*. The residue was resuspended in TFA (1 ml) and transferred to a 50 ml centrifuge tube, degassed -80°C diethyl ether (45 ml) was added to precipitate the product. The product was precipitated by centrifugation (4000 rpm, 15 min). The supernatant was removed to leave the peptide pellet which was dissolved in DMF (2ml), acetonitrile (1.5 ml) and water (3 ml) was added. The solution was added and filtered and purified by reverse-phase HPLC (HPLC program 1). The product containing fractions were combined, and the acetonitrile was removed *in vacuo*, and water was removed by lyophilisation to leave a LFVYSL **24** as a colourless solid (30 mg, 16%). **Prep HPLC (280 nm)** HPLC program 1 *rt***=** 8.121 minutes; **m/z** (ESI^+, MeCN) 741.4 ($[M+H]^+$, 100%), 763.4 ($[M+Na]^+$, 41.1%).

*5.3.3.26 Synthesis of Cyclo-SLLFVY **10***

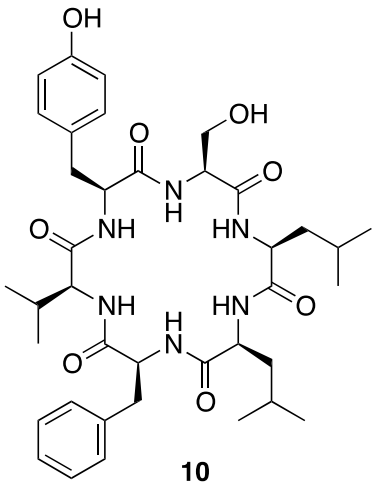
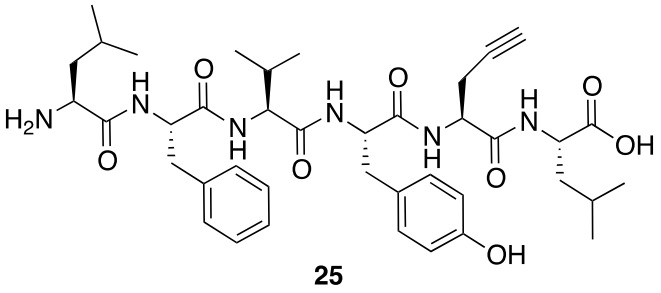


Figure 109 Structure of *cyclo*-SLLFVY 10.

LFVYSL (30 mg, 3.5 x 10^{-5} mol) was cyclised (section 5.3.2.7) to leave c*yclo*-SLLFVY **10** as a colourless solid (9.7 mg, 38%). **Prep HPLC (280 nm)** HPLC program 1 *rt*= 10.734 minutes; **m/z** (ESI^+, MeCN) 745.4 $(M+Na]^+$, 100%); **HRMS** (ESI^+, MeOH) 723.4083 found $C_{38}H_{55}N_6O_8$ required 723.4076 (err 1.0 ppm).

*5.3.3.27 Synthesis of LFVYLPr **25***



25

Figure 110 Structure of LFVYLPr 25.

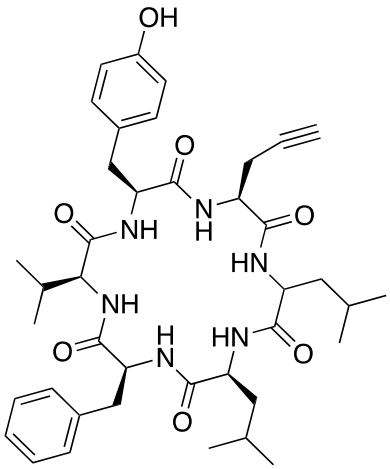
Pr= propargylglycine.

Reagent	**Equivalents**	**Quantity**
Fmoc-Leu-Wang resin	1 (0.5 mM)	0.3-0.8 mmol/g resin loading
Fmoc-propargylglycine-OH	3	0.50 g
Fmoc-Tyr(tBu)-OH	3	0.69 g
Fmoc-Val-OH	3	0.51 g
Fmoc-Phe-OH	3	0.51 g
Fmoc-Leu-OH	3	0.53 g

Table 63 Amino acids used in the synthesis of LFVYPrL. Masses of protected Fmoc amino acids used in the manual solid phase peptide synthesis.

On completion of peptide coupling (Table 63) (section 5.3.2.2) the Fmoc group was deprotected (section 5.3.2.1) and the linear peptide was cleaved from the resin and the side chain groups were deprotected (section 5.3.2.5). The deprotected peptide was dissolved in DMF and the solution was then purified using HPLC program 1 (Table 53). Product containing fractions were combined and acetonitrile was removed *in vacuo,* the water lyophilised to leave LPrFVYL **25** (102 mg, 27.3%). **Analytical HPLC (280 nm)** HPLC program 3 *rt*= 20.271 minutes*;* **Prep HPLC (280 nm)** HPLC program 1 *rt*=8.568 minutes; **m/z** (ESI^+, MeOH) 749.5 ($[M+H]^+$, 68.67%), 771.5 ($[M+H]^+$, 68.02%); **HRMS** (ESI^+, MeOH) 749.4228 found $C_{40}H_{57}N_6O_8$ required 749.4234 (err 0.6 ppm).

*5.3.3.28 Synthesis of cyclo-PrLLFVY **11***



11

Figure 111 Structure of *cyclo*-PrLLFVY 11.

Pr= propargylglycine**.**

LPrFVYL (100 mg, 1.34 x 10^{-4} mol) was cyclised (section 5.3.2.7) to leave cyclo-PrLLFVY **11** as a colourless solid (19 mg, 19.4%). **Analytical HPLC (280 nm)** HPLC program 3 *rt*= 25.731 minutes; **Prep HPLC (280 nm)** HPLC program 1 *rt*=10.370 minutes; **m/z** (ESI^+, MeOH) 753.4 ($[M+H]^+$, 100%); **HRMS** (ESI^+, MeOH) 753.3939 found $C_{40}H_{54}N_6NaO_7$ required 753.3946 (err 1.0 ppm).

Appendix A: ITC analysis of alanine analogue binding to HIF-1α-PAS-B$_{238-349}$

The binding of each Tat-tagged alanine analogue to His-HIF-1α-PAS-B$_{238-349}$ was investigated by ITC. To ensure a buffer match during the assay the protein was dialysed into EMSA buffer and the inhibitors dissolved into the dialysis solution. Experiments were carried out using a Micro-ITC calorimeter (Microcal, GE Healthcare, UK) at 25°C. For ITC measurement HIF-1α-PAS-B$_{238-349}$ was loaded into the measurement cell and the inhibitor was added in small increments under continuous stirring. The concentration of inhibitor used was 168 μM, a 10-fold excess to the protein solution. With exception of Tat-*cyclo*-CLLFVA, all concentrations were calculated by absorbance measurement at 280 nm. Data was analysed using Microcal Origin software.

The heat of dilution was measured for each and was taken from the ligand peptide results to give a final scatter graph. No conclusive binding for any inhibitor was observed (Figure 111-116).

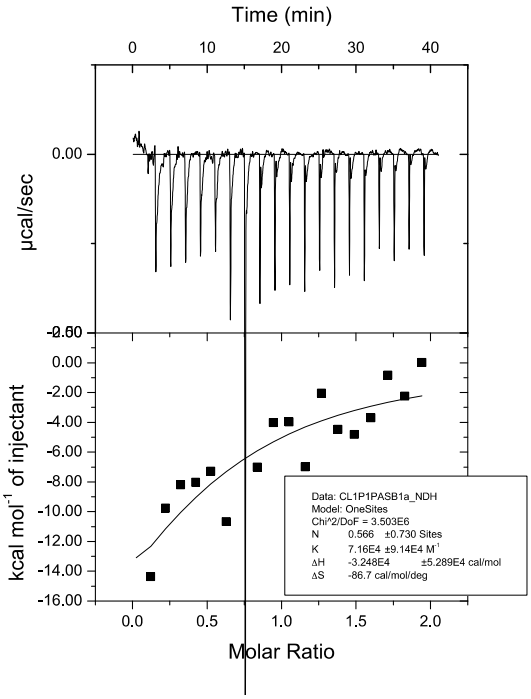


Figure 112 Graph of ITC of His-HIF-1α-PAS-B$_{238-349}$ with P1.

P1 (168 μM) was titrated with His-HIF-1α-PAS-B$_{238-349}$ (16.8 μM) at 25°C. The graph was adjusted to account for the measured heat of dilution of inhibitor addition to buffer.

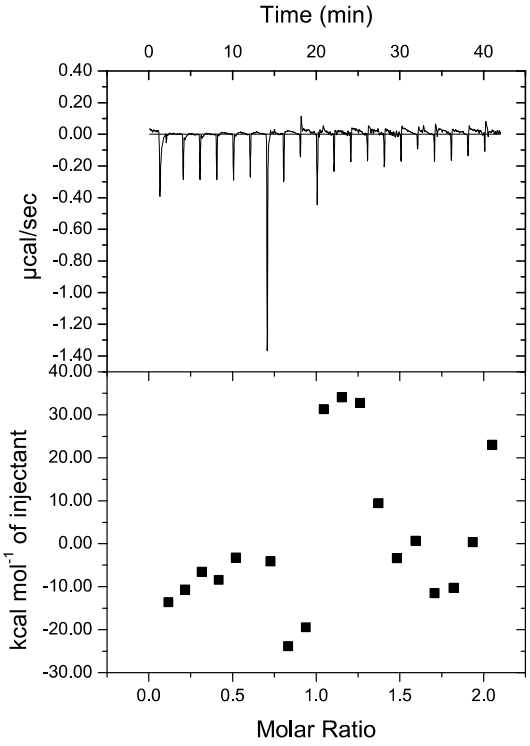


Figure 113 Graph of ITC of His-HIF-1α-PAS-$B_{238-349}$ with Tat-*cyclo*-CALFVY.

Tat-*cyclo*-CALFVY (168 µM) was titrated with His-HIF-1α-PAS-$B_{238-349}$ (16.8 µM) at 25°C. The graph was adjusted to account for the measured heat of dilution of inhibitor addition to buffer. The trace was not as expected and no binding was observed.

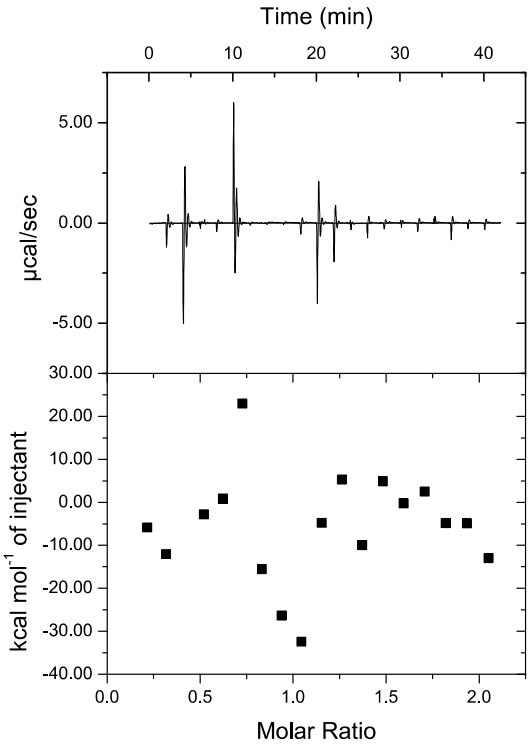


Figure 114 Graph of ITC of His-HIF-1α-PAS-$B_{238-349}$ with Tat-*cyclo*-CLAFVY.

Tat-*cyclo*-CLAFVY (168 μM) was titrated with His-HIF-1α-PAS-$B_{238-349}$ (16.8 μM) at 25°C. The graph was adjusted to account for the measured heat of dilution of inhibitor addition to buffer. The trace was not as expected and no binding was observed.

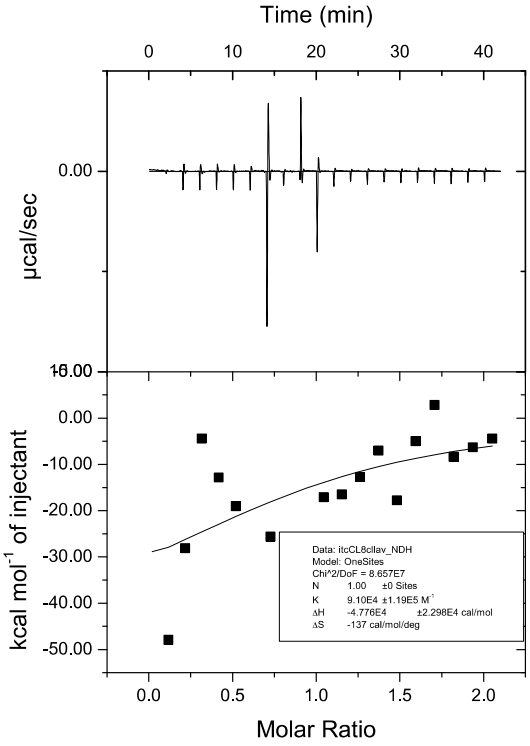


Figure 115 Graph of ITC of His-HIF-1α-PAS-$B_{238-349}$ with Tat-*cyclo*-CLLAVY.

Tat-*cyclo*-CLLAVY (168 µM) was titrated with His-HIF-1α-PAS-$B_{238-349}$ (16.8 µM) at 25°C. The graph was adjusted to account for the measured heat of dilution of inhibitor addition to buffer. The trace was not as expected and no binding was observed.

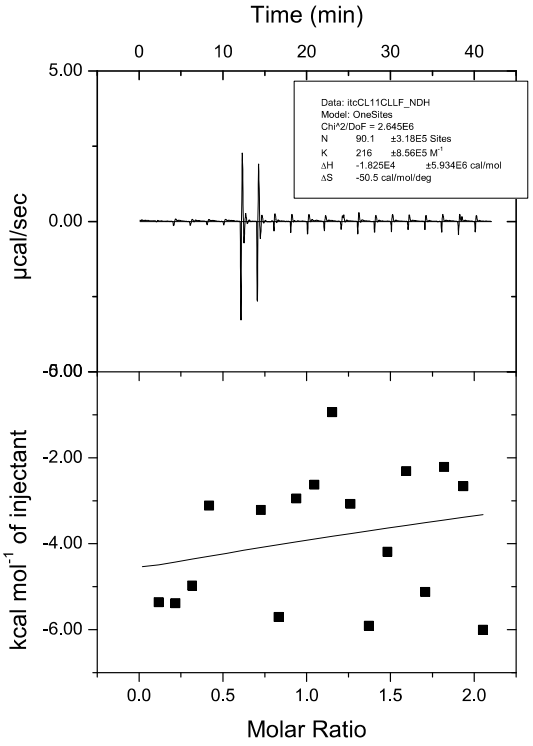


Figure 116 Graph of ITC of His-HIF-1α-PAS-$B_{238-349}$ with Tat-*cyclo*-CLLFAY. Tat-*cyclo*-CLLFAY (168 µM) was titrated with His-HIF-1α-PAS-$B_{238-349}$ (16.8 µM) at 25°C. The graph was adjusted to account for the measured heat of dilution of inhibitor addition to buffer. The trace was not as expected and no binding was observed.

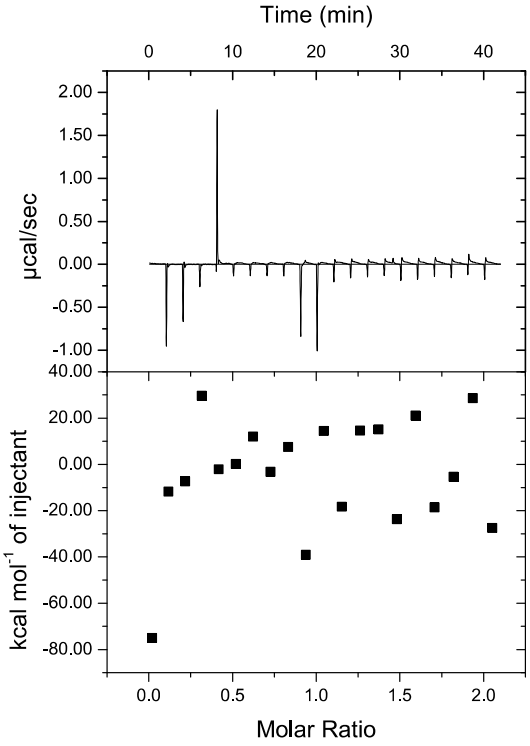


Figure 117 Graph of ITC of His-HIF-1α-PAS-$B_{238-349}$ with Tat-*cyclo*-CLLFVA.

Tat-*cyclo*-CLLFVA (168 µM) was titrated with His-HIF-1α-PAS-$B_{238-349}$ (16.8 µM) at 25°C. The graph was adjusted to account for the measured heat of dilution of inhibitor addition to buffer. The trace was not as expected and no binding was observed.

Appendix B: RT-PCR

Initial set-up of a RT-PCR assay to test the alanine analogues effect on the expression of HIF-1 gene (VEGF, GLUT1) was carried out. Initially, MCF-7 cells were cultured under both Hypoxia (1% oxygen in a hypoxia workstation) and normoxia. RT-PCR was then carried out using probes that amplified the VEGF and GLUT1 genes. The results were normalised to housekeeper genes; as expected and in agreement with the literature hypoxia led to up-regulation of these reporter genes.[1, 127, 189, 190] These results were repeatable and lay the foundations fro the development of the assay (Figure 118).

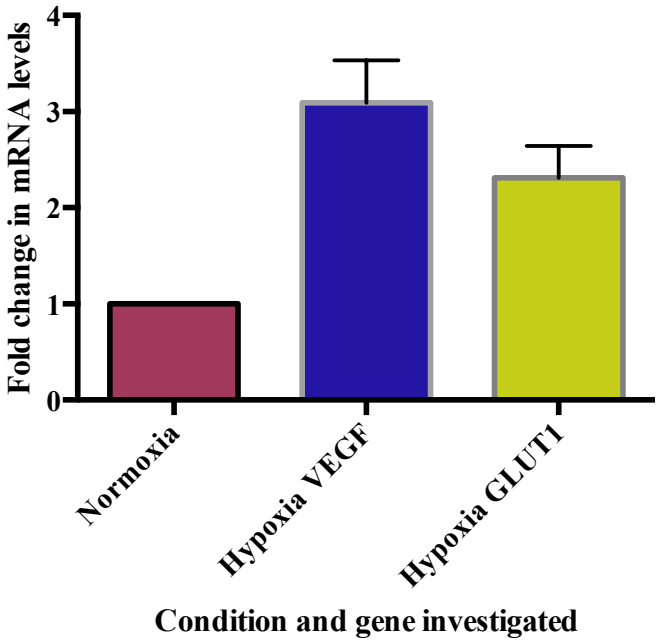


Figure 118 RT-PCR of VEGF and GLUT1 level in response to hypoxia and normoxia.
The results are shown as fold change from normoxia. Results are from n = 3.

RT-PCR methodology

TaqMan probes (life technology, UK)	Manufacturer's code
VEGFA	Hs00900055_m1
SLC2A1 (GLUT-1)	Hs00892681_m1
18S	Hs99999901_s1
TBP	Hs99999903_m1
ActB (β Actin)	Hs00427621_m1

Table 64 List of TaqMan probes used for RT-PCR assays.

MCF-7 cells were split, counted and plated 450,000 cells per well of a 6 cm round plate (as in section 5.2.3.1). The cells were incubated at 37°C for 24 hours (approximately 60% confluency reached). At this time the plates were split into normoxic and hypoxic conditions and either placed in a 'normal' 20% O_2 incubator or in a 1% hypoxia workstation (H35 Hypoxystation, don whitley scientific, UK), respectively. The cells were incubated for a further 16 hours. After this time the media was aspirated and the cells washed with PBS (2 ml) and detached with trypsin (1 ml, for hypoxic samples this was undertaken under hypoxia). The trypsin was deactivated by the addition of DMEM/FBS (3 ml). The cell suspension was then centrifuged (1000 rpm, 4 minutes). The supernatant was removed and the cell pellet resuspended in Lysis buffer RLT (400 μl). This buffer was from the RNeasy kit (QIAGEN, UK), used as per the manufacturer's instructions.

The mRNA was then reverse transcribed. The reverse transcription samples were made as followed using iScript reverse transcriptase (Bio-Rad, UK, Table 65).

Component	Volume (μl)
5 X iScript reverse transcription reaction mix (Bio-Rad, USA)	4
iScript reverse transcriptase (Bio-Rad, USA)	1
RNA template (1 μg)	Variable
Nuclease free water	Make up to 20 μl

Table 65 Composition of iScript reverse transcription reaction mix for 1 μg of RNA template.

The reaction mixture can be scaled up for larger quantities of RNA.

The reverse transcription mixtures were then run on the following program in a PCR machine (Table 66).

Step	Temperature	Time
Priming	25°C	5 min
Reverse transcription	42°C	30 min
Reverse transcriptase inactivation	85°C	5 min

Table 66 iScript reverse transcriptase program

These cDNA samples were then used for the RT-PCR. The TaqMan gene expression assay (life technologies, UK) was utilised following the manufacturer's instructions. Each sample was analysed in triplicate well repeats, each well contained 20 μl samples incorporating 2 μl of the cDNA solution.

Initially, a master mix for each TaqMan (life technologies, UK) gene expression probe was made (Table 67). To avoid a lack of sample an extra 15% of master mix was made.

Component	Volume required for assay (μl)	Volume for master mix (+15%, μl)
2 X TaqMan Univerisal PCR master mix (life technologies, UK)	10	(10 x n) x 15%
20 X TaqMan gene expression probes (specific to each gene, life technologies, UK)	1	(1 x n) x 15%
Nuclease free water	7	(7 x n) x 15%

Table 67 Composition of TaqMan RT-PCR master mixes for each gene amplification mix.

n = number of wells required for RT-PCR.

The master mix was then split into smaller aliquots for each triplicate ((18 μl x n) + 10%), where n = 3 for triplicate well repeats). The cDNA was then added for each sample ((2 μl x n) + 10%). Each aliquot was 66 μl (20 μl per well x triplicate + 10 % extra). The samples were set up in the dark to avoid deactivation of sensitive fluorescent probes. 20 μl of each

sample was loaded into the 96-well RT-PCR plate, again in triplicate. The plate was briefly centrifuged to remove any bubbles (1000 rpm, 30 seconds) and then sealed.

Two control samples were done on each run, a no reverse transcriptase sample and a no template. This is was to ensure no contamination had occurred during experimental set up.

The RT-PCR was then run in CFX connect Real-Time System (Bio-Rad, USA), using a RT-PCR program based on manufacturer's instructions (Table 68).

Step	**Temperature**	**Time**	**Number of cycles**
Hold	50°C	2 min	1
Initial denaturation	95°C	10 min	1
Denaturation	95°C	15 s	30
Hold + plate read	60°C	1 min	
Melt curve			
Hold	95°C	10 s	1
Melt + plate read	95°C to 65°C	Increment 0.5°C per 5 s	1
End			

Table 68 RT-PCR program.

Relative quantification of gene expression was calculated using the comparative C_T methods $\Delta\Delta C_T$ and normalised to the expression of the housekeeper genes (TBP and 18 S).

Appendix C: Publication from this work

A Cyclic Peptide Inhibitor of HIF-1 Heterodimerisation That Inhibits Hypoxia Signalling in Cancer Cells.

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, *Journal of the American Chemical Society*, 2013, **135**, 10418-10425.

JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

pubs.acs.org/JACS

A Cyclic Peptide Inhibitor of HIF-1 Heterodimerization That Inhibits Hypoxia Signaling in Cancer Cells

Elena Miranda,[†,#] Ida K. Nordgren,[†,#] Abigail L. Male,[†] Charlotte E. Lawrence,[†] Franciane Hoakwie,[†] Francesco Cuda,[†] William Court,[∥] Keith R. Fox,[‡,§] Paul A. Townsend,[§,⊥,▽] Graham K. Packham,[§,⊥] Suzanne A. Eccles,[∥] and Ali Tavassoli*[,†,§,⊥]

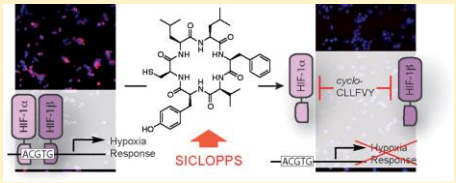
[†]Chemistry, [‡]Centre for Biological Sciences, [§]Institute for Life Sciences, University of Southampton, Southampton SO17 1BJ, United Kingdom

[∥]Cancer Research UK Cancer Therapeutics Unit, The Institute of Cancer Research, Sutton SM2 5NG, United Kingdom

[⊥]Cancer Sciences, Faculty of Medicine, University of Southampton, Southampton SO16 6YD, United Kingdom

Supporting Information

ABSTRACT: Hypoxia inducible factor-1 (HIF-1) is a heterodimeric transcription factor that acts as the master regulator of cellular response to reduced oxygen levels, thus playing a key role in the adaptation, survival, and progression of tumors. Here we report *cyclo*-CLLFVY, identified from a library of 3.2 million cyclic hexapeptides using a genetically encoded high-throughput screening platform, as an inhibitor of the HIF-1α/HIF-1β protein−protein interaction in vitro and in cells. The identified compound inhibits HIF-1 dimerization and transcription activity by binding to the PAS-B domain of HIF-1α, reducing HIF-1-mediated hypoxia response signaling in a variety of cell lines, without affecting the function of the closely related HIF-2 isoform. The reported cyclic peptide demonstrates the utility of our high-throughput screening platform for the identification of protein−protein interaction inhibitors, and forms the starting point for the development of HIF-1 targeted cancer therapeutics.



INTRODUCTION

Homeostasis of oxygen, a key metabolite, is critical for mammalian cell survival. This necessitates a robust network that senses and rapidly responds to hypoxia (low oxygen levels). The key component of this hypoxia response network is hypoxia-inducible factor (HIF), a heterodimeric transcription factor composed of an oxygen-regulated α-subunit and a ubiquitously expressed β-subunit (also known as the aryl nuclear transcription factor or ARNT). Mammals possess three isoforms of HIF-α; the ubiquitously expressed HIF-1α mounts the immediate response to reductions in cellular oxygen,[1] while HIF-2α (also known as EPAS1) and HIF-3α are thought to regulate the response to prolonged hypoxia. While the intricate interplay between HIF-α isoforms in cancer is complicated and yet to be fully deciphered,[2] the role of HIF-1 activity in angiogenesis, tumor growth, and metastasis is well established.[3,4] HIF-1α is overexpressed in many cancers,[5] and oncogene activation and loss of tumor suppressor function is shown to be associated with HIF-1 activation.[6]

HIF-1α is negatively regulated at the protein level by oxygen via prolyl hydroxylase enzymes that use oxygen as a substrate for the hydroxylation of residues 402 and 564 of HIF-1α, marking it for ubiquitination by an E3 ubiquitin ligase complex and rapid proteolysis.[7,8] Reduced oxygen levels lead to the stabilization and nuclear translocation of HIF-1α,[9] where it binds HIF-1β to form the HIF-1 transcription factor complex. HIF-1 rapidly mounts a transcriptional response to hypoxia[1,10] by directing the expression of a wide variety of hypoxia response genes.[11,12] By utilizing changes in the substrate concentration of a continuously occurring enzymatic reaction (hydroxylation of HIF-1α), the cellular response to hypoxia is near instantaneous,[13] with HIF-1α acting as both the sensor and a key component of the hypoxia response machinery.

Inhibition of HIF-1 has long been known to hold much potential for cancer therapy;[14] there are multiple possible points for therapeutic intervention in the hypoxia response network, and molecules that inhibit various components of this diverse pathway have been reported,[15−22] but the absolute requirement for the dimerization of HIF-1α and HIF-1β for DNA binding and transcription activity of the HIF-1 complex makes this protein−protein interaction a seemingly optimal point of interception. Several high-throughput screens have been conducted in the effort to identify HIF-1 inhibitors,[15,19,21,22] but there are currently no selective inhibitors of the HIF-1α/HIF-1β protein−protein interaction. The only reported inhibitor of HIF-1 dimerization is the heteroaromatic

Received: March 25, 2013
Published: June 24, 2013

acridine derivative acriflavine, which nonselectively inhibits both HIF-1 and HIF-2.[19]

A compound that specifically inhibits HIF-1 dimerization in cells will not only serve as a chemical tool to decipher the mechanism of hypoxia response, but would also form the starting point for the development of HIF-1-directed therapeutic agents. Here we report an inhibitor of the HIF-1α/HIF-1β protein−protein interaction, identified from a genetically encoded library of 3.2 million cyclic peptides. The most potent identified cyclic peptide (cyclo-CLLFVY) inhibits HIF-1 dimerization in vitro and in cells by binding to the PAS-B domain of HIF-1α, and prevents HIF-1- but not HIF-2-mediated hypoxia signaling in a variety of cell lines.

RESULTS AND DISCUSSION

Identification of HIF-1 Heterodimerization Inhibitors Using a Genetically Encoded High-Throughput Screening Platform. We used our genetically encoded high-throughput screening platform[23−25] to identify inhibitors of the HIF-1α/HIF-1β protein−protein interaction. We built and verified the function of a HIF-1 bacterial reverse two-hybrid system (RTHS; for a detailed description see the Supporting Information and Figures S1−S3). The HIF-1 RTHS was used to screen a plasmid-encoded SICLOPPS (split intein circular ligation of peptides and proteins)[26,27] library of 3.2 million cyclic hexapeptides for HIF-1 dimerization inhibitors.

After the first round of screening, 120 surviving colonies were observed and subjected to several rounds of secondary screening to eliminate false positives and nonselective inhibitors, leaving 12 potential HIF-1 dimerization inhibitors that were ranked by drop-spotting. Four of these peptides were significantly more active than the others; the SICLOPPS plasmids encoding these 4 peptides were sequenced to reveal the identity of the HIF-1 inhibitors as cyclo-CLLFVY (encoded by two SICLOPPS plasmids), cyclo-CRLMVL, and cyclo-CLLRMY (Figure 1, R = H). Interestingly, the two isolated

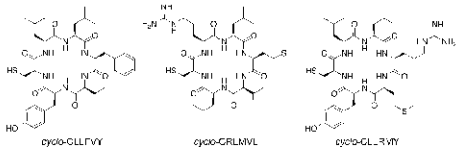


Figure 1. Cyclic peptide HIF-1 inhibitors. Cyclic peptide HIF-1 inhibitors identified from a SICLOPPS library of 3.2 million cyclic hexapeptides (R = H).

cyclo-CLLFVY plasmids encoded the peptide via different codons. To ensure that the cyclic peptides, not their unspliced peptide aptamers, were responsible for the observed inhibition of HIF-1, we utilized mutant SICLOPPS dnaE C-terminal inteins (H24A, F26A) that do not splice.[28] The unspliced peptides lost the ability of their parent cyclic equivalents to disrupt HIF-1α/HIF-1β dimerization (Figure S3B), demonstrating that HIF-1 inhibition by these peptides is dependent on their cyclic form.

The three identified HIF-1 inhibitors were synthesized and tagged with Tat peptide (via a disulfide bond between the set cysteine of the cyclic peptide and a cysteine introduced to the start of Tat)[25] to aid the translocation across the plasma membrane of mammalian cells. In the following experiments, Tat-cyclo-CLLFVY is referred to as P1, Tat-cyclo-CRLMVL as P2, and Tat-cyclo-CLLRMY as P3 (Figure 1, R = CGRKKRRQRRRPPQ).

P1 Inhibits HIF-1 Activity in a Mammalian Cell Luciferase Reporter Assay. The ability of P1−P3 to disrupt HIF-1 function in cells was assessed using a HIF-1-dependent luciferase reporter assay.[15,21] The assay uses human osteosarcoma U2OS cells, stably transfected with a HIF-dependent luciferase reporter construct (U2OS-HRE-luc), where activation of HIF results in an increase in luciferase expression.[15]

Hypoxia (1% O_2) results in a ~12-fold increase in the luciferase signal, which is inhibited in a dose-dependent manner by P1 (IC_{50} of 19 ± 2 μM); P1 did not alter basal luciferase activity in normoxia (Figure 2A). P2 and P3 did not show an effect on the luciferase reporter in hypoxia or normoxia. P1−P3

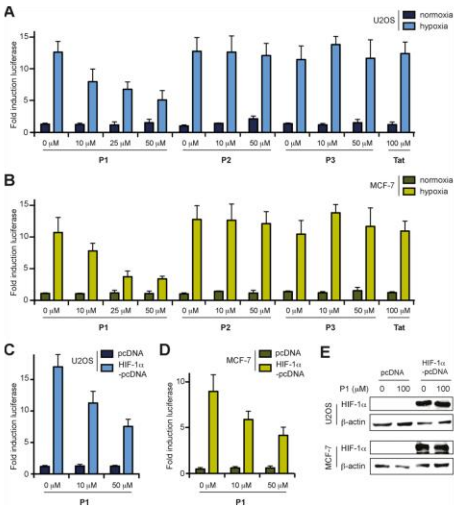


Figure 2. Effect of P1−P3 in a HIF-1 luciferase-reporter assay. Data shows fold-increase of the luciferase signal compared to untreated cells in normoxia. (A) P1 causes a dose-dependent reduction in the HIF-1-mediated luciferase signal in hypoxic U2OS cells. (B) Assay in (A) repeated in MCF-7 cells. (C) P1 inhibits the HIF-1-mediated luciferase signal in a normoxic reporter assay in U2OS cells. (D) Assay in (C) repeated in MCF-7 cells. (E) Representative blots showing that 100 μM P1 does not alter HIF-1α protein levels in the cells in (C) and (D).

had no effect in the U2OS-luc control cell line[15] stably expressing luciferase (Figure S4), indicating that P1 does not inhibit endogenous cellular processes such as transcription or translation. To assess the cell-specificity of P1, the experiment was repeated in MCF-7 breast cancer cells with similar results (P1 IC_{50} of 16 ± 1 μM) (Figure 2B). Tat-tag alone (100 μM) did not affect the luciferase signal in these assays (Figure 2A and 2B).

HIF-1 has been shown to directly promote the expression of HIF-1α in hypoxia by binding to a hypoxia-response element (HRE) upstream of the HIF-1α.[29] This has been shown to be dependent on the methylation state of a cytosine in the HIF-1 binding site (A<u>C</u>GTG) upstream of HIF-1α; a methylated

cytosine prevents HIF-1 binding and inhibits the autotransactivation of HIF-1α in hypoxia.[29] Bisulfite sequencing of this region of MCF-7 and U2OS cells revealed this HRE to be unmethylated, and thus, an inhibitor of HIF-1α/HIF-1β dimerization would potentially reduce the HIF-1-promoted upregulation of HIF-1α mRNA and protein levels in these cell lines. The reduction of luciferase signal observed in P1-treated U2OS and MCF-7 cells (Figure 2A and B) could therefore be partially a result of a reduction in hypoxic HIF-1α levels (this would be the case for any inhibitor of HIF-1 dimerization in these cell lines). To decouple P1's effect on HIF-1α transactivation from its effect on HIF-1α/HIF-1β dimerization, a normoxic luciferase-reporter assay was devised where HIF-1α is expressed from a transiently transfected vector, resulting in continuously elevated levels of HIF-1α in normoxic cells. P1 continued to inhibit the luciferase signal in both U2OS and MCF-7 cells in this assay (Figure 2C and D), with HIF-1α protein levels not being altered by P1 (Figure 2E).

The observed discrepancy between the activity of *cyclo*-CRLMVL and *cyclo*-CLLRMY in the HIF-1 RTHS (Figure S3B), and the lack of activity of P2 and P3 in the luciferase assay (Figure 2) is likely due to the ±10-fold range in each step of the drop-spotting assay; P2 and P3 may be up to 10-fold less active than P1 and still result in the same drop-spotting pattern. P1 was therefore taken forward for further assessment of its activity. P2, a Tat-tagged cyclic hexapeptide that differs from P1 by two amino acids, was used as a negative control in the following experiments.

***cyclo*-CLLFVY Disrupts HIF-1, but Not HIF-2 Dimerization in vitro by Binding to the PAS-B Domain of HIF-1α.** We next probed the effect of *cyclo*-CLLFVY on the interaction of HIF-1α with HIF-1β in vitro. Recombinant His-HIF-1α_{1-350} and GST-HIF-1β_{1-474} were produced and purified. Electrophoretic mobility shift assay (EMSA) was used to demonstrate that the recombinant HIF-1 proteins form functional heterodimers (Figure S5A). Unfortunately, the positively charged Tat-tag of P1 is incompatible with EMSA (interferes with the bandshift of DNA). We therefore assessed the ability of P1 to disrupt HIF-1 dimerization by enzyme-linked immunosorbent assay (ELISA); P1 was found to disrupt the protein−protein interaction of His-HIF-1α_{1-350} and GST-HIF-1β_{1-474} with an IC_{50} of 1.3 μM (Figure 3A). The control compound P2 had no effect on HIF-1 dimerization in this assay (Figure S6). To determine the HIF-1-specificity of *cyclo*-CLLFVY, we assessed its ability to disrupt HIF-2 dimerization by ELISA. Recombinant His-HIF-2α_{1-351} was produced and purified, and shown to form functional heterodimers with GST-HIF-1β_{1-474} by EMSA (Figure S5B). P1 had no effect on the dimerization of His-HIF-2α_{1-351} with GST-HIF-1β_{1-474} (Figure 3B).

To identify the target (HIF-1α or HIF-1β) of our HIF-1 inhibitor, we synthesized a biotinylated derivative of *cyclo*-CLLFVY for use as bait in pull-down assays by replacing the cysteine residue (present in all members of the SICLOPPS library to allow intein splicing) with propargylalanine. This compound was linked to biotin-PEG-azide by click-chemistry (copper-catalyzed alkyne azide reaction) to give biotin-PEG-triazole-*cyclo*-ALLFVY as the bait molecule. The bait was immobilized onto streptavidin-coated beads and mixed with recombinant His-HIF-1α_{1-350} and GST-HIF-1β_{1-474}. The pulled-down protein(s) were analyzed by Western blot, revealing His-HIF-1α_{1-350} as the target of *cyclo*-CLLFVY (Figure 3C, lane 3). Streptavidin beads did not pull down either HIF-1 subunit in the absence of the bait molecule (Figure 3C, lane 2), and propargylalanine click-linked to biotin-PEG-azide did not pull down either HIF-1 subunit (Figure 3C, lane 4). To verify binding of *cyclo*-CLLFVY to HIF-1α, we used the propargylalanine derivative of this molecule and azide-Megastoke dye 673 to synthesize a fluorescent derivative by click-chemistry. His-HIF-1α_{1-350} and His-HIF-2α_{1-351} were immobilized onto Ni^{2+}-coated 96-well plates, and the fluorescent analogue of *cyclo*-CLLFVY was washed over these proteins. Binding of the fluorescent derivative to these proteins was monitored via increased fluorescence at 680 nm. We observed binding of the fluorescent derivative of *cyclo*-CLLFVY to His-HIF-1α_{1-350}, with binding to His-HIF-2α_{1-351} at close to background levels (Figure 3D).

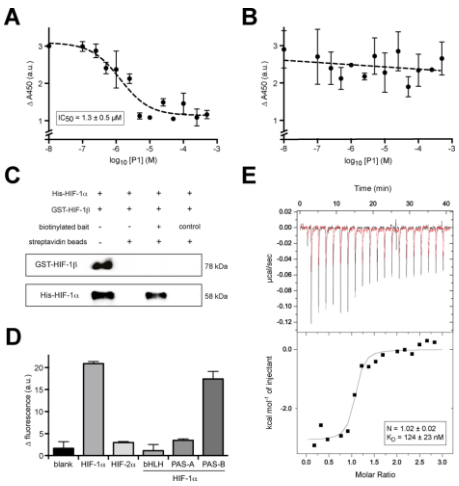


Figure 3. Assessing the activity of *cyclo*-CLLFVY in vitro. (A) Effect of 10 nM to 500 μM P1 on the heterodimerization of His-HIF-1α_{1-350} with GST-HIF-1β_{1-474} analyzed by ELISA; P1 disrupts this interaction with an IC_{50} of 1.3 ± 0.5 μM. (B) Effect of 10 nM to 500 μM P1 on the interaction of His-HIF-2α_{1-351} with GST-HIF-1β_{1-474}; P1 does not affect HIF-2 heterodimerization. (C) His-HIF-1α_{1-350} is selectively pulled-down by streptavidin beads coated with biotin-PEG-triazole-*cyclo*-ALLFVY (lane 3); lane 1 is the loading control. Neither protein is pulled-down by streptavidin beads alone (lane 2), or streptavidin beads coated with a biotin-linked control (lane 4). (D) Fluorescent binding assay showing a Megastoke 673-labeled derivative of *cyclo*-CLLFVY binding to His-HIF-1α_{1-350}, while its binding to His-HIF-2α_{1-351} is close to background levels. The fluorescent derivative of *cyclo*-CLLFVY binds the PAS-B domain of HIF-1α, whereas binding to bHLH and PAS-A domains are close to background level. (E) ITC shows P1 binding to the PAS-B domain of HIF-1α with 1:1 stoichiometry and 124 ± 23 nM affinity. Red line shows control P1 injection into buffer only.

We next sought to identify the domain of HIF-1α bound by *cyclo*-CLLFVY. The HIF-1α protein used in our RTHS is composed of the DNA-binding basic-helix−loop−helix domain (bHLH, amino acids 1−80) and the protein−protein interaction Per-ARNT-SIM-A (PAS-A, amino acids 90−155) and Per-ARNT-Sim-B (PAS-B, amino acids 235−350) domains. Recombinant His-bHLH, His-PAS-A, and His-PAS-B domains were immobilized onto Ni^{2+}-coated 96-well plates, and the fluorescent derivative of *cyclo*-CLLFVY was washed over the bound proteins. We observed an increase in fluorescence (at 680 nm) of the PAS-B domain,

whereas bHLH and PAS-A remained close to background, indicating binding of our inhibitor to the PAS-B domain of HIF-1α (Figure 3D). We next used P1 in isothermal titration calorimetry (ITC) to verify these observations and quantify binding affinities; P1 bound the PAS-B domain of HIF-1α in 1:1 stoichiometry and with a K_D of 124 nM (Figure 3E). P1 did not bind to the bHLH or PAS-A domain of HIF-1α (Figures S7A and B), and P2 did not bind the HIF-1α PAS-B domain (Figure S7C). P1 did not bind HIF-1β_{1-474} (Figure S7D). The observed selectivity of P1 for HIF-1 over HIF-2 in vitro (Figure 3A versus B, and 3D) suggests the possibility of selectivity for HIF-1 over HIF-2 in cells.

P1 Disrupts HIF-1 Dimerization in MCF-7 and U2OS Cells. The effect of P1 on the endogenous HIF-1α/HIF-1β interaction in intact cells was directly probed using an in situ proximity ligation assay (PLA);[30] primary antibodies (HIF-1α and HIF-1β here) raised in different species are bound by specific secondary antibodies that are tagged with a short DNA strand. The DNA on the interacting PLA probes forms a mini-plasmid that is amplified and bound by a red fluorescent dye. The HIF-1α/HIF-1β interaction is thus visualized as red dots in the DAPI-stained nuclei of MCF-7 and U2OS cells (Figure 4).

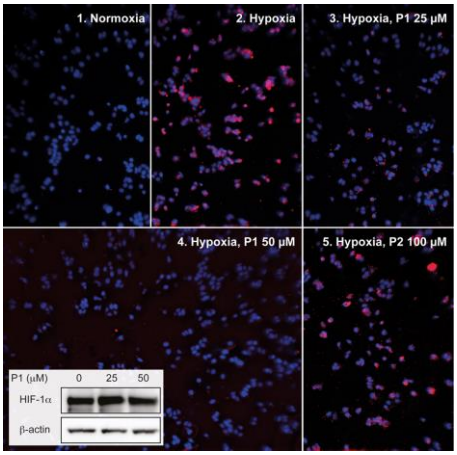


Figure 4. Inhibition of HIF-1 dimerization by *cyclo*-CLLFVY assessed by immunofluorescence detection of endogenous HIF-1α/HIF-1β dimerization by in situ PLA in MCF-7 cells. The PLA signal is absent in normoxia (panel 1) but readily observed in hypoxia (panel 2). P1-treatment (25 or 50 μM) of hypoxic cells results in a loss of the PLA signal (panels 3 and 4, respectively), whereas 100 μM P2 shows no effect (panel 5). Inset: Western blot analysis of cell lysates show that P1 does not affect HIF-1α levels in this assay.

We observed an increase in PLA signal in MCF-7 cells incubated in hypoxia (1% O_2) for 4 h, indicating the expected formation of the HIF-1 dimer in hypoxia (Figure 4, panel 1 versus panel 2). Hypoxic MCF-7 cells dosed with 25 and 50 μM P1 showed a reduction in the HIF-1 PLA signal (Figure 4, panels 3 and 4). P2 at 100 μM did not have any effect on the PLA signal (Figure 4, panel 5). The observed effect is not due to a reduction in HIF-1α by P1 (Western blot inset in Figure 4) and thus may be solely attributed to disruption of HIF-1α/HIF-1β dimers.

P1 Inhibits HIF-1-Mediated Cellular Hypoxia Response. The effect of P1's disruption of the HIF-1α/HIF-1β heterodimer on cellular hypoxia response was next characterized. We measured the effect of P1 on vascular endothelial growth factor (VEGF), a HIF-1 regulated gene that stimulates vasculogenesis and angiogenesis in hypoxia.[31] The transcription of VEGF increased ~3-fold after incubation for 16 h in hypoxia in untreated MCF-7 and U2OS cells; P1-treated cells showed a dose-dependent reduction in VEGF mRNA in both cell lines, while the control peptide P2 (100 μM) had no effect (Figure 5A). VEGF protein levels increased 3- to 5-fold in both MCF-7 and U2OS cells after 16 h of incubation in hypoxia, with P1-treatment resulting in a dose-dependent reduction of VEGF protein (measured by a quantitative immunoassay) in both cell lines; pretreatment with 50 μM of P1 fully inhibited the hypoxic induction of VEGF protein, resulting in normoxic VEGF levels in hypoxic MCF-7 and U2OS cells (Figures 5B). As VEGF is a regulator of angiogenesis, we next probed the effect of P1 on HIF-1 mediated tubule formation in hypoxic human umbilical vein endothelial cells (HUVEC),[32] and observed a dose-dependent reduction of HUVEC tubularization in P1-treated cells, with no effect from P2 (Figure 5C; for representative images, see Figure S8). The effect of P1 on HIF-1 signaling was further assessed via carbonic anhydrase IX (CAIX), an extracellular metalloenzyme whose expression is significantly upregulated (~25-fold) by HIF-1 after 16 h incubation in hypoxia.[33] Treatment of MCF-7 and U2OS cells with 50 μM P1 resulted in a 5-fold reduction of CAIX mRNA in hypoxia, with no effect from 100 μM P2 (Figure 5D). A sulforhodamine B (SRB) cytotoxicity assay[34] was used to demonstrate that the observed effects are not due to toxicity of P1 (Figure 5E). The cytotoxicity and effect of P1 on cell proliferation was further probed by measuring the effect of increasing doses of P1 (1–100 μM) on MCF-7 cells over 72 h, with 1 μM 4-hydroxytamoxifen used as a positive control. P1 at 100 μM did not affect the viability of MCF-7 cells, whereas 1 μM 4-hydroxytamoxifen inhibited proliferation as expected (Figure 5F).

P1 Does Not Affect HIF-2-Mediated Hypoxia Signaling. The cellular HIF-1-specificity of P1 was probed using 786-O cells, a VHL-defective renal cell carcinomal line that does not express detectable levels of HIF-1α, but instead expresses HIF-2α at a high constitutive level.[35] This results in regulation of hypoxia response genes such as VEGF and lysyl oxidase (LOX) in 786-O cells by HIF-2 instead of HIF-1.[36] We reasoned that if P1 also inhibits the dimerization of HIF-2 in cells, a dose-dependent reduction in the mRNA and protein products of these hypoxia response genes would be observed in hypoxic 786-O cells, whereas a specific HIF-1 inhibitor would not affect hypoxia-response in this cell line. In contrast to MCF-7 and U2OS cells, P1 at concentrations up to 100 μM had no effect on a variety of HIF-2-promoted genes in 786-O cells (Figure 6A). In addition, P1 had no effect on VEGF protein levels in hypoxic 786-O cells (Figure 6B). To further probe the effect of P1 on HIF-2, a luciferase reporter assay was developed and used. The assay was derived from the HIF-1 luciferase reporter assay detailed above (Figure 2C and D) and adapted for HIF-2 by replacing the plasmid encoding HIF-1α with the equivalent plasmid encoding HIF-2α. P1 did not affect the luciferase reporter signal in MCF-7 or U2OS cells (Figure 6C and D), providing additional evidence that P1 inhibits HIF-1, but not HIF-2 signaling in cells. P2 and P3 also had no effect on HIF-2 in this assay (Figure S9).

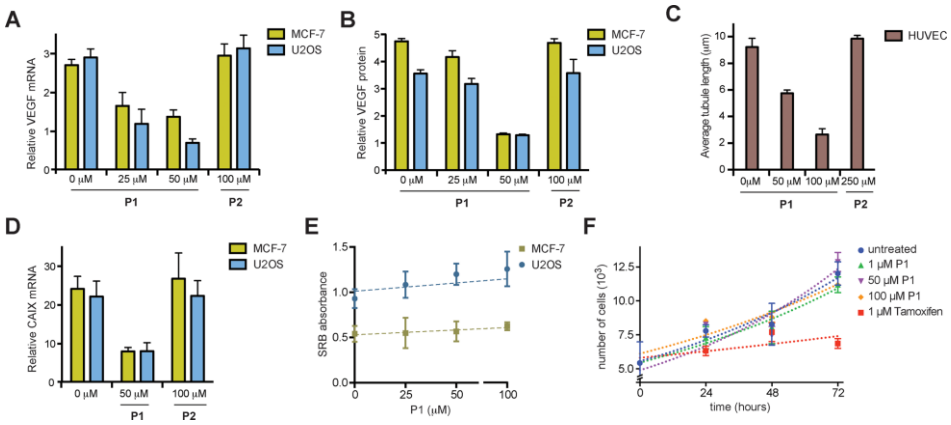


Figure 5. Inhibition of hypoxia response in MCF-7 and U2OS cells by P1. (A) qPCR analysis shows the dose-dependent reduction of VEGF mRNA, elevated by ~3-fold in hypoxia (1 = normoxic levels), by P1 in hypoxic MCF-7 and U2OS cells; 100 μM P2 (control) has no effect in either cell line. (B) VEGF protein level, elevated by 3- to 5-fold in hypoxia (1 = normoxic levels), is also reduced by P1 in hypoxic MCF-7 and U2OS cells; 100 μM P2 (control) has no effect in either cell line. (C) P1 pretreatment causes the dose-dependent reduction of tubule length in hypoxic HUVECs; 250 μM P2 (control) has no effect. For representative images, see Figure S8. (D) qPCR analysis shows 50 μM P1 significantly reduces CAIX mRNA, which is elevated by ~25-fold in hypoxia (1 = normoxic levels); 100 μM P2 (control) has no effect. (E) SRB assay results reveal that 100 μM P1 is not cytotoxic to MCF-7 or U2OS cells. (F) Cell proliferation assays show that 100 μM P1 has no effect on MCF-7 proliferation over 72 h, whereas 1 μM 4-hydroxytamoxifen (control) arrests cell division.

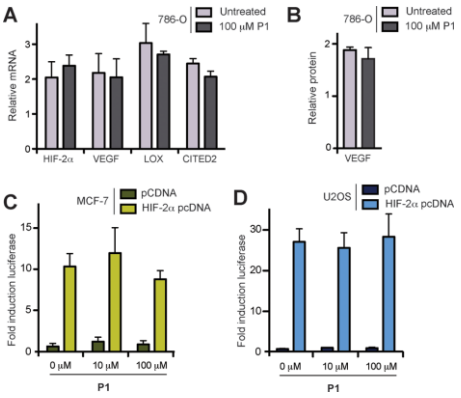


Figure 6. P1 does not affect HIF-2 mediated hypoxia signaling. (A) 100 μM P1 does not affect the transcription of HIF-2α, or HIF-2-promoted VEGF, LOX, CITED2 in hypoxic 786-O cells. (B) 100 μM P1 does not affect VEGF protein levels in hypoxic 786-O cells. (C) P1 does not affect the reporter signal in a HIF-2 luciferase reporter assay in MCF-7 cells. (D) P1 does not affect the reporter signal in a HIF-2 luciferase reporter assay in U2OS cells.

■ CONCLUSIONS

There is extensive evidence verifying HIF-1 as key in multiple processes critical to cancer progression, and thus a target of significant potential for cancer therapy.[3−5,37] Recent findings such as HIF-1 regulating the survival of tumor cells that escape radiation therapy[38] provides additional evidence for its significance as a target. But the challenges of identifying protein−protein interaction inhibitors in the absence of structural data,[39] combined with the difficulties associated with uncovering transcription factor inhibitors, are substantial. Peptides and macromolecules are increasingly viewed as the optimal scaffold for protein−protein interaction inhibitors,[40] and the high-throughout screening platform employed here has previously proven robust for the identification of cyclic peptide inhibitors of a variety of protein−protein interactions.[24,25,41] When employed for the identification of HIF-1 inhibitors, the platform identified *cyclo*-CLLFVY from a plasmid encoded library of 3.2 million cyclic peptides, via two independent plasmids. Verification of the function of this compound in vitro and in cells revealed that it disrupts HIF-1 dimerization by binding the PAS-B domain of HIF-1α, without affecting HIF-2.

Interestingly, the PAS-B domain of HIF-1α (and HIF-2α) is also targeted by acriflavine.[19] Furthermore, a recently reported heteroaromatic HIF-2 dimerization inhibitor also functions by binding a cavity on HIF-2α PAS-B;[42] although this compound was specifically developed to target HIF-2α PAS-B, the screen for acriflavine and our screen were not biased toward a single domain, yet both identified compounds binding to HIF-α PAS-B. This suggests the PAS-B domain of HIF-α as the optimal point of intervention for a HIF dimerization inhibitor. A homology model of HIF-1α PAS-B mapped onto the NMR structure of HIF-2α PAS-B suggests that the PAS-B cavity is substantially smaller in HIF-1α than HIF-2α.[42] This postulated difference in the cavity size between the two isoforms may be the source of the HIF-1 selectivity observed with *cyclo*-CLLFVY.

As well as its potential for cancer therapy, a specific inhibitor of HIF-1 serves as a chemical tool to verify hypotheses about the unique and sometimes opposing cellular function of HIF-1

and HIF-2.[2] The compound identified here will also serve to further illuminate the recently reported nontranscriptional function of HIF-1α.[43] A key challenge is the availability of tools to separate the transcriptional function of HIF-1α (or any other transcription factor) from its nontranscriptional function in cells.[41,44] Current approaches that knockdown or knockout the target protein (e.g., siRNA) will not suffice, as they will equally deplete both functions of the protein. In contrast, a HIF-1 dimerization inhibitor only affects the transcriptional function of HIF-1α.

We are currently conducting structural and SAR studies to guide the design of our second-generation HIF-1 dimerization inhibitors. We have previously demonstrated the development of potent, cell-permeable, small molecule protein−protein interaction inhibitors from similarly identified cyclic peptides.[45] The evolution of *cyclo*-CLLFVY, the first example of a molecule that selectively inhibits HIF-1 dimerization in cells, to a small molecule by a similar approach is currently underway in our laboratory.

■ EXPERIMENTAL SECTION

Oligonucleotides used in this study are detailed in Table S1. DNA synthesis and sequencing was carried out by Eurofins MWG Operon (Germany). All restriction endonucleases were purchased from New England Biolabs; all other molecular biology reagents were purchased from New England Biolabs, Fisher Scientific, or Promega and were used as directed by the manufacturer. Chemical reagents were purchased from Sigma Aldrich, Fisher Scientific, or Merck and were used as received. Amino acids and peptide coupling reagents were obtained from Novabiochem or Matrix Innovations. DNA purification carried out using QIAquick PCR Purification Kit, and plasmid purification was carried out using QIAGEN Plasmid Mini Kit. The CRIM plasmids pAH68 and pAH69 were obtained from the *E. coli* Genetic Stock Centre, Yale University. All peptides were synthesized using a Liberty 1 microwave peptide synthesizer (CEM), and purified on a Waters HPLC system using a Waters C18 Atlantis T3, or a Waters C18 Atlantis Prep OBD column. ITCs were conducted using a MicroCal iTC200 (GE Healthcare). All cell lines were maintained in DMEM (Life Technologies) containing 10% FBS; for aerobic incubation, cells were cultured at 37 °C in 5% CO_2. For hypoxic treatment, cells were cultured and manipulated (DNA, RNA, and protein extraction) in a H35 hypoxia workstation (Don Whitley Scientific) in 1% O_2, 5% CO_2 and 94% N_2. Luminecence was measured in a GloMAX-96 microplate luminometer (Promega). All assays were conducted in triplicate. Data was analyzed in Excel (Microsoft) or Prism (GraphPad Software).

SICLOPPS Screening for HIF-1 Dimerization Inhibitors. The HIF-1 RTHS, associated control RTHS, and SICLOPPS library were constructed as detailed in the Supporting Information. Electrocompetent cells of the HIF-1 RTHS were prepared and transformed with the C+5 SICLOPPS plasmid library. Transformation efficiency, assessed by plating 10-fold serial dilutions of the recovery solution on LB agar supplemented with chloramphenicol (35 μg/mL), was consistently found to be ∼5 × 10^7, thus ensuring adequate coverage of the 3.2 × 10^6 member cyclic peptide library. Transformants were washed with minimal media and plated onto minimal media agar plates supplemented with ampicillin (50 μg/mL), spectinomycin (25 μg/mL), kanamycin (50 μg/mL), 3-AT (7.5 μM), IPTG (100 μM), L-arabinose (6.5 μM), and chloramphenicol (35 μg/mL). The plates were incubated for 2−3 days at 37 °C until individual colonies were visible. Colonies were picked and restreaked onto LB agar plates containing ampicillin (50 μg/mL), spectinomycin (25 μg/mL), and chloramphenicol (35 μg/mL) and incubated overnight at 37 °C.

Surviving colonies from these plates were grown overnight and assessed by drop-spotting 10-fold serial dilutions onto minimal media plates, supplemented with antibiotics, IPTG and 3-AT as above, with and without 6.5 μM L-arabinose. Plasmids from strains showing a growth advantage in the presence of arabinose were isolated and retransformed into the original selection strain and reassessed for IPTG-dependent inhibition of growth, and arabinose growth rescue. SICLOPPS plasmids from colonies demonstrating the expected phenotypes were assessed for their HIF-1 specificity by transformation into two identical RTHS, except for the replacement of HIF-1 with unrelated proteins (ATIC, a homodimeric enzyme used in purine biosynthesis, and P6/UEV, a heterodimeric interaction required for the budding of HIV from infected cells).[24,25] Plasmids that caused a growth-advantage in the ATIC or P6/UEV RTHS were discarded for being nonspecific. The activity of the cyclic peptides encoded by the remaining SICLOPPS plasmids was ranked by retransforming into the HIF-1 RTHS and drop spotting of 10-fold dilutions. The identity of the variable insert regions encoding the active cyclic peptides was revealed by DNA sequencing.

Peptide Synthesis. Cyclic peptides were synthesized and characterized as detailed in the Supporting Information.

HIF Luciferase Reporter Assays. Endogenous HIF-1 luciferase reporter assays were conducted as previously reported in U2OS-HRE-luc[15] and MCF-7 cells.[46] For plasmid-expressed HIF-α luciferase reporter assays, MCF-7 and U2OS cells were transiently transfected with plasmids expressing HIF-1α, HIF-2α, or a blank control (pcDNA3.1-HIF-1α, pcDNA3.1-HIF-2α, or pcDNA3.1), a renilla-encoding control (phRL-TK), and a HIF-dependent firefly luciferase reporter construct (pGL2-TK-HRE), using Transfast (Promega) according to the manufacturer's instructions. After 24 h, cells were recovered and plated (4000 cells/well) in 96-well plates (Perkin-Elmer) and incubated for 5 h before either hypoxic or aerobic incubation in presence or absence of cyclic peptide inhibitors. Firefly and renilla activities were determined using Dual-Glo Reagent (Promega) according to the manufacturer's instructions. The luciferase signal was normalized using the corresponding renilla values.

Recombinant Production of HIF-1α and HIF-1β. HIF-1, HIF-2, HIF-1, bHLH, PAS-A, PAS-B, and PAS-B were expressed in *E. coli* (BL21.DE3) as detailed in the Supporting Information.

In Vitro Assays. Pull downs, ELISA, fluorescent binding assays, and ITC were conducted as detailed in the Supporting Information.

Dosing Cells with Inhibitors. Cells were treated with the stated concentrations of inhibitor (P1, P2, or P3) and incubated in normoxia for 4 h, followed by incubation in a hypoxic environment. All manipulation of cell pellets (e.g., lysis, mRNA, and protein extraction) was conducted in a hypoxic environment.

Duolink Proximity Ligation Assay. Duolink proximity ligation assay was conducted using the in situ PLA Kit (O-Link Bioscience, Uppsala, Sweden) according to the manufacturer's instructions. The antibodies used were rabbit monoclonal anti-HIF-1α (NB100-449, Novus Biologicals) and mouse monoclonal anti-HIF-1β (H00000405-B01P, Abnova). Cells were dosed with inhibitors as above and incubated in a hypoxic environment for 4 h, after which they were fixed with cold methanol for 10 min and permeabilized with 0.2% Triton (diluted in PBS) for 10 min. After preincubation with the Duolink Blocking Reagent for 1 h, samples were incubated overnight with the primary antibodies to HIF-1α (1:500) and HIF-1β (1:500). Duolink PLA probes and reagents were added as recommended by the manufacture's instructions.

Quantitative Reverse-Transcription Polymerase Chain Reaction. Total RNA was extracted from MCF-7 and U2OS cells using RNeasy Mini Kit (74104, QIAGEN) and quantified using a Nanodrop ND-1000 spectrophotometer. Complementary DNA was synthesized in a 20 μL reaction containing 1 μg of total RNA, using qScript cDNA SuperMix (95048-100, Quanta Biosciences) according to the manufacturer's instructions. Quantitative real-time PCRs were performed using Universal Taqman PCR master mix (Applied Biosystems) >and the Taqman gene expression assay of interest (Applied Biosystems) on an ABI StepOnePlus Real-Time PCR system (Applied Biosystems). Expression assays used in this study were HIF-1α (00936376_m1), VEGF-A (Hs00173626_m1), CAIX (Hs00154208_m1), LOX (Hs00942480_m1), HIF2α (Hs01026149_m1), CITED2 (Hs01897804_s1), and 18S (Hs99999901_m1). All expression values were normalized using expression of 18S.

Western Immunoblotting. Cells were washed with ice-cold PBS and lysed by incubation on ice in radioimmunoprecipitation assay buffer (50 mM Tris (pH 7.4), 150 μM NaCl, 1 mM EDTA, 1% v/v Triton X-100), and protease inhibitor cocktail (Sigma) for 20 min. Lysates were centrifuged at 14 500 rpm for 15 min at 4 °C, and the protein concentration in the supernatant quantified by Bradford assay. Proteins were separated on precast NuPAGE 4% to 12% polyacrylamide gradient Bis−Tris gels (Invitrogen) under denaturing conditions, transferred to PVDF membranes (Invitrogen), and subjected to immunoblot analysis. Mouse monoclonal anti-HIF-1α (610958, BD Biosciences) and rabbit anti-β-actin (ab8226, Abcam) antibodies were diluted (1:250 and 1:10 000, respectively) in PBS containing 5% nonfat powdered milk and 0.1% Tween-20 and then incubated with the membrane overnight at 4 °C. Horseradish peroxidase conjugated secondary antibodies (Cell Signaling) were used. Bound immunocomplexes were detected using ECL prime Western blot detection reagent (RPN2232, GE Healthcare).

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Supporting Information includes Figures S1−S9, Tables S1 and S2, and supplemental methods. This material is available free of charge via the Internet at http://pubs.acs.org

■ AUTHOR INFORMATION

Corresponding Author

a.tavassoli@soton.ac.uk

Present Address

[▽]P.A.T.: Institute for Cancer Sciences, University of Manchester, St. Mary's Hospital, Manchester M13 9WL, UK.

Author Contributions

[#]E.M. and I.K.N. contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We gratefully acknowledge funding of this work by Cancer Research UK (Career Establishment Award A10263 to A.T., and Drug Discovery Grant A9984 to A.T.), the Gerald Kerkut Charitable Trust (Ph.D. studentship for C.E.L. to A.T. and P.A.T.), and the Biotechnology and Biological Sciences Research Council (Doctoral Training Award for I.K.N. to A.T.). We thank Peter Gimeson and GE Healthcare for assistance with ITC, and Prof. Christopher Pugh, Dr. Christopher Bradfield, and Dr. Oliver Hankinson for providing HIF-1β cDNA. We are grateful to Prof. Margaret Ashcroft for U2OS-HRE-luc and U2OS-luc control cells. We also thank Dr. Daniel Asby, Alex Hoose, Dr. David Rusling, Dr. Alison Donlevy, Dr Samantha Larkin, and Dr. Josephine Corsi for helpful discussions and technical assistance.

■ REFERENCES

(1) Wang, G. L.; Jiang, B. H.; Rue, E. A.; Semenza, G. L. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 5510.

(2) Keith, B.; Johnson, R. S.; Simon, M. C. *Nat. Rev. Cancer* **2012**, *12*, 9.

(3) Harris, A. L. *Nat. Rev. Cancer* **2002**, *2*, 38.

(4) Semenza, G. L. *Nat. Rev. Cancer* **2003**, *3*, 721.

(5) Giaccia, A.; Siim, B. G.; Johnson, R. S. *Nat. Rev. Drug Discovery* **2003**, *2*, 803.

(6) Bardos, J. I.; Ashcroft, M. *Bioessays* **2004**, *26*, 262.

(7) Jaakkola, P.; Mole, D. R.; Tian, Y. M.; Wilson, M. I.; Gielbert, J.; Gaskell, S. J.; Kriegsheim, A.; Hebestreit, H. F.; Mukherji, M.; Schofield, C. J.; Maxwell, P. H.; Pugh, C. W.; Ratcliffe, P. J. *Science* **2001**, *292*, 468.

(8) Ivan, M.; Kondo, K.; Yang, H.; Kim, W.; Valiando, J.; Ohh, M.; Salic, A.; Asara, J. M.; Lane, W. S.; Kaelin, W. G., Jr. *Science* **2001**, *292*, 464.

(9) Chilov, D.; Camenisch, G.; Kvietikova, I.; Ziegler, U.; Gassmann, M.; Wenger, R. H. *J. Cell Sci.* **1999**, *112* (Pt 8), 1203.

(10) Ebert, B. L.; Bunn, H. F. *Mol. Cell. Biol.* **1998**, *18*, 4089.

(11) Manalo, D. J.; Rowan, A.; Lavoie, T.; Natarajan, L.; Kelly, B. D.; Ye, S. Q.; Garcia, J. G.; Semenza, G. L. *Blood* **2005**, *105*, 659.

(12) Schodel, J.; Oikonomopoulos, S.; Ragoussis, J.; Pugh, C. W.; Ratcliffe, P. J.; Mole, D. R. *Blood* **2011**, *117*, e207.

(13) Jewell, U. R.; Kvietikova, I.; Scheid, A.; Bauer, C.; Wenger, R. H.; Gassmann, M. *FASEB J.* **2001**, *15*, 1312.

(14) Kung, A. L.; Wang, S.; Klco, J. M.; Kaelin, W. G.; Livingston, D. M. *Nat. Med.* **2000**, *6*, 1335.

(15) Chau, N. M.; Rogers, P.; Aherne, W.; Carroll, V.; Collins, I.; McDonald, E.; Workman, P.; Ashcroft, M. *Cancer Res.* **2005**, *65*, 4918.

(16) Escuin, D.; Simons, J. W.; Giannakakou, P. *Cancer Biol. Ther.* **2004**, *3*, 608.

(17) Jones, D. T.; Harris, A. L. *Mol. Cancer Ther.* **2006**, *5*, 2193.

(18) Kung, A. L.; Zabludoff, S. D.; France, D. S.; Freedman, S. J.; Tanner, E. A.; Vieira, A.; Cornell-Kennon, S.; Lee, J.; Wang, B.; Wang, J.; Memmert, K.; Naegeli, H. U.; Petersen, F.; Eck, M. J.; Bair, K. W.; Wood, A. W.; Livingston, D. M. *Cancer Cell* **2004**, *6*, 33.

(19) Lee, K.; Zhang, H.; Qian, D. Z.; Rey, S.; Liu, J. O.; Semenza, G. L. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 17910.

(20) Nordgren, I. K.; Tavassoli, A. *Chem Soc Rev* **2011**, *40*, 4307.

(21) Rapisarda, A.; Uranchimeg, B.; Scudiero, D. A.; Selby, M.; Sausville, E. A.; Shoemaker, R. H.; Melillo, G. *Cancer Res.* **2002**, *62*, 4316.

(22) Welsh, S.; Williams, R.; Kirkpatrick, L.; Paine-Murrieta, G.; Powis, G. *Mol. Cancer Ther.* **2004**, *3*, 233.

(23) Horswill, A. R.; Savinov, S. N.; Benkovic, S. J. *Proc. Natl. Acad. Sci. U.S..A* **2004**, *101*, 15591.

(24) Tavassoli, A.; Benkovic, S. J. *Angew. Chem., Int. Ed.* **2005**, *44*, 2760.

(25) Tavassoli, A.; Lu, Q.; Gam, J.; Pan, H.; Benkovic, S. J.; Cohen, S. N. *ACS Chem. Biol.* **2008**, *3*, 757.

(26) Scott, C. P.; Abel-Santos, E.; Wall, M.; Wahnon, D. C.; Benkovic, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 13638.

(27) Tavassoli, A.; Benkovic, S. J. *Nat Protoc* **2007**, *2*, 1126.

(28) Ding, Y.; Xu, M. Q.; Ghosh, I.; Chen, X.; Ferrandon, S.; Lesage, G.; Rao, Z. *J. Biol. Chem.* **2003**, *278*, 39133.

(29) Koslowski, M.; Luxemburger, U.; Tureci, O.; Sahin, U. *Oncogene* **2011**, *30*, 876.

(30) Soderberg, O.; Gullberg, M.; Jarvius, M.; Ridderstrale, K.; Leuchowius, K. J.; Jarvius, J.; Wester, K.; Hydbring, P.; Bahram, F.; Larsson, L. G.; Landegren, U. *Nat. Methods* **2006**, *3*, 995.

(31) Forsythe, J. A.; Jiang, B. H.; Iyer, N. V.; Agani, F.; Leung, S. W.; Koos, R. D.; Semenza, G. L. *Mol. Cell. Biol.* **1996**, *16*, 4604.

(32) Calvani, M.; Rapisarda, A.; Uranchimeg, B.; Shoemaker, R. H.; Melillo, G. *Blood* **2006**, *107*, 2705.

(33) Wykoff, C. C.; Beasley, N. J.; Watson, P. H.; Turner, K. J.; Pastorek, J.; Sibtain, A.; Wilson, G. D.; Turley, H.; Talks, K. L.; Maxwell, P. H.; Pugh, C. W.; Ratcliffe, P. J.; Harris, A. L. *Cancer Res.* **2000**, *60*, 7075.

(34) Skehan, P.; Storeng, R.; Scudiero, D.; Monks, A.; McMahon, J.; Vistica, D.; Warren, J. T.; Bokesch, H.; Kenney, S.; Boyd, M. R. *JNCI, J. Natl. Cancer Inst.* **1990**, *82*, 1107.

(35) Maxwell, P. H.; Wiesener, M. S.; Chang, G. W.; Clifford, S. C.; Vaux, E. C.; Cockman, M. E.; Wykoff, C. C.; Pugh, C. W.; Maher, E. R.; Ratcliffe, P. J. *Nature* **1999**, *399*, 271.

(36) Carroll, V. A.; Ashcroft, M. *Cancer Res.* **2006**, *66*, 6264.

(37) Melillo, G. *Mol. Cancer Res.* **2006**, *4*, 601.

(38) Harada, H.; Inoue, M.; Itasaka, S.; Hirota, K.; Morinibu, A.; Shinomiya, K.; Zeng, L.; Ou, G.; Zhu, Y.; Yoshimura, M.; McKenna, W. G.; Muschel, R. J.; Hiraoka, M. *Nat. Commun.* **2012**, *3*, 783.

(39) Wells, J. A.; McClendon, C. L. *Nature* **2007**, *450*, 1001.

(40) Mullard, A. *Nat. Rev. Drug Discovery* **2012**, *11*, 173.

(41) Birts, C. N.; Nijjar, S. K.; Mardle, C. A.; Hoakwie, F.; Duriez, P. J.; Blaydes, J. P.; Tavassoli, A. *Chem. Sci.* **2013**, *4*, 3046.
(42) Scheuermann, T. H.; Li, Q.; Ma, H. W.; Key, J.; Zhang, L.; Chen, R.; Garcia, J. A.; Naidoo, J.; Longgood, J.; Frantz, D. E.; Tambar, U. K.; Gardner, K. H.; Bruick, R. K. *Nat. Chem. Biol.* **2013**, *9*, 271.
(43) Hubbi, M. E.; Kshitiz; Gilkes, D. M.; Rey, S.; Wong, C. C.; Luo, W.; Kim, D. H.; Dang, C. V.; Levchenko, A.; Semenza, G. L. *Sci. Signaling* **2013**, *6*, ra10.
(44) Huang, L. E. *Science* **2013**, *339*, 1285.
(45) Spurr, I. B.; Birts, C. N.; Cuda, F.; Benkovic, S. J.; Blaydes, J. P.; Tavassoli, A. *ChemBioChem* **2012**, *13*, 1628.
(46) Wang, X. H.; Cavell, B. E.; Syed Alwi, S. S.; Packham, G. *Biochem. Pharmacol.* **2009**, *78*, 261.

References

1. 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, *Journal of the American Chemical Society*, 2013, **135**, 10418-10425.
2. M. P. H. Stumpf, T. Thorne, E. de Silva, R. Stewart, H. J. An, M. Lappe and C. Wiuf, *Proceedings of the National Academy of Sciences of the United States of America*, 2008, **105**, 6959-6964.
3. S. Byrum, S. K. Smart, S. Larson and A. J. Tackett, in *Chromatin Remodeling: Methods and Protocols*, ed. R. H. Morse, Humana Press Inc, Totowa, Editon edn., 2012, vol. 833, pp. 143-152.
4. A. J. Levine, W. Hu and Z. Feng, *Cell Death and Differentiation*, 2006, **13**, 1027-1036.
5. J. A. Wells and C. L. McClendon, *Nature*, 2007, **450**, 1001-1009.
6. L. T. Vassilev, B. T. Vu, B. Graves, D. Carvajal, F. Podlaski, Z. Filipovic, N. Kong, U. Kammlott, C. Lukacs, C. Klein, N. Fotouhi and E. A. Liu, *Science*, 2004, **303**, 844-848.
7. T. Zhang, Y. Li, Y. Yu, P. Zou, Y. Jiang and D. Sun, *Journal of Biological Chemistry*, 2009, **284**, 35381-35389.
8. A. J. Wilson, *Chemical Society Reviews*, 2009, **38**, 3289-3300.
9. C. N. Birts, S. K. Nijjar, C. A. Mardle, F. Hoakwie, P. J. Duriez, J. P. Blaydes and A. Tavassoli, *Chemical Science*, 2013, **4**, 3046-3057.
10. A. W. White, A. D. Westwell and G. Brahemi, *Expert Reviews in Molecular Medicine*, 2008, **10**, 1-14.
11. D. C. Fry and L. T. Vassilev, *Journal of Molecular Medicine*, 2005, **83**, 955-963.
12. G. L. Semenza and G. L. Wang, *Molecular and Cellular Biology*, 1992, **12**, 5447-5454.
13. B. H. Jiang, E. Rue, G. L. Wang, R. Roe and G. L. Semenza, *Journal of Biological Chemistry*, 1996, **271**, 17771-17778.
14. B. H. Jiang, G. L. Semenza, C. Bauer and H. H. Marti, *American Journal Physiology*, 1996, **271**, C1172-1180.
15. G. L. Semenza, M. K. Nejfelt, S. M. Chi and S. E. Antonarakis, *Proceedings of the National Academy of Sciences of the United States of America*, 1991, **88**, 5680-5684.
16. F. Coulet, S. Nadaud, M. Agrapart and F. Soubrier, *Journal of Biological Chemistry*, 2003, **278**, 46230-46240.
17. Y. Xia, H.-K. Choi and K. Lee, *European Journal of Medicinal Chemistry*, 2012, **49**, 24-40.
18. G. L. Semenza, *Current Opinion in Genetics & Development*, 1998, **8**, 588-594.
19. D. J. Manalo, A. Rowan, T. Lavoie, L. Natarajan, B. D. Kelly, S. Q. Ye, J. G. N. Garcia and G. L. Semenza, *Blood*, 2005, **105**, 659-669.
20. G. L. Wang and G. L. Semenza, *Journal of Biological Chemistry*, 1995, **270**, 1230-1237.
21. G. L. Wang, B. H. Jiang, E. A. Rue and G. L. Semenza, *Proceedings of the National Academy of Sciences of the United States of America*, 1995, **92**, 5510-5514.
22. R. J. Kewley, M. L. Whitelaw and A. Chapman-Smith, *International Journal of Biochemistry & Cell Biology*, 2004, **36**, 189-204.

23. P. B. Card, P. J. A. Erbel and K. H. Gardner, *Journal of Molecular Biology*, 2005, **353**, 664-677.
24. J. S. Yang, L. Zhang, P. J. A. Erbel, K. H. Gardner, K. Ding, J. A. Garcia and R. K. Bruick, *Journal of Biological Chemistry*, 2005, **280**, 36047-36054.
25. J. L. Ruas, L. Poellinger and T. Pereira, *Journal of Biological Chemistry*, 2002, **277**, 38723-38730.
26. P. Jaakkola, D. R. Mole, Y. M. Tian, M. I. Wilson, J. Gielbert, S. J. Gaskell, A. von Kriegsheim, H. F. Hebestreit, M. Mukherji, C. J. Schofield, P. H. Maxwell, C. W. Pugh and P. J. Ratcliffe, *Science*, 2001, **292**, 468-472.
27. I. K. Nordgren and A. Tavassoli, *Chemical Society Reviews*, 2011, **40**, 4307-4317.
28. J. Zhu, M. Martinez-Yamout, R. Cardoso, J. Yan, R. A. Love, N. Grodsky, A. Brooun and H. J. Dyson, *Journal of Physical Chemistry B*, 2012, **116**, 6960-6965.
29. R. Cardoso, R. Love, C. L. Nilsson, S. Bergqvist, D. Nowlin, J. L. Yan, K. K. C. Liu, J. Zhu, P. Chen, Y. L. Deng, H. J. Dyson, M. J. Greig and A. Brooun, *Protein Science*, 2012, **21**, 1885-1896.
30. T. H. Scheuermann, D. R. Tomchick, M. Machius, Y. Guo, R. K. Bruick and K. H. Gardner, *Proceedings of the National Academy of Sciences of the United States of America*, 2009, **106**, 450-455.
31. J. Key, T. H. Scheuermann, P. C. Anderson, V. Daggett and K. H. Gardner, *Journal of the American Chemical Society*, 2009, **131**, 17647-17654.
32. J. L. Rogers, L. Bayeh, T. H. Scheuermann, J. Longgood, J. Key, J. Naidoo, L. Melito, C. Shokri, D. E. Frantz, R. K. Bruick, K. H. Gardner, J. B. MacMillan and U. K. Tambar, *Journal of Medicinal Chemistry*, 2013, **56**, 1739-1747.
33. P. J. A. Erbel, P. B. Card, O. Karakuzu, R. K. Bruick and K. H. Gardner, *Proceedings of the National Academy of Sciences of the United States of America*, 2003, **100**, 15504-15509.
34. C. A. Amezcua, S. M. Harper, J. Rutter and K. H. Gardner, *Structure*, 2002, **10**, 1349-1361.
35. T. H. Scheuermann, Q. Li, H. W. Ma, J. Key, L. Zhang, R. Chen, J. A. Garcia, J. Naidoo, J. Longgood, D. E. Frantz, U. K. Tambar, K. H. Gardner and R. K. Bruick, *Nature Chemical Biolology*, 2013, **9**, 271-276.
36. M. Ema, S. Taya, N. Yokotani, K. Sogawa, Y. Matsuda and Y. FujiiKuriyama, *Proceedings of the National Academy of Sciences of the United States of America*, 1997, **94**, 4273-4278.
37. I. Flamme, T. Frohlich, M. vonReutern, A. Kappel, A. Damert and W. Risau, *Mechanisms of Development*, 1997, **63**, 51-60.
38. H. Tian, S. L. McKnight and D. W. Russell, *Genes & Development*, 1997, **11**, 72-82.
39. Y. Makino, R. H. Cao, K. Svensson, G. R. Bertilsson, M. Asman, H. Tanaka, Y. H. Cao, A. Berkenstam and L. Poellinger, *Nature*, 2001, **414**, 550-554.
40. J. B. Hogenesch, W. K. Chan, V. H. Jackiw, R. C. Brown, Y. Z. Gu, M. PrayGrant, G. H. Perdew and C. A. Bradfield, *Journal of Biological Chemistry*, 1997, **272**, 8581-8593.
41. A. Loboda, A. Jozkowicz and J. Dulak, *Molecules and Cells*, 2010, **29**, 435-442.
42. K. Lisy and D. J. Peet, *Cell Death and Differentiation*, 2008, **15**, 642-+.
43. B. Keith, R. S. Johnson and M. C. Simon, *Nature Reviews Cancer*, 2012, **12**, 9-22.
44. C. J. Hu, L. Y. Wang, L. A. Chodosh, B. Keith and M. C. Simon, *Molecular and Cellular Biology*, 2003, **23**, 9361-9374.
45. T. D. Eubank, J. M. Roda, H. Liu, T. O'Neil and C. B. Marsh, *Blood*, 2011, **117**, 323-332.

46. K. H. Vousden and C. Prives, *Cell*, 2009, **137**, 413-431.
47. Y. Pan, P. R. Oprysko, A. M. Asham, C. J. Koch and M. C. Simon, *Oncogene*, 2004, **23**, 4975-4983.
48. D. L. Chen, M. Y. Li, J. Y. Luo and W. Gu, *Journal of Biological Chemistry*, 2003, **278**, 13595-13598.
49. R. Ravi, B. Mookerjee, Z. M. Bhujwalla, C. H. Sutter, D. Artemov, Q. W. Zeng, L. E. Dillehay, A. Madan, G. L. Semenza and A. Bedi, *Genes & Development*, 2000, **14**, 34-44.
50. J. Yang, A. Ahmed, E. Poon, N. Perusinghe, A. de Haven Brandon, G. Box, M. Valenti, S. Eccles, K. Rouschop, B. Wouters and M. Ashcroft, *Molecular and Cellular Biology*, 2009, **29**, 2243-2253.
51. J. A. Bertout, A. J. Majmundar, J. D. Gordan, J. C. Lam, D. Ditsworth, B. Keith, E. J. Brown, K. L. Nathanson and M. C. Simon, *Proceedings of the National Academy of Sciences of the United States of America*, 2009, **106**, 14391-14396.
52. C. V. Dang, J. W. Kim, P. Gao and J. Yustein, *Nature reviews. Cancer*, 2008, **8**, 51-56.
53. EMBL-EBI, ClustalW and ClustalX version 2 (2007), M. A. Larkin, G. Blackshields, N.P Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M Wallace, A. Wilm, R. Lopez, J. D Thompson, T. J. Gibson, D.G Higgins, *Bioinformatics*, 2007, **23(21)**, 2947-2948
54. V. A. Carroll and M. Ashcroft, *Cancer Research*, 2006, **66**, 6264-6270.
55. Y. Z. Gu, S. M. Moran, J. B. Hogenesch, L. Wartman and C. A. Bradfield, *Gene Expression*, 1998, **7**, 205-213.
56. S. Hara, J. Hamada, C. Kobayashi, Y. Kondo and N. Imura, *Biochemical and Biophysical Research Communications*, 2001, **287**, 808-813.
57. Y. Makino, A. Kanopka, W. J. Wilson, H. Tanaka and L. Poellinger, *Journal of Biological Chemistry*, 2002, **277**, 32405-32408.
58. A. Augstein, D. M. Poitz, R. C. Braun-Dullaeus, R. H. Strasser and A. Schmeisser, *Cellular and Molecular Life Sciences*, 2011, **68**, 2627-2642.
59. J. I. Bardos and M. Ashcroft, *Biochimica et Biophysica Acta-Reviews on Cancer*, 2005, **1755**, 107-120.
60. A. Chapman-Smith, J. K. Lutwyche and M. L. Whitelaw, *Journal of Biological Chemistry*, 2004, **279**, 5353-5362.
61. K. Gradin, J. McGuire, R. H. Wenger, I. Kvietikova, M. L. Whitelaw, R. Toftgard, L. Tora, M. Gassmann and L. Poellinger, *Molecular and Cellular Biology*, 1996, **16**, 5221-5231.
62. Z. Arany, L. E. Huang, R. Eckner, S. Bhattacharya, C. Jiang, M. A. Goldberg, H. F. Bunn and D. M. Livingston, *Proceedings of the National Academy of Sciences of the United States of America*, 1996, **93**, 12969-12973.
63. P. J. Kallio, K. Okamoto, S. O'Brien, P. Carrero, Y. Makino, H. Tanaka and L. Poellinger, *The EMBO Journal*, 1998, **17**, 6573-6586.
64. J. Gu, J. Milligan and L. E. Huang, *The Journal of Biological Chemistry*, 2001, **276**, 3550-3554.
65. R. H. Wenger, *FASEB Journal*, 2002, **16**.
66. P. Carrero, K. Okamoto, P. Coumailleau, S. O'Brien, H. Tanaka and L. Poellinger, *Molecular and Cellular Biology*, 2000, **20**, 402-415.
67. M.Y. Koh, T. R. Spivak-Kroizman, G. Powis, *Trends in Biochemical Sciences*, 2008, **33**, 526.
68. R. K. Bruick and S. L. McKnight, *Science*, 2001, **294**, 1337-1340.

69. P. C. Mahon, K. Hirota and G. L. Semenza, *Genes and Development*, 2001, **15**, 2675-2686.
70. D. Lando, D. J. Peet, D. A. Whelan, J. J. Gorman and M. L. Whitelaw, *Science*, 2002, **295**, 858-861.
71. D. Lando, D. J. Peet, J. J. Gorman, D. A. Whelan, M. L. Whitelaw and R. K. Bruick, *Genes and Development*, 2002, **16**, 1466-1471.
72. Y. V. Liu, J. H. Baek, H. Zhang, R. Diez, R. N. Cole and G. L. Semenza, *Molecular Cell*, 2007, **25**, 207-217.
73. J. H. Baek, Y. V. Liu, K. R. McDonald, J. B. Wesley, H. F. Zhang and G. L. Semenza, *Journal of Biological Chemistry*, 2007, **282**, 33358-33366.
74. S. H. Bae, J. W. Jeong, J. A. Park, S. H. Kim, M. K. Bae, S. J. Choi and K. W. Kim, *Biochemical and Biophysical Research Communications*, 2004, **324**, 394-400.
75. R. Shao, F. P. Zhang, F. Tian, P. Anders Friberg, X. Wang, H. Sjoland and H. Billig, *FEBS letters*, 2004, **569**, 293-300.
76. M. A. Berta, N. Mazure, M. Hattab, J. Pouyssegur and M. C. Brahimi-Horn, *Biochemical and Biophysical Research Communications*, 2007, **360**, 646-652.
77. J. Cheng, X. Kang, S. Zhang and E. T. Yeh, *Cell*, 2007, **131**, 584-595.
78. P. H. Maxwell, M. S. Wiesener, G. W. Chang, S. C. Clifford, E. C. Vaux, M. E. Cockman, C. C. Wykoff, C. W. Pugh, E. R. Maher and P. J. Ratcliffe, *Nature*, 1999, **399**, 271-275.
79. J. W. Jeong, M. K. Bae, M. Y. Ahn, S. H. Kim, T. K. Sohn, M. H. Bae, M. A. Yoo, E. J. Song, K. J. Lee and K. W. Kim, *Cell*, 2002, **111**, 709-720.
80. G. L. Semenza, *Oncogene*, 2010, **29**, 625-634.
81. J. M. Brown and W. R. William, *Nature Reviews Cancer*, 2004, **4**, 437-447.
82. G. L. Semenza, *Critical Reviews in Biochemistry and Molecular Biology*, 2000, **35**, 71-103.
83. G. L. Semenza, *Current Opinion in Genetics and Development*, 2010, **20**, 51-56.
84. G. L. Semenza, *Nature Reviews Cancer*, 2003, **3**, 721-732.
85. B. Onnis, A. Rapisarda and G. Melillo, *Journal of Cellular and Molecular Medicine*, 2009, **13**, 2780-2786.
86. W. Zundel, C. Schindler, D. Haas-Kogan, A. Koong, F. Kaper, E. Chen, A. R. Gottschalk, H. E. Ryan, R. S. Johnson, A. B. Jefferson, D. Stokoe and A. J. Giaccia, *Genes and Development*, 2000, **14**, 391-396.
87. P. Birner, M. Schindl, A. Obermair, C. Plank, G. Breitenecker and G. Oberhuber, *Cancer Research*, 2000, **60**, 4693-4696.
88. H. Zhong, A. M. De Marzo, E. Laughner, M. Lim, D. A. Hilton, D. Zagzag, P. Buechler, W. B. Isaacs, G. L. Semenza and J. W. Simons, *Cancer Research*, 1999, **59**, 5830-5835.
89. A. Horiuchi, T. Hayashi, N. Kikuchi, A. Hayashi, C. Fuseya, T. Shiozawa and I. Konishi, *International Journal of Cancer*, 2012, **131**, 1755-1767.
90. T. Imamura, H. Kikuchi, M. T. Herraiz, D. Y. Park, Y. Mizukami, M. Mino-Kenduson, M. P. Lynch, B. R. Rueda, Y. Benita, R. J. Xavier and D. C. Chung, *International Journal of Cancer*, 2009, **124**, 763-771.
91. S. S. Zheng, X. H. Chen, X. Yin and B. H. Zhang, *PloS one*, 2013, **8**, e65753.
92. S. Rasheed, A. L. Harris, P. P. Tekkis, H. Turley, A. Silver, P. J. McDonald, I. C. Talbot, R. Glynne-Jones, J. M. Northover and T. Guenther, *British Journal of Cancer*, 2009, **100**, 1666-1673.
93. Y. Yamamoto, M. Ibusuki, Y. Okumura, T. Kawasoe, K. Kai, K. Iyama and H. Iwase, *Breast Cancer Research and Treatment*, 2008, **110**, 465-475.
94. T. E. Walshe and P. A. D'Amore, *Annual Review of Pathology*, 2008, **3**, 615-643.

95. K. Sarkar, Z. Cai, R. Gupta, N. Parajuli, K. Fox-Talbot, M. S. Darshan, F. J. Gonzalez and G. L. Semenza, *Proceedings of the National Academy of Sciences of the United States of America*, 2012, **109**, 10504-10509.
96. H. Shi, *Current Medicinal Chemistry*, 2009, **16**, 4593-4600.
97. G. Melillo, *Cancer Metastasis Reviews*, 2007, **26**, 341-352.
98. J. S. Isaacs, Y. J. Jung, E. G. Mimnaugh, A. Martinez, F. Cuttitta and L. M. Neckers, *Journal of Biological Chemistry*, 2002, **277**, 29936-29944.
99. S. E. Jackson, *Topics in Current Chemistry*, 2013, **328**, 155-240.
100. O. Alqawi, M. Moghaddas and G. Singh, *Prostate Cancer and Prostatic Diseases*, 2006, **9**, 126-135.
101. D. Zagzag, M. Nomura, D. R. Friedlander, C. Y. Blanco, J. P. Gagner, N. Nomura and E. W. Newcomb, *Journal of Cellular Physiology*, 2003, **196**, 394-402.
102. S. Kummar, M. E. Gutierrez, E. R. Gardner, X. Chen, W. D. Figg, M. Zajac-Kaye, M. Chen, S. M. Steinberg, C. A. Muir, M. A. Yancey, Y. R. Horneffer, L. Juwara, G. Melillo, S. P. Ivy, M. Merino, L. Neckers, P. S. Steeg, B. A. Conley, G. Giaccone, J. H. Doroshow and A. J. Murgo, *European Journal of Cancer*, 2010, **46**, 340-347.
103. S. Modi, A. Stopeck, H. Linden, D. Solit, S. Chandarlapaty, N. Rosen, G. D'Andrea, M. Dickler, M. E. Moynahan, S. Sugarman, W. Ma, S. Patil, L. Norton, A. L. Hannah and C. Hudis, *Clinical Cancer Research*, 2011, **17**, 5132-5139.
104. L. Xiang, D. M. Gilkes, P. Chaturvedi, W. Luo, H. Hu, N. Takano, H. Liang and G. L. Semenza, *Journal of Molecular Medicine*, 2014, **92**, 151-164.
105. A. L. Kung, S. D. Zabludoff, D. S. France, S. J. Freedman, E. A. Tanner, A. Vieira, S. Cornell-Kennon, J. Lee, B. Wang, J. Wang, K. Memmert, H. U. Naegeli, F. Petersen, M. J. Eck, K. W. Bair, A. W. Wood and D. M. Livingston, *Cancer Cell*, 2004, **6**, 33-43.
106. G. M. Burslem, H. F. Kyle, A. L. Breeze, T. A. Edwards, A. Nelson, S. L. Warriner and A. J. Wilson, *ChemBioChem*, 2014, **15**, 1083-1087.
107. L. K. Henchey, S. Kushal, R. Dubey, R. N. Chapman, B. Z. Olenyuk and P. S. Arora, *Journal of the American Chemical Society*, 2010, **132**, 941-943.
108. K. M. Block, H. Wang, L. Z. Szabo, N. W. Polaske, L. K. Henchey, R. Dubey, S. Kushal, C. F. Laszlo, J. Makhoul, Z. Song, E. J. Meuillet and B. Z. Olenyuk, *Journal of the American Chemical Society,*, 2009, **131**, 18078-18088.
109. S. A. Dames, M. Martinez-Yamout, R. N. De Guzman, H. J. Dyson and P. E. Wright, *Proceedings of the National Academy of Sciences of the United States of America*, 2002, **99**, 5271-5276.
110. D. J. Asby, F. Cuda, F. Hoakwie, E. Miranda and A. Tavassoli, *Molecular Biosystems*, 2014, **10**, 2505-2508.
111. W. Jeong, A. Rapisarda, S. R. Park, R. J. Kinders, A. Chen, G. Melillo, B. Turkbey, S. M. Steinberg, P. Choyke, J. H. Doroshow and S. Kummar, *Cancer Chemotherapy and Pharmacology*, 2014, **73**, 343-348.
112. H. Zhang, D. Z. Qian, Y. S. Tan, K. Lee, P. Gao, Y. R. Ren, S. Rey, H. Hammers, D. Chang, R. Pili, C. V. Dang, J. O. Liu and G. L. Semenza, *Proceedings of the National Academy of Sciences of the United States of America*, 2008, **105**, 19579-19586.
113. M. Creighton-Gutteridge, J. H. Cardellina, 2nd, A. G. Stephen, A. Rapisarda, B. Uranchimeg, K. Hite, W. A. Denny, R. H. Shoemaker and G. Melillo, *Clinical Cancer Research*, 2007, **13**, 1010-1018.
114. S. Welsh, R. Williams, L. Kirkpatrick, G. Paine-Murrieta and G. Powis, *Molecular Cancer Therapeutics*, 2004, **3**, 233-244.

115. J. J. Lou, Y. L. Chua, E. H. Chew, J. Gao, M. Bushell and T. Hagen, *PloS one*, 2010, **5**, e10522.
116. N. G. Nickols, C. S. Jacobs, M. E. Farkas and P. B. Dervan, *ACS Chemical Biology*, 2007, **2**, 561-571.
117. D. H. Kong, E. J. Park, A. G. Stephen, M. Calvani, J. H. Cardellina, A. Monks, R. J. Fisher, R. H. Shoemaker and G. Melillo, *Cancer Research*, 2005, **65**, 9047-9055.
118. K. Lee, H. F. Zhang, D. Z. Qian, S. Rey, J. O. Liu and G. L. Semenza, *Proceedings of the National Academy of Sciences of the United States of America*, 2009, **106**, 17910-17915.
119. E. J. Park, D. Kong, R. Fisher, J. Cardellina, R. H. Shoemaker and G. Melillo, *Cell Cycle*, 2006, **5**, 1847-1853.
120. D. T. Jones and A. L. Harris, *Molecular Cancer Therapeutics*, 2006, **5**, 2193-2202.
121. N. M. Chau, P. Rogers, W. Aherne, V. Carroll, I. Collins, E. McDonald, P. Workman and M. Ashcroft, *Cancer Research*, 2005, **65**, 4918-4928.
122. O. Soderberg, M. Gullberg, M. Jarvius, K. Ridderstrale, K. J. Leuchowius, J. Jarvius, K. Wester, P. Hydbring, F. Bahram, L. G. Larsson and U. Landegren, *Nature Methods*, 2006, **3**, 995-1000.
123. C. A. Lipinski, F. Lombardo, B. W. Dominy and P. J. Feeney, *Advanced Drug Delivery Reviews*, 1997, **23**, 3-25.
124. I. B. Spurr, C. N. Birts, F. Cuda, S. J. Benkovic, J. P. Blaydes and A. Tavassoli, *ChemBioChem*, 2012, **13**, 1628-1634.
125. A. Tavassoli and S. J. Benkovic, *Angewandte Chemie-International Edition*, 2005, **44**, 2760-2763.
126. A. R. Horswill, S. N. Savinov and S. J. Benkovic, *Proceedings of the National Academy of Sciences of the United States of America*, 2004, **101**, 15591-15596.
127. F. D. Bushman and M. Ptashne, *Proceedings of the National Academy of Sciences of the United States of America*, 1986, **83**, 9353-9357.
128. K. Struhl and R. W. Davis, *Proceedings of the National Academy of Sciences of the United States of America*, 1977, **74**, 5255-5259.
129. S. P. Hilchey, L. Wu and G. B. Koudelka, *Journal of Biological Chemistry*, 1997, **272**, 19898-19905.
130. C. P. Scott, E. Abel-Santos, M. Wall, D. C. Wahnon and S. J. Benkovic, *Proceedings of the National Academy of Sciences of the United States of America*, 1999, **96**, 13638-13643.
131. A. Tavassoli and S. J. Benkovic, *Nature Protocols*, 2007, **2**, 1126-1133.
132. H. Iwai, S. Zuger, J. Jin and P. H. Tam, *FEBS Letters*, 2006, **580**, 1853-1858.
133. R. B. Merrifield, *Journal of the American Chemical Society*, 1963, **85**, 2149-2154.
134. G. B. Fields and R. L. Noble, *International Journal of Peptide and Protein Research*, 1990, **35**, 161-214.
135. W. Konig and R. Geiger, *Chemische Berichte-Recueil*, 1970, **103**, 788-and.
136. L. A. Carpino, *Journal of the American Chemical Society,*, 1993, **115**, 4397-4398.
137. E. Kaiser, Colescot.Rl, Bossinge.Cd and P. I. Cook, *Analytical Biochemistry*, 1970, **34**, 595-and.
138. J. Coste, D. Lenguyen and B. Castro, *Tetrahedron Letters*, 1990, **31**, 205-208.
139. A. Tavassoli, Q. Lu, J. Gam, H. Pan, S. J. Benkovic and S. N. Cohen, *ACS Chemical Biology*, 2008, **3**, 757-764.
140. P. A. Wender, D. J. Mitchell, K. Pattabiraman, E. T. Pelkey, L. Steinman and J. B. Rothbard, *Proceedings of the National Academy of Sciences of the United States of America*, 2000, **97**, 13003-13008.
141. A. D. Frankel and C. O. Pabo, *Cell*, 1988, **55**, 1189-1193.

142. A. Rapisarda, B. Uranchimeg, D. A. Scudiero, M. Selby, E. A. Sausville, R. H. Shoemaker and G. Melillo, *Cancer Research*, 2002, **62**, 4316-4324.
143. S. S. S. Alwi, B. E. Cavell, U. Telang, M. E. Morris, B. M. Parry and G. Packham, *British Journal of Nutrition*, 2010, **104**, 1288-1296.
144. X.-H. Wang, B. E. Cavell, S. S. S. Alwi and G. Packham, *Biochemical Pharmacology*, 2009, **78**, 261-272.
145. M. Chiani, A. Akbarzadeh, A. Farhangi, M. Mazinani, Z. Saffari, K. Emadzadeh and M. R. Mehrabi, *Pakistan Journal of Biological Sciences*, 2010, **13**, 546-550.
146. J. C. Matthews, K. Hori and M. J. Cormier, *Biochemistry*, 1977, **16**, 85-91.
147. S. S. I. o. Bioinformatics, *ExPASY Bioinformatics Resource Portal ProtParam tool*.
148. D. F. Swinehart, *Journal of Chemical Education*, 1962, **39**, 333.
149. G. Chachami, E. Paraskeva, E. Georgatsou, S. Bonanou and G. Simos, *Biochemical and Biophysical Research Communications*, 2005, **331**, 464-470.
150. F. Baneyx, *Current Opinion in Biotechnology*, 1999, **10**, 411-421.
151. H. P. Sorensen and K. K. Mortensen, *Microbial Cell Factories*, 2005, **4**.
152. F. W. Studier, *Protein Expression and Purification*, 2005, **41**, 207-234.
153. N. H. Tolia and L. Joshua-Tor, *Nature Methods*, 2006, **3**, 55-64.
154. N. S. Holden and C. E. Tacon, *Journal of Pharmacological and Toxicological Methods*, 2011, **63**, 7-14.
155. C. Byrne, P. A. McEwan, J. Emsley, P. M. Fischer and W. C. Chan, *Chemical Communications*, 2011, **47**, 2589-2591.
156. V. V. Rostovtsev, L. G. Green, V. V. Fokin and K. B. Sharpless, *Angewandte Chemie-International Edition*, 2002, **41**, 2596-+.
157. M. Jerabek-Willemsen, C. J. Wienken, D. Braun, P. Baaske and S. Duhr, *Assay and Drug Development Technologies*, 2011, **9**, 342-353.
158. S. Duhr and D. Braun, *Proceedings of the National Academy of Sciences of the United States of America*, 2006, **103**, 19678-19682.
159. P. Baaske, C. J. Wienken, P. Reineck, S. Duhr and D. Braun, *Angewandte Chemie-International Edition*, 2010, **49**, 2238-2241.
160. Z. Q. Tan, M. Wortman, K. L. Dillehay, W. L. Seibel, C. R. Evelyn, S. J. Smith, L. H. Malkas, Y. Zheng, S. Lu and Z. Y. Dong, *Molecular Pharmacology*, 2012, **81**, 811-819.
161. A. L. Kung, S. Wang, J. M. Klco, W. G. Kaelin and D. M. Livingston, *Nature Medicine*, 2000, **6**, 1335-1340.
162. H. Zhong, K. Chiles, D. Feldser, E. Laughner, C. Hanrahan, M. M. Georgescu, J. W. Simons and G. L. Semenza, *Cancer Research*, 2000, **60**, 1541-1545.
163. D. T. Jones and A. L. Harris, *Expert Opinion on Therapeutic Targets*, 2012, **16**, 463-480.
164. A. S. Savio, O. R. Acosta, H. G. Perez, Y. R. Alvarez, A. Chico, M. O. Ojeda, C. A. A. Aguero, M. Estevez and G. G. Nieto, *Journal of Peptide Science*, 2012, **18**, 25-29.
165. Y. P. Chang, R. Mahadeva, W. S. Chang, A. Shukla, T. R. Dafforn and Y. H. Chu, *American Journal of Respiratory Cell and Molecular Biology*, 2006, **35**, 540-548.
166. V. Macri, H. Nazzari, E. McDonald and E. A. Accili, *Journal of Biological Chemistry*, 2009, **284**, 15659-15667.
167. Z. Li, K. Migita, D. S. Samways, M. M. Voigt and T. M. Egan, *Journal of Neuroscience*, 2004, **24**, 7378-7386.
168. S. Datta, M. E. Bucks, D. Koley, P. X. Lim and S. N. Savinov, *Bioorganic and Medicinal Chemistry*, 2010, **18**, 6099-6108.

169. C. P. Scott, E. Abel-Santos, A. D. Jones and S. J. Benkovic, *Chemistry and Biology*, 2001, **8**, 801-815.
170. T. C. Evans, Jr., D. Martin, R. Kolly, D. Panne, L. Sun, I. Ghosh, L. Chen, J. Benner, X. Q. Liu and M. Q. Xu, *Journal of Biological Chemistry*, 2000, **275**, 9091-9094.
171. T. A. Naumann, S. N. Savinov and S. J. Benkovic, *Biotechnology and Bioengineering*, 2005, **92**, 820-830.
172. F. B. Perler and E. Adam, *Current Opinion in Biotechnology*, 2000, **11**, 377-383.
173. S. Kjelstrup, P. M. Hansen, L. E. Thomsen, P. R. Hansen and A. Lobner-Olesen, *PloS one*, 2013, **8**, e72273.
174. D. Kozakov, D. R. Hall, G. Y. Chuang, R. Cencic, R. Brenke, L. E. Grove, D. Beglov, J. Pelletier, A. Whitty and S. Vajda, *Proceedings of the National Academy of Sciences of the United States of America*, 2011, **108**, 13528-13533.
175. A. A. Bogan and K. S. Thorn, *Journal of Molecular Biology*, 1998, **280**, 1-9.
176. I. S. Moreira, P. A. Fernandes and M. J. Ramos, *Proteins*, 2007, **68**, 803-812.
177. R. Bollhagen, M. Schmiedberger, K. Barlos and E. Grell, *Journal of the Chemical Society-Chemical Communications*, 1994, 2559-2560.
178. Novabiochem, ed. M. Millipore, Nottingham, UK, Editon edn., 2012-2013.
179. M. M. Condron, B. H. Monien and G. Bitan, *Open Biotechnology Journal*, 2008, **2**, 87-93.
180. M. Gustafsson, T. Curstedt, H. Jornvall and J. Johansson, *Biochemical Journal*, 1997, **326 (Pt 3)**, 799-806.
181. M. Ryba, *Chromatographia*, 1982, **15**, 227-230.
182. S. M. Manohar, A. A. Padgaonkar, A. Jalota-Badhwar, V. Sonawane, M. J. Rathos, S. Kumar and K. S. Joshi, *BMC cancer*, 2011, **11**, 338.
183. H. P. Sorensen and K. K. Mortensen, *Journal of Biotechnology*, 2005, **115**, 113-128.
184. C. Gustafsson, S. Govindarajan and J. Minshull, *Trends in Biotechnology*, 2004, **22**, 346-353.
185. T. Wiseman, S. Williston, J. F. Brandts and L. N. Lin, *Analytical Biochemistry*, 1989, **179**, 131-137.
186. W. B. Turnbull and A. H. Daranas, *Journal of the American Chemical Society*, 2003, **125**, 14859-14866.
187. A. Viegas, J. Manso, F. L. Nobrega and E. J. Cabrita, *Journal of Chemical Education*, 2011, **88**, 990-994.
188. A. Haldimann and B. L. Wanner, *Journal of Bacteriology*, 2001, **183**, 6384-6393.
189. Z. Zhang, D. Yin and Z. Wang, *Animal Science Journal*, 2011, **82**, 244-250.
190. J. Lu, K. Zhang, S. Chen and W. Wen, *Carcinogenesis*, 2009, **30**, 636-644.