

TARGETING THE NRF2/KEAP1 INTERACTION

Richard James Steel

A thesis submitted for the degree of Doctor of Philosophy

January 2014

School of Pharmacy

University of East Anglia

© “This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with the author and that no quotation from the thesis, nor any information derived there-from may be published without the author's prior, written consent.”

Declaration

This thesis is submitted to the University of East Anglia for the Degree of Doctor of Philosophy and has not been previously submitted at this or any university assessment or for any other degree. Except where stated, and reference and acknowledgment is given, this work is original and has been carried out by the author alone.

Richard James Steel

Acknowledgements

My supervisors Prof. Mark Searcey and Dr. Maria O'Connell have given me the opportunity to work on a project with great freedom and to try my hand at an incredible range of techniques. I am immensely grateful for your support, advice and most of all, trust. They are certainly an amazing team.

For training me in the art of tissue culture, Jon Cowan deserves my deepest thanks and also for contributing a large part of the cell work on TAT-14 in Chapter 2. Those Western blots will always stick in my mind. For taking the FP assay and turning it from something that works to something that works well, a subtle but important difference, I have Dr. Tony Blake to thank. For providing the Keap1 plasmid I am grateful to Dr. Mark Hannink and also to Dr. Alex Roberts for teaching me the necessary microbiology to produce the protein. Last but by no means least are Patricia De Souza Fonseca, for completing my final ELISAs and Dr. Vasily Oganessian, who was kind enough to let me use his computers for docking calculations.

I would like to thank Dr. Lesley Howell, Dr. Estelle Payerne and Dr. Sunil Sharma for their support, help and training. In addition, all the groups in both the School of Pharmacy and School of Chemistry, and in particular the Medicinal Chemistry groups, who lent me chemicals, equipment and a sympathetic ear. Thank you all for making my time at UEA so enjoyable.

My final thanks go to my family and friends who have always been so supportive. Especially Liz, we made this journey together and I can't imagine a better companion.

Abstract

The Nrf2/Keap1 protein-protein interaction (PPI) regulates activity of the Nrf2 antioxidant and anti-inflammatory pathway. The transcription factor Nrf2 has been found to be a key mediator in the resolution of inflammation and the progression of chronic diseases. Most known inducers of the Nrf2 pathway act by covalent modification of Keap1 via electrophilic functional groups. Controlled induction of the Nrf2 pathway via specific disruption of the Nrf2/Keap1 interaction is an attractive therapeutic target.

This work describes a cell penetrating, TAT-Nrf2 peptide which targets the Nrf2/Keap1 interaction *in vitro*. Induction of downstream genes is both sequence and dose dependent. In an established model of bacterial sepsis, the peptide reduces pro-inflammatory mediators. Investigation of both cell penetrating and Keap1 binding sequences has identified the requirements for effective Nrf2 induction in cell based assays. An *in vitro* purified protein, fluorescence polarisation (FP) assay was established in order to rapidly characterise these peptides.

Based on the secondary structure of the Keap1 binding portion of Nrf2, further peptides were designed to constrain the conformation and mimic the full protein, while reducing overall size. Synthesis of cyclic peptides has identified the minimal sequence required for efficient binding and provides significant improvement in affinity over linear sequences. Several macrocyclisation techniques were explored in an attempt to retain biological activity, without the need for cell penetration sequences. Initially, disulfide bridge formation was used to produce peptides with affinities for Keap1 similar to the TAT-Nrf2 peptides at considerably reduced size. Subsequently, both head-to-tail cyclisation and peptide stapling were examined in order to restore potency in cell based assays.

Finally, an alternative method for identification of Nrf2/Keap1 disruptors was explored. *In silico* docking calculations were used to identify potential novel PPI disruptors through library screening. Extracted hits were assessed using the FP assay, validating its use for high throughput screening.

Published Work Within This Thesis

Steel, R.; Cowan, J.; Payerne, E.; O'Connell, M. A.; Searcey, M. "Anti-inflammatory Effect of a Cell-Penetrating Peptide Targeting the Nrf2/Keap1 Interaction" *ACS Med. Chem. Lett.* **2012**, *3*, 407–410.

Table of Contents

Acknowledgements	3
Abstract	4
Published Work Within This Thesis	4
Table of Contents	5
List of Figures	9
List of Schemes	13
List of Tables	14
Abbreviations	15
Chapter 1: Inflammation, Nrf2 and Protein-Protein Interactions	18
1.1 - Inflammation	18
1.1.1 - The Inflammatory Response	18
1.1.2 - Inflammatory Disease	19
1.1.3 - Anti-inflammatory Drugs	20
1.2 - Nrf2	22
1.2.1 - Nrf2 Antioxidant and Anti-inflammatory Response	22
1.2.2 - Nrf2 Protein Structure	23
1.2.3 - Keap1 Regulator of Nrf2	25
1.2.4 - Activation of Nrf2	29
1.2.5 - Inducers of Nrf2 Activity	30
1.2.6 - Small Molecule Inducers	30
1.3 - Protein-Protein Interactions	37
1.3.1 - Protein-Protein Interaction Disruptors	37
1.3.2 - Recent Developments	38
1.3.3 - Development of Nrf2 Based Peptides as Nrf2/Keap1 PPI Disruptors	40
Chapter 2: Anti-inflammatory Effects of Cell Penetrating Peptides	42
2.0 - Introduction	42

2.0.1 - Cell Penetrating Peptides	42
2.0.2 - Solid Phase Peptide Synthesis	45
2.1 - TAT-Nrf2 Peptides	53
2.1.1 - TAT-Nrf2 Peptide Synthesis	53
2.1.2 - TAT-Nrf2 Peptide Synthesis Optimisation	55
2.1.3 - TAT-Nrf2 Peptide <i>In Vitro</i> Assays	59
2.2 - Polyarginine-Nrf2 Peptides	65
2.2.1 - Polyarginine-Nrf2 Peptide Synthesis	65
2.2.2 - Polyarginine-Nrf2 Peptide <i>In Vitro</i> Assays	65
2.3 - Fluorescently Tagged CPPs	68
2.3.1 - Fluorescently Tagged CPP Synthesis	68
2.3.2 - Fluorescently Tagged CPP <i>In Vitro</i> Assays	68
2.4 - Conclusions	70
Chapter 3: Fluorescence Polarisation	71
3.0 - Introduction	71
3.0.1 - Fluorescence Polarisation	71
3.1 - Fluorescence Polarisation Assay	73
3.1.1 - Setup of Fluorescence Polarisation Assay	73
3.1.2 - Fluorescence Polarisation Inhibition Assays	77
3.2 - Conclusions	82
Chapter 4: Design and Synthesis of Cyclic Peptides	84
4.0 - Introduction	84
4.0.1 - Peptide Macrocyclisation	84
4.0.2 - Cyclic Peptide Natural Products	84
4.0.3 - Synthetic Cyclic Peptides	86
4.1 - Disulfide Cyclised Peptides	90
4.1.1 - Synthesis of Disulfide Cyclised Peptides	90
4.1.2 - Fluorescence Polarisation Inhibition	92

4.1.3 - HO-1 Protein Induction	95
4.2 - Head-to-tail Cyclised Peptides	97
4.2.1 - Synthesis of Fmoc-Glu-ODMAB	97
4.2.2 - Synthesis of Head-to-tail Cyclised Peptides	98
4.2.3 - Fluorescence Polarisation Inhibition	103
4.2.4 - HO-1 Protein Induction	103
4.3 - Aryl Stapled Peptides	105
4.3.1 - Synthesis of Stapled Peptides	105
4.3.2 - Fluorescence Polarisation Inhibition	106
4.3.3 - HO-1 Protein Induction	108
4.4 - Conclusions	110
Chapter 5: <i>In Silico</i> Screening	112
5.0 - Introduction	112
5.0.1 - <i>In Silico</i> Screening	112
5.1 - <i>In Silico</i> Screening	114
5.1.1 - <i>In Silico</i> Screening Validation	114
5.1.2 - NCI Diversity Set II	117
5.1.3 - ChemBridge Building Blocks	120
5.1.4 - Synthesis and <i>In Vitro</i> Screening	123
5.1.5 - Cross-Receptor Screening	125
5.2 - Conclusions	129
Chapter 6: Conclusions and Future Work	130
6.1 - General Conclusions	130
6.2 - Future work	132
Chapter 7: Experimental	134
7.1 - Chapter 2	134
7.1.1 - Peptide Synthesis	134
7.1.2 - Cell Biology	137

7.2 - Chapter 3	141
7.2.1 - Peptide Synthesis	141
7.2.2 - Fluorescence Polarisation	143
7.3 - Chapter 4	146
7.3.1 - Peptide Synthesis	146
7.3.2 - Small Molecule Synthesis	150
7.3.3 - Fluorescence Polarisation	152
7.3.4 - Cell Biology	153
7.4 - Chapter 5	155
7.4.1 - <i>In Silico</i> Screening	155
7.4.2 - Small Molecule Synthesis	157
Chapter 8: References	160
Appendix 1: Published Work	171

List of Figures

Chapter 1

Figure 1.1: Functional domains of the Nrf2 protein

Figure 1.2: Keap1 Kelch domain X-ray crystal structure viewed from Nrf2 binding face, comprising six anti-parallel β -sheets forming a β -propeller. Coloured by secondary structure progression from blue (N-term) to red (C-term). PDB I.D. 1U6D

Figure 1.3: Hydrogen bonding interactions of the ETGE motif β -turn of the Nrf2 Neh2 domain with the Keap1 Kelch domain

Figure 1.4: Schematic showing Keap1 dimerisation via its broad complex - tramtrack - bric-a-brac (BTB) domain, binding of Nrf2 Neh2 domain via DLG and ETGE β -hairpin motifs and display of seven Lys residues along an intervening α -helix

Figure 1.5: Degradation and induction of Nrf2. A) Ubiquitously expressed Nrf2 is bound by Keap1 via two distinct motifs. Acting as a substrate adaptor, Keap1 facilitates binding of Cul3 and polyubiquitination of Nrf2. B) Following modification of Keap1 sulfhydryls, degradation of Nrf2 is suppressed. Newly synthesised Nrf2 translocates to the nucleus, heterodimerises with Maf proteins and triggers gene transcription by binding to antioxidant response element (ARE) sequences

Figure 1.6: Selected small molecule inducers of the Nrf2/Keap1 pathway of the isothiocyanate, organosulfur, leaving group and indole categories

Figure 1.7: Selected small molecule inducers of the Nrf2/Keap1 pathway of the phenolic, and Michael acceptor categories

Figure 1.8: Recently identified inhibitors of the Nrf2/Keap1 interaction

Chapter 2

Figure 2.1: Simplified Peptide Synthesis A) Solution phase block synthesis B) Solid phase stepwise synthesis

Figure 2.2: Solvation and aggregation of peptide chains and polymer supports. A) Fully solvated B) Intra-molecular peptide chain aggregation C) Polymer support aggregation D) Inter-molecular peptide chain aggregation

Figure 2.3: Hydropathy plots of native 14 mer peptide and designed scrambled sequence showing distribution of polar side chains

Figure 2.4: Crude HPLC trace of the TAT-14 peptide synthesised on Wang resin

Figure 2.5: Crude HPLC trace of the TAT-14 peptide synthesised using triple couplings

Figure 2.6: Crude HPLC trace of the TAT-14 peptide synthesised using microwave irradiation for coupling steps

Figure 2.7: Sequence of 14 mer peptide indicating coupling requirements and onset of aggregation

Figure 2.8: Crude HPLC trace of the 14 mer peptide synthesised manually to assess aggregation

Figure 2.9: Crude HPLC trace of the TAT-14 peptide synthesised on Nova Syn TGA resin

Figure 2.10: HO-1 mRNA induction by TAT-10, TAT-14 and TAT-16 peptides, Mean \pm SEM, n = 3, ***p < 0.001

Figure 2.11: Hairpin sizes for 10, 14 and 16 amino acid binding sequence peptides

Figure 2.12: Nrf2 protein levels following treatment with TAT-14 or TAT-14Sc peptides

Figure 2.13: HO-1 mRNA induction by 14 mer, TAT-14Sc and TAT-14, Mean \pm SEM, n = 3, p*** < 0.001

Figure 2.14: HO-1 protein levels following treatment with TAT-14 or TAT-14Sc peptides

Figure 2.15: Dose Response of TAT-14 induced HO-1 mRNA levels. Mean \pm SEM, n = 3, **p < 0.01, ***p < 0.001

Figure 2.16: Attenuation of lipopolysaccharide (LPS) induced TNF α mRNA levels by TAT-14 compared to TAT-14Sc. Mean \pm SEM, n = 3, **p < 0.01

Figure 2.17: Cell viability of THP-1 cells incubated with varying concentrations of the R₈-14 (circles) and R₄-14 (triangles) peptides for 24 h. Mean \pm SEM, n = 3

Figure 2.18: Nrf2 protein levels following treatment with R₈-14 or R₄-14 peptide

Figure 2.19: ELISA of HO-1 protein levels following treatment with TAT-14, R₈-14 and R₄-14 peptide. Mean \pm SEM, n = 3, ***p < 0.001, *p < 0.05

Figure 2.20: Fluorescence microscopy images of the F-TAT-14 peptide in live THP-1 cells. 1×10^7 cells/mL in PBS, Ex 490 nm, Em 520 nm

Figure 2.21: Representative overlay of 30 min sample showing internalisation of peptide within cells

Chapter 3

Figure 3.1: Effect of rotational correlation time on polarisation of emitted light.

Figure 3.2: Fluorescence anisotropy of the F-14 peptide with varying concentrations of Keap1, Mean \pm SEM, n = 3, K_d 338 \pm 231 nM

Figure 3.3: Fluorescence anisotropy of the F-14 peptide with varying concentrations of Keap1, Mean \pm SEM, n = 3, equilibration 20 min, K_d 128.4 \pm 67.7 nM

Figure 3.4: Fluorescence Polarisation inhibition by the 14 mer peptide, IC₅₀ 9.4 nM, 95% CI [2.6, 29.2], Mean \pm SEM, n = 3, 200 nM Keap1, 5 nM F-14

Figure 3.5: Fluorescence anisotropy of F-14 with varying concentrations of Keap1 and 0.1% Tween20 additive, Mean \pm SEM, n = 3, average over 30 min, 11 readings, K_d 42.1 \pm 7.2 nM

Figure 3.6: Fluorescence anisotropy of the F-14 peptide with varying concentrations of Keap1 and 0.1% Tween20, Mean \pm SEM, n = 3, K_d 14.6 ± 1.1 nM

Figure 3.7: Fluorescence Polarisation inhibition by the 14 mer peptide, K_i 12.3 nM, 95% CI [10.7, 14.1], Mean \pm SEM, n = 3, 30 nM Keap1, 5 nM F-14

Figure 3.8: Hairpin backbone loop sizes for TAT-Nrf2 peptides, excess TAT chain truncated for clarity

Figure 3.9: Z' test, wells 1-48: 5 nM F-14, 30 nM Keap1, wells: 49-96 5 nM F-14, 30 nM Keap1, 2 μ M 14 mer

Chapter 4

Figure 4.1: Selected cyclisation motifs found in natural peptides and employed in synthetic strategies (adapted from Liskamp *et al.*¹⁴⁷)

Figure 4.2: Cyclic peptide natural products A) Vancomycin B) Cyclosporin A C) Chlorofusin

Figure 4.3: Synthetic cyclic peptides A) Stapled p53 derived α -helix B) Stapled SH2 domain binding peptide C) Disulfide cyclised oestrogen receptor binding peptide D) RGD derived head-to-tail cyclised peptide

Figure 4.4: Fluorescence Polarisation inhibition by the Ds8 peptide, K_i 95.0 nM, 95% CI [73.2, 123.3], Mean \pm SEM, n = 3, 30 nM Keap1, 5 nM F-14

Figure 4.5: Induction of HO-1 protein by disulfide cyclised peptides, Mean \pm SEM, n = 3

Figure 4.6: Induction of HO-1 protein by head to tail cyclised peptides, Mean \pm SEM, n = 3, 24 h, 100 μ M peptide

Figure 4.7: Fluorescence Polarisation inhibition by the Ar8P peptide, K_i 6.1 nM, 95% CI [4.4, 8.5], Mean \pm SEM, n = 3, 30 nM Keap1, 5 nM F-14

Figure 4.8: Induction of HO-1 protein by aryl stapled peptides, Mean \pm SEM, n = 3, 24 h, 100 μ M peptide

Chapter 5

Figure 5.1: Docked conformation of the DEETGE binding sequence of Nrf2 into the Keap1 Kelch domain calculated by Vina

Figure 5.2: Overlay of DEETGE crystal structure (pink) and conformation calculated by Vina (grey)

Figure 5.3: Top 10 hits identified by NCI Diversity Set II screen. NCI number, binding energy in brackets (kcal/mol) hydrogen bonding, pi-pi and pi-cation interactions shown in grey

Figure 5.4: Conformation of the highest scoring hit from NCI Diversity Set II screen (NCI_61610)

Figure 5.5: Top 10 hits identified by ChemBridge Building Blocks screen. ChemBridge number, binding energy in brackets (kcal/mol), hydrogen bonding, pi-pi and pi-cation interactions shown in grey

Figure 5.6: Conformation of the highest scoring hit from the ChemBridge Building Blocks screen (CB_5560378)

Figure 5.7: Fluorescence Polarisation inhibition screen of NCI_61610 and CB_5560378, Mean \pm SEM, n = 3, 30 nM Keap1, 5 nM F-14, 1000 nM compound

Figure 5.8: Top 10 hits from receptor screening, intermolecular interactions with highest affinity receptor in grey, ChemBridge ID. Frequency, affinity (kcal/mol), PDB ID of highest affinity receptor, in brackets

Figure 5.9: Representative calculated binding modes for CB_6571942 with Keap1 A) benzimidazole moiety binds in central channel, PDB I.D. 3ADE B) partially blocked central channel causes benzimidazole to bind into a cationic pocket, PDB I.D. 1ZGK

List of Schemes

Chapter 2

Scheme 2.1: Typical Fmoc solid phase peptide synthesis (SPPS) procedure, using a benzyl alcohol functionalised resin, symmetrical anhydride loading and HBTU coupling

Scheme 2.2: Coupling of an amino acid to a resin bound peptide using HBTU

Scheme 2.3: Kaiser's ninhydrin test for the detection of primary amines

Scheme 2.4: Removal of N-terminal Fmoc protecting group using piperidine

Scheme 2.5: Synthesis of F-TAT-14 on NovaSyn TGA resin by Fmoc SPPS

Chapter 4

Scheme 4.1: Synthesis of DMAB protecting group

Scheme 4.2: Synthesis of Fmoc-Glu-ODMAB

Scheme 4.3: Initial synthesis of a head-to-tail cyclised peptide, resin loading, chain elongation and attempted cyclisation

Scheme 4.4: Synthesis of the Ht10 peptide, resin loading, chain elongation and on-resin cyclisation

Scheme 4.5: Mechanism of N-terminal guanidinium capping by HATU

Scheme 4.6: Perfluoroarylation of thiols via nucleophilic aromatic substitution

Chapter 5

Scheme 5.1: Synthesis of NCI_61610

Scheme 5.2: Synthesis of CB_5560378

Chapter 7

Scheme 7.1: Synthesis of the DMAB-OH protecting group

Scheme 7.2: Synthesis of DMAB protected glutamic acid

List of Tables

Chapter 1

Table 1.1: Examples of inflammatory diseases (adapted from Nathan *et al.*²)

Table 1.2: Target genes of Nrf2 (Adapted from Suzuki *et al.*³²)

Chapter 2

Table 2.1: Selected cell penetrating peptides and their sequences

Table 2.2: TAT-Nrf2 peptide sequences

Table 2.3: HO-1 mRNA levels by qPCR following treatment with TAT-10, TAT-14 and TAT-16 peptides in THP-1 cells, 75 μ M peptide, Mean \pm SEM, n = 3

Table 2.4: Peptide sequences based on the 14 mer peptide

Chapter 3

Table 3.1: Fluorescence Polarisation inhibition by CPP-Nrf2 peptides, n = 3, 30 nM Keap1, 5 nM F-14

Chapter 4

Table 4.1: Disulfide cyclised peptides and linear controls based on key Nrf2/Keap1 binding motif

Table 4.2: Fluorescence Polarisation inhibition by disulfide cyclised peptides, n = 3, 30 nM Keap1, 5 nM F-14

Table 4.3: Fluorescence Polarisation inhibition by disulfide cyclised peptides, n = 3, 30 nM Keap1, 5 nM F-14

Table 4.4: Peptide sequences for head-to-tail cyclised Nrf2 peptides, linear sequence as synthesised in italics

Table 4.5: Fluorescence Polarisation inhibition by the head-to-tail cyclised peptides, n = 3, 30 nM Keap1, 5 nM F-14

Table 4.6: Perfluoroaryl stapled peptide sequences

Table 4.7: Fluorescence Polarisation inhibition by the aryl stapled peptides, n = 3, 30 nM Keap1, 5 nM F-14

Chapter 7

Table 7.1: Antibodies used in Western blotting, manufacturer and dilution

Abbreviations

Å	Ångstrom
ADT	5-(4-Methoxyphenyl)-3H-1,2-dithiole-3-thione
AI-1	ARE-inducer-1
Ar	Aromatic
ARE	Antioxidant Response Element
BHA	Butylated hydroxyanisole
Bn	Benzyl
Bnlm	Benzylimidazole
Boc	<i>tert</i> -Butoxycarbonyl
BTB	Broad complex - Tramtrack - Bric-a-brac
bZip	Basic Leucine Zipper
CI	Confidence Interval
COX	Cyclooxygenase
CPP	Cell Penetrating Peptide
Cul3	Cullin 3
D3T	3H-1,2-Dithiole-3-thione
Da	Dalton
Dde	Dimedone
DGR	Double Glycine Repeat
DIC	Diisopropylcarbodiimide
DIM	3,3'-Diindolylmethane
DIPEA	<i>N,N</i> -Diisopropylethylamine
DMAB	4-{N-[1-(4,4-Dimethyl-2,6-dioxocyclohexylidene)-3-methylbutyl]-amino} benzyl
DMAP	4-Dimethylaminopyridine
DMF	<i>N,N</i> -Dimethylformamide
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic Acid
EDCI	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
EDT	1,2-Ethanedithiol
ELISA	Enzyme-linked Immunosorbent Assay
EM	Electron Microscopy
FAM	6-Carboxyfluorescein
Fmoc	Fluorenylmethyloxycarbonyl

FRET	Förster Resonance Energy Transfer
GC	Glucocorticoid
GST	Glutathione S-transferase
HATU	1-[bis(Dimethylamino)methylene]-1 <i>H</i> -1,2,3-triazolo[4,5- <i>b</i>]pyridinium 3-oxide hexafluorophosphate
HBTU	<i>o</i> -Benzotriazole- <i>N,N,N',N'</i> -tetramethyl-uronium-hexafluoro-phosphate
HDAC2	Histone Deacetylase 2
Hmb	2-Hydroxy-4-methoxybenzyl
HO-1	Hemeoxygenase 1
HOAt	1-Hydroxy-7-azabenzotriazole
HOBt	Hydroxybenzotriazole
HPLC	High Performance Liquid Chromatography
HSPG	Heparan Sulfate Proteoglycan
I3C	Indole-3-carbinol
IAB	<i>N</i> -Iodoacetyl- <i>N</i> -biotinylhexylenediamine
IC_{50}	Median Inhibition Concentration
IL-6	Interleukin 6
ITC	Isothermal Titration Calorimetry
Iv	Isovaleryl
IvDde	Isovaleryl dimedone
IVR	Intervening Region
Keap1	Kelch-like ECH-associated protein 1
K_d	Dissociation Constant
K_i	Inhibition Constant
LPS	Lipopolysaccharide
m	Multiplet
mABA	<i>m</i> -Aminobenzoic acid
Maf	V-maf Musculoaponeurotic Fibrosarcoma Oncogene Homolog
MALDI	Matrix-assisted Laser Desorption/Ionisation
MDM2	Murine Double Minute 2
mRNA	Messenger Ribonucleic Acid
MTS	3-(4,5-Dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium
Nap	Naphthyl

Neh	Nrf2-ECH Homolog
NMP	<i>N</i> -Methyl-2-pyrrolidone
NMR	Nuclear Magnetic Resonance Spectroscopy
NQO1	NAD(P)H:Quinone Oxidoreductase 1
Nrf2	Nuclear Factor Erythroid 2 Related Factor 2
NSAID	Non-steroidal Anti-inflammatory Drug
PAGE	Polyacrylamide Gel Electrophoresis
PBS	Phosphate Buffered Saline
PEG	Polyethylene Glycol
PGE2	Prostaglandin E2
PI3K	Phosphoinositide 3-Kinase
Pip	Piperonyl
PPI	Protein-Protein Interaction
PyBOP	(Benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate
PyAOP	(7-Azabenzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate
qPCR	Quantitative Polymerase Chain Reaction
ROS	Reactive Oxidant Species
s	Singlet
SD	Standard Deviation
SDS	Sodium dodecyl sulfate
SH2	Src Homology 2
SPPS	Solid Phase Peptide Synthesis
SPR	Surface Plasmon Resonance
TAT	Transcription-Transactivating
TBHQ	<i>tert</i> -Butylhydroquinone
^t Bu	<i>tert</i> -Butyl
TFA	Trifluoroacetic acid
TIPS	Triisopropylsilane
TNF α	Tumour Necrosis Factor α
TP	Triterpenoid
TRIS	Trisaminomethane
Trpt	Terephthaloyl
ZINC	ZINC Is Not Commercial
γ GCS	γ -Glutamylcysteine Synthetase

Chapter 1: Inflammation, Nrf2 and Protein-Protein Interactions

1.1 - Inflammation

1.1.1 - The Inflammatory Response

The body's inflammatory response acts as a first line of defence against invasion. Known as acute inflammation, the objective is to return the affected tissue to its pre-injury state.¹ The second form of inflammation, termed chronic, is caused by persistent engagement of the innate and acquired immune system. Both forms of inflammation can cause significant damage due to dysregulation, for example, anaphylactic shock and sepsis in the case of acute inflammation and a wide range of conditions, from cardiovascular and neurodegenerative diseases to various forms of cancer as a result of chronic inflammation.² Whereas the acute inflammatory response has a defined purpose in defence, it is still unclear whether chronic inflammation may have a beneficial counterpart.³

Whether the result of bacterial, viral, chemical or physical trauma, the initial stages of acute inflammation are the same.¹ Initial damage is detected by tissue resident macrophages and mast cells, which produce a variety of pro-inflammatory mediators. This leads to activation of blood vessel endothelial cells near the site, which release cytokines and chemokines into the bloodstream to attract leukocytes.⁴ At the same time, the endothelial cells display adhesion molecules on their surfaces, for the leukocytes to attach to. The tight junctions between endothelial cells reversibly open to allow plasma protein and fluid to enter the tissue. This is the cause of the characteristic swelling and pain associated with inflammation, the redness and heat coming from increased blood flow to the area. Once the leukocytes, predominantly neutrophils, have bound to the endothelium, they migrate into the tissue through the tight junctions. At the site of inflammation, the neutrophils release granules containing reactive oxygen and nitrogen species alongside proteases to destroy the invading agent.³ During this process, there is no control over the cells targeted by these toxic compounds, resulting in collateral damage to the host tissue. However, the damage caused by neutrophils does not continue unchecked. Once neutrophils have entered the tissue, they switch from production of pro-inflammatory to anti-inflammatory mediators.⁵ As the balance of signals shifts from induction to reduction of inflammation, the recruitment of leukocytes ceases and apoptosis of neutrophils in the tissue is triggered. Macrophages clear apoptotic neutrophils and other debris by phagocytosis and release repair cytokines. The macrophages then leave the tissue by draining into the lymphatic system and eventually return to the blood. The repair phase can result in total regeneration of the tissue

as seen with the liver, or to formation of scar tissue when repair is not possible, as with the myocardium.¹

1.1.2 - Inflammatory Disease

If the invading agent is not cleared, for instance in severe bacterial infections, the inflammatory response persists and can escalate to dangerous levels. Sepsis is one result of uncontrolled acute inflammation and can lead to respiratory or renal failure and death in around 30% of cases.⁶ Treatment of sepsis is still a challenge, despite recent advances in reduction of mortality rates. As well as escalation of the acute response to dangerous levels, changes in the cells recruited to the site of injury can lead to damage through a switch to chronic inflammation. Despite the association of “chronic” with duration, chronic inflammation is defined by the types of leukocyte present in the tissue.³ The exact cause of chronic inflammation is still unclear, however it appears to be due to tissue malfunction rather than specific infection or injury. A change in the adhesion molecules displayed by endothelial cells promotes binding of lymphocytes and monocytes rather than neutrophils. Once these have migrated into the tissue, the monocytes differentiate into macrophages over a number of days.¹ The lymphocytes and macrophages release pro-inflammatory mediators which trigger fibroblasts to produce scar tissue. These mediators also activate further macrophages and lymphocytes which perpetuate the response leading to long term damage.

Diseases in which inflammation plays an important role	
Alzheimer's disease	Pemphigus
Anaphylaxis	Periodic fever syndromes
Ankylosing spondylitis	Psoriasis
Asthma	Rheumatoid arthritis
Atherosclerosis	Sarcoidosis
Atopic dermatitis	Systemic lupus erythematosus
Chronic obstructive pulmonary disease	Type I diabetes mellitus
Crohn's disease	Ulcerative colitis
Gout	Vasculitides (Wegener's syndrome, Goodpasture's syndrome, giant cell arteritis, polyarteritis nodosa)
Hashimoto's thyroiditis	Xenograft rejection
Ischemia-reperfusion injury	
Multiple sclerosis	
Osteoarthritis	
Diseases in which inflammation may be as harmful as the underlying condition	
Bacterial dysentery	Influenza virus pneumonia
Chagas disease	Leprosy (tuberculoid form)
Cystic fibrosis pneumonitis	Neisserial or pneumococcal meningitis
Filariasis	Post-streptococcal glomerulonephritis
<i>Helicobacter pylori</i> gastritis	Sepsis syndrome
Hepatitis C	Tuberculosis
Diseases in which scarring following inflammation causes the majority of damage	
Bleomycin-induced pulmonary fibrosis	Hepatic cirrhosis
Chronic allograft rejection	Radiation-induced pulmonary fibrosis
Idiopathic pulmonary fibrosis	Schistosomiasis

Table 1.1: Examples of inflammatory diseases (adapted from Nathan *et al.*²)

Uncontrolled inflammation has been linked to a wide range of diseases (Table 1.1). In the case of many of these conditions, the underlying cause is unknown and control of the inflammation is the only available course of action.² As the pathogens are identified, methods of treatment may increase, as was the case for gastric ulcers caused by *Helicobacter pylori*.⁷ However, until this happens, more effective treatments than those currently available for inflammation need to be developed. In any case, the inflammatory response may be more harmful than the pathogen causing it.

1.1.3 - Anti-inflammatory Drugs

Currently there are two categories of drugs commonly used to treat inflammation, glucocorticoids (GCs) and non-steroidal anti-inflammatory drugs (NSAIDs).⁸ In addition, some success has been found in the form of antibodies targeting specific pro-inflammatory cytokines. GCs are one of the most effective anti-inflammatory treatments available. The main effect of GCs is through binding to the glucocorticoid receptor, which then translocates to the nucleus.⁹ It binds to the glucocorticoid responsive element in the promoter region of various genes and causes the

production of anti-inflammatory proteins. Activation of the glucocorticoid receptor also reduces inflammation indirectly by suppressing pro-inflammatory gene expression. Modelled on the hormone cortisol, the GCs were heralded as a solution for treatment of chronic inflammatory disorders. Their wide spread efficacy led to Philip Hench, Edward Kendall and Tadeus Reichstein receiving the 1950 Nobel Prize in medicine, however, common and severe side effects from their use means they are now used only in low doses or to treat life threatening conditions.^{10,11} Side effects from short term use, such as suppressed resistance to infection, are reversible, however conditions such as osteoporosis, atherosclerosis and even psychosis from long term use can be permanent.

While not as potent as the GCs, the NSAIDs achieve fewer side effects by targeting a single pro-inflammatory enzyme, cyclooxygenase (COX). COX produces PGE₂, a prostaglandin, which lowers the body's pain threshold.⁸ Aspirin, the archetypal NSAID, inhibits both COX-1 and COX-2, whereas more modern NSAIDs have been designed to be more selective to COX-2. This has two benefits, COX-2 produces around 100 fold more PGE₂ than COX-1, making them more effective, and inhibition of COX-2 has fewer associated side effects. Compared to the GCs, side effects are minor, including gastric irritation, reduced renal blood flow and skin reactions.¹¹ However, while NSAIDs reduce swelling and pain, they do little to treat the underlying disease. In conditions such as rheumatoid arthritis, vasculitis and nephritis where NSAIDs are used to relieve symptoms, the underlying tissue damage is unaffected.

The use of antibodies as therapeutic agents offers a new avenue for treatment of inflammation. Because the concentration of cytokines is very low, 10-20 pM in some cases, targeting them directly requires similarly small concentrations of drug.⁸ By using antibodies specific for TNF α or the pro-inflammatory interleukins, effective therapies have been developed for rheumatoid arthritis, inflammatory bowel disease, psoriasis and multiple sclerosis. Antibody therapies have little to no organ toxicity, however they suppress the body's innate immune response, increasing the likelihood of opportunistic infections.¹² A more worrying side effect seen with some antibody therapies is progressive multifocal leukoencephalopathy, which causes demyelination of nerves and can be fatal. Why this occurs is unknown.

1.2 - Nrf2

1.2.1 - Nrf2 Antioxidant and Anti-inflammatory Response

An area that has only recently been explored therapeutically, seeks to treat a host of diseases by harnessing the body's own antioxidant and anti-inflammatory defences. Nuclear factor erythroid 2-related factor 2 (Nrf2) is widely acknowledged as the master regulator of the body's innate antioxidant response.¹³ It is a ubiquitously expressed transcription factor of the cap'n'collar family, which is responsible for basal and induced expression of proteins involved in metabolism, resolution of oxidative stress and cytoprotection.¹⁴ This system is tightly controlled however. The negative regulator of Nrf2, Kelch-like ECH-associated protein 1 (Keap1), rapidly sequesters Nrf2 in the cytosol, facilitating its ubiquitination. As a result, Nrf2 has a half life of just 10 to 20 minutes in unstressed cells.

The close connection between inflammation and generation of reactive oxidant species (ROS) is already highlighting areas where induction of Nrf2 regulated genes could be beneficial. Continued oxidative stress can lead to chronic inflammation which can in turn mediate cancer, diabetes, cardiovascular and neurological diseases.¹⁵ In atherosclerosis, where low shear stress in vessels is associated with formation of lesions, Nrf2 has been found to be responsible for inducing a number of ARE mediated protective enzymes in areas of high shear stress.¹⁶ Under the same conditions, Nrf2 was found to reduce TNF α levels and expression of inflammatory response adhesion molecules. There is also evidence to suggest that Nrf2 induction could be used in combination with current treatments. Adenuga *et al.* have shown that histone deacetylase 2 (HDAC2) levels are regulated by Nrf2.¹⁷ Nrf2 knockout mice were found to have lower levels of HDAC2 which led to increased resistance to steroidal treatment of inflammation. Some interactions between Nrf2 mediation and the inflammatory response appear to be more complex. The cytokine interleukin 6 (IL-6) is usually pro-inflammatory, however IL-6 deficient mice show increased oxidative stress and neurodegeneration.¹⁸ Nrf2 knockout mice were used to show that Nrf2 is a potent inducer of IL-6. This is unusual as IL-6 has neither an antioxidant effect nor is it a detoxification enzyme. However, it appears to have a protective function in cells subjected to oxidative stress. One particularly promising area of investigation, is the role of Nrf2 mediated hemeoxygenase 1 (HO-1) expression in inflammation.¹⁹ The HO-1 enzyme converts heme to the antioxidant biliverdin and carbon monoxide. Biliverdin is itself converted to the more potent bilirubin by biliverdin reductase. The anti-inflammatory properties of HO-1 have been demonstrated in mouse models as well as a case of human HO-1 deficiency.^{20,21} The anti-inflammatory effects of HO-1 come not only from degradation of pro-inflammatory heme, but

also from production of anti-inflammatory carbon monoxide. The results of HO-1 deficiency have been demonstrated in HO-1 knockout mice which develop chronic inflammatory diseases and are highly vulnerable to sepsis.²²

Despite the potential therapeutic benefits of Nrf2 activation, there are a few cautionary findings which would prevent the use of Nrf2 induction as a panacea. Permanent upregulation of Nrf2 activity has been found to be detrimental in mouse models. In Keap1 null mice, where Nrf2 is not suppressed, thickening of the skin in the oesophagus results in mortality by three weeks of age.²³ Similarly, in mice with a constitutively activated Nrf2 gene mutation, various skin abnormalities were observed, including thickening, as well as small body size and low weight.²⁴

While Nrf2 induction has potential as a chemopreventive, use in the treatment of cancer has recently been called into question. Several Keap1 mutations have been identified in various cancer cell lines and cancer tissues.²⁵ These cause permanent activation of Nrf2, which confers an advantage to the cancerous cells. In addition, these cells are better able to detoxify chemotherapeutic drugs, which makes them harder to treat. This effect has been confirmed by pretreatment of cells with an Nrf2 inducer, reducing cisplatin toxicity and by siRNA knockdown of Nrf2 in cancerous cells, which restores sensitivity.^{26,27} In spite of these short comings, harnessing the Nrf2 response for the treatment of inflammatory disease is an exciting prospect.

1.2.2 - Nrf2 Protein Structure

The Nrf2 protein is comprised of six highly conserved Nrf2-ECH homolog (Neh) domains (Figure 1.1), which form the various interactions responsible for its activity.¹⁴ The Neh3-5 domains are involved in transactivation though their exact role is poorly understood. The Neh6 domain may be involved in an alternate degradation mechanism which operates under conditions of cell stress. The two most extensively studied domains are the Neh1 and Neh2 domains, which are responsible for the transcriptional activity and negative regulation of Nrf2 respectively.

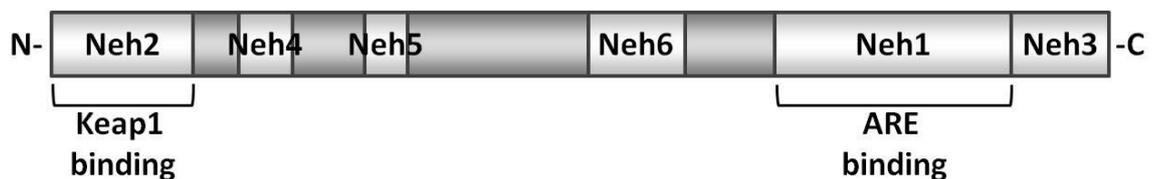


Figure 1.1: Functional domains of the Nrf2 protein

The Neh1 domain contains a basic leucine zipper (bZip) motif which allows the protein to interact with small Maf proteins, which also feature a bZip motif, to form a heterodimer.²⁸ Once formed, this dimer can interact with DNA sequences and activate gene transcription.²⁹ The complex binds

selectively to antioxidant response element (ARE) sequences in the promoter region of a large number of antioxidant and anti-inflammatory genes. Venugopal *et al.* first described the effects of Nrf2 overexpression on the ARE mediated gene expression of two detoxifying enzymes, NAD(P)H:quinone oxidoreductase 1 (NQO1) and glutathione S-transferase (GST).³⁰ Subsequently Alam *et al.* found that induction of HO-1 protein expression was suppressed in Nrf2 deficient mutants.³¹ They noted that most inducers of HO-1 expression stimulate production of ROS or deplete glutathione levels. As HO-1 catalyses the first and rate limiting step in heme catabolism, it was proposed that HO-1 presented an important aspect of the cellular defence mechanism. Since this time, a wide range of antioxidant and metabolic enzymes have been found to be regulated by Nrf2, a selection of which are presented in Table 1.2.

Function	Name	Abbreviation
Synthesis and conjugation of glutathione	Glutamate cysteine ligase	GCL
	Glutathione reductase	GSR
	Glutathione S-transferase (family)	GST
Antioxidant	Peroxiredoxin 1	PRDX1
	Sulfiredoxin 1	SRXN1
	Thioredoxin	TXN
	Thioredoxin reductase 1	TXNRD1
Drug metabolism and transport	NAD(P)H dehydrogenase, quinone 1	NQO1
	Epoxide hydrolase 1, microsomal	EPHX1
	Aldehyde oxidase 1	AOX1
	ATP-binding cassette (family)	ABC
	Solute carrier (family)	SLC
Metabolic enzymes	Transketolase	TKT
	Glucose-6-phosphate dehydrogenase	G6PD
	Isocitrate dehydrogenase 1	IDH1
	2-Aminoethanthiol dioxygenase	ADO
	Lipoprotein lipase	LPL
Heme and iron metabolism	Heme oxygenase 1	HO-1
	Biliverdin reductase B	BLVRB
	Ferrochelatase	FECH
	Ferritin, heavy polypeptide 1	FTH1
	Ferritin, light polypeptide	FTL
Transcription factors	Aryl hydrocarbon receptor	AHR
	CCAAT/enhancer binding protein, alpha	CEBPA
	Hairy and enhancer of split 1	HES1
	v-maf musculoaponeurotic fibrosarcoma oncogene homolog	MAFG
	Retinoid X receptor, alpha	RXRA
	YY1 associated factor 2	YAF2

Table 1.2: Target genes of Nrf2 (Adapted from Suzuki *et al.*³²)

The Neh2 domain is responsible for repression of Nrf2 activity. Itoh *et al.* noted that removal of this region resulted in increased Nrf2 activity.³³ The binding partner for this region was then identified using an artificial Neh2 containing protein to capture it in a yeast based assay. As the protein shared similar homology with the *Drosophila* protein Kelch they named it Kelch-like ECH-associated protein 1 (Keap1).

1.2.3 - Keap1 Regulator of Nrf2

Keap1 is a cytoplasmic actin binding protein, comprising a globular double glycine repeat (DGR) Kelch domain at its C-terminus, an intervening region (IVR) and a broad complex - tramtrack - bric-a-brac (BTB) domain.³⁴ The BTB domain at its N-terminus is found among many Kelch proteins and usually mediates protein dimerisation.³⁵ This is also the case for Keap1 which possesses a highly conserved Ser104 residue in the BTB domain found to be essential for homodimerisation.³⁶ The wild type protein was found to form a high molecular weight complex and bind Nrf2, however a S104A mutant did not dimerise and crucially was unable to suppress Nrf2 activity.

In the C-terminal region of Keap1, the Kelch domain binds to actin and sequesters the Neh2 domain of Nrf2, repressing its activity. By screening for Nrf2 mutants that failed to interact with Keap1, Kobayashi *et al.* identified several single point mutations which significantly affected binding.³⁷ All the important mutations were found in the C-terminal region, corresponding to the ⁷⁹ETGE⁸² motif, which they noted was conserved in all vertebrate Nrf2 sequences. The importance of this motif was confirmed by the inability of Keap1 to suppress ETGE mutant Nrf2 activity in an ARE-luciferase reporter assay.

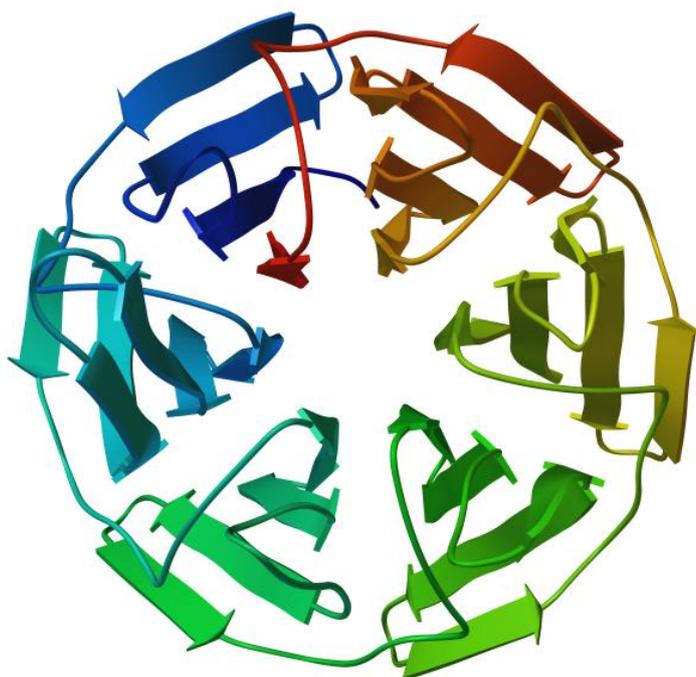


Figure 1.2: Keap1 Kelch domain X-ray crystal structure viewed from Nrf2 binding face, comprising six anti-parallel β -sheets forming a β -propeller. Coloured by secondary structure progression from blue (N-term) to red (C-term). PDB I.D. 1U6D

The crystal structure of the human Keap1 Kelch domain was first reported and characterised by Li *et al.* in 2004 (Figure 1.2).³⁸ The domain consists of six copies of the conserved Kelch repeat motif, forming a six bladed β -propeller. Each blade is a twisted β -sheet composed of four anti-parallel beta strands, with the C-terminus forming the first strand in the first blade. The centre of the propeller forms a channel which is exposed to the solvent and runs through the entire domain. Large portions of the surface of the protein were found to be positively charged, indicating a possible binding site for the negatively charged ETGE motif. Subsequently, murine Kelch domain was crystallised with a Neh2-based 9 mer peptide corresponding to residues ⁷⁶LDEETGEFL⁸⁴.³⁹ The peptide was found to bind into a positively charged pocket at one end of the β -propeller, containing multiple arginine residues. The key ETGE motif formed part of a tight four residue β -turn, stabilised by three intramolecular hydrogen bonds in the Asp77-Glu82 sequence. The side chains of Asp77 and Thr80 contribute two of these hydrogen bonds. In total 13 interactions were identified between the Kelch protein and the peptide localised on one face of the β -propeller and mainly within two discreet pockets. The most significant of these being between the side chain of Glu79 and Arg415, Arg483 and Ser508 in one pocket and between the side chain of Glu82 and Arg380, Asn382 and Ser363 in a second pocket (Figure 1.3). These findings were confirmed for human Kelch domain and a 16 mer peptide later the same year.⁴⁰

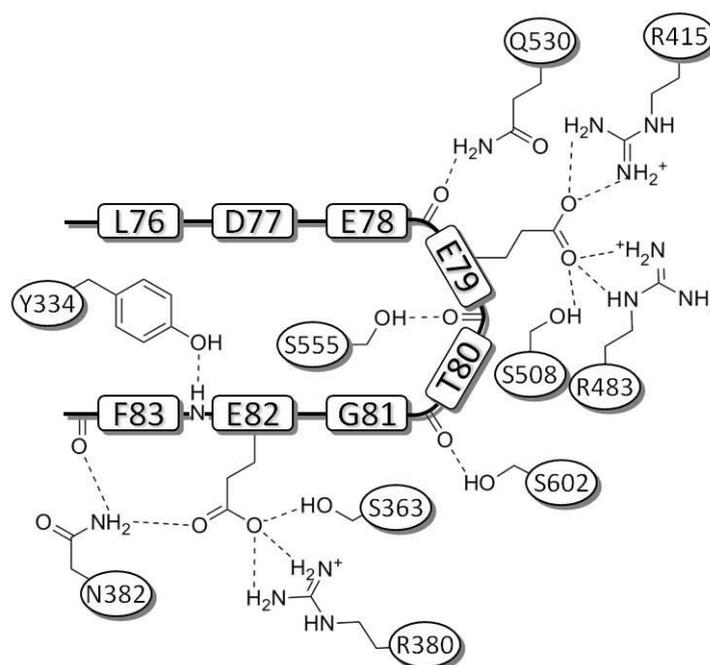


Figure 1.3: Hydrogen bonding interactions of the ETGE motif β -turn of the Nrf2 Neh2 domain with the Keap1 Kelch domain

In addition to the ETGE binding motif, a second highly conserved sequence in the Neh2 domain has been found to interact with Keap1. The DLG motif, in the N-terminal region of the Neh2 domain, has been shown to be important for Keap1 mediated ubiquitination of Nrf2.⁴¹ However, the DLG motif is not necessary for Neh2 domain binding to Keap1. Using NMR studies, Tong *et al.* have determined that the ETGE motif resides in a short antiparallel β -sheet and that both the DLG and ETGE motifs bind to the same region of the Keap1 Kelch domain.⁴² They also found that a 33 residue α -helix with a high lysine content separates the DLG and ETGE motifs. This helix has no affinity for Keap1, but the seven lysine residues are essential for polyubiquitination and Nrf2 degradation.⁴³ In addition, using isothermal titration calorimetry (ITC), Tong *et al.* found that one Neh2 domain is bound by two Keap1 units, with a high affinity to the ETGE motif and a low affinity to the DLG motif.⁴² They have proposed that the ETGE motif allows Keap1 to sequester Nrf2 in the cytoplasm and subsequently, interaction of the DLG motif with the other Kelch domain of the Keap1 dimer locks Nrf2 in place. This places the α -helix between the two Keap1 Kelch domains, in a favourable position for ubiquitination of the lysine residues.

This model has been built upon using electron microscopy (EM) data of Keap1 dimer structures.⁴⁴ Ogura *et al.* have shown that the globular Kelch domain is enclosed in the IVR domain, while a short linker leads to the BTB dimerisation domain. Overlay of Kelch domain crystal structures onto the globular domain shows alignment of the central channel of the β -propeller in both the X-ray and EM structures. Interestingly, the predicted volume of each of the Keap1 domains indicates

that the linker between BTB and globular portions of the protein is in fact part of the BTB domain rather than the IVR as previously thought. Using the EM data, the distance between the two Kelch binding pockets was estimated to be 80 Å. By combining the length of the α -helix between the DLG and ETGE motifs with the remaining length of amino acids in the Neh2 domain, Ogura *et al.* predict an overall distance of 98 Å between binding motifs. They propose that the distance between the binding pockets of the Keap1 dimer plays a critical role in regulation of Nrf2.

The current understanding of the *in vivo* structure of Keap1 and the Nrf2/Keap1 interaction is summarised by the schematic in Figure 1.4.

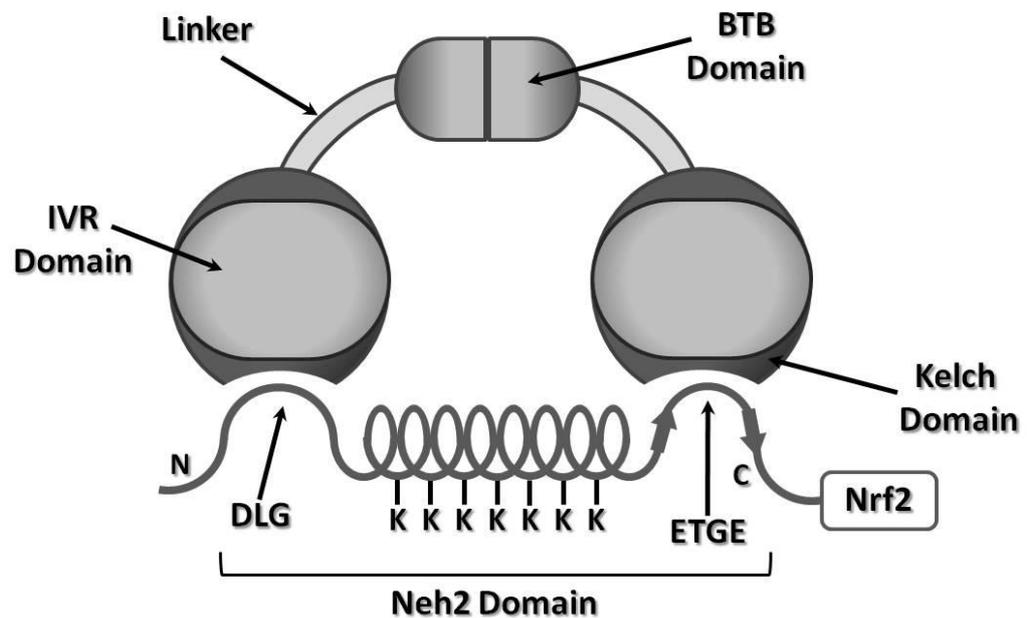


Figure 1.4: Schematic showing Keap1 dimerisation via its broad complex - tramtrack - bric-a-brac (BTB) domain, binding of Nrf2 Neh2 domain via DLG and ETGE β -hairpin motifs and display of seven Lys residues along an intervening α -helix

Once formed, the complex binding interaction between the Keap1 dimer and Nrf2 allows the efficient degradation of Nrf2 via ubiquitination (Figure 1.5A). Zhang *et al.* determined that Keap1 functions as a substrate adaptor protein for the formation of a Cul3-dependent E3 ubiquitin ligase complex.⁴³ Using immunoprecipitation with several cullin proteins, Cul3 was identified as the main binding partner for the Nrf2/Keap1 complex. The cullin family of proteins are known to act as scaffolds for ubiquitin ligase. Formation of this complex with E3 ligase facilitates the polyubiquitination of Nrf2 and subsequent degradation. Zhang *et al.* also found that inhibition of Nrf2 ubiquitination decreased the association of Keap1 and Cul3 but did not affect Keap1/Nrf2 binding.

1.2.4 - Activation of Nrf2

While Nrf2 is being degraded, it cannot exert its protective effect on the cell. As a result when damage occurs, the Nrf2 suppression system must detect the stress and stop Nrf2 degradation. This sensor of cellular stress takes the form of a series of cysteine residues arrayed over the surface of Keap1.³⁴ The protein is unusually rich in cysteine, comprising 4.3% of the total residues, which is double the average for proteins.⁴⁵ Current evidence suggests that one or more of these residues reacts with electrophiles or oxidants present in the cytosol, triggering a conformational change in Keap1 that prevents Nrf2 ubiquitination. Of the 27 cysteine residues, there are certain patterns of reactivity. While no single residue has been found to react with all electrophiles, certain residues appear to be more reactive than others. Residue 151 in the BTB domain and residues 273 and 288 in the IVR domain are considered to be critical for switching off Nrf2 ubiquitination.¹⁴

Following reaction with an activator, the fate of the bound Nrf2 is uncertain. The theory initially proposed by Dinkova-Kostova *et al.* provided the first direct evidence of covalent cysteine modification, alongside evidence to suggest that Nrf2 is released from Keap1.⁴⁶ Once released, Nrf2 would be free to translocate to the nucleus where it could activate ARE mediated genes. This is not supported by later work by Tong *et al.* who proposed the hinge and latch mechanism of binding.⁴² In this case, where the DLG motif acts as a latch to lock Nrf2 in place after ETGE binding, only the DLG portion of Nrf2 is released from Keap1. This results in build up of Nrf2 protein in the cell because Keap1 is unable to facilitate its ubiquitination. As Nrf2 is still bound to Keap1 in this case, activation of ARE mediated genes is achieved by newly translated Nrf2 protein. Recently an alteration to this model of activation has been proposed. Förster Resonance Energy Transfer (FRET) based measurements of the Nrf2/Keap1 interaction indicate that neither of the two binding motifs are released by Keap1 modification.⁴⁷ Instead, a conformational change in Keap1 is proposed that prevents binding of Cul3 and subsequent ubiquitination of Nrf2. This is somewhat supported by the work of Zhang *et al.* who noted that inhibition of ubiquitination reduced association of Keap1 and Cul3 but did not affect the Nrf2/Keap1 interaction.⁴³ In Baird's model of activation, the existence of two binding sites for Nrf2 allows activation of Nrf2 by direct inhibition of the Nrf2/Keap1 interaction. If this is the case, it offers an interesting prospect for controlled induction of the Nrf2 pathway. A generalised scheme for Nrf2 induction is presented in Figure 1.5B.

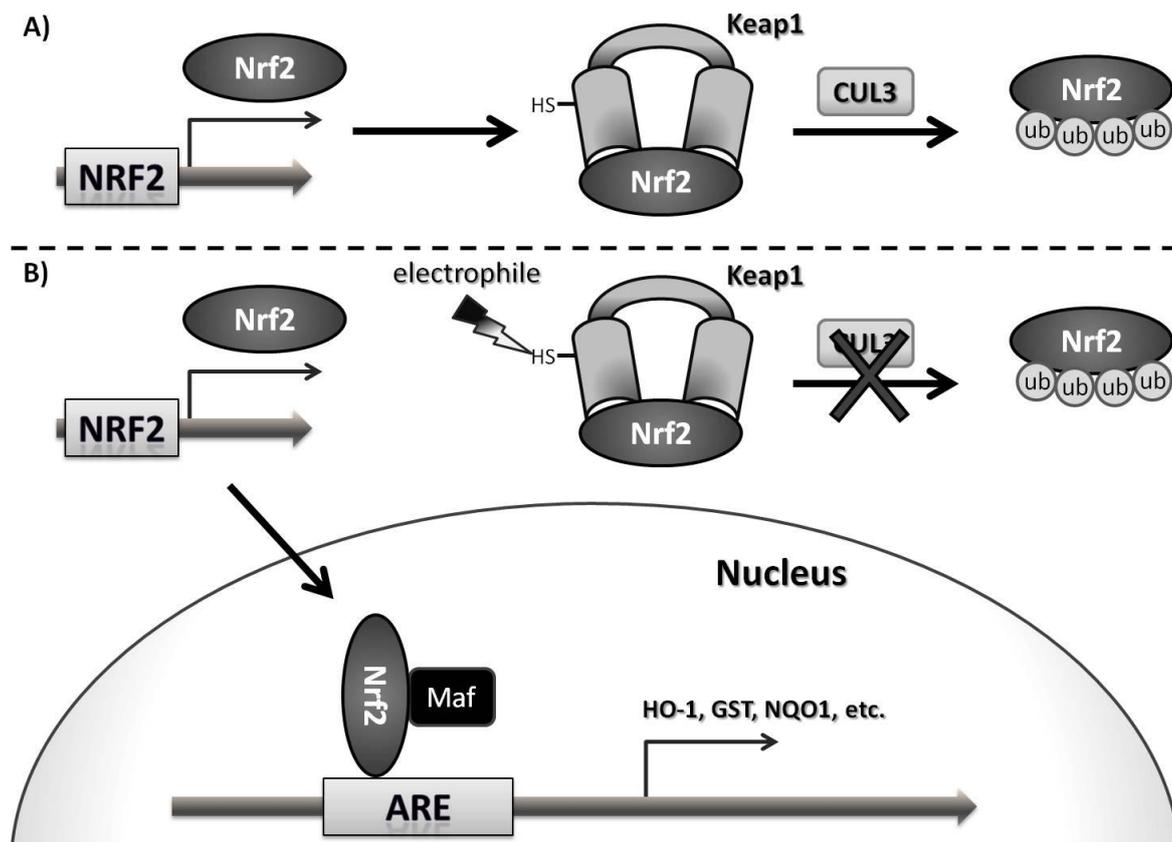


Figure 1.5: Degradation and induction of Nrf2. A) Ubiquitously expressed Nrf2 is bound by Keap1 via two distinct motifs. Acting as a substrate adaptor, Keap1 facilitates binding of Cul3 and polyubiquitination of Nrf2. B) Following modification of Keap1 sulfhydryls, degradation of Nrf2 is suppressed. Newly synthesised Nrf2 translocates to the nucleus, heterodimerises with Maf proteins and triggers gene transcription by binding to antioxidant response element (ARE) sequences

1.2.5 - Inducers of Nrf2 Activity

Given the information currently available about the Keap1/Nrf2/ARE pathway, two methods of activating Nrf2 appear valid. The first is to mimic cellular stress by reacting with the cysteine residues of Keap1 using compounds with low toxicity. There are an increasing number of compounds which may fulfil this role, however, due to their reactive nature, preventing off target reactions is likely to be very challenging. Alternatively, mimicking the Keap1 binding portion of Nrf2 could be used block Nrf2 binding and prevent ubiquitination. There is a smaller background of research to support this approach, however the specificity required by protein-protein interactions means inhibitors are far less likely to suffer from off target effects.

1.2.6 - Small Molecule Inducers

A vast number of small molecules are now known to induce genes in an Nrf2 dependent manner, the majority of which are believed to act via the modification of key cysteine residues of Keap1. While these molecules are highly diverse in structure, they can be categorised by the specific

functionalities determining the way in which they react with thiols. The compounds fall into one of six categories depending on reactive functional groups: isothiocyanates, organosulfur compounds, compounds with a leaving group, indoles, phenolic compounds and Michael acceptors.

A range of biologically active isothiocyanates are found in cruciferous vegetables such as cabbage and broccoli. While they have no direct antioxidant capacity due to relatively low reactivity, they are able to prevent degradation of Nrf2 by Keap1. The best known of these compounds is sulforaphane (Figure 1.6), which has been shown to be a potent inducer of HO-1, NQO1 and GST in an ARE dependent manner.⁴⁸ The involvement of an ARE in the activity of sulforaphane has been confirmed through the use of an ARE regulated GFP construct.⁴⁹ Induction is both time and dose dependent, with significant efficacy at just 2 μ M. Examining the cysteine adducts of sulforaphane is problematic as they readily revert to sulforaphane and free cysteine. Using carefully controlled conditions however, the most reactive cysteine residues of Keap1 with sulforaphane have been identified as Cys489, Cys513, Cys518 and Cys583 all of which are located in the Kelch domain. Due to its naturally high levels in cruciferous vegetables, sulforaphane has seen significant attention as a dietary supplement as well as in clinical trials for cancer chemoprevention.⁵⁰

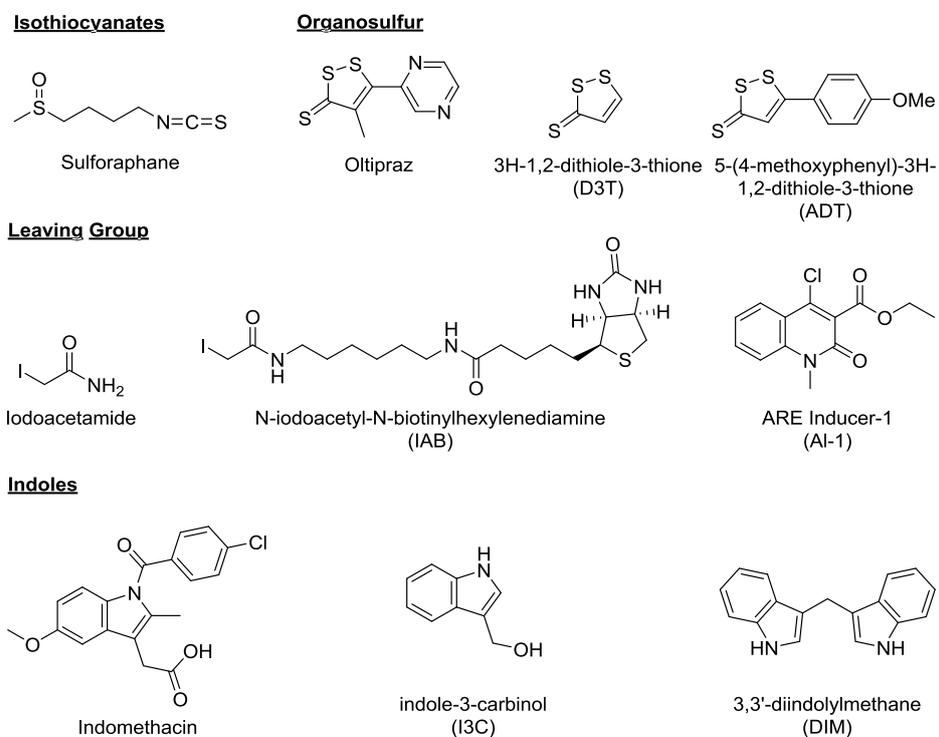


Figure 1.6: Selected small molecule inducers of the Nrf2/Keap1 pathway of the isothiocyanate, organosulfur, leaving group and indole categories

The organosulfur compounds (Figure 1.6) have had limited success in clinical trials as Nrf2 inducers. Oltipraz, a synthetic 1,2-dithiole-3-thione, has been shown to increase expression of several phase II enzymes *in vivo* in an Nrf2 dependent manner.⁵¹ Direct reaction of Oltipraz and Keap1 has not yet been shown however. Despite several clinical trials in humans examining the possible chemopreventative properties of Oltipraz, no consistent effect has been observed. During one trial in which participants were given up to 1 g/m² as a single dose, no side effects were apparent and several proteins were upregulated.⁵² However in trials where Oltipraz was given regularly over a longer period, fatigue, numbness, tingling and pain in the extremities have been observed with no beneficial activity.^{53,54}

While less extensively studied, other compounds in this category seem to have greater efficacy in trials with fewer side effects. 5-(4-methoxyphenyl)-3H-1,2-dithiole-3-thione (**ADT**), has been shown to elevate GST and NQO1 levels in rats as well as protect against multiple carcinogens.⁵⁵ In a trial where smokers with bronchial dysplasia were given 25 mg **ADT** three times per day over six months, a decrease in the progression of preexisting dysplastic lesions was seen.⁵⁶ In this case the only side effects were mild gastrointestinal symptoms. Another compound with promising activity is 3H-1,2-dithiole-3-thione (**D3T**), which has been shown to be a highly potent inducer of phase II enzymes *in vitro*.⁵⁷ However, a recent review has highlighted some discrepancies, suggesting that *in vitro* assays for these compounds do not accurately reflect their activity *in vivo*.⁵⁸

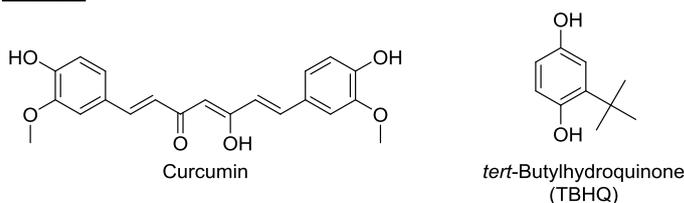
Compounds containing leaving groups (Figure 1.6) have primarily been utilised to examine modification of Keap1 cysteine residues, though progress has been made in developing a more drug-like inducer. Reaction of iodoacetamide with Keap1 has provided key structural data about the nature of cysteine adducting Nrf2 inducers.⁵⁹ *N*-iodoacetyl-*N*-biotinylhexylenediamine (**IAB**) was used as a probe, identifying 6 cysteine residues which were adducted after incubation with Keap1. The sites of modification determined by LC-MS-MS were Cys196, Cys226, Cys241, Cys257, Cys288 and Cys319, all located within the IVR domain, showing little overlap with sulforaphane adducts.

The discovery of a drug-like inducer incorporating a leaving group is the result of a high-throughput cellular screen.⁶⁰ The compound, dubbed ARE-Inducer-1 (**AI-1**), was screened from a library of 1.2 million compounds using **TBHQ** as a positive control. Induction of NQO1 was found to be concentration dependent and not due to oxidative stress. ARE activation was shown to be dependent on the activity of the phosphorylase PI3K however, suggesting several mechanisms of activation may be involved. Using biotinylated **AI-1** and LC-MS-MS, Cys151 was identified as a key modified residue.

The indole, indomethacin (Figure 1.6), an NSAID, has been shown to induce antioxidant genes in an Nrf2 dependent manner.⁶¹ Stimulation of HepG2 cells with indomethacin showed an increase in glutathione levels and prevention of oxidation by diethyl maleate. Removal of the indolic *N*-aromatic substituent and replacement with a methyl group or proton abolishes activity.⁶² Modification of Keap1 cysteines by these compounds was examined by isoelectric focusing, which shows Keap1 is adducted, though which residues are affected was not determined.

In addition to sulforaphane, other components of cruciferous vegetables have been shown to induce Nrf2 activity. One of these, indole-3-carbinol (**I3C**), has shown interesting activity *in vitro*. When tested alongside several other inducers it was found to modestly induce luciferase activity in HepG2-C8 cells.⁶³ However, this was not translated into an increase in HO-1 protein levels. In contrast, activity *in vivo* has been known for some time, including anti-proliferative and pro-apoptotic effects in various cancers.⁶⁴ It is proposed that it is in fact metabolites of **I3C**, which under acidic conditions will form a series of oligomeric products, that are responsible for the observed activity. Recently this has been clarified in the case of Nrf2, using a luciferase reporter in addition to mRNA and protein measurements.⁶⁵ Specifically, it was shown that while **I3C** was unable to induce activity, its major metabolite 3,3'-diindolylmethane (**DIM**) showed significant dose dependent induction of HO-1, NQO1 and γ GCS.

Phenolic



Michael Acceptors

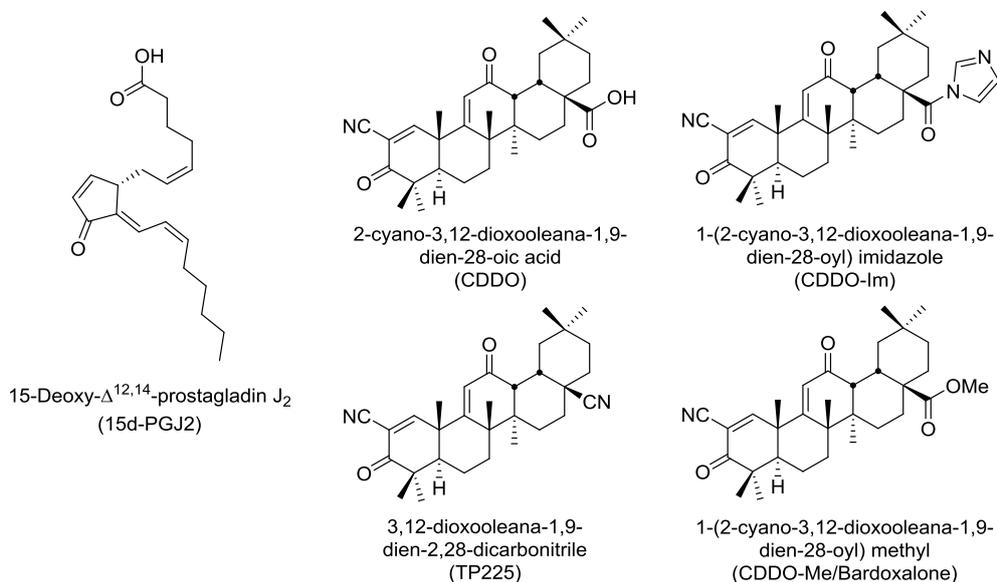


Figure 1.7: Selected small molecule inducers of the Nrf2/Keap1 pathway of the phenolic, and Michael acceptor categories

Polyphenolic compounds possess intrinsic antioxidant and anti-inflammatory properties. After undergoing oxidation, however, they are able to react with Keap1 by acting as Michael acceptors. One of the most extensively studied of these compounds, Curcumin, a major component of the spice turmeric, has been shown to induce Nrf2 in a time and concentration dependent manner by inactivation of the Keap1/Nrf2 complex.⁶⁶ Protein quantification experiments in LLC-PK₁ cells show a maximal HO-1 protein induction of 12.4 fold with 20 μ M curcumin, though at greater concentrations this decreases. This induction has been shown to be ARE dependent using luciferase fusion constructs.⁶⁷

Butylated hydroxyanisole (**BHA**) is a synthetic preservative used in foods and cosmetics. It is converted by cytochrome p450 into the widely studied Nrf2 inducer *tert*-butylhydroquinone (**TBHQ**).⁶⁸ Once oxidised, **TBHQ** prevents degradation of Nrf2 and causes its translocation to the nucleus. While generation of ROS by **TBHQ** is observed and may increase Nrf2 induction, usage of PEG-Catalase as a peroxide scavenger has shown that ROS are not necessary for Nrf2 induction.⁶⁹ Following this, the specific cysteine residues modified by **TBHQ** were determined. In total four residues are modified, Cys23, Cys151, Cys226 and Cys368 some of which correlate with data from reaction of Keap1 with IAB mentioned previously.

Unlike the phenolic compounds, Michael acceptors have no intrinsic antioxidant capacity but act via induction of Nrf2. Of these compounds two specific subsets are of particular interest. The first are the endogenous cyclopentenone prostaglandins, which are able to form cysteine adducts due to the presence of an electrophilic α,β -unsaturated carbonyl in their cyclopentenone ring.⁷⁰ Removal of this double bond eliminates their activity. The most extensively characterised of these, 15-Deoxy- $\Delta^{12,14}$ -prostaglandin J₂ (**15d-PGJ₂**), has been shown to induce both nuclear Nrf2 accumulation and induction of several phase II genes.⁷¹ In addition, binding of **15d-PGJ₂** to the IVR of Keap1 has been shown using biotinylated **15d-PGJ₂** and various Keap1 mutants.⁷² Removal of the IVR results in only weak binding of **15d-PGJ₂** compared to natural Keap1 or Keap1 lacking either the BTB or DGR domains. Within the IVR, point mutation of the cysteine residues to alanine prevents **15d-PGJ₂** binding. This suggests direct reaction of **15d-PGJ₂** with cysteine residues of Keap1 via addition to the Michael acceptor functionalities.

The second sub category of Michael acceptors which has seen significant research are the triterpenoids (**TP**). These synthetic compounds, based on the natural products oleanolic and ursolic acid, are probably the most potent anti-inflammatory and anti-carcinogenic compounds known.⁷³ The **TPs** 2-cyano-3,12-dioxooleana-1,9-dien-28-oic acid (**CDDO**) and 1-(2-cyano-3,12-dioxooleana-1,9-dien-28-oyl) imidazole (**CDDO-Im**), have shown rapid induction of HO-1 both *in vitro* and *in vivo*, at nanomolar concentrations.⁷⁴ While removal of the ARE containing sequence from gene promoters nullifies activity, addition of kinase inhibitors reduces efficacy as well, indicating several mechanism are involved in induction.

The reactions of **TPs** with Keap1 have been explored in detail using a library of synthetic oleanolic acid analogues in order to determine a structure activity relationship.⁷⁵ This identified the Michael acceptor functionality as a requirement for activity and subsequently used the characteristic UV absorbance of this group to show direct interaction of the compound **TP225** with purified recombinant Keap1. As these compounds contain two discreet Michael acceptor functionalities further work has identified the contribution of each of these to the overall potency, though not which cysteine residues are modified.⁷⁶

To date the most successful Nrf2 inducer is the triterpenoid 1-(2-cyano-3,12-dioxooleana-1,9-dien-28-oyl) methyl (**CDDO-Me**) branded as bardoxalone methyl. A recent phase 3, randomised, double blind clinical trial, sought to determine the efficacy of bardoxalone in treating patients with type 2 diabetes.⁷⁷ 2185 patients with type 2 diabetes mellitus and stage 4 chronic kidney disease were recruited and assigned to either the bardoxalone (20 mg per day) or placebo group. Participants in the bardoxalone group were found to have a significant increase in kidney

function. However, the bardoxalone group also had a significantly higher incidence of heart failure, myocardial infarction, stroke and death from cardiovascular causes. The trial was terminated less than half way through, on the recommendation of an independent data and safety monitoring committee. A total of 96 patients in the bardoxalone group were hospitalised for or died from heart failure, compared to 55 in the placebo group. Whether the increased incidence of heart failure is due to Nrf2 induction or other effects of bardoxalone is unknown.

1.3 - Protein-Protein Interactions

1.3.1 - Protein-Protein Interaction Disruptors

An alternative to inducing the Nrf2 pathway via the reactive cysteines of Keap1 is to target the Nrf2/Keap1 interaction directly. The concept of protein-protein interactions (PPIs) as therapeutic targets has only recently been accepted as viable.⁷⁸ There are a number of factors unique to PPIs which make targeting them more challenging than small molecule enzyme interactions. The area encompassed by a PPI is typically between 1500 to 3000 Å² as compared to 300 to 1000 Å² for small molecule interactions.⁷⁹ As well as sheer size, PPIs tend to have flat surfaces which lack grooves and pockets for molecules to bind into.⁸⁰ Where these pockets do exist, they may be separated by great distances. In spite of this, a number of PPIs have emerged where smaller binding pockets contribute the majority of binding. This has allowed the development of peptides, peptidomimetics and small molecules to target these so called 'hot spots'. These PPIs span a wide range of diseases, including inflammation (interleukins⁸¹ and TNFα⁸²), signal transduction⁸³ and HIV.⁸⁴ Undoubtedly though, it is in the treatment of cancer that targeting PPIs has seen the most progress.

Two pathways currently have drugs in clinical trials which act by disruption of PPIs. Abbot Laboratories have developed a series of small molecule binders for the Bcl-XI protein, inhibition of which promotes apoptotic cell death.⁸⁵ The best of these has a sub nanomolar affinity for the protein and a mass of just 813 Da. The drug, branded as Navitoclax, is currently in phase II clinical trials for small cell lung cancer. The other targets the MDM2/p53 interaction, which is probably the most widely studied of all PPIs and as a result has been a testing ground for a variety of inhibition approaches.⁸⁶ The p53 protein is a transcription factor which responds to cellular stress and DNA damage by triggering cell cycle arrest or apoptosis. Interest in targeting the PPI followed publication of the MDM2 crystal structure with a bound p53 based peptide.⁸⁷ The structure identified three hydrophobic side chains which were key to the interaction. The viability of MDM2 as a target was confirmed using a variety of peptide antagonists *in vitro* and *in vivo*.⁸⁸ Following this, a number of inhibitors have been reported, including the natural peptide chlorofusin,⁸⁹ small molecules identified by high throughput screening (benzodiazapinediones⁹⁰ and nutlins⁹¹) and compounds produced by structure based design (spiro-oxindoles⁹² and aromatic bicyclics⁹³). The nutlins, developed by Roche, deserve particular mention, as they are currently in clinical trials for acute myeloid leukemia.⁹⁴ The large body of work surrounding targeting PPIs provides a framework for approaching a new interaction, first through validation of the target, followed by development of peptidic and non-peptidic inhibitors.

A number of key components in the work towards producing Keap1 inhibitors have already been established. In order to study the interaction between Keap1 and Nrf2, Lo *et al.* measured the ability of Nrf2 peptide fragments based on the ETGE motif of the Neh2 region to disrupt Nrf2/Keap1 binding *in vitro*.⁴⁰ Using isothermal titration calorimetry (ITC), a 16 mer peptide (⁶⁹AFFAQLQLDEETGEFL⁸⁴), was assessed for its ability to bind Keap1. A K_d value of 20 nM was determined, which is in line with previously reported values for Nrf2.^{42,95} Two further peptides were tested, a 14 mer which was found to bind as effectively as the 16 mer and a 10 mer which was found to have considerably weaker binding. The 16 mer peptide was crystallised with Keap1 Kelch domain, confirming the key side chain binding interactions of Glu79 and Glu82 seen previously with murine Keap1.³⁹

1.3.2 - Recent Developments

Since beginning the work presented in this thesis, a number of developments have been reported in the area. A peptide sequence containing the ETGE motif of Nrf2 and TAT, the cell transduction domain of HIV, linked by a Calpain cleavage sequence has been tested for its ability to increase Nrf2 mediated gene expression in mice suffering from brain injury.⁹⁶ It was found that without the calpain cleavage sequence, there was no significant increase in gene expression and therefore a linker cleaved under the specific conditions found in injured brain tissue was incorporated. Expression of GPx1, Catalase and GSTm1 was evaluated by quantitative polymerase chain reaction (qPCR) and upon treatment with the TAT-CAL-Nrf2 peptides a 2 fold increase in the genes of interest was observed. In addition, the integrity of the blood brain barrier was far greater in comparison to controls. In mice without brain injury the peptide had no effect, due to lack of cleavage by calpain. While no direct evidence is given, it is likely that the mode of action is via disruption of the Nrf2/Keap1 interaction.

In order to facilitate the hunt for a PPI specific inducer of Nrf2, Hancock *et al.* have developed a high throughput fluorescence polarization (FP) assay using the Kelch domain of Keap1 and a fluorescently tagged peptide based upon the binding sequence of Nrf2.⁹⁷ The fluorescein tagged peptide (FITC- β DEETGEF) was chosen from a selection of Nrf2 binding sequence peptides, based upon the maximum observed change in polarization upon binding rather than the lowest observed K_d . With this assay in place they were able to explore the minimum peptide sequence required to displace the fluorescently tagged peptide. Removal of the leucine residues either side of the DEETGEF binding sequence was seen to reduce activity, suspected to be due to their ability to stabilise the β -hairpin structure of the peptide. Further removal of Asp77 and Phe83 was shown to abolish activity, though whether both these amino acids are required for activity was

not explored. All further modifications to the peptide sequence were found to be disadvantageous with the exception of replacement of Glu78 with proline which was found to improve binding.

Hancock *et al.* have utilised this assay to further develop peptide inhibitors.⁹⁸ By modifying the termini of peptides, they have developed a 7 mer peptide with a stearylated N-terminus which has cellular activity (Figure 1.8). The addition of a long alkyl chain overcomes the multiple acidic residues, allowing cell membrane permeation. In their FP assay, the peptide was found to have an IC_{50} of 22 nM, comparable to values determined for other Neh2 based peptides.⁴⁰ The activity of the peptide was also examined in a cellular assay measuring the concentration required to double NQO1 activity. The concentration required for doubling of activity for the stearylated peptide was 30-80 μ M compared to 0.3 μ M for sulforaphane. The development of this assay provides a powerful tool for screening potential PPI disruptors and the analysis of modified peptides sheds further light on the required properties of any target molecule.

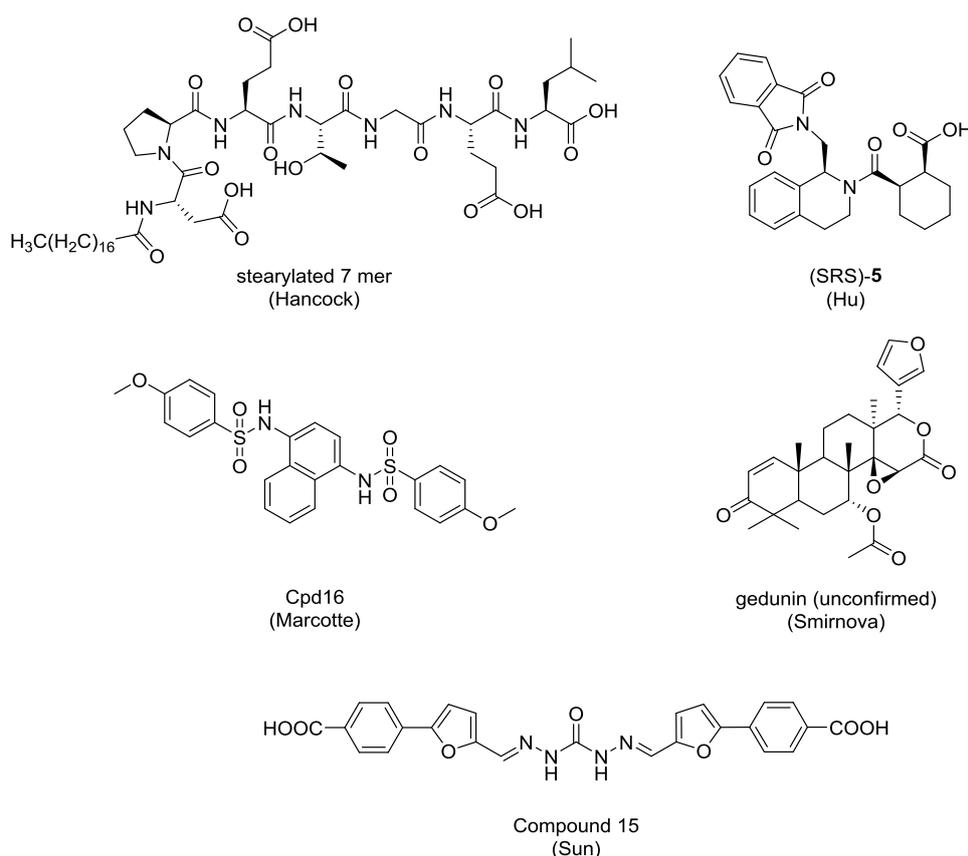


Figure 1.8: Recently identified inhibitors of the Nrf2/Keap1 interaction

To date three small molecules have been identified as acting via disruption of the Nrf2/Keap1 interaction (Figure 1.8). The first was identified using a high throughput FP screen of 337,116 compounds.⁹⁹ The compound, named (SRS)-5 due to its three chiral centres, was found to have a

K_d of 1 μM by surface plasmon resonance and an EC_{50} under 20 μM in cellular assays. Docking calculations suggest that the compound binds into both of the arginine rich pockets on the face of the Kelch domain. A second small molecule, also identified by a high throughput FP assay, was found to have activity at 100 μM in an ARE luciferase reporter assay.¹⁰⁰ Selected from a screen of 269,462 molecules, the interaction of Cpd16 with Keap1 was confirmed by NMR and X-ray crystallography. The third small molecule disruptor (Compound 15) was identified via a process of structure-based virtual screening followed by validation in an FP assay.¹⁰¹ The initial screen of 21,199 structures from the Specs library, identified 17 structures which were validated by FP. Of these, Compound 15, with an IC_{50} of 9.8 μM , was also found to have activity in HepG2 cells. Induction was found to be dose dependent in a stably transfected ARE luciferase reporter assay, with a maximal 10 fold induction at 200 μM .

A fourth molecule has been proposed to act by PPI disruption though this has not been confirmed. While developing a luciferase based assay for screening potential Nrf2 inducers, a novel inducer was identified.¹⁰² The natural product gedunin (Figure 1.8) was identified, alongside others, from a screen of 2000 biologically active compounds as an inducer of Nrf2. While similar to other known inducers, Smirnova *et al.* suggest that the rate and shape of the induction curve indicates a competitive binding of Keap1 rather than the typical cysteine modification mechanism. Computer modelling shows this is a favourable binding, the shape of which is similar to the ⁸³FETGE⁷⁹ section of Nrf2 responsible for binding to Keap1. However, the molecule contains a Michael acceptor which would also allow it to act via cysteine modification.

1.3.3 - Development of Nrf2 Based Peptides as Nrf2/Keap1 PPI Disruptors

Given the work by Lo *et al.* showing strong binding of Nrf2 based peptides to the Keap1 Kelch domain, it was proposed that a peptide containing one of the sequences tested, with the ability to cross cell membranes, could affect Nrf2 levels in cells. Determination of the efficacy of these peptides could be achieved by assessing their ability to increase Nrf2 protein levels. Development of such a peptide would comprise the first stage of this work. The peptide would act as a first validation of disrupting the Nrf2/Keap1 interaction without covalent modification of Keap1. Using this tool it would be possible to examine the effects of disruption of the Nrf2/Keap1 interaction on downstream genes and ultimately reduction of inflammation. With validation of the approach completed, work would then focus on developing inhibitors more suitable for use as drug leads, either through modification of the peptide framework or using small molecules to mimic specific functionalities. It was expected that this work would take the form of several approaches,

including rational drug design and high throughput screening, culminating in the identification of highly potent, novel Nrf2 inducers.

The work presented in the following four chapters describes several approaches to targeting the Nrf2/Keap1 interaction. Chapter 2 aims to validate the inhibition of the Nrf2/Keap1 interaction for the reduction of inflammation *in vitro* using Nrf2 based peptides, conjugated to cell penetration sequences. Design, synthesis and biological evaluation of several peptides is described. One sequence which was shown to have high potency in inducing Nrf2 downstream genes, was found to have anti-inflammatory activity in a model of sepsis. Following initial validation of the approach, a fluorescence polarisation assay was established to allow more rapid *in vitro* screening of compounds as compared to cellular techniques. Development of the assay and characterisation of the Nrf2 based cell penetrating peptides using the assay is described in Chapter 3. With a screening assay in place, developments towards smaller peptide inhibitors which retain potency through formation of macrocycles are presented in Chapter 4. Three cyclisation strategies were explored, disulfide bridge formation, head-to-tail cyclisation and peptide stapling. From this work, the minimum sequence for binding was identified and the benefits of cyclisation assessed. Finally, approaches toward identification of novel PPI inhibitors through high throughput *in silico* screening are presented in Chapter 5.

Chapter 2: Anti-inflammatory Effects of Cell Penetrating Peptides

2.0 - Introduction

2.0.1 - Cell Penetrating Peptides

While the ability to synthesise peptide sequences to target a specific interaction is an enticing prospect, the delivery of these peptides to their intended target is often far from trivial. In utilising larger and often complex molecules to ensure specificity, stability and rapid cellular uptake are often negatively affected. Despite this, the potency and specificity of these molecules has driven research into new methods of drug delivery.¹⁰³ Cell penetrating peptides (CPPs) offer a promising avenue for delivery, with over 100 peptide sequences now identified and an extensive amount of background data on stability, toxicity and efficacy now available (Table 2.1).¹⁰⁴ In addition, their popularity has led to several therapies utilising CPPs entering clinical trial for applications ranging from scarring to myocardial infarction and from hearing loss to cancer.¹⁰⁵ Of these Xigen's XG-102 peptide has recently completed a phase 2 trial as has Capstone Therapeutics' AZX100 peptide for keloid scarring.

Cationic:		
	TAT	YGRKKRRQRRR
	Penetratin	RQIKIWFQNRRMKWKK
	Octa-arginine	RRRRRRRR
Anionic:		
	SAP(E)	VELPPPVELPPPVELPPP
Non-polar:		
	PFVYLI	PFVYLI

Table 2.1: Selected cell penetrating peptides and their sequences

While a diverse range of peptides fall into the category of CPPs, most share a few common features. Typically they consist of fewer than 30 residues with amphipathic characteristics and most often a net positive charge.¹⁰³ The term "protein transduction domain" is sometimes used interchangeably with "cell penetrating peptide" highlighting their origin in larger proteins which were found to translocate across cell membranes. The first of these to be identified and consequently, the most studied is the transcription-transactivating (TAT) protein of HIV-1. While trying to develop an assay for measuring the activity of the protein, Frankel *et al.* found TAT was able to enter cells and translocate to the nucleus.¹⁰⁶ They noticed that when purified TAT protein

was added to cultured cells the protein was taken up by them. However, their note that the protein was found to translocate to the nucleus, may have been an artefact that was only revealed much later.^{107,108} Subsequently a second protein showing similar abilities to cross cell membranes was discovered in *Drosophila*.¹⁰⁹ A synthetic peptide corresponding to 60 amino acids of the antennapedia homeobox peptide was found to enter both live and fixed neurons, a fact that was confirmed using confocal microscopy. Subsequent research has shortened these sequences to the more manageable peptides currently in use.^{110,111} More recently, Schwarze *et al.* clarified the potential of CPPs by using a β -galactosidase-TAT fusion protein to deliver the 120 kDa protein across the blood brain barrier in mice.¹¹² Early work on CPPs had focussed on reducing the length of natural peptides while retaining efficacy, little research had been carried out to elucidate the residues crucial for transduction across the membrane. More recently, such a structure activity relationship was carried out for the TAT peptide, which identified arginine as the key residue.¹¹³ In fact, it was found that a peptide consisting solely of nine arginine residues was 20 fold more potent than the TAT peptide itself. Inverso sequences were found to increase potency further, with up to 100 fold greater activity for D-nona-arginine compared to the TAT peptide. From this work Wender *et al.* concluded that the guanidine moiety was having the key effect and subsequently designed a series of guanidine containing peptoids which retained potency while increasing resistance to proteolysis. While the polyarginine peptides show greater efficacy than natural CPPs there is a definite dependence on length. By examining polyarginine peptides of between 4 and 16 amino acids it was determined that approximately 8 amino acids offers optimal transduction properties.¹¹⁴ Above and below this length, potency decreases with size. The potency of polyguanidino sequences is believed to be due to their ability to form bidentate hydrogen bonds, either with the phospholipid bilayer directly or through surface glycans.¹¹⁵

Despite the large body of work investigating the properties of CPPs, their exact mechanism of action is still unknown. In a large part this is due to an artefact found in all early data which suggested that the penetration was a zero energy process. Up until 2003 it was believed that all CPPs entered the cell by some form of passive diffusion, as experiments which depleted ATP supply or were conducted at low temperature still resulted in internalised peptide.¹⁰⁶ In these cases, the peptides were found to penetrate the membrane and translocate to the nucleus within 5 minutes of peptide addition. In 2003, two separate studies simultaneously determined that, when imaging live cells, the data for CPPs was quite different to previous work where cells had been fixed using formaldehyde.^{107,108} Both groups noted that the peptides adsorb to the surface of the cell and are then internalised by classical endocytosis type mechanisms. In live cells, no

nuclear transport was observed and it was concluded that the fixing process of previous work had permeabilised both the cell membrane and endosomes allowing the CPP to leak into the cytosol. In order to visualise the cells efficiently Richard *et al.* introduced a trypsin digest step to sample preparation, stripping membrane adsorbed peptides from the cells and allowing accurate measurement of internalised peptide.¹⁰⁸ The uptake process was found to be significantly slower than previously thought and dependent on both temperature and ATP availability, characteristic of endocytosis mechanisms. They could not, however, rule out the possibility that a small portion of the peptide was entering by an endocytosis independent mechanism.

The most common technique to examine translocation is via fluorescent labelling of the peptide, usually at one of the termini. While fluorescent labelling is convenient and has provided a lot of valuable data, its limitations must be appreciated. The properties of the fluorophore itself must be considered. There is a very real possibility that the cell entry mechanism may be altered by the fluorescent tag. In addition, due to the tendency of CPPs to adsorb to the cell membrane, fluorescence intensity may not be a valid measure of cellular peptide concentration. This can be overcome by trypsin digest prior to sampling (though this introduces another variable) or by utilising confocal microscopy, which can differentiate depth within the cell, at the cost of reduced sampling size.

Current understanding suggests the following sequence of events are important or essential for the majority of CPP uptake. Firstly the CPP binds to anionic surface receptors, most likely membrane associated proteoglycans, particularly heparan sulfate proteoglycan (HSPG).¹¹⁶ This leads to activation of Rac resulting in filamentous actin organisation and macropinocytosis.¹¹⁷ There are conflicting reports which, for TAT at least, suggest that clathrin-mediated endocytosis is equally important, though it is acknowledged that the situation may be different when TAT is attached to a cargo.¹¹⁸ These authors also note that the typical end stage of endocytosis results in hydrolysis of the contents by lysosomes and that the mechanism by which CPPs escape this fate is unclear.

In selecting a cell penetrating sequence to conjugate to the Nrf2 binding sequence peptides, it was important to consider several factors. While there are a wide range of sequences known, including both anionic and non-polar (Table 2.1) as well as the more typical cationic peptides, only a few have been extensively studied. Among these is the TAT sequence and recent work by Sugita *et al.* has shown that of the CPPs tested, TAT is the most versatile with regards to attached cargo.¹¹⁹ They do note that ideally several sequences should be examined in order to determine the most effective for the specific application. In addition, the use of TAT in the delivery of several

compounds currently undergoing clinical trials makes it a prudent choice for first validation of the approach.¹⁰⁵

In addition to the advantages of CPPs mentioned previously, by utilising a cell penetrating peptide for delivery of a synthetic peptide binding sequence, synthesis can be achieved via a single route rather than synthesis followed by formulation for delivery. Over the last 50 years, interest in peptide synthesis has exploded, in a large part thanks to the relative ease of synthesis using solid phase techniques.

2.0.2 - Solid Phase Peptide Synthesis

The synthesis of peptides can be achieved using two principle approaches. Typically in the solution phase approach peptides are constructed by forming blocks of amino acids which are then assembled into the full chain (Figure 2.1A). Alternatively, in the solid phase approach, the C-terminal amino acid is immobilised on a solid support and the peptide chain is extended in a stepwise fashion (Figure 2.1B). Stepwise synthesis in solution can be performed, as can block synthesis on the solid phase, however, for the majority of syntheses it is not considered practical. Both methods of synthesis have advantages. In general solution phase is preferred for short peptides or peptides containing precious non-natural amino acids due to the excesses of reagents involved in solid phase synthesis. On the other hand for longer sequences, comprising natural amino acids, the increase in yield and rapidity of assembly favours the use of solid phase techniques. Due to the length of the proposed cell penetrating Nrf2 binding sequence peptides, solid phase synthesis is the more practical choice.

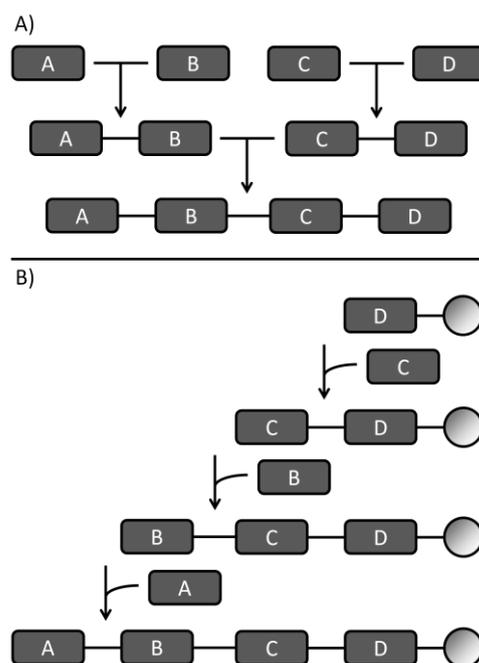
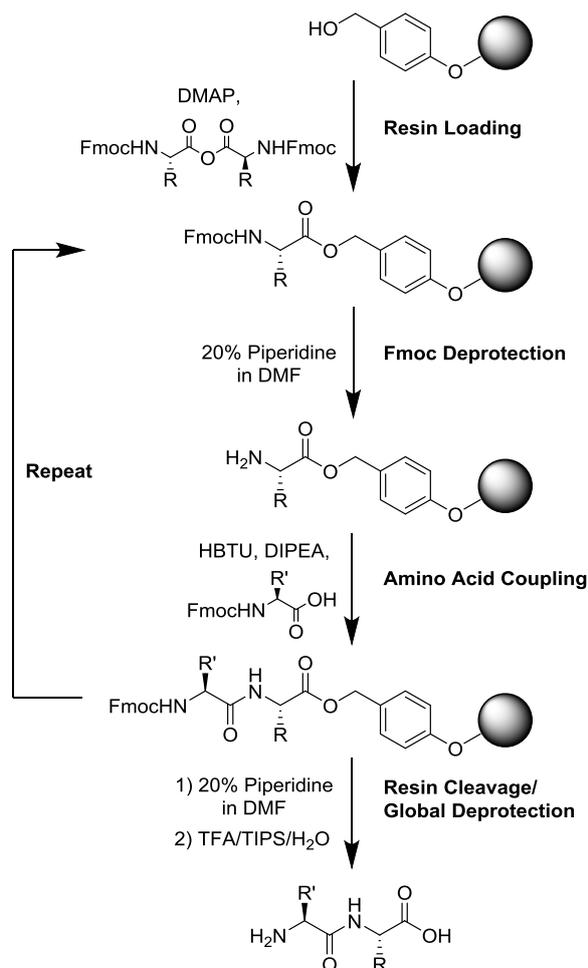


Figure 2.1: Simplified Peptide Synthesis A) Solution phase block synthesis B) Solid phase stepwise synthesis

In its most basic form, solid phase peptide synthesis uses a chloromethylated polystyrene polymer bead as an insoluble but porous support on which to construct a peptide chain. While it is referred to as "solid phase" synthesis, due to the swelling properties of the polymer it is more accurate to think of it as solvated gel phase, which is closer to solution than to solid.¹²⁰ The initial development of this technique is attributable to Merrifield, a contribution which earned him the 1984 Nobel Prize in Chemistry.¹²¹ In his 1963 paper, Merrifield describes the synthesis of a tetrapeptide by carbodiimide coupling, however, the difficulty in driving the reactions to completion led to numerous by products.¹²² Had it not been for a subsequent communication the same year in which the synthesis of the nona-peptide bradykinin was described, the technique may have been entirely forgotten.¹²³ The peptide was synthesised in just four days and isolated in another five with an overall yield of 68%, an achievement far beyond the capabilities of solution phase synthesis at the time. This method of solid phase synthesis has come to be referred to as Boc/benzyl, due to the protecting group strategy, or Merrifield synthesis. The major advantage of this system over solution phase synthesis is reduced stepwise purification, requiring only washing of the resin between couplings rather than lengthy chromatography. This also allows the use of excess reagents to drive the couplings as separation is easy. The global deprotection of side chain protecting groups along with cleavage from the resin in a single step also reduces synthesis time. The acid labile nature of both the N-terminal and side chain protecting groups used in Merrifield's original synthesis meant a small portion of side chains were deprotected during each N-terminal deprotection, leading to undesirable side products. More significantly, the more acid stable benzyl

groups and resin linker require even more strongly acidic conditions than the trifluoroacetic acid (TFA) used to deprotect the N-terminus. As a result, final deprotection and resin cleavage is achieved via use of anhydrous hydrogen fluoride, a highly dangerous chemical requiring specialist equipment and protocols. In the 50 years following this initial report much progress has been made in expanding both the scope and ease of solid phase peptide synthesis.

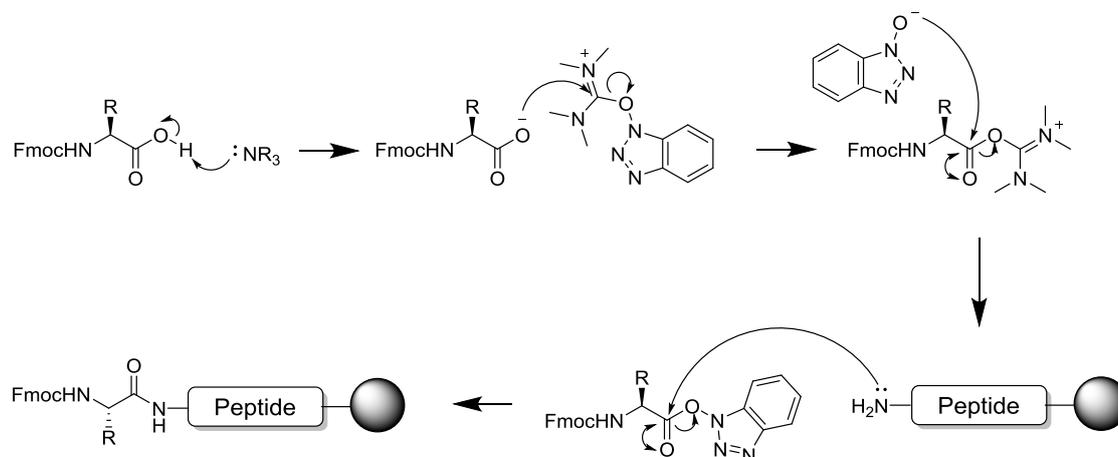
Attempts to develop a protecting group strategy which could be removed under more mild conditions were initially hampered by the availability of appropriately protected amino acids. However, Carpino's fluorenylmethoxy carbonyl (Fmoc) protecting group, which had seen limited usage in solution due to the reactivity of and difficulty in separating its cleavage side products, became the N-terminal protecting group of choice in solid phase synthesis.¹²⁴ Fmoc deprotection under basic conditions is truly orthogonal in nature to the tertiary butyl (^tBu) side chain protecting groups, which are removed by acid. This meant that there was no concern about unwanted side chain deprotection during synthesis. In addition, the reactive dibenzofulvene from Fmoc removal could simply be washed off the resin before it had time to react. The commercial availability of Fmoc protected amino acids secured its place as a viable alternative to Merrifield synthesis. Combined with ^tBu side chain protection, global deprotection and resin cleavage can now be achieved using the comparatively much safer TFA method previously used for N-terminal deprotection.



Scheme 2.1: Typical Fmoc solid phase peptide synthesis (SPPS) procedure, using a benzyl alcohol functionalised resin, symmetrical anhydride loading and HBTU coupling

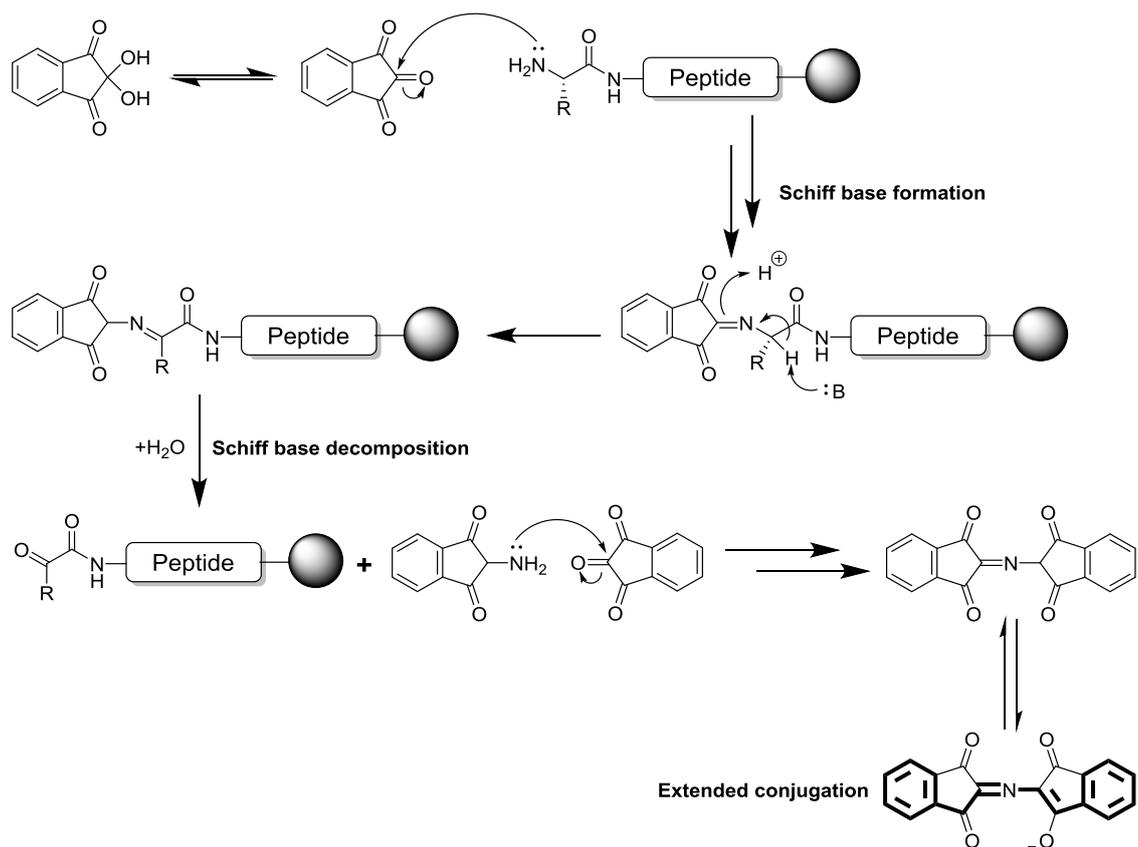
In general, current peptide synthesis follows a standard procedure of N-terminal deprotection, followed by amino acid coupling and free amine detection, to ensure the reaction is complete, before repetition of the cycle (Scheme 2.1). For hydroxy functionalised resins, which yield a free acid C-terminus, resin loading is typically achieved via formation of a symmetrical anhydride. The first amino acid to be attached is converted to its anhydride by a condensation reaction with diisopropylcarbodiimide (DIC), and then introduced to the resin with a catalytic amount of dimethylaminopyridine (DMAP). Keeping the equivalence of DMAP low reduces racemisation of the first residue. The ester produced is stable to the stepwise deprotection and coupling reactions involved in chain elongation. Couplings are achieved via formation of an active ester followed by attack of the peptide chain's terminal amine (Scheme 2.2). Many coupling agents are available, including carbodiimides (DIC, EDCI), uronium salts (HBTU, HATU) and phosphonium salts (PyBOP, PyAOP). All of these are used regularly in SPPS, however, HBTU is generally preferred as it offers a balance of efficiency and cost while minimising racemisation. Due to ease of excess reagent and

side product removal by washing the resin, it is typical to use both coupling agents and amino acids in a 4 to 8 fold excess to drive the reactions to completion.



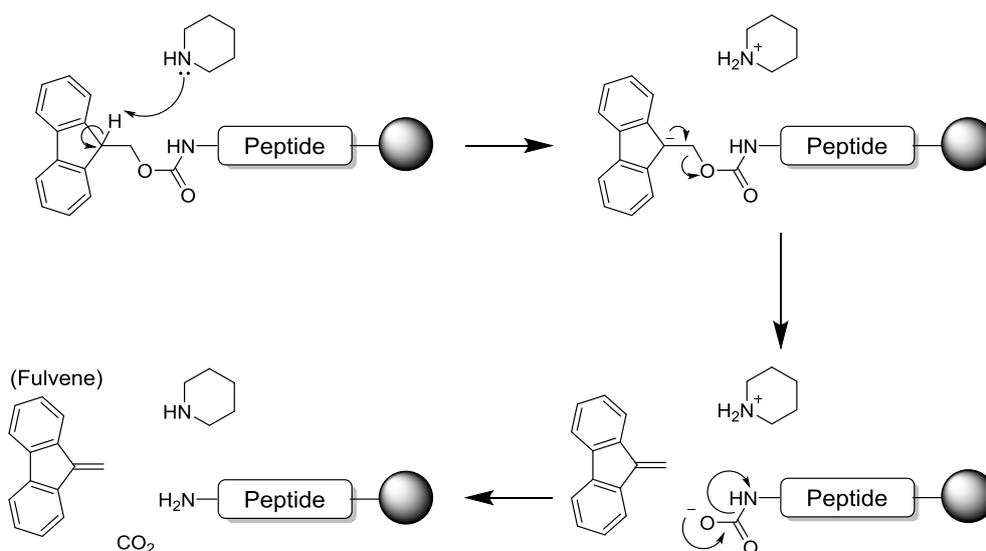
Scheme 2.2: Coupling of an amino acid to a resin bound peptide using HBTU

Even with these excess reagents, it is not uncommon for the coupling to be incomplete after a single treatment. To assess this, the colourimetric ninhydrin test described by Kaiser *et al.* is employed (Scheme 2.3).¹²⁵ A small sample of washed resin is treated with ninhydrin, phenol and potassium cyanide and heated to 120°C for a few minutes. Reaction of the ninhydrin with primary amines generates an extended conjugation network, resulting in a deep blue colour to the resin beads and solution. This decreases as the number of free amines is reduced by coupling, leaving a yellow colour when all amines are reacted. There are a few notable drawbacks to the Kaiser test, firstly, it is destructive, so the amount of resin used should be minimised. Secondly, it is unable to detect secondary amines or primary amines adjacent to a tertiary carbon due to the imine formation and deprotonation steps respectively. Despite this, its sensitivity to primary amines and ease of use make it the most commonly used amine test in SPPS.



Scheme 2.3: Kaiser's ninhydrin test for the detection of primary amines

Once coupling is completed, the N-terminal Fmoc protecting group is removed using piperidine as a hindered base (Scheme 2.4). Coupling can then be repeated to add another amino acid to the chain. This process is repeated until the peptide is complete, at which point the chain is cleaved from the resin and all side chain protecting groups are removed.



Scheme 2.4: Removal of N-terminal Fmoc protecting group using piperidine

An unexpected result of adopting Fmoc based chemistry was the ability to monitor deprotection reactions by UV spectroscopy. The fulvene (Scheme 2.4) species generated have a strong 280 nm absorption which can be measured to calculate an estimate of resin loading as well as to assess the aggregation of so called 'difficult sequences'. This phenomenon is characterised by a sudden decrease in reactivity of the growing peptide chain, both to deprotection and coupling. While free amine may be detected by Kaiser test, repeated coupling steps have no effect. The gel-phase properties of SPPS require a fully solvated matrix, which the reagent can penetrate freely. Intra and inter-molecular hydrogen bonding can disrupt this accessibility, leading to the effects of aggregation (Figure 2.2). Under ideal conditions of the fully solvated gel phase, both the peptide chains and polymer support are evenly distributed and exposed to solvent (Figure 2.2A). Aggregation can manifest in a number of ways as a result of less than ideal solvation. Firstly, individual peptide chains can form intra-molecular bonds, preventing access to the N-terminus for elongation (Figure 2.2B). Alternatively, the polymer support may form intra-molecular interactions with itself, contracting the resin bead and preventing access to the peptide chains (Figure 2.2C). Lastly, proximity of peptide chains to one another may allow formation of inter-molecular interactions, again preventing access to the N-terminus for elongation (Figure 2.2D).

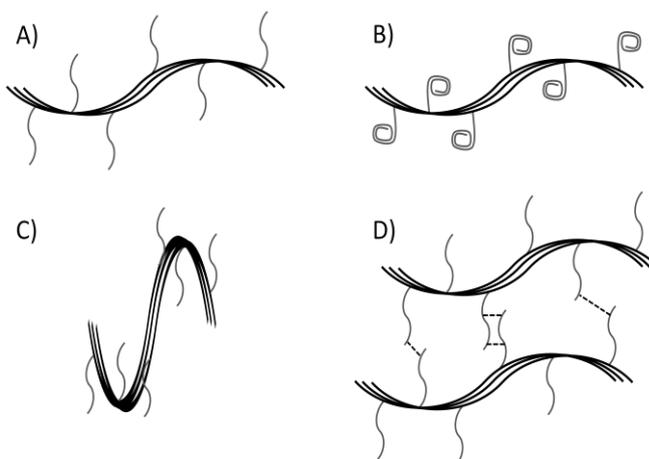


Figure 2.2: Solvation and aggregation of peptide chains and polymer supports. A) Fully solvated B) Intra-molecular peptide chain aggregation C) Polymer support aggregation D) Inter-molecular peptide chain aggregation

The earliest attempts to combat aggregation, introduced by Sheppard, included switching to polyamide based resins to increase solvation and utilising more polar aprotic solvents than DCM which was standard at the time.¹²⁶ The highly polar aprotic amide solvent dimethylformamide is now the solvent of choice for most syntheses. Although this alone has not completely solved the problem of aggregation, there are several other alterations that can be made if aggregation is suspected. During a synthesis it may be possible to reverse aggregation by addition of chaotropic salts to the reaction. Addition of 0.4 M lithium chloride or potassium thiocyanate has had some

success.¹²⁷ Alternative solvents may also be employed to increase solvation, including N-methylpyrrolidone and dimethylsulfoxide.^{128,129} However, both these methods attempt to reverse aggregation rather than prevent its occurrence. Two methods exist to physically prevent intermolecular backbone hydrogen bonding, the first utilises PEGylated resin to increase solvation and place growing peptide chains at a greater distance from one another.¹³⁰ The second requires more planning and knowledge about the onset of aggregation in the sequence but provides the only definite solution to aggregation. By selectively alkylating the amine of the amino acid, its ability to hydrogen bond is removed. The 2-hydroxy-4-methoxybenzyl (Hmb) group has been developed for this purpose and insertion of an Hmb protected amino acid prior to the onset of aggregation is the most reliable method for its prevention.¹³¹ While still far from trivial, all these developments have made synthesis of peptides accessible to a wide audience and equipped them to deal with the challenges it presents.

2.1 - TAT-Nrf2 Peptides

2.1.1 - TAT-Nrf2 Peptide Synthesis

The design of the proposed cell penetrating Nrf2 binding sequence peptides was based upon previous work by Lo *et al.* who examined the ability of three peptides, based on the key DEETGE binding motif of Nrf2, to disrupt the Nrf2/Keap1 interaction.⁴⁰ These peptides, a 16 mer, 14 mer and 10 mer were all found to be able to displace tagged Nrf2 protein *in vitro*. Using isothermal titration calorimetry they were able to determine a K_d for the binding of the 16 mer peptide of 20 nM. This was in good agreement with other data for the Nrf2/Keap1 interaction.^{42,95} While no data was provided for the shorter peptides, they state that the 14 mer peptide had comparable binding while the 10 mer peptide's binding was considerably weaker. Based on this information, initially three peptides were devised comprising those described by Lo *et al.* with the addition of the TAT cell penetrating sequence to their N-terminus (Table 2.2). It was predicted that with the addition of TAT, the peptides would enter cells and bind to Keap1, preventing Nrf2 degradation. All three sequences were tested simultaneously in order to compare efficacy between them. It was anticipated that potency would follow the same pattern as seen by the previous author.

Peptide	Sequence
TAT-10	YGRKKRRQRRR ⁷⁶ LDEETGEFLP ⁸⁵
TAT-14	YGRKKRRQRRR ⁷⁴ LQLDEETGEFLPIQ ⁸⁷
TAT-16	YGRKKRRQRRR ⁶⁹ AFFAQLQLDEETGEFL ⁸⁴

Table 2.2: TAT-Nrf2 peptide sequences

The initial three TAT-Nrf2 peptides were synthesised using standard Fmoc solid phase peptide synthesis techniques. In order to produce C-terminal acid peptides the hydroxybenzyl linker of Wang resin was used to immobilise amino acids on the resin via ester formation. For the TAT-10 peptide however, 3-chlorotrityl chloride resin was employed in order to prevent diketopiperazine formation. This is a common occurrence in sequences featuring C-terminal proline attached to the resin via benzyl esters. The much more hindered trityl linker used in this case prevents the cyclisation reaction, suppressing side product formation. Purity of the two longest peptides following cleavage was low, resulting in the need for extensive purification of both. Initial gross purification was achieved using reverse phase flash chromatography, yielding mixtures in which the desired product was the majority component. Subsequent purification by reverse phase HPLC increased the purity to >90%, suitable for further usage. The TAT-10 peptide had far higher crude

purity and required only reverse phase flash purification to yield pure peptide. After purification, the peptides were lyophilised from 0.05% TFA in water, yielding the products as TFA salts.

Peptide	HO-1 fold induction (6h)	SEM
TAT-10	2.6	0.63
TAT-14	18.9	3.92
TAT-16	6.2	1.59

Table 2.3: HO-1 mRNA levels by qPCR following treatment with TAT-10, TAT-14 and TAT-16 peptides in THP-1 cells, 75 μ M peptide, Mean \pm SEM, n = 3

Peptide concentrations were calculated assuming all basic residues were protonated and associated with a trifluoroacetate counter ion. Assessment of these peptides' ability to stabilise Nrf2 and induce downstream genes via qPCR (Table 2.3) and Western blotting, identified the TAT-14 peptide as the lead sequence (see below, Section 2.1.3). Following which a second set of peptides were synthesised.

Peptide	Sequence
TAT-14	YGRKKRRQRRRLQLDEETGEFLPIQ
TAT-14Sc	YGRKKRRQRRREFGTDIQLLEPQLE
14mer	LQLDEETGEFLPIQ

Table 2.4: Peptide sequences based on the 14 mer peptide

Having identified TAT-14 as a strong inducer of Nrf2 regulated genes, it was necessary to design control molecules in order to demonstrate both the specific binding of the 14 mer sequence and the enhanced cellular uptake provided by TAT (Table 2.4). To this end, the 14 mer sequence without TAT was synthesised to demonstrate that uptake of TAT-14 into cells was dependent on the CPP. In addition a scrambled 14 mer sequence was designed to demonstrate that increase in Nrf2 regulated genes was due to the specific sequence chosen i.e. one from the native Nrf2 protein and therefore disruption of the Nrf2/Keap1 interaction as the most likely mode of action. The TAT-14Sc sequence was created from the same amino acids as the 14 mer peptide but ensuring that none of the residues remained in the same position or order, resulting in a peptide with distributed polarity rather than clustered as in the natural sequence (Figure 2.3). In particular, the glutamic acid residues believed to play a crucial role in the interaction were spaced so as to prevent similar binding patterns forming to those found in the native peptide. It was also ensured that no inverse sequence remained intact or similarly charged residues were adjacent.

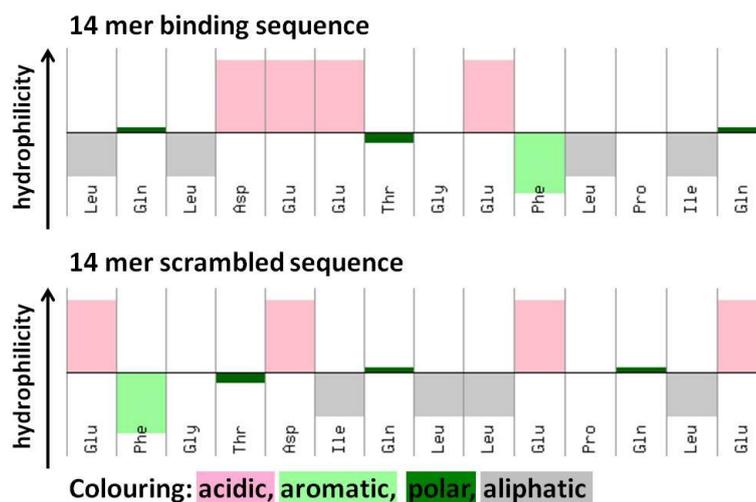


Figure 2.3: Hydropathy plots of native 14 mer peptide and designed scrambled sequence showing distribution of polar side chains

2.1.2 - TAT-Nrf2 Peptide Synthesis Optimisation

Before proceeding further with generation of analogues, the synthesis of the TAT-14 peptide was revisited in order to address some shortcomings of the original method. Initial synthesis on Wang resin was extremely low yielding, resulting in the need for extensive purification and as a consequence, low quantities of pure material. Analysis of crude peptide mixtures by MALDI revealed that a series of amino acid deletions were responsible for the multiple peaks observed by HPLC (Figure 2.4).

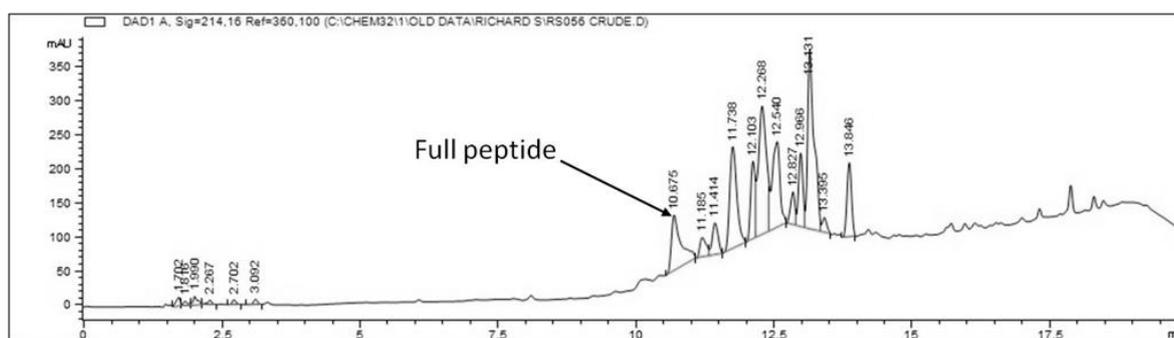


Figure 2.4: Crude HPLC trace of the TAT-14 peptide synthesised on Wang resin

The occurrence of multiple deletions indicates consistently low coupling efficiencies. Had a single residue or sequence been the source of the problem it would be more likely to see a small number of by-products rather than the wide range observed. In order to attempt to counteract this, couplings were repeated in triplicate, it was hoped that this would force each step to completion and thus yield pure peptide. This was not the case however and while the crude material contained far fewer peaks, the desired peptide was not the majority (Figure 2.5).

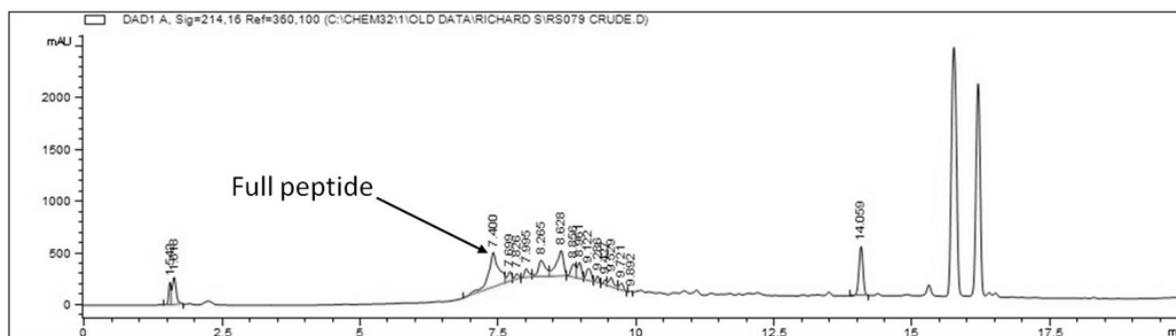


Figure 2.5: Crude HPLC trace of the TAT-14 peptide synthesised using triple couplings

The shift in retention times for the peaks is due to a change in mobile phase from water/methanol to water/acetonitrile. The two large peaks at 15.7 and 16.2 minutes are side chain protection by products as a result of simultaneous deprotection during cleavage. Following this synthesis, the opportunity arose to trial a microwave peptide synthesiser. It was decided that the TAT-14 peptide would be a good test as it should be easy to see any improvements afforded by the technique. The synthesis was carried out using the same reagents and conditions as the first synthesis with the exception of the coupling steps where 25 min agitation at room temperature was replaced with 10 min microwave irradiation. In this case there appears to have been little benefit to microwave irradiation for couplings. While the coupling steps were significantly shortened, the purity of peptide is comparable to the synthesis utilising triple couplings (Figure 2.6).

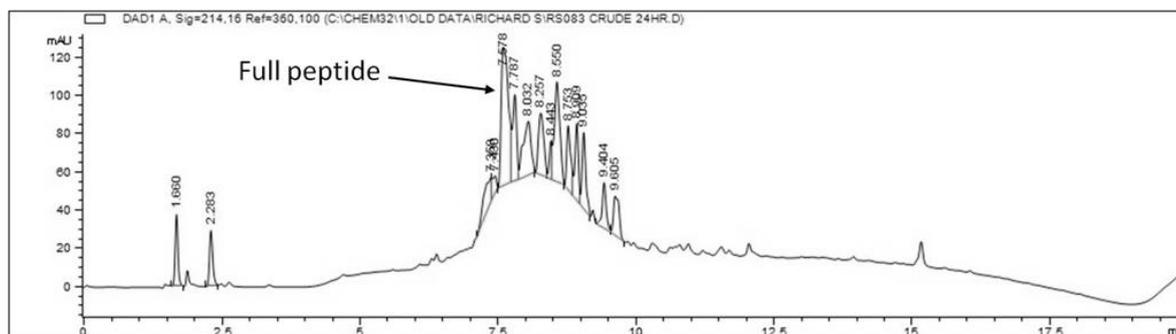


Figure 2.6: Crude HPLC trace of the TAT-14 peptide synthesised using microwave irradiation for coupling steps

It was suspected that aggregation was the cause of the poor coupling efficiency, despite the lack of residues commonly associated with the problem. In order to identify the point at which aggregation was beginning and hopefully find a way to continue the synthesis, the peptide was synthesised manually.

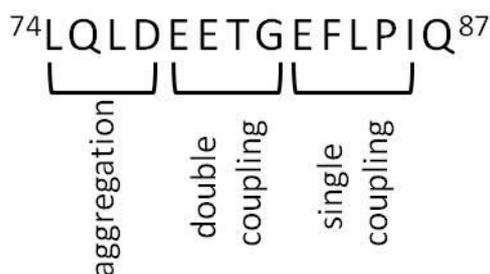


Figure 2.7: Sequence of 14 mer peptide indicating coupling requirements and onset of aggregation

To this end it was decided to synthesise the 14 mer peptide and subsequently assess the viability of continuing to the full TAT-14 peptide (Figure 2.7). The first amino acid, glutamine, was loaded and capped as normal using the symmetrical anhydride method and acetic anhydride respectively. The following five amino acid additions, ⁸²EFLPI⁸⁶, were complete by Kaiser test with a single 25 min coupling. Subsequently three of the next four amino acids, ⁷⁸EETG⁸¹, required double coupling to drive them to completion. Couplings were extended to 45 min to try and improve efficiency. This seems to have had little effect as the next two amino acids, Asp77 and Leu76, required four and three couplings respectively. It was fairly clear that this was the point at which aggregation became insurmountable. For the penultimate amino acid addition, Gln75, several attempts were made to ease the coupling. Initially the coupling agent was switched from HBTU to PyBOP in the hope that a more potent agent would benefit coupling, though this was not seen to have an effect. Several approaches were assessed to reduce the aggregation itself, firstly lithium chloride was added to the DMF used in the reaction at a concentration of 0.4 M, with the hope that the additional ions would prevent the association of peptide chains. The reaction solvent was then changed to NMP still containing 0.4 M LiCl as there is evidence to suggest improved solvation compared to DMF. This finally led to completion of the reaction, though it took seven couplings to do so. The final amino acid addition was achieved in just two couplings though each lasted 24 hours. The crude HPLC of this peptide contains one major product, aside from cleavage by-products at 16 min, however this is only 14 amino acids of the full 25 in the TAT-14 peptide (Figure 2.8).

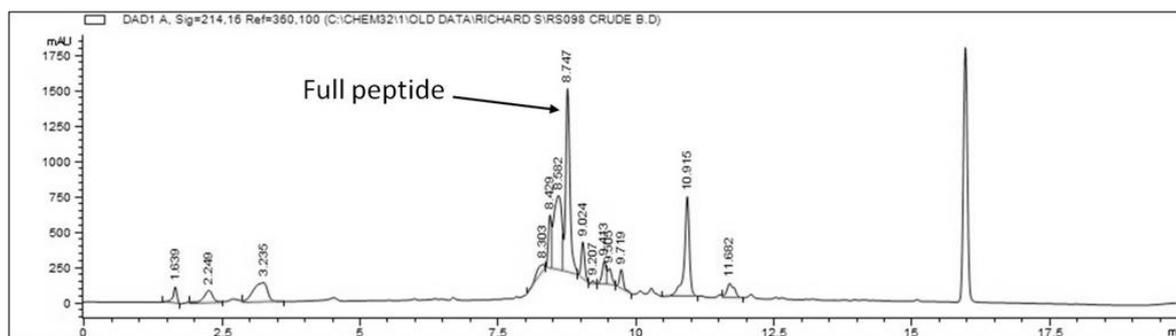


Figure 2.8: Crude HPLC trace of the 14 mer peptide synthesised manually to assess aggregation

Due to the multiple extended couplings required to add additional amino acids to this chain, continuing the synthesis to the full TAT-14 peptide and analogues was not viable. Instead, options for preventing the onset of aggregation were examined. One option was to introduce Hmb protected amino acids into the sequence prior to the onset of aggregation. With good knowledge of where aggregation starts to affect coupling efficiency this would have been possible but it was felt that it would be prohibitively expensive, especially when considering synthesis of analogues. An alternative approach was to employ a resin with increased solvation properties and lower loading with the hope that this would prevent aggregation by increasing the physical distance between immobilised peptide chains. This was believed to be a far more practical approach as it would be applicable to any further analogues synthesised. The resin chosen for the synthesis was NovaSyn TGA, a PEG functionalised polystyrene resin bearing the same benzyl alcohol functional group as Wang resin. This affords the free acid C-terminus when cleaved and can be handled under the same conditions as Wang resin, making it an easy transition with our current syntheses. The first amino acid was loaded as normal using the symmetrical anhydride method and chain elongation was carried out using the Syro I automated peptide synthesiser, with each amino acid attached by double coupling. The benefits of this resin are clear from the crude HPLC trace, note a return to the original water/methanol mobile phase (Figure 2.9). There is a single major product, aside from cleavage by products between 17 and 18 min, which was isolated in a single purification step.

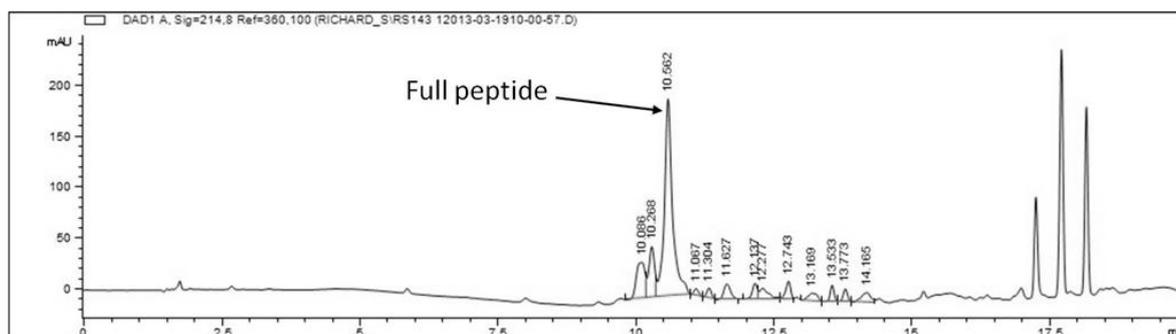


Figure 2.9: Crude HPLC trace of the TAT-14 peptide synthesised on Nova Syn TGA resin

2.1.3 - TAT-Nrf2 Peptide *In Vitro* Assays

Following purification, the peptides were taken up in sterile PBS to a concentration of 7.5 mM and stored at -80°C. *In vitro* testing was carried out in THP-1 cells, a leukemic human monocytic cell line.¹³² As monocytes play a vital role in the inflammatory response, the ability to modulate their behaviour with Nrf2 based peptides could have a large influence on resolution of inflammation.³ Assessment of the peptides' ability to affect Nrf2 levels by qPCR and Western blot was carried out in conjunction with Jon Cowan. None of the peptides were found to be toxic up to 150 µM when assessed for cytotoxicity by MTS assay over 72 hours. In all subsequent experiments, aside from determining dose response, the peptides were tested at a final concentration of 75 µM.

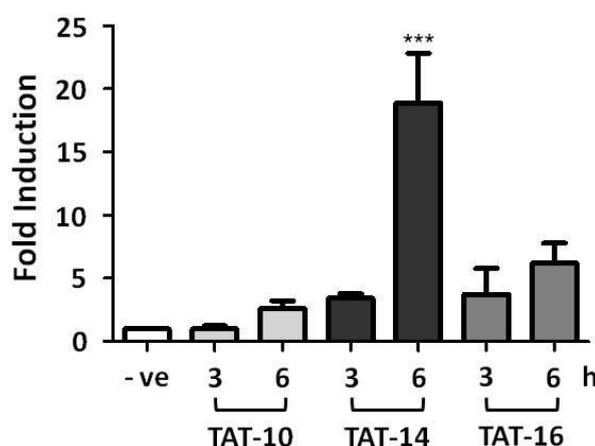


Figure 2.10: HO-1 mRNA induction by TAT-10, TAT-14 and TAT-16 peptides, Mean ± SEM, n = 3, ***p < 0.001

The TAT-10, TAT-14 and TAT-16 peptides were expected to enter cells, disrupt the Nrf2/Keap1 interaction and cause build up of Nrf2 protein. This would result in activation of downstream genes which could then be detected in order to measure the potency of the peptides. Hemeoxygenase-1 (HO-1) transcription is particularly strongly affected by Nrf2 induction, which makes it ideal for monitoring the effects of the peptides.³¹ Initially, HO-1 mRNA expression was

examined by qPCR for each peptide, by comparison to a standard curve and normalised against the housekeeping gene glyceraldehyde-3-phosphate dehydrogenase (GAPDH) (Figure 2.10). Cells were treated for either three or six hours with 75 μ M peptide. The cells were then pelleted and taken up in TRI reagent and the RNA isolated using a modified phenol/chloroform extraction method.¹³³ The RNA was then converted into the corresponding DNA by reverse transcription and the DNA amplified by PCR, monitored in real time by measurement of SYBR Green fluorescence. Relative concentrations were calculated by comparison to a standard curve. Neither TAT-10 nor TAT-16 showed significant induction, however the induction by TAT-14 was strong, averaging an 18 fold increase at 6 hours. The reasons for the lack of induction by the other peptides is unclear, however the binding of TAT-10 may have been affected by the presence of an arginine residue within one amino acid of the key DEETGE binding sequence. A similar result was seen by Inoyama *et al.* who noted that acetylation of the N-terminus of short peptide sequences increased binding.¹³⁴ It was also noted that extending the peptide by a further two amino acids had a similar effect. This would explain why the TAT-14 peptide was not affected as there are three amino acids in the sequence between the binding region and the TAT peptide. The inability of TAT-16 to induce was a more puzzling result as Lo *et al.* had previously reported that the 14 and 16 amino acid peptides showed similar binding to the Keap1 Kelch domain by ITC. While still not a significant increase to HO-1 mRNA levels, it is clear that the increase induced by the TAT-16 peptide is greater than that for the TAT-10 peptide.

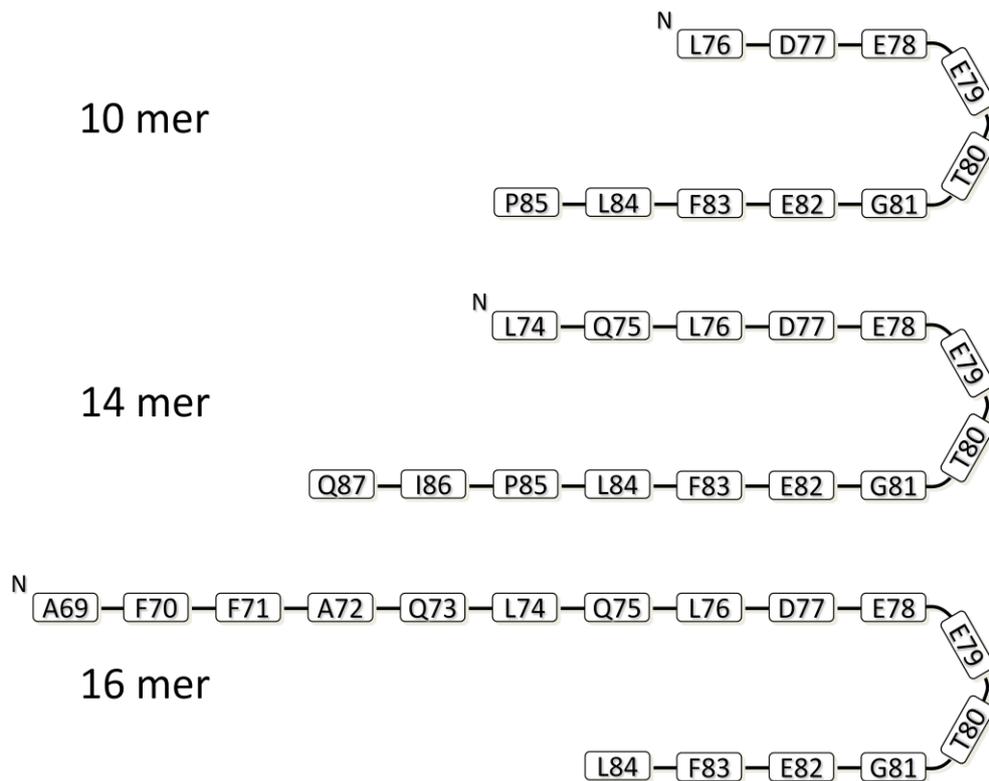


Figure 2.11: Hairpin sizes for 10, 14 and 16 amino acid binding sequence peptides

It is possible then, that the ability of the peptides to disrupt the Nrf2/Keap1 interaction in a cellular environment is partially due to the length of hairpin as well as overall length. Figure 2.11 shows the proposed folding pattern for each of the binding sequences tested. While the 16 mer peptide is the longest it has only the second longest loop, with four amino acids each side of the turn region. In comparison, the 14 mer peptide has five amino acids each side of the turn and the 10 mer peptide has only three. If an alternate 10 amino acid peptide, removing Pro85 and adding Glu75, showed increased binding it would support this hypothesis. Alternate 14 and 16 amino acid peptides could be constructed along the same lines to further explore this effect.

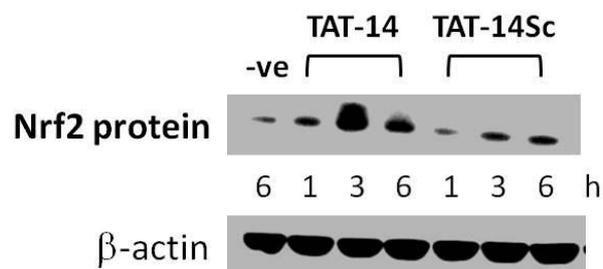


Figure 2.12: Nrf2 protein levels following treatment with TAT-14 or TAT-14Sc peptides

Having synthesised the TAT-14Sc and 14 mer peptides (Table 2.4), more detailed data could be collected on the effects of the TAT-14 peptide. In order to confirm that the TAT-14 peptide was

increasing Nrf2 levels, protein expression was examined via Western blot. Cells were treated with both TAT-14 and TAT-14Sc peptides for the time indicated (Figure 2.12). Following stimulation, the cells were pelleted and lysed in SDS buffer. Proteins in the sample were denatured by boiling under reducing conditions and then separated according to size by gel electrophoresis.¹³⁵ After transfer to a nitrocellulose membrane, Nrf2 protein levels were measured by immunoblotting using Nrf2 specific antibodies and compared to the structural protein β -actin to ensure even loading. With addition of the TAT-14 peptide, protein levels were found to peak after 3 hours while induction by the TAT-14Sc peptide was negligible. The induction by the TAT-14 peptide is strong and persists at 6 hours though it is greatly reduced. In comparison to the negative control and the 3 hour stimulation for the TAT-14Sc peptide, the 1 hour stimulation by TAT-14 may be showing an effect. If this is the case then there may be a slight induction at 3 and 6 hours with the TAT-14Sc peptide though this may be due to cellular stress as a result of penetration of the peptide. This shows a clear relationship between sequence and ability to disrupt the Nrf2/Keap1 interaction which cannot be attributed to cellular stress alone. In addition, a preliminary experiment examining Nrf2 mRNA levels by qPCR showed no increase in transcription upon stimulation with TAT-14, suggesting disruption of Nrf2 ubiquitination and not increased protein production. As Nrf2 production is linked to cell stress, this is also an indication that the effects are not simply due to stress.

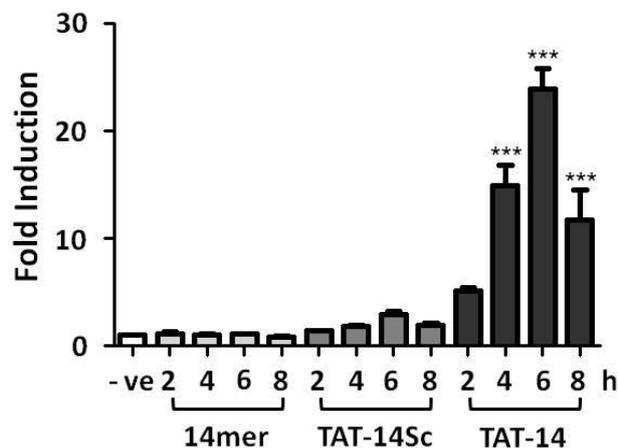


Figure 2.13: HO-1 mRNA induction by 14 mer, TAT-14Sc and TAT-14, Mean \pm SEM, n = 3, p* < 0.001**

In order to demonstrate the progression of induction by the TAT-14 peptide from increased Nrf2 levels to downstream genes, a time course was devised to measure HO-1 mRNA by qPCR as described previously (Figure 2.13). The increase in HO-1 mRNA over a time course from 2 to 8 hours was compared to that of the non-cell penetrating 14 mer peptide and the TAT-14Sc peptide. A clear pattern of induction was seen in cells treated with the TAT-14 peptide. mRNA

levels were already elevated after 2 hours reaching an average peak induction of 24 fold after 6 hours which quickly decreased to 12 fold after 8 hours. Conversely the effects of the 14 mer and TAT-14Sc peptides were insignificant; it is possible that any slight induction by the TAT-14Sc peptide was due to cellular stress as a result of cell penetration. The inability of the 14 mer peptide to induce HO-1 mRNA clearly suggests that addition of the TAT cell penetrating peptide is necessary for internalisation of the binding sequence peptides.

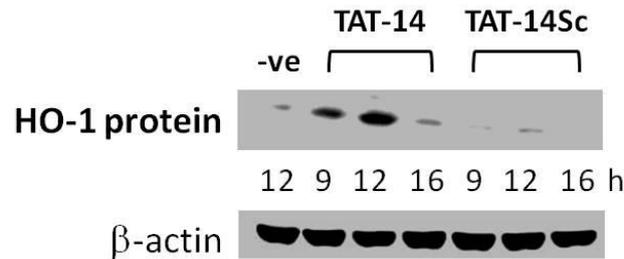


Figure 2.14: HO-1 protein levels following treatment with TAT-14 or TAT-14Sc peptides

Examining the translation of HO-1 mRNA into protein was important in ensuring that any induction seen in mRNA levels would result in increased antioxidant and anti-inflammatory response (Figure 2.14). Therefore, HO-1 protein levels were assessed by Western blot in the same way as previously described for Nrf2 protein. The results clearly show a progression from maximal HO-1 mRNA production at 6 hours to maximal protein levels at 12 hours. Again the TAT-14Sc peptide was seen to have no effect. As with Nrf2 protein and HO-1 mRNA the induction is sharp, while there is some increase in protein at 9 hours it has dropped back to basal levels by 16 hours. The TAT-14 peptide clearly has a transient effect which suggests it is not permanently altering the function of the cells. This is a desirable characteristic as permanent elevation of Nrf2 levels may lead to damage as seen in Keap1 null mice, which do not survive to maturity.³⁴

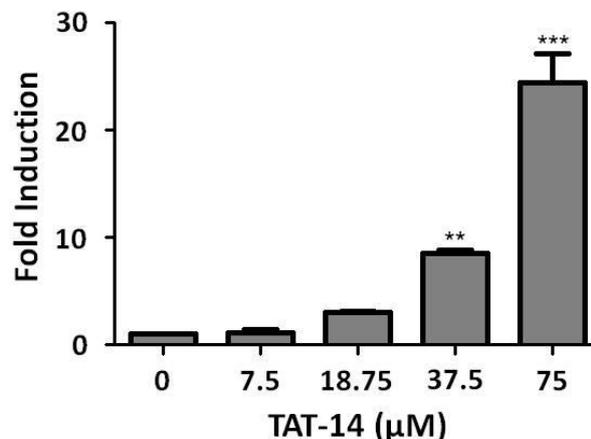


Figure 2.15: Dose Response of TAT-14 induced HO-1 mRNA levels. Mean ± SEM, n = 3, **p < 0.01, ***p < 0.001

A dose response experiment was performed in order to demonstrate both the concentration dependence of induction and identify the lowest usable level of peptide for future work. As HO-1 mRNA induction was so strong with 75 μ M TAT-14, it was decided that measuring transcription of this gene would provide a clear dose response (Figure 2.15). The dose dependent nature of the induction is obvious, and while only the 37.5 μ M stimulation is significant it is likely that with a greater number of replicates the significance of the lower concentrations would increase. Due to the sharp decrease in potency, it was decided to keep the concentration of peptide at 75 μ M in further experiments.

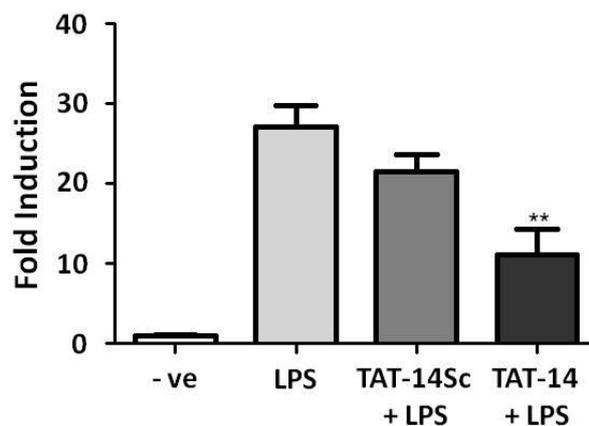


Figure 2.16: Attenuation of lipopolysaccharide (LPS) induced TNF α mRNA levels by TAT-14 compared to TAT-14Sc.
Mean \pm SEM, n = 3, **p < 0.01

To relate these results directly to reduction in inflammation, one of the ultimate goals of this approach, it was decided to measure the effect of the TAT-14 peptide on a key pro-inflammatory cytokine, tumour necrosis factor alpha (TNF α). Previous work within the group has shown that stimulation of cells with lipopolysaccharide (LPS) induces TNF α production as a model of sepsis.¹³⁶ Therefore, cells were treated with either TAT-14 or TAT-14Sc for 4 hours followed by stimulation with LPS for 3 hours following which TNF α mRNA levels were measured by qPCR (Figure 2.16). The cells treated with TAT-14 show a significant reduction in TNF α mRNA, averaging 61%, compared to cells treated with LPS alone. This is a clear demonstration of the utility of the TAT-14 peptide and disruption of the Nrf2/Keap1 interaction in general as an anti-inflammatory target. From this basis it was hoped to develop both more active peptide sequences and small molecules that target the protein-protein interaction as well as using the TAT-14 peptide as a tool for examining the extent of Nrf2 induction without increasing cellular stress.

2.2 - Polyarginine-Nrf2 Peptides

2.2.1 - Polyarginine-Nrf2 Peptide Synthesis

It was suspected that potency of the TAT-14 peptide could be increased by optimising the CPP sequence. A wide range of cell penetrating peptides have now been identified, among the most studied are TAT, transportan, penetratin and octa-arginine. The most potent of these are the polyarginine peptides, where efficiency of penetration increases with chain length up to 10 residues beyond which efficiency has been seen to decrease.¹¹⁵ The result is that octa- and nona-arginine have very similar properties. As reduction in chain length where possible is desirable and the penetration efficiency of octa-arginine is generally far greater than that of TAT it was decided that the octa-arginine peptide conjugated to the 14 mer sequence would be an attractive target. In addition, as penetration efficiency decreases with chain length, a second peptide was designed with tetra-arginine conjugated to the N-terminus of the 14 mer peptide. It has been reported that tetra-arginine is unable to internalise cargo and therefore should show no induction of Nrf2 regulated genes. This provides a simple control to show that uptake is as a result of cell penetrating peptide length. These R₄-14 and R₈-14 peptides were synthesised using the method developed for synthesis of the TAT-14 peptide, utilising PEGylated polystyrene resin. They were purified by reverse phase chromatography and lyophilised before being taken up in sterile PBS at a concentration of 7.5 mM.

2.2.2 - Polyarginine-Nrf2 Peptide *In Vitro* Assays

Before testing the efficacy of the compounds, they were both assessed for cytotoxicity by MTS assay to determine the maximum concentration at which the peptides could be tested without inducing Nrf2 through toxicity. Both R₄-14 and R₈-14 peptides were tested between 150 μ M and 7.5 nM in THP-1 cells for 24 h. At high concentrations cell viability is clearly reduced by addition of the R₈-14 peptide, with almost 60% cell death at 150 μ M (Figure 2.17). As the highest concentrations do not completely prevent cell growth it was not possible to calculate an *IC*₅₀ value for this compound. However, this data was very important in ensuring the correct concentration was used for further testing. Cell viability began to decrease at concentrations between 18.75 and 37.5 μ M, therefore the concentration for further testing was set at 25 μ M to ensure the greatest response without decreasing cell viability. In contrast, the R₄-14 peptide showed no toxicity even at 150 μ M, as with the TAT-14 peptide. This indicates that the source of toxicity is the extended arginine chain on increase from four to eight units. The cytotoxicity of octa-arginine has been noted previously as a drawback of its increased efficacy.¹³⁷ It was reasoned that as the TAT-14

peptide was potent at 75 μM the more potent $\text{R}_8\text{-14}$ peptide would still stabilise Nrf2 at lower concentrations. The ability to be effective while using less peptide is a significant advantage for the $\text{R}_8\text{-14}$ peptide.

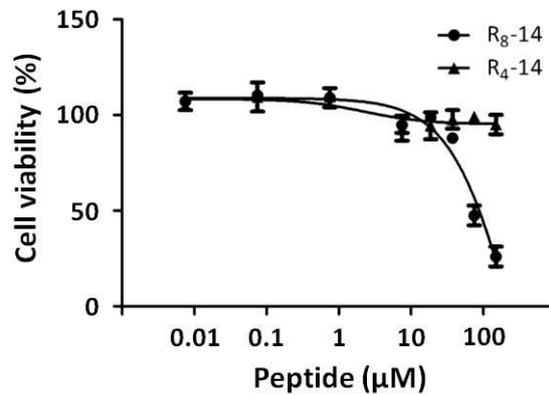


Figure 2.17: Cell viability of THP-1 cells incubated with varying concentrations of the $\text{R}_8\text{-14}$ (circles) and $\text{R}_4\text{-14}$ (triangles) peptides for 24 h. Mean \pm SEM, n = 3

In order to test if this was correct, THP-1 cells were treated with the $\text{R}_8\text{-14}$ and $\text{R}_4\text{-14}$ peptides at 25 μM and protein levels of Nrf2 were assessed. Cells were treated at 1, 3 and 6 hours as for the TAT peptides and Nrf2 levels measured by Western blot (Figure 2.18). As with the TAT-14 peptide, maximal stabilisation of Nrf2 by the $\text{R}_8\text{-14}$ peptide occurs after 3 hours with a sharp decrease by 6 hours. As expected, the $\text{R}_4\text{-14}$ peptide shows no activity at any of the time points tested. As a qualitative technique and without the TAT-14 peptide in parallel it is not possible to assess the relative potency of the TAT-14 and $\text{R}_8\text{-14}$ peptides, though as the $\text{R}_8\text{-14}$ peptide was tested at one third of the concentration of the TAT-14 peptide it would appear to be more potent.

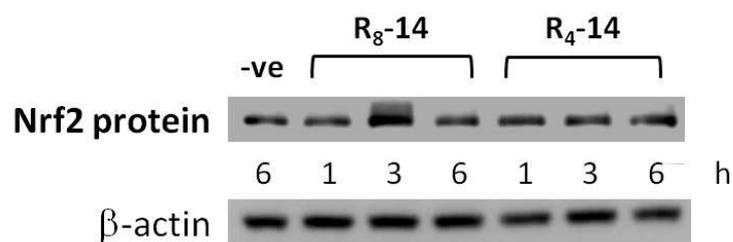


Figure 2.18: Nrf2 protein levels following treatment with $\text{R}_8\text{-14}$ or $\text{R}_4\text{-14}$ peptide

In order to determine the relative efficacy of the TAT-14 and $\text{R}_8\text{-14}$ peptides as well as the induction of HO-1 protein by the $\text{R}_8\text{-14}$ peptide an HO-1 ELISA was employed. Cells were treated with either the $\text{R}_8\text{-14}$ or $\text{R}_4\text{-14}$ peptide at 25 μM or the TAT-14 peptide at 75 μM for the time specified. Cells were treated for 9, 12, 16 and 24 hours, the same time points as used for TAT-14 HO-1 induction assessed by Western blot with the addition of a 24 hour time point to determine if

induction decreased after 16 hours. HO-1 levels were assessed using an R&D Systems DuoSet IC Human Total HO-1 kit by comparison of sample concentration to a standard curve. Total concentration of HO-1 was determined by analysing the standard dilution with a four parameter logistic curve fit. Induction of HO-1 was then expressed as fold induction compared to a non-treated control.

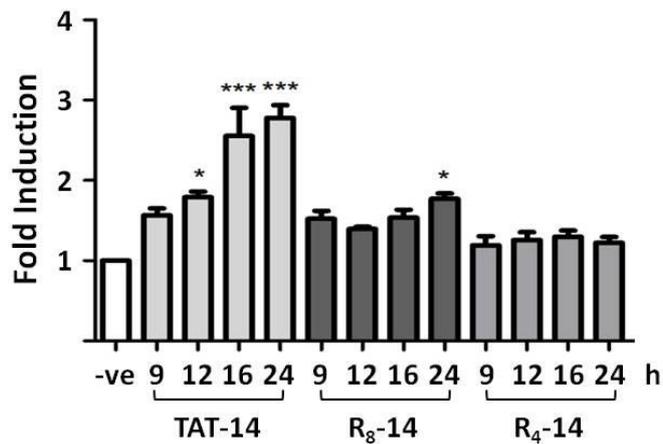


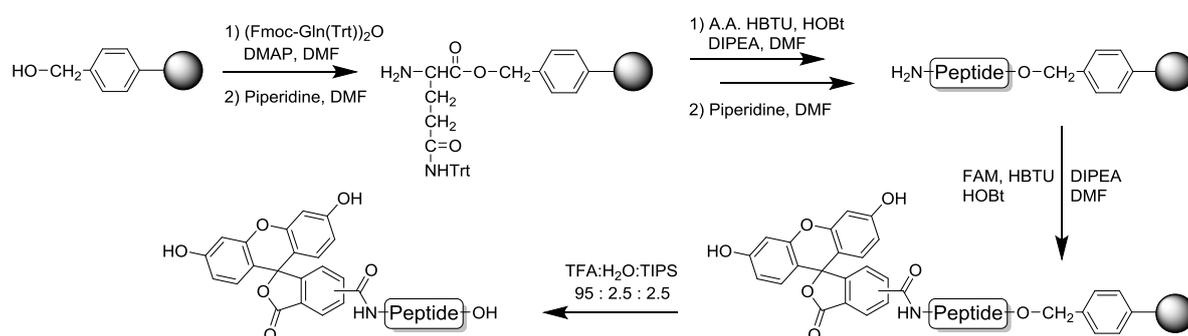
Figure 2.19: ELISA of HO-1 protein levels following treatment with TAT-14, R₈-14 and R₄-14 peptide. Mean ± SEM, n = 3, *p < 0.001, *p < 0.05**

Unlike the data from Western blot experiments, the induction by the TAT-14 peptide does not decrease after 12 hours and in fact continues to increase up to 24 hours (Figure 2.19). The reasons for this are unclear, though with some optimisation it should be possible to use samples prepared for ELISA in Western blotting experiments to clarify what is occurring. While not as strong an effect, the R₈-14 peptide also displays a similar pattern of induction to the TAT-14 peptide, though only the 24 hour time point is a significant induction over the control. Conversely, the R₄-14 peptide shows no induction at any time point. As with the Nrf2 Western data, this is a clear indication that the ability of the peptides to disrupt the Nrf2/Keap1 interaction is dependent on the cell penetrating sequence. The decreased induction of HO-1 protein by the R₈-14 peptide compared to the TAT-14 peptide is likely due to the stimulation being one third the concentration of the TAT-14 peptide. This highlights the importance of cytotoxicity assessment, as while the R₈ cell penetrating peptide may be more potent in general, its cytotoxicity outweighs any benefit when compared to the TAT peptide which shows very low toxicity.

2.3 - Fluorescently Tagged CPPs

2.3.1 - Fluorescently Tagged CPP Synthesis

Up to this point the cell penetration of these peptides was assumed to be effective due to the inability of the 14 mer and R₄-14 peptides to induce Nrf2 stabilisation or HO-1 induction. In order to confirm that the peptides were indeed entering the cells and not simply binding to surface receptors and acting via an unknown pathway, a fluorescently labelled cell penetrating peptide was designed. It was envisaged that a fluorophore attached to the N-terminus of the TAT-14 peptide would not interfere with binding to Keap1 due to the remoteness from the key binding sequence. Therefore, the linear sequence of the peptide was assembled as previously on NovaSyn TGA resin and following final Fmoc deprotection, 5(6)-carboxyfluorescein (FAM) was coupled to the N-terminal amine using HBTU/HOBt and DIPEA in DMF (Scheme 2.5). The only difference to the standard coupling procedure being the use of just two equivalents of FAM, HBTU and HOBt and four equivalents of DIPEA, which was found to be sufficient to produce complete functionalisation in 30 minutes as assessed by Kaiser test. The peptide was then cleaved from the resin and purified as for the TAT-14 peptide.



Scheme 2.5: Synthesis of F-TAT-14 on NovaSyn TGA resin by Fmoc SPPS

2.3.2 - Fluorescently Tagged CPP *In Vitro* Assays

The ability of this peptide to enter cells was monitored by fluorescence microscopy over a period of three hours. It was assumed that as maximal Nrf2 induction was seen by this time that the peptide must enter the cell in under three hours. THP-1 cells were treated with fluorescently labelled peptide as previously, for 180, 120, 60, 30 and 5 minutes, following which cells were pelleted and washed to remove excess peptide. In order to prevent the appearance of artefacts, cells were imaged live following a trypsin digest step as detailed by Richard *et al.*¹⁰⁸ The 5 min sample was split in two before trypsin digest in order to provide a control.

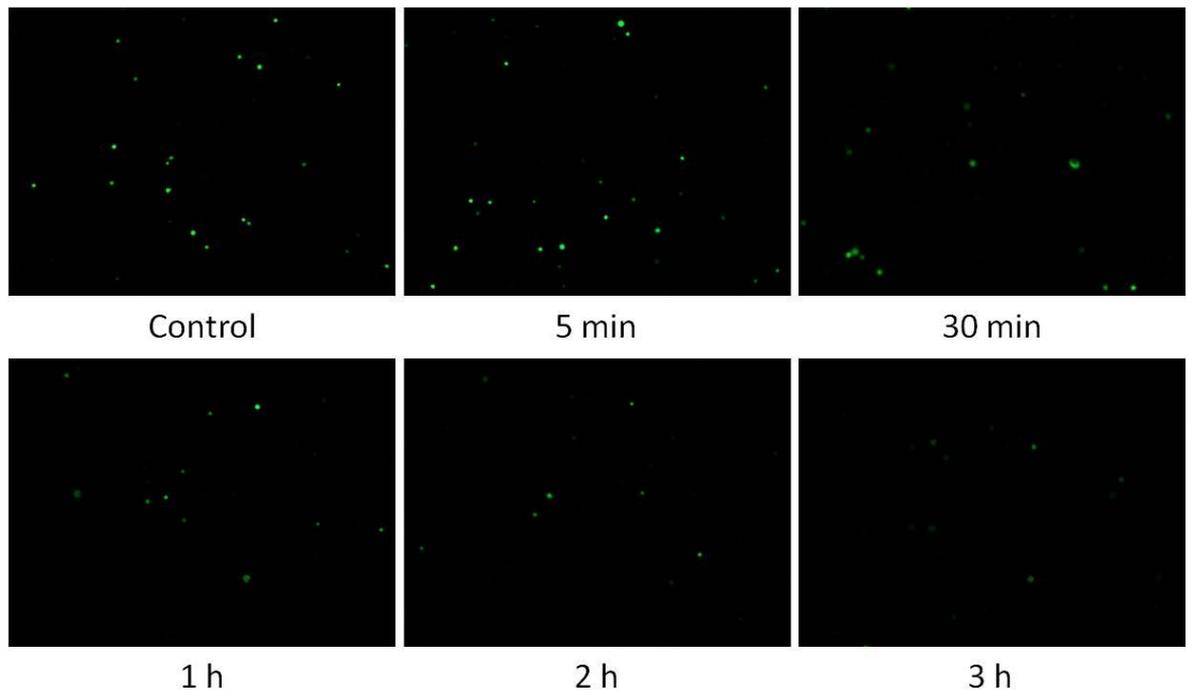


Figure 2.20: Fluorescence microscopy images of the F-TAT-14 peptide in live THP-1 cells. 1×10^7 cells/mL in PBS, Ex 490 nm, Em 520 nm

The most significant difference between the samples is the decreasing fluorescence with time. The cause of this is unclear, though it may be due to degradation of the peptide, which could explain the decrease in Nrf2 protein levels after three hours. Whether the similar levels of fluorescence in the control and 5 min samples is due to ineffective trypsinisation, or simply that levels of membrane associated peptide are low in THP-1 cells is unclear. There is evidence to show that circulating monocytes express very low levels of HSPGs, which are responsible for membrane adhesion of CPPs.¹³⁸ Whether this is the case or not, the higher magnification overlay indicates that the fluorescence is restricted to the inside of the cells (Figure 2.21).

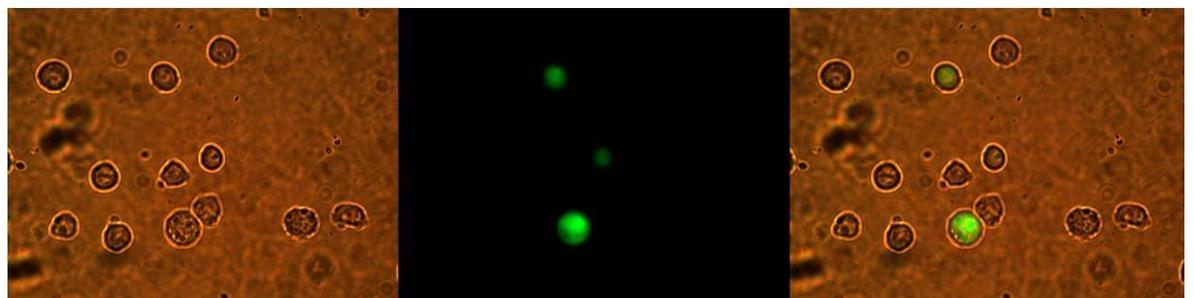


Figure 2.21: Representative overlay of 30 min sample showing internalisation of peptide within cells

While there appears to be significant quantities of peptide within some cells, the uneven distribution is puzzling. More detailed experiments need to be conducted in order to determine the causes of uneven uptake and decreasing fluorescence over time.

2.4 - Conclusions

Of the three peptides assessed by Lo *et al.* in purified protein assays, only one was found to be active in cell based assays when conjugated to a cell penetrating sequence. The TAT-14 peptide has been shown to be a promising tool for investigation of Nrf2 regulated genes and the role of Nrf2 in inflammation. Perhaps most significantly, the peptide validates targeting the Nrf2/Keap1 PPI for reduction of pro-inflammatory mediators. Optimising the synthesis of the TAT-14 peptide has allowed more rapid production and subsequently easier synthesis of similar peptides.

Investigation of polyarginine peptides as alternative CPP sequences for delivery of Nrf2 based peptides has highlighted the importance of assessing multiple transduction sequences. While the R₈-14 peptide shows some efficacy in cellular assays, the toxicity of the R₈ sequence cannot be ignored. It is possible that cell types other than monocytes would be more tolerant of the R₈-14 peptide, which may allow its increased cell penetration properties to be determined more easily.

The fluorescence microscopy experiments seem to indicate the F-TAT-14 peptide is entering cells rapidly, however, several questions are raised. The uneven uptake has no clear explanation and neither does the speed of uptake which would seem contrary to other recent work. As a tool however, the F-TAT-14 peptide provides another way in which disruption of the Nrf2/Keap1 interaction can be examined.

While the peptides presented in this chapter serve as useful tools and proof of concept for targeting the Nrf2/Keap1 interaction, they do not fulfil the role of lead-like compounds required to develop a small molecule inducer. In addition, the cellular assays required to characterise each of the peptides produced to this point were not conducive to rapid screening of new molecules or libraries. With this in mind, a fluorescence polarisation assay was established, the development of which is described in the following chapter. Using this assay, the CPP-Nrf2 peptides were further characterised.

Chapter 3: Fluorescence Polarisation

3.0 - Introduction

3.0.1 - Fluorescence Polarisation

Analysis of protein ligand interactions, where the ligand is either a peptide or small molecule, can be conducted in a number of ways, for example using fluorescence polarisation (FP), isothermal titration calorimetry (ITC), surface plasmon resonance (SPR) or radioligand assays.¹³⁹ Fluorescence polarisation has a number of advantages over other techniques, specifically, unlike radioligand assays there is no need to separate bound and unbound ligand before measurements are taken. This allows measurements to be taken at equilibrium and as a result readings can be taken from the same samples under varied conditions, e.g. temperature and time.¹⁴⁰ In addition the equipment required for high throughput fluorescence polarisation screens is relatively inexpensive and has uses outside the single technique. As a side product of using fluorescently tagged ligands rather than radioligands there is no need to handle radioactive compounds for prolonged periods which reduces both the cost and safety concerns related with the assay. As with any technique there are important limitations, potentially the most significant being the possibility of reduced binding of the fluorescently tagged ligand compared to the natural peptide. This can be limited somewhat by ensuring the fluorescent tag is introduced away from the key binding sequence and by measuring competition with the un-tagged sequence. The binding data obtained through fluorescence polarisation is typically not as reliable as ITC, though for screening purposes this should not be a problem.¹⁴¹ As a practical consideration, due to the need to titrate receptor into a fixed concentration of ligand, fluorescence polarisation uses more protein than an equivalent radioligand technique.

The principles behind fluorescence polarisation are based on rotational correlation time and absorption and emission of polarised light by rigid dipoles.¹⁴² In a fluorescence polarisation experiment, light emitted by an excited fluorophore is measured in two planes, in the same plane as the excitation and perpendicular to it. As a rigid dipole will emit in the same plane as it absorbed, given sufficient fluorescence lifetime, the angle of emission will be determined by the speed of rotation of the molecule. Assuming a spherical molecule, and identical viscosity of solution, rotational correlation time is proportional to volume. This means that a small unbound fluorescent ligand will rotate quickly, whereas a ligand bound to a large receptor will rotate slowly. As a result the emitted light from an unbound fluorescent ligand in solution will have no polarisation; detected light is equal in both parallel and perpendicular planes. In contrast a ligand

bound to a large receptor will have a far higher proportion of emission in the parallel plane than the perpendicular (Figure 3.1). This results in a polarisation signal which is proportional to ligand binding.

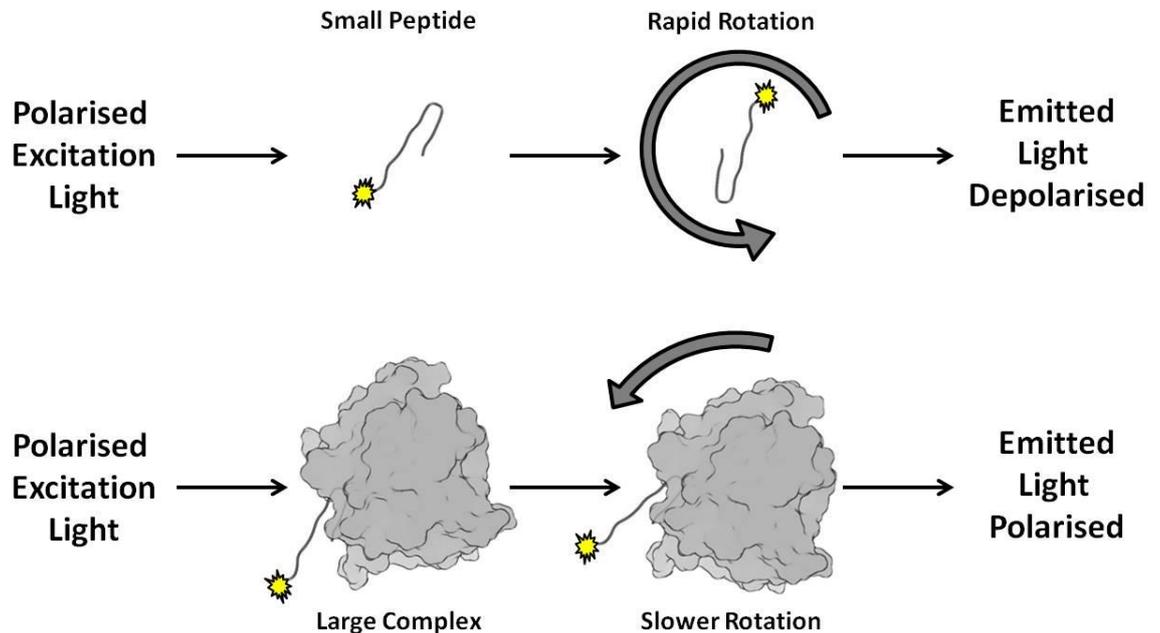


Figure 3.1: Effect of rotational correlation time on polarisation of emitted light.

Previous work by Hancock *et al.* has produced detailed sequence activity relationships for multiple fluoroscein tagged Nrf2 based peptides.⁹⁷ Peptides based on both the ETGE and DLG binding sequences were assessed along with sequences from sequestrome-1 which contains a sequence similar to the ETGE binding region of Nrf2. The DLG sequence peptides and ETGE sequence peptides less than 7 amino acids in length were unable to induce polarisation at any concentration tested, however 7, 9 and 10 mer peptides were found to induce strongly. Strength of binding was found to increase with chain length as may be expected with the 10 mer having a K_d of 51.0 nM. Despite this increase in binding the 7 mer peptide was selected for further screening assays as it produced the greatest change in polarisation on binding. This was expected to benefit the sensitivity of the assay due to the increased signal-to-noise ratio. Having established the assay, a screen of several peptide sequences was conducted. Notably, N-acetyl capping of an LDEETGEFL peptide was found to significantly increase binding, presumably due to removal of the N-terminal positive charge. Elongation of the peptide by 2 amino acids was found to have a similar effect. The lack of binding of TAT-10 may be related to this due to the proximity of the positively charged arginine residues to the binding sequence. Further modification of the peptide sequence using an alanine scan found that mutation of Ac-DEETGEF to Ac-DAETGEF increased binding from 5.39 μ M to 0.730 μ M. Sequences from sequestrome-1 indicated that

changing from DEETGEF to DPETGEL would increase binding further. With this in mind a fluorescence polarisation assay utilising the 14 mer peptide was developed in order to screen small molecules and assess the binding of peptides used previously in cells to explain their observed activities.

3.1 - Fluorescence Polarisation Assay

3.1.1 - Setup of Fluorescence Polarisation Assay

Development of an Nrf2/Keap1 based fluorescence polarisation assay for the screening of peptides and small molecules, as well as measurement of detailed binding data was setup based on those established by Rossi *et al.*, Hancock *et al.* and Cox *et al.*^{97,142,143} In order to construct the assay, a receptor-ligand pair was required. Isolation of the complete Keap1 protein is not trivial, however, analysis of binding to the Kelch region was expected to be sufficient as this is the location of the key binding interactions.³⁹ For this reason, Keap1 Kelch plasmid and protein provided by Hannink, as used to produce the crystal structure of the Kelch domain bound to the 16 mer peptide, was selected as the receptor.⁴⁰ For the ligand it seemed reasonable to select a peptide sequence for which we have strong positive data, therefore the 14 mer sequence was chosen as the basis for the fluorescently tagged ligand. As the N-terminal Leu residue of the 14 mer peptide plays no part in the interaction, and is positioned far from the binding site in the crystal structure (PDB I.D. 2FLU), it was decided that N-terminal functionalisation of the peptide to create a fluorescent ligand would be optimal. This was achieved by reaction of the fully deprotected and purified 14 mer peptide with fluorescein isothiocyanate in DMSO using triethylamine as a base. As there are no other nucleophiles present in the peptide, side products, such as double tagging were not expected, however, two major products were seen by RP-HPLC. One of these was identified by MALDI-TOF as having the correct mass and isolated by semi-preparative RP-HPLC.

The concentration of the N-terminally fluorescently tagged 14 mer peptide (F-14) was determined spectroscopically by dissolving the compound in DMSO, which was found to quench fluorescence, and diluting in water. The absorbance at 495 nm was measured using a Thermo NanoDrop 100 spectrophotometer and the concentration calculated using the Beer-Lambert law (Equation 1) with the extinction coefficient (ϵ) of fluorescein being $63000 \text{ M}^{-1} \text{ cm}^{-1}$ and path length of 0.1 cm to give a stock concentration of 425 μM from which all subsequent dilutions were made.

$$A = \epsilon cl$$

Equation 1: Beer-Lambert law. A: absorbance at 495 nm, ϵ : molar extinction coefficient ($M^{-1} cm^{-1}$), c: concentration (M), l: path length (cm)

Initial experiments sought to determine the K_d of the F-14/Keap1 interaction. In order to do this an estimate of the K_d was needed as a starting point for binding curves. The calorimetry data given by Lo *et al.* for binding of the 16 mer Nrf2 peptide indicated that 20 nM would be a reasonable concentration.⁴⁰

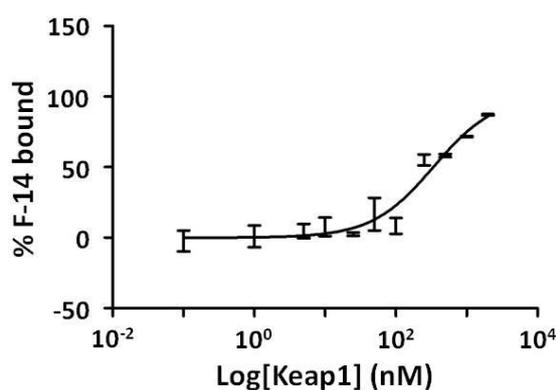


Figure 3.2: Fluorescence anisotropy of the F-14 peptide with varying concentrations of Keap1, Mean \pm SEM, n = 3, K_d 338 \pm 231 nM

Initially the only components of the system were the fluorescent peptide ligand and the protein receptor. Unlike typical radioligand assays, due to the build up of background fluorescence, the fluorescent ligand concentration must be fixed while the receptor is titrated in. Work by Hancock *et al.* indicated that 50 mM sodium phosphate buffer at pH 7.4 would provide the best signal while maintaining buffering capacity.⁹⁷ With the concentration of the F-14 peptide fixed at 20 nM a binding curve was constructed with Keap1 concentrations from 0 nM to 2000 nM (Figure 3.2). Analysis by non-linear regression gave a K_d of 338.4 \pm 231.3 nM, far higher than the expected value even taking into account reduced binding caused by the attachment of fluorescein. The large range of this value is mostly due to the lack of data points in the high concentration range. While stock concentrations of Keap1 would not allow higher data points, optimisation to decrease the K_d would have a similar effect. In order to optimise the assay the concentration of the F-14 peptide which produced the greatest change in polarisation upon binding to Keap1 was determined. A titration of 1000 nM Keap1 with between 1 and 15 nM F-14 suggested that with a 100 μ L well volume 5 nM F-14 peptide produced the greatest change in polarisation.

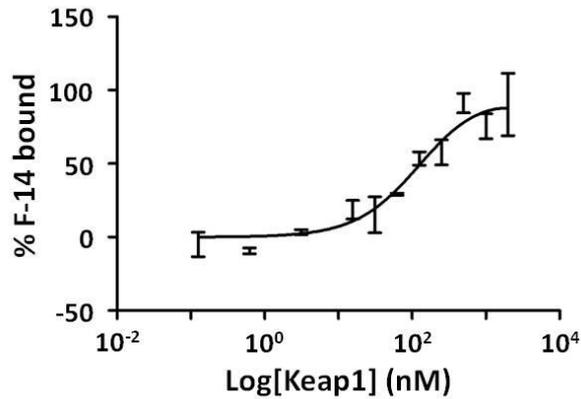


Figure 3.3: Fluorescence anisotropy of the F-14 peptide with varying concentrations of Keap1, Mean \pm SEM, n = 3, equilibration 20 min, K_d 128.4 \pm 67.7 nM

At this point it was noted that successive readings within an experiment showed steadily increasing polarisation, indicating that the peptide protein interaction had yet to reach equilibrium. To test this, a more detailed triplicate binding curve was performed, with readings taken immediately after mixing and then with increasing time intervals up to 6 hours. The resulting data shows that the K_d is stable for around 30 minutes after mixing; after which the K_d increases rapidly, suggesting a breakdown of the interaction. Analysis of the data after 20 minutes equilibration using non-linear regression produced a K_d value of 128.4 \pm 67.7 nM (Figure 3.3). Despite the relatively short window in which readings could be made and the still higher than expected K_d it was decided that collection of preliminary data for inhibition by the untagged peptide would be beneficial to ensure the assay was working as expected.

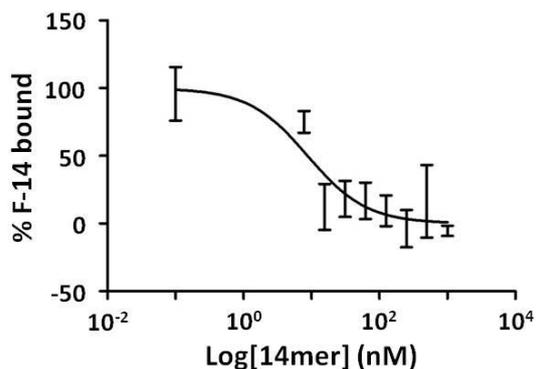


Figure 3.4: Fluorescence Polarisation inhibition by the 14 mer peptide, IC_{50} 9.4 nM, 95% CI [2.6, 29.2], Mean \pm SEM, n = 3, 200 nM Keap1, 5 nM F-14

Displacement of the F-14 peptide by the unlabelled 14 mer peptide was assessed as a function of concentration. The concentration of Keap1 was fixed at 200 nM so as to produce a fully bound signal of between 0.5 and 0.8 of saturation and F-14 peptide concentration was fixed at 5 nM as before. The concentration of the 14 mer peptide was then varied from 1000 to 0 nM in triplicate.

Readings were taken after 15 min incubation and show a clear decrease in polarisation with increasing 14 mer peptide concentration (Figure 3.4). Analysis of inhibition by non-linear regression resulted in an IC_{50} of 9.4 nM though a greater number of data points in the low concentration range would be needed for a robust figure. Conversion of IC_{50} to K_d via the Cheng-Prusoff equation was not thought to be valid at this point as an accurate K_d value was yet to be determined.¹⁴⁴

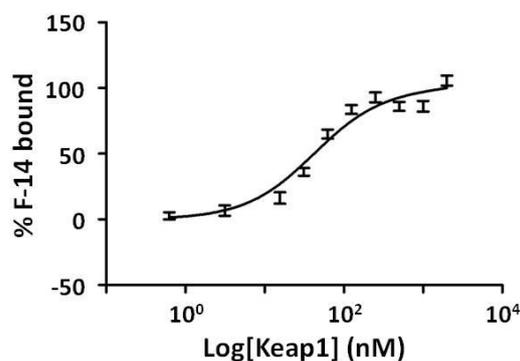


Figure 3.5: Fluorescence anisotropy of F-14 with varying concentrations of Keap1 and 0.1% Tween20 additive, Mean \pm SEM, n = 3, average over 30 min, 11 readings, K_d 42.1 \pm 7.2 nM

It was noted that Hancock *et al.* had examined the effect of surfactant on their similar assay and found it to have little effect on either binding affinity or assay stability. As surfactant is known to both decrease non-specific binding and increase stability of FP assays it was decided that the addition of 0.1% Tween20 may be beneficial.¹⁴¹ To this end an F-14/Kelch binding curve was produced both with and without the added surfactant. It was found that with surfactant the assay was stable for far longer and it appeared the K_d was reduced. In order to confirm this, a triplicate experiment was performed and the K_d measured over a period of 7 hours. Analysis by non-linear regression on all of the data collected over the first 30 minutes produced a K_d of 42.1 \pm 7.2 nM, far better than any previous experiment (Figure 3.5). In addition stability was increased dramatically, with the K_d rising to only to 55 nM after 7 hours.

The fluorescence polarisation assay was further optimised within the group, primarily in terms of simplifying the setup and reducing error due to pipetting where possible. The microplates for the assay were also changed to ones with very low surface binding. In addition it was found that in order to reduce the error of a single reading it was necessary to average multiple readings from the same wells. As a result each plate was subsequently read 10 times and values averaged before comparison to replicates. This reduced the observed error to within the expected limits for the plate reader and resulted in far more reliable data. After this process, a final binding curve was

performed to determine an accurate K_d value from which the parameters for inhibition experiments could be determined.

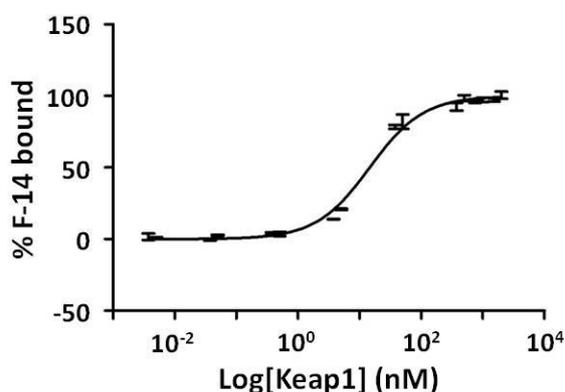


Figure 3.6: Fluorescence anisotropy of the F-14 peptide with varying concentrations of Keap1 and 0.1% Tween20, Mean \pm SEM, n = 3, K_d 14.6 \pm 1.1 nM

The optimised binding curve provides a K_d value that is in line with literature values for similar systems (Figure 3.6).⁹⁷ The assay was assessed for stability after 18 hours with no change in K_d over this time. With an accurate value for K_d determined and the stability of the assay confirmed, it was possible to define the parameters for the inhibition assay. For ideal conditions, the concentration of receptor should be fixed at between K_d and 0.8 maximal binding (B_{max}). With 80% of B_{max} being between 40 and 50 nM, 30 nM was chosen as a reasonable concentration.

3.1.2 - Fluorescence Polarisation Inhibition Assays

Following optimisation of the FP assay, the peptides screened previously for activity by cell culture techniques were assessed. The *in vitro* experiments had shown that ability to increase HO-1 mRNA expression and translation as well as stabilisation of Nrf2 protein was highly sequence specific. It was hypothesised that the inability of the TAT-10 peptide to induce these effects was due to disruption of binding to the Keap1 Kelch domain by the proximity of multiple positively charged arginine residues to the key negatively charged DEETGE binding sequence. In comparison to the TAT-14 peptide which has three residues between the cell-penetrating and binding sequences, the TAT-10 peptide has only a single residue between the two. For comparison, the 14 mer and 10 mer peptides were also tested. The 10 mer peptide was synthesised on 2-chlorotrityl chloride resin in an analogous manner to the TAT-10 peptide (See Chapter 2). It was expected that without the addition of the TAT sequence the potency would be increased due to greater ease of binding.

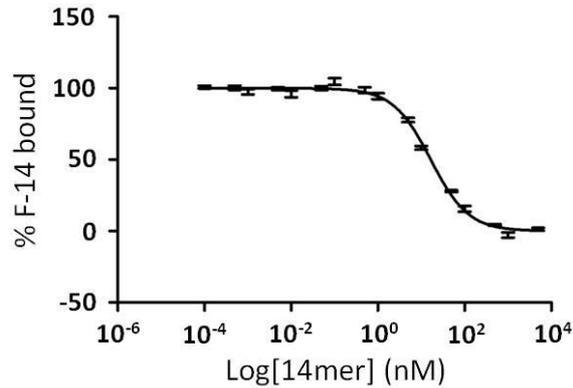


Figure 3.7: Fluorescence Polarisation inhibition by the 14 mer peptide, K_i 12.3 nM, 95% CI [10.7, 14.1], Mean \pm SEM, n = 3, 30 nM Keap1, 5 nM F-14

As a base line, the 14 mer peptide was used to displace the fluorescently tagged version (Figure 3.7). With a K_i of 12.3 nM (95% CI [10.7, 14.2]) the affinity is very similar to the tagged peptide, which had a K_d of 14.6 nM in the assay. This means that addition of the fluorescein has not affected the binding of the peptide. As a result the K_i values for other peptides and small molecules should be relevant to the full PPI. Subsequently the remaining peptides screened *in vitro* were tested for their ability to disrupt the assay.

The first surprise in these experiments came with the assessment of TAT-14 binding (Table 3.1). The 3.6 nM (95% CI [2.9, 4.6]) K_i was unexpected and indicates that the addition of the cell-penetrating sequence increases binding affinity rather than reducing as may be expected. The second surprise was that the TAT-10 and 10 mer peptides follow the same pattern of binding affinity.

Lo *et al.* previously determined that the binding affinity of the 10 mer peptide was an order of magnitude weaker than the 14 mer peptide which correlates with the 146.8 nM K_i (95% CI [102.3, 210.8]) by FP (Table 3.1). The TAT-10 peptide however has a far lower K_i of 6.9 nM (95% CI [4.7, 10.1]), which is more potent than the 14 mer peptide alone (Table 3.1). This suggests that the TAT peptide sequence contributes to binding significantly.

Peptide	K_i (nM)	95% CI (nM)	R^2
14 mer	12.3	10.7-14.2	0.9931
TAT-14	3.6	2.9-4.6	0.9809
10 mer	146.8	102.3-210.8	0.9499
TAT-10	6.9	4.7-10.1	0.9478
16 mer	88.1	67.4-115.3	0.9703
TAT-16	38.9	15.6-96.9	0.8857
R ₈ -14	5.8	4.7-7.1	0.9838
R ₄ -14	5.3	4.6-6.2	0.9918

Table 3.1: Fluorescence Polarisation inhibition by CPP-Nrf2 peptides, n = 3, 30 nM Keap1, 5 nM F-14

With these results in mind it was decided that analysis of the TAT-16 peptide and its non-TAT counterpart would provide more details about the relationship between the addition of the cell penetrating peptide and increase in binding affinity. The TAT-16 and 16 mer peptide were synthesised using the revised PEG-polystyrene resin synthesis producing the peptides in good crude purity. It is interesting to note that the 16 mer peptide was retained far more strongly on the reverse phase column than the other non-TAT peptides. This proved problematic for purification, as methanol was not sufficient to elute the peptide. Instead acetonitrile was used, as a more non-polar solvent. The importance of this finding is that it indicates the peptide is far more non-polar than the 14 mer peptide. This is significant when considering the effect of hydrophobic residues, particularly phenylalanine, on the potency of cell penetrating peptides. Addition of two Phe residues to the terminus of polyarginine peptides increases their uptake efficiency, which is thought to be due to increased hydrophobic membrane interactions.¹⁴⁵

The binding affinity of these two peptides was assessed using the FP assay as before. The TAT-16 peptide however, caused interference with the assay at concentrations above 100 nM, requiring the points above this concentration to be excluded (Table 3.1). In spite of this a K_i value of 38.9

nM (95% CI [15.6, 96.9]) could be calculated, though it is not as accurate as the values determined for the other peptides. In comparison to the other TAT conjugated peptides, there is a clear reduction in affinity even taking this into account.

As may be expected from looking at the degree of expected intra-molecular backbone hydrogen bonding of the Nrf2 binding sequence peptides, the 16 mer peptide falls in between the 16 mer and 14 mer peptides in binding affinity (Table 3.1). With 4 backbone hydrogen bonds and a K_i of 88.1 nM (95% CI [67.4, 115.3]), it fits with the pattern of the 14 mer peptide having 5 hydrogen bonds and the strongest binding affinity and with the 16 mer peptide having 3 hydrogen bonds and the weakest affinity.

For comparison, the polyarginine conjugated peptides were also tested. In this case it was found that the length of the cell penetrating peptide has little effect on the binding affinity (Table 3.1). The K_i of 5.8 nM for the R₈-14 and 5.3 nM for the R₄-14 peptides (95% CI [4.7, 7.1] and [4.6, 6.2] respectively) are identical within experimental error. Interestingly they are both more potent than the 14 mer sequence alone, suggesting that it is the first few amino acids before the key binding sequence which increases potency. The potency is decreased compared to that of TAT-14 which perhaps indicates that there are additional interactions as a result of the increased length of the TAT peptide. It must be noted however that the Keap1 Kelch domain expressed for use in this assay comprises only a portion of the entire Keap1 protein. Recent electron microscopy experiments indicate that the IVR domain of Keap1 is wrapped around the Kelch domain in the native protein.⁴⁴ As a result, if the cell penetrating portion of these peptides is interacting with the outer face of the Kelch domain this may not be an accurate representation of the interaction *in vivo*.

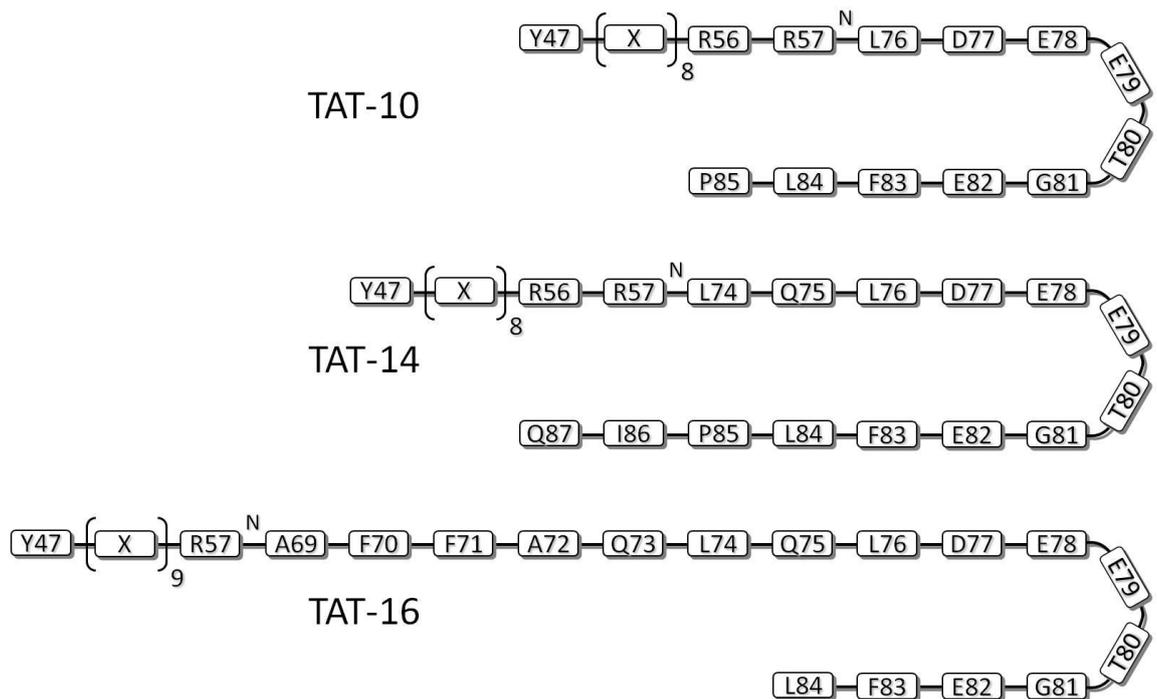


Figure 3.8: Hairpin backbone loop sizes for TAT-Nrf2 peptides, excess TAT chain truncated for clarity

With this new information, the hypothesis presented in Chapter 2 regarding levels of HO-1 induction and hairpin length may need to be revised. The factors affecting cellular systems are vastly more complex than those in a purified protein assay, as a result there may be entirely unrelated causes which result in similar patterns of binding affinity and gene induction. Taking this into account, the FP data indicates that the CPP portion of the peptide can affect its properties beyond the expected increase in cell membrane permeability. This might be explained by examining the secondary structure of the peptides when bound to Keap1. Taking into account the additional backbone interactions that could be formed between the binding and cell penetrating sequences the relative loop sizes between the TAT-10, TAT-14 and TAT-16 peptides change compared to those proposed in Chapter 2 (Figure 3.8). The TAT-14 peptide still has the largest loop, but because of the positioning of the amino acids in the TAT-16 peptide relative to the β -turn, the addition of the TAT sequence does not increase the loop size. This means that the TAT-10 peptide has the second longest loop when including the TAT peptide. While this would explain why the TAT-14 and R₈-14 peptides have increased binding compared to the 14 mer peptide it would not follow for the TAT-10 and TAT-16 peptides as by qPCR the TAT-16 peptide is more potent than the TAT-10 peptide. If however, the increased hydrophobicity of the 16 mer sequence has an enhancing effect on cell penetration of the TAT-16 peptide it could explain the observed activity despite reduced binding to the target.

$$Z' = 1 - \frac{3SD_{Bound} - 3SD_{Free}}{mAu_{Bound} - mAu_{Free}}$$

Equation 2: Z' test for well-to-well variation

Before embarking on further inhibition screens to assess the ability of unknown compounds to inhibit the interaction, the suitability of the assay to high throughput screening was determined. In order to do this, the Z' test was employed which measures the signal to noise ratio for inhibition assays.¹⁴⁶ Multiple data points with fully bound protein are taken, in this case 48, and compared to the same number with a sufficient concentration of inhibitor to fully displace the fluorescent ligand. The standard deviations of these measurements are then compared according to Equation 2, where SD is standard deviation of the anisotropy and mAu is the average anisotropy, to determine a value for variability. A value above 0.5 is considered to be suitable for usage as a high throughput assay. In this case, 48 wells containing 5 nM F-14 peptide and 30 nM Keap1 for the fully bound state and 48 wells with an additional 2000 nM 14 mer peptide for the fully dissociated state were compared (Figure 3.9). The calculated Z' value of 0.86 indicates that the assay is suitable for usage in high throughput screening.

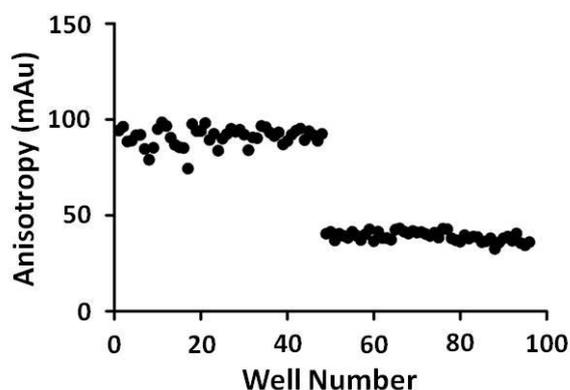


Figure 3.9: Z' test, wells 1-48: 5 nM F-14, 30 nM Keap1, wells: 49-96 5 nM F-14, 30 nM Keap1, 2 μ M 14 mer

3.2 - Conclusions

Having developed the FP assay, two avenues of investigation are now open. The assay can be used to gather detailed information about the binding affinities of further Nrf2 based peptides. Alternatively, as shown by the Z' test, the assay is suitable for high throughput screening, which could be used to identify entirely novel binders of Keap1 through analysis of compound libraries.

The binding affinities of the CPP-Nrf2 peptides determined using the FP assay have shed light on the relationship between loop length and binding affinity. It appears that additional residues added evenly each side of the turn increases binding to Keap1, possibly through stabilisation of

the β -hairpin motif. This would correlate with the intramolecular backbone hydrogen bonding seen in crystal structures, though more work is needed to confirm this. The binding affinities of the TAT-14 and R_x-14 peptides show that these are particularly potent inhibitors of the Nrf2/Keap1 interaction, further highlighting their potential as tools for investigating the Nrf2/ARE pathway.

With these potent tools in hand, development of peptides progressed towards more lead-like molecules. Key to this was minimisation of size while maintaining potency. As stabilisation of the backbone hairpin appeared advantageous, synthesis of peptide macrocycles was explored via a number of routes. These are described in the following chapter.

Chapter 4: Design and Synthesis of Cyclic Peptides

4.0 - Introduction

4.0.1 - Peptide Macrocyclisation

Cyclic peptides offer a number of benefits over their linear counterparts. The main effect of cyclisation is to restrict free rotation of the peptide's bonds which results in a more rigid, ordered structure that loses less entropy upon binding to its target than the linear analogue.¹⁴⁷ The decreased entropy penalty means binding affinity is increased. A side effect of cyclisation is decreased proteolytic degradation, particularly by exoproteases, this in turn increases the half-life of the peptide and therefore bioavailability.¹⁴⁸ There are several motifs for forming macrocycles in peptides, these are summarised in Figure 4.1. In addition to those shown, mixtures of strategies are possible, for example side chain to backbone or side chain to terminus. While many examples exist of each motif, a few examples, specifically of head-to-tail and sidechain-sidechain cyclisation, are highlighted.

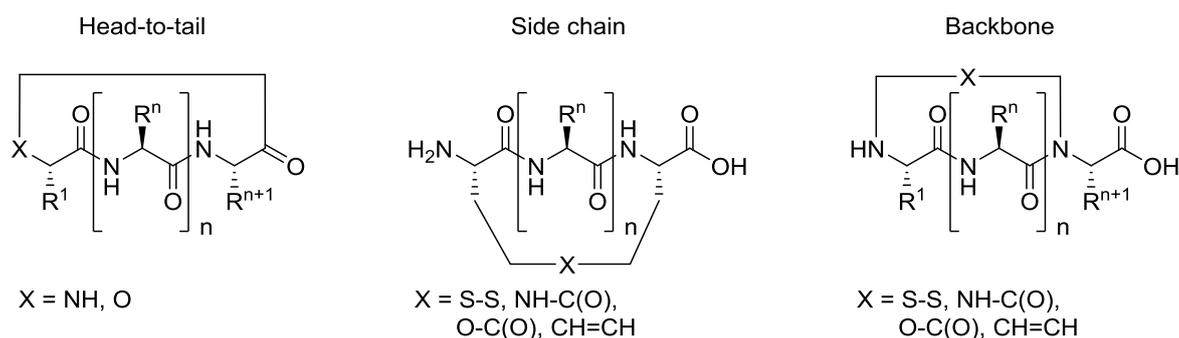


Figure 4.1: Selected cyclisation motifs found in natural peptides and employed in synthetic strategies (adapted from Liskamp *et al.*¹⁴⁷)

Cyclic peptides can be divided into two broad categories, natural products and protein derived macrocycles. As is often the case with drugs, the natural cyclic peptides have a long and detailed history from initial isolation through structure elucidation to total synthesis. The protein derived cyclic peptides are a more recent development, based on the observation that restriction of the conformation of a protein derived peptide can result in a compound with similar or increased activity to the entire protein.¹⁴⁹

4.0.2 - Cyclic Peptide Natural Products

A number of cyclic peptide natural products have found applications in areas such as antibiotics and immunosuppressants. Of these, vancomycin and ciclosporin are probably the most successful.

Aside from their cyclic structure, there are several features commonly found in cyclic peptides, which can be exploited when designing synthetic macrocycles. The cyclic heptapeptide portion of the glycopeptide vancomycin (Figure 4.2A) forms three macrocycles via modified sidechains.¹⁵⁰ These introduce atropisomerism to the molecule, which adopts a single atropisomer in the natural product. This produces a highly structured three dimensional conformation in the glycopeptide that affords high affinity binding to its target dipeptide, preventing bacterial cell wall biosynthesis. In spite of several approaches to total synthesis and development of analogues, the complexity of the compound means large scale synthetic production is still impractical.¹⁴⁷

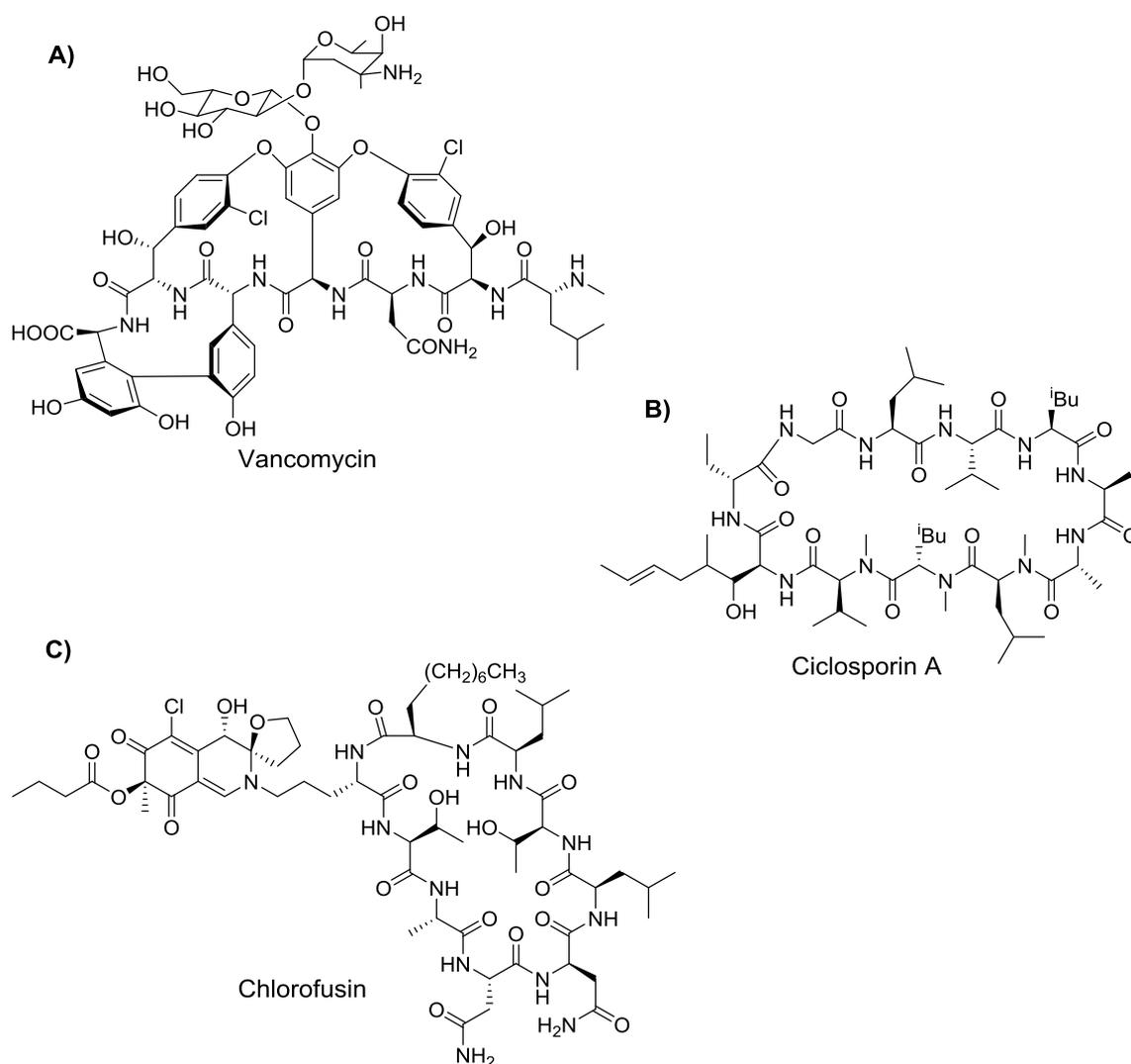


Figure 4.2: Cyclic peptide natural products A) Vancomycin B) Cyclosporin A C) Chlorofusin

The immunosuppressant cyclosporin (Figure 4.2B) is one of the main drugs used to prevent graft rejection following transplant surgery.¹⁵¹ The peptide is head-to-tail cyclised and consists of a mixture of L and D amino acids, a common feature in cyclic peptides as D amino acids allow

formation of a turn in the peptide backbone. Of the 11 residues, six are N-methylated and all are hydrophobic, important features in allowing the peptide to cross cell membranes.¹⁵²

Naturally occurring cyclic peptides have also been found to target PPIs. The head-to-tail cyclised nonapeptide chlorofusin (Figure 4.2C), binds to the N-terminal domain of MDM2, the negative regulator of the tumour suppressor p53.¹⁵³ Overexpression of MDM2 has been implicated in several cancers and represents an attractive target for development of chemotherapeutics. The cyclic peptide portion of chlorofusin consists of nine head-to-tail cyclised residues of which four are D amino acids. Synthesis of chlorofusin analogues has shown that both the peptide and chromophore portions are necessary for activity.¹⁵⁴

4.0.3 - Synthetic Cyclic Peptides

The second class of cyclic peptides are those based upon truncated protein structures, which seek to replicate the efficacy of the full protein while reducing overall size and increasing bioavailability. These peptides tend to be far smaller than the natural cyclic peptides, and in general feature less complex structures. Bernal *et al.* have developed an alternate method for targeting MDM2 by using truncated p53 based peptides.¹⁵⁵ In their strategy, non-natural amino acids bearing alkene side chains are inserted into the peptide sequence in place of residues that do not take part in the PPI. Following synthesis of the peptide chain, the two hydrocarbon side chains are linked by ruthenium-catalysed olefin metathesis, generating a so-called "stapled" peptide (Figure 4.3A). By spacing these amino acids seven residues apart, the α -helical structure of the peptide is stabilised. Variation of the position of these residues within the peptide chain generated a series of analogues with varying α -helix structure. This was measured using circular dichroism (CD) and the affinity of the peptide for MDM2 assessed by fluorescence polarisation (FP) binding experiments, using fluorescently-labelled peptides. The most efficient MDM2 binder of these peptides was found to restore p53 protein levels in SJS-1 cells.

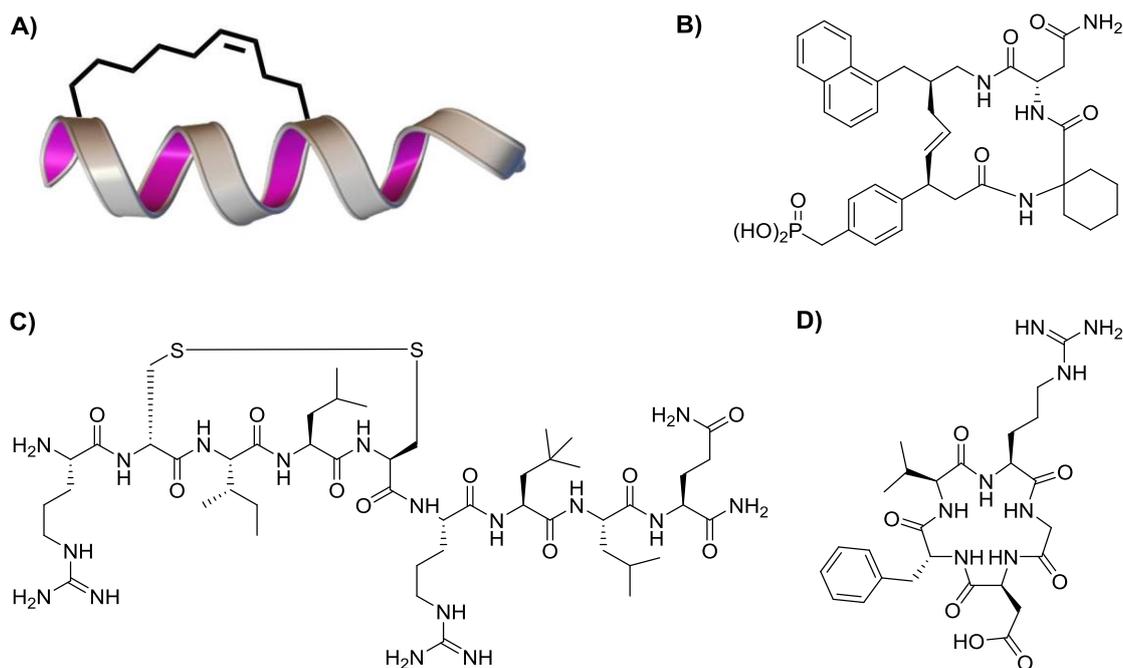


Figure 4.3: Synthetic cyclic peptides A) Stapled p53 derived α -helix B) Stapled SH2 domain binding peptide C) Disulfide cyclised oestrogen receptor binding peptide D) RGD derived head-to-tail cyclised peptide

Src homology 2 (SH2) domain binding peptides represent another area where small cyclic peptides have been designed based upon proteins with a specific recognition sequence. SH2 domains are found in a variety of regulatory and signalling proteins which recognise phosphotyrosine, with the specificity of each protein determined by the flanking residues.¹⁵⁶ Gao *et al.* have developed a stapled peptide using olefin metathesis for macrocyclisation in addition to a 1-aminocyclohexanecarboxylic acid residue which induces a turn in the peptide chain (Figure 4.3B).¹⁵⁷ Further work by the same group has generated analogues through side chain variation, resulting in a lead compound which shows picomolar affinity for the Grb2 protein and anti-mitogenic effects against breast cancer cells.¹⁵⁸

Disulfide bridges exist throughout natural proteins as a method of inducing structure.¹⁵⁹ This strategy has also been used to generate small cyclic peptides to target PPIs. The alpha and beta forms of oestrogen receptors and their cofactors are an attractive target for inhibition due to their role in skeletal and cardiovascular regulation as well as being implicated in certain forms of cancer.¹⁶⁰ Galande *et al.* have used the LXXLL recognition motif found in these receptors to synthesise a large library of cyclic helical peptides stabilised by disulfide bridges between *i* and *i*+3 residues (Figure 4.3C). By variation of side chains and disulfide position within the peptides, they were able to identify structure activity data for both affinity and selectivity between alpha and beta forms of the oestrogen receptor protein.

Head-to-tail cyclisation has also been employed in synthetic peptides. Integrins comprise a large family of cell surface adhesion and signalling receptors. The RGD based peptides aim to disrupt cell surface signalling by binding integrins which contain a highly conserved Arg-Gly-Asp (RGD) binding motif.¹⁶¹ Through an iterative design-synthesis-activity approach, a series of head-to-tail cyclised hexa- and pentapeptides were produced in order to probe the interaction.¹⁶² Of the peptides synthesised, Gurrath *et al.* were able to identify a c(RGD Δ FV) pentapeptide as a potent inhibitor of cell adhesion to both laminin and vitronectin (Figure 4.3D). Through analysis of X-ray crystal structures, the conformation of these peptides has been confirmed to be the same as for endogenous RGD motifs.¹⁶³

The cyclic peptide natural products and synthetic strategies highlighted above provide a selection of examples for methods of peptide macrocyclisation. When designing a novel cyclic peptide structure, however, the specific conditions for each transformation must be considered. For every strategy there are a variety of approaches, each with their own benefits dependent on application.

For synthetic peptides, two main categories for cyclisation exist, either solution phase or on-resin. There are two key restrictions for solution phase cyclisation. Dilutions must be kept high to prevent intermolecular reactions as cyclisations tend to be slow and coupling agents must be carefully selected to allow separation of side products from the cyclic peptide.¹⁶⁴ On resin cyclisation avoids these issues due to the pseudo dilution effect of peptide chain immobilisation and the ease of side product removal by washing the resin. Unlike solution phase reactions however, on resin cyclisations tend to be more difficult to monitor due to the qualitative nature of protocols such as the Kaiser test.¹²⁵ While resin tests exist for free thiols and amines, for cyclisation strategies not involving these moieties and for following removal of orthogonal protecting groups, cleavage of small portions of resin is often the only way to be certain about the progress of reactions.¹⁶⁵ As a result, choice of conditions for macrocyclisation largely depends on the specific application.

Formation of a disulfide bridge is one of the most straightforward methods of macrocyclisation, with numerous oxidation methods available.¹⁵⁹ Developed alongside these techniques are a range of orthogonal protecting groups that allow not only disulfide formation in otherwise fully protected peptides, but selective formation of multiple bridges for the construction of polycyclic peptides.¹⁶⁶ In its most basic form however, where only a single disulfide bridge is formed, the oxidation can be carried out in the presence of a number of other functional groups without

concern for side reactions. This makes the use of cysteine disulfides an attractive starting point for investigation of cyclic peptide sequences.

While slightly more complex than disulfide formation, lactam cyclisation offers a number of distinct advantages. The amide bond produced is more stable to reduction and in head-to-tail cyclisation, the peptide has no N-terminus for exopeptidases to attach to. Whether in solution or on-resin, in many cases orthogonal protection for functional groups not involved in cyclisation is required, which has resulted in a number of specialist protecting groups including allyl/alloxycarbonyl and DMAB.^{167,168} These are stable to the standard deprotection conditions used in chain elongation, but can be selectively removed prior to cyclisation. In situations where head-to-tail cyclisation is not desirable, lactam formation can be achieved via careful placement of amine and carboxylic acid bearing amino acids. Where this is not possible, for example due to side chain interactions, methods have been developed to allow backbone cyclisation via N-alkylation to introduce amines and carboxylic acids.¹⁶⁹

For strategies involving cyclisation by formation of a lactam, careful selection of coupling agents is necessary to avoid C-terminal racemisation and optimise coupling time. The highly active HOAt catalyst used in conjunction with carbodiimide condensating agents has become a popular solution for this cyclisation.¹⁷⁰ More recently, introduction of phosphonium coupling agents has further expanded the options available for slow couplings.¹⁷¹

Peptide stapling is a more recent development which allows covalent macrocyclisation through non-peptidic coupling methods. While this strategy has most often been applied to the stabilisation of α -helical peptides, the principle is applicable generally.¹⁷² This method uses a side chain hydrocarbon bridge by olefin metathesis between two alkene bearing side chains.¹⁷³ The linkers produced in this way are not only more stable than disulfide and amide bonds, but confer more cell permeability to the resulting peptide.¹⁷⁴ This is believed to be due to the more lipophilic nature of the hydrocarbon chain. However, this method does require specialist non-natural amino acids, making it one of the more costly techniques.

4.1 - Disulfide Cyclised Peptides

One reason for the TAT-14 peptide's increased potency in *in vitro* testing over the 10 and 16 amino acid sequences may be the increased hairpin stabilisation by backbone hydrogen bonding (see Chapter 2). If this is the case then a peptide with a constrained conformation, for example by cyclisation, would not require as long a chain to achieve the same affinity. Based on this assumption, a cyclic peptide consisting of just the key DEETGE motif should be capable of disrupting the Nrf2/Keap1 interaction. As formation of disulfide bridges is one of the most simple methods for forming cyclic peptides, this was chosen as a starting point.

4.1.1 - Synthesis of Disulfide Cyclised Peptides

In order to maintain the six key residues while making the peptide cyclic, it was necessary to add a Cys residue to either end of the sequence. This provided the basic eight amino acid sequence for the first disulfide cyclised peptides. Previous work by Hancock *et al.* has shown that N-terminal acetylation of short binding sequence peptides increases their binding affinity.⁹⁷ This is thought to be due to the positive charge of the N-terminus interfering with the cationic binding pockets of Keap1. In order to determine if N-terminal acetylation was necessary for these peptides, sequences were synthesised both with and without N-terminal modification. From a synthetic perspective, C-terminal cysteine presents a specific challenge. When attached to benzyl alcohol functionalised resins, a high degree of racemisation occurs. To prevent this, two options are available. First, the solid support can be changed for a trityl functionalised resin, which is resistant to racemisation of the first residue due to the bulky trityl group. The drawback with 2-chlorotrityl chloride resin is that it is moisture sensitive, requiring strictly anhydrous conditions during the resin loading step. The second option is to use Rink amide resin which has the benefit of allowing loading under standard coupling conditions, though the resulting peptide is C-terminally amidated. In order to determine whether either of these methods would affect binding, peptides using both resins were synthesised. With these N- and C-terminal modifications in mind, three peptides were synthesised to test the effect on binding (Ds8, H₂N-Ds8, Ds8-OH, Table 4.1).

Ds8	
H ₂ N-Ds8	
Ds8-OH	
Ln8	Ac-SDEETGES-NH ₂
Ds10	
Ln10	Ac-SLDEETGEFS-NH ₂
Ds6	
Ln6	Ac-SETGES-NH ₂

Table 4.1: Disulfide cyclised peptides and linear controls based on key Nrf2/Keap1 binding motif

The linear sequences for the two C-terminally amidated peptides, Ds8 and H₂N-Ds8, were synthesised on Rink Amide resin, using standard coupling procedures. Following chain elongation, the Ds8 peptide was N-terminally acetylated using acetyl chloride and DIPEA. The linear sequence for the Ds8-OH peptide was synthesised using preloaded 2-chlorotrityl resin and standard conditions for chain elongation. The peptide was then N-terminally acetylated using acetyl chloride and DIPEA. All three peptides were cleaved with a cocktail containing EDT to prevent reaction of cysteine with cleavage side products, and purified by reverse phase flash chromatography. Oxidation of the cysteine residues to form the cyclic peptides was achieved using iodine in acetic acid, as this method does not require as high dilutions as other methods. Monitoring by HPLC showed a slight shift in retention time and mass spectrometry showed a decrease in mass by two Da. Following completion of oxidation, iodine was removed by washing with carbon tetrachloride and the peptides purified by reverse phase flash chromatography. Purity was confirmed by RP-HPLC and identity by MALDI-TOF MS.

Following analysis of the first three peptides by (FP) (see Section 4.1.2), N-terminal acetylation and C-terminal amidation were selected as the optimum configuration for further analogues. Two new sets of peptides were designed to test the effects of loop size and the contribution of cyclisation on binding affinity. In order to test the influence of loop size, two further peptides

were synthesised. The first increased the peptide to 10 amino acids, adding one residue to either end of the key binding sequence (Ds 10, Table 4.1). With an increased size, the peptide should be able to adopt an optimal binding configuration more easily and in addition allow an increased number of backbone interactions. The second peptide decreased the size to six amino acids, though in order to achieve this, the amino acids were not removed symmetrically from each end. This maintained the presence of Glu82, which forms a side chain interaction with one of the cationic binding pockets of Keap1. This peptide features just the ETGE core, with the two residues involved in side chain interactions with Keap1 and the two residues allowing formation of the β -turn flanked by two Cys residues (Ds6, Table 4.1). Alongside these two peptides, a further three peptides were designed to be linear counterparts to the six, eight and ten residue cyclic peptides (Ln6, Ln8, Ln10, Table 4.1). By substituting cysteine for serine, similar physical properties were maintained, while removing the ability to cyclise. The peptides were synthesised in the same manner as Ds8, with the exception that the serine analogues were not subjected to the oxidation step. The compounds were purified by reverse phase chromatography and characterised by MALDI-TOF MS.

4.1.2 - Fluorescence Polarisation Inhibition

Following synthesis of the initial three peptides, Ds8, H₂N-Ds8 and Ds8-OH, their ability to disrupt the Nrf2/Keap1 interaction was assessed by FP. Based on previous binding experiments (see Chapter 3), fluorescent peptide concentration was fixed at 5 nM and the concentration of Keap1 at 30 nM, consistent with the inhibition assays using the CPP-Nrf2 peptides.

The Ds8 peptide, having both an acetylated N-terminus and amidated C-terminus was expected to bind well to Keap1. This was found to be the case, with a K_i of 95.0 nM (95% CI [73.2, 123.3]), indicating a strong interaction with the receptor (Figure 4.4). Interactions in the nanomolar range are consistent with work by Lo *et al.*⁴⁰ In comparison however, the Ds8 peptide is half the length of the 16 mer peptide for which Lo *et al.* determined a K_d value of 20 nM. Linear peptides of a similar length to the Ds8 peptide, analysed by Hancock *et al.*, were found to have IC_{50} values in the micromolar range.⁹⁷ This was a clear indication that cyclisation could improve affinity for Keap1 significantly. The Ds8 peptide formed a baseline to which the other peptides synthesised could be compared.

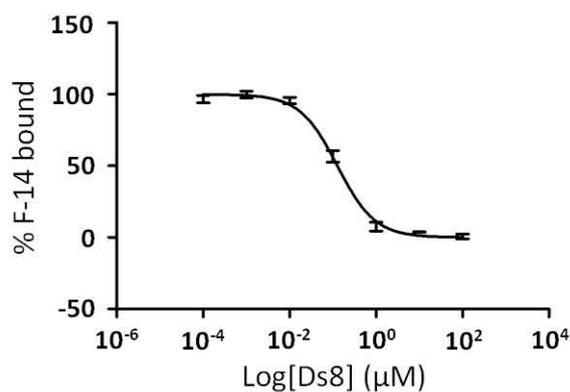


Figure 4.4: Fluorescence Polarisation inhibition by the Ds8 peptide, K_i 95.0 nM, 95% CI [73.2, 123.3], Mean \pm SEM, n = 3, 30 nM Keap1, 5 nM F-14

As expected, removal of the N-terminal acetyl group from the Ds8 peptide severely affected binding. The H₂N-Ds8 peptide had a K_i of 274.9 nM (95% CI [182.6, 413.8]), threefold weaker than the Ds8 peptide (Table 4.2). As noted previously this is likely due to the interference of the positively charged N-terminus with the cationic arginine binding pocket of Keap1. Despite this, binding is still in the nanomolar range, indicating the interaction is strong.

Peptide	K_i (nM)	95% CI (nM)	R ²
Ds8	95.0	73.2-123.3	0.9899
H ₂ N-Ds8	274.9	182.6-413.8	0.9770
Ds8-OH	47.7	35.5-64.1	0.9875

Table 4.2: Fluorescence Polarisation inhibition by disulfide cyclised peptides, n = 3, 30 nM Keap1, 5 nM F-14

Replacing the C-terminal amide with the free acid had a similarly potent result, increasing the binding affinity to 47.7 nM (95% CI [35.5, 64.1]) (Table 4.2). The cause of this improved binding is unclear, though may be due to an additional hydrogen bonding interaction between the C-terminus and the receptor.

From the initial screen, it was clear that protection of the N-terminus was necessary to increase binding efficiency. As a result all subsequent peptides were N-terminally acetylated. Masking of the C-terminus however, reduces affinity. This presented a problem in that the C-terminal acid peptides were significantly more difficult to synthesise in high purity and yield than the C-terminal amide analogues. In addition, the binding sequence peptides already bear four acidic residues,

which make them less likely to diffuse across cell membranes. The C-terminal acid results in a peptide with an extra negative charge in solution. Before examining these peptides in cells it was not possible to know how significant this would be, however reduction of charge was considered more important than binding affinity in this case. It was decided that as the Ds8 peptide showed strong binding and was easily accessed, further peptides should be based on C-terminal amides, despite the possible reduction in binding affinity.

This led to the production of two peptides with differing sizes of cycle and three as linear counterparts (Table 4.1). The linear peptides were expected to show weaker binding to Keap1 as they lacked the constrained conformation of the cyclic peptides. This is evident from the Ln8 peptide binding data, with a K_i of 1271 nM (95% CI [1046, 1545]), binding is an order of magnitude weaker than the cyclic Ds8 peptide (Table 4.3), similar to a 7 mer peptide synthesised by Hancock *et al.*⁹⁷

Peptide	K_i (nM)	95% CI (nM)	R^2
Ln8	$1.3 * 10^3$	$1.1 * 10^3$ - $1.5 * 10^3$	0.9942
Ds10	47.9	33.1-69.2	0.9807
Ln10	176.9	135.6-230.6	0.9900
Ds6	$151.5 * 10^3$	$88.8 * 10^3$ - $258.7 * 10^3$	0.9522
Ln6	$280.4 * 10^3$	$116.2 * 10^3$ - $676.2 * 10^3$	0.9246

Table 4.3: Fluorescence Polarisation inhibition by disulfide cyclised peptides, n = 3, 30 nM Keap1, 5 nM F-14

In comparison, increasing the length of the peptide to 10 amino acids, improves the binding dramatically. While still not as potent as the Ds8 peptide, the Ln10 peptide was found to have a K_i of 176.9 nM (95% CI [135.6, 230.6]) which is stronger than the non-acetylated H₂N-Ds8 peptide (Table 4.3).

This trend in increased binding with peptide length was continued with the Ds10 cyclic peptide. Comprising 10 amino acids, the peptide's K_i was found to be 48.0 nM (95% CI [33.2, 69.2]) which is half that of the Ds8 peptide (Table 4.3). Encouragingly, this is as potent as the C-terminal acid

Ds8-OH peptide, the most potent cyclic peptide so far. Again the increase in binding afforded by cyclisation is clear in comparison to the Ln10 peptide.

Moving in the opposite direction, by decreasing the number of amino acids to six, it was hoped that the minimum binding sequence could be identified. As expected the affinity of the Ds6 peptide (Table 4.3) is weaker (K_i of $151.5 * 10^3$ nM (95% CI [$88.8 * 10^3$, $258.7 * 10^3$])), however the extent to which binding was decreased was surprising. While the weak binding made determining an accurate K_i difficult, as indicated by the range of the 95% CI, the fact that binding is three orders of magnitude weaker than the Ds8 peptide is a strong indication that decreasing beyond the DEETGE sequence is not viable.

As would be expected from the previous data, displacement of the F-14 peptide by the Ln6 peptide required very high concentrations. As with the Ds6 peptide, the $280.4 * 10^3$ nM K_i has a very wide confidence interval (95% CI [$116.2 * 10^3$, $676.2 * 10^3$]), as a result of the high concentrations required for displacement, though it appears that binding is decreased compared to the cyclic peptide (Table 4.3). This would follow the pattern seen with the other peptides in the series.

From the fluorescence polarisation data it was clear that the degree to which peptides bind to Keap1, is strongly dependent on both length and conformation. However, as the FP assay uses purified protein and a fluorescently tagged peptide, rather than intact cells, the data gives no indication as to how easily the peptides cross cell membranes.

4.1.3 - HO-1 Protein Induction

In order to assess the ability of the peptides to disrupt the Nrf2/Keap1 interaction in cells, their effect on HO-1 protein induction was measured by ELISA. As the linear peptides and non-acetylated N-terminus peptide showed very weak binding in the FP assay, they were not included in cell testing. Despite the very weak binding of the Ds6 peptide, it was included in case its small size and removal of two negatively charged side chains allowed cell entry, resulting in weak induction. Each of the peptides was tested at a concentration of 10 mM in THP-1 cells between 9 and 16 hours.

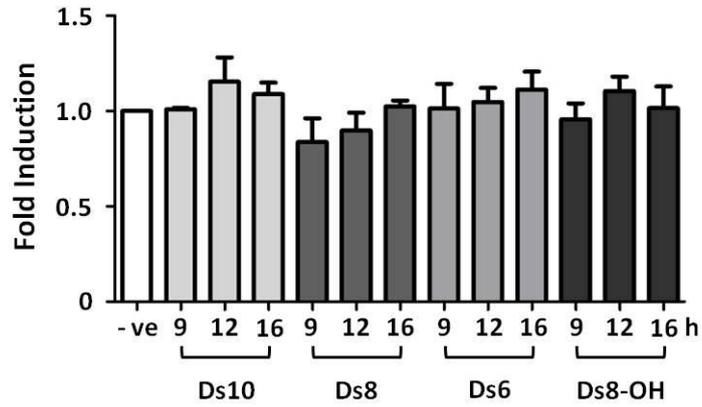


Figure 4.5: Induction of HO-1 protein by disulfide cyclised peptides, Mean ± SEM, n = 3

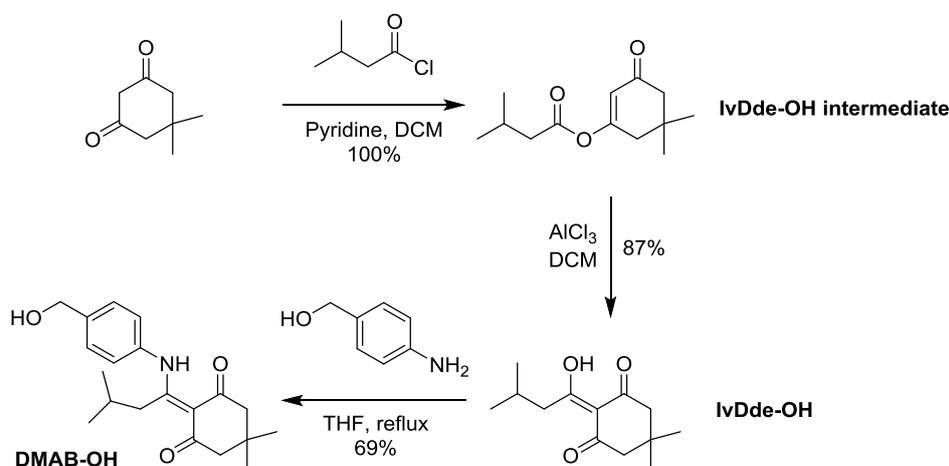
As can be seen from Figure 4.5, none of the peptides increased HO-1 protein levels at the time points tested. This may be due to their inability to enter cells, which would not be unreasonable, as despite being cyclic and having uncharged termini, they bear several negatively charged side chains.

4.2 - Head-to-tail Cyclised Peptides

As an alternate cyclisation strategy to the disulfide peptides, head-to-tail cyclisation by lactam formation was explored. By attaching an amino acid to the resin by its side chain, it was hoped that both chain elongation and cyclisation could be achieved on the resin. This would reduce the number of purification steps and remove the need for high dilutions during cyclisation.

4.2.1 - Synthesis of Fmoc-Glu-ODMAB

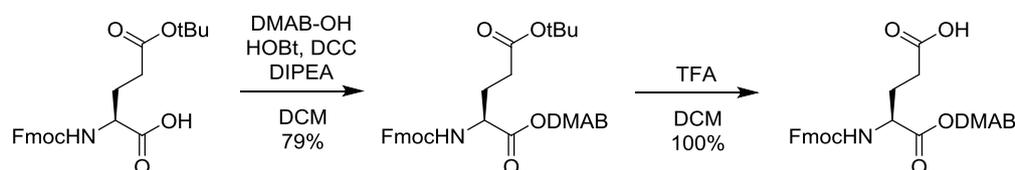
Before synthesising the linear precursor to a head-to-tail cyclised peptide, it was necessary to produce an orthogonally protected amino acid which could be selectively deprotected to allow on-resin cyclisation. Because the key Nrf2 binding motif contains multiple acidic residues it was decided that production of a C-terminally DMAB protected amino acid would provide the largest number of options for the peptide synthesis.¹⁶⁸ While commercially available, the DMAB protecting group can be prepared on a large scale from inexpensive reagents in just three steps (Scheme 4.1).^{175,176} As production of the DMAB protecting group has been reported, each synthetic step was confirmed by comparison to published NMR data. The final compound, Fmoc-Glu-ODMAB, was characterised by NMR and mass spectrometry.



Scheme 4.1: Synthesis of DMAB protecting group

The first step in the synthesis was production of isovaleryl dimedone (IvDde) via an ester intermediate. This was obtained by reaction of dimedone with 1.2 equivalents isovaleryl chloride in the presence of pyridine. Reaction for 18 hours is necessary to ensure quantitative conversion. Aqueous workup was sufficient to produce the pure product as a yellow oil. Conversion of the ester into IvDde via a Fries type rearrangement, was achieved by reacting the material with aluminium chloride at room temperature. At a 10 mmol scale, rearrangement required only 1.2 equivalents aluminium chloride. However when scaled up to 50 mmol 2.2 equivalents aluminium

chloride are necessary, along with the use of anhydrous solvent. It was suspected that both the starting material and product were coordinating to aluminium, resulting in the need for an excess of reagent. Under these conditions however, the reaction proceeded in 2 hours and required only filtration through silica to remove baseline impurities to afford the product in an 87% yield as a yellow oil.



Scheme 4.2: Synthesis of Fmoc-Glu-ODMAB

Reaction of p-amino benzylalcohol with IvDde to produce DMAB by refluxing in THF was relatively slow, requiring 60 hours to complete. However, after purification by flash chromatography the product was isolated as a yellow solid in 69% yield. DMAB was then used to C-terminally protect Fmoc-Glu(^tBu)-OH using standard condensation conditions (Scheme 4.2). Flash chromatography afforded the product as a white solid in 79% yield. Side chain deprotection of Fmoc-Glu(^tBu)-ODMAB was achieved by treatment with 50% TFA in DCM for 1 hour. Removal of solvent under reduced pressure and coevaporation with methanol to remove residual t-butanol produced Fmoc-Glu-ODMAB in quantitative yield as a yellow foam.

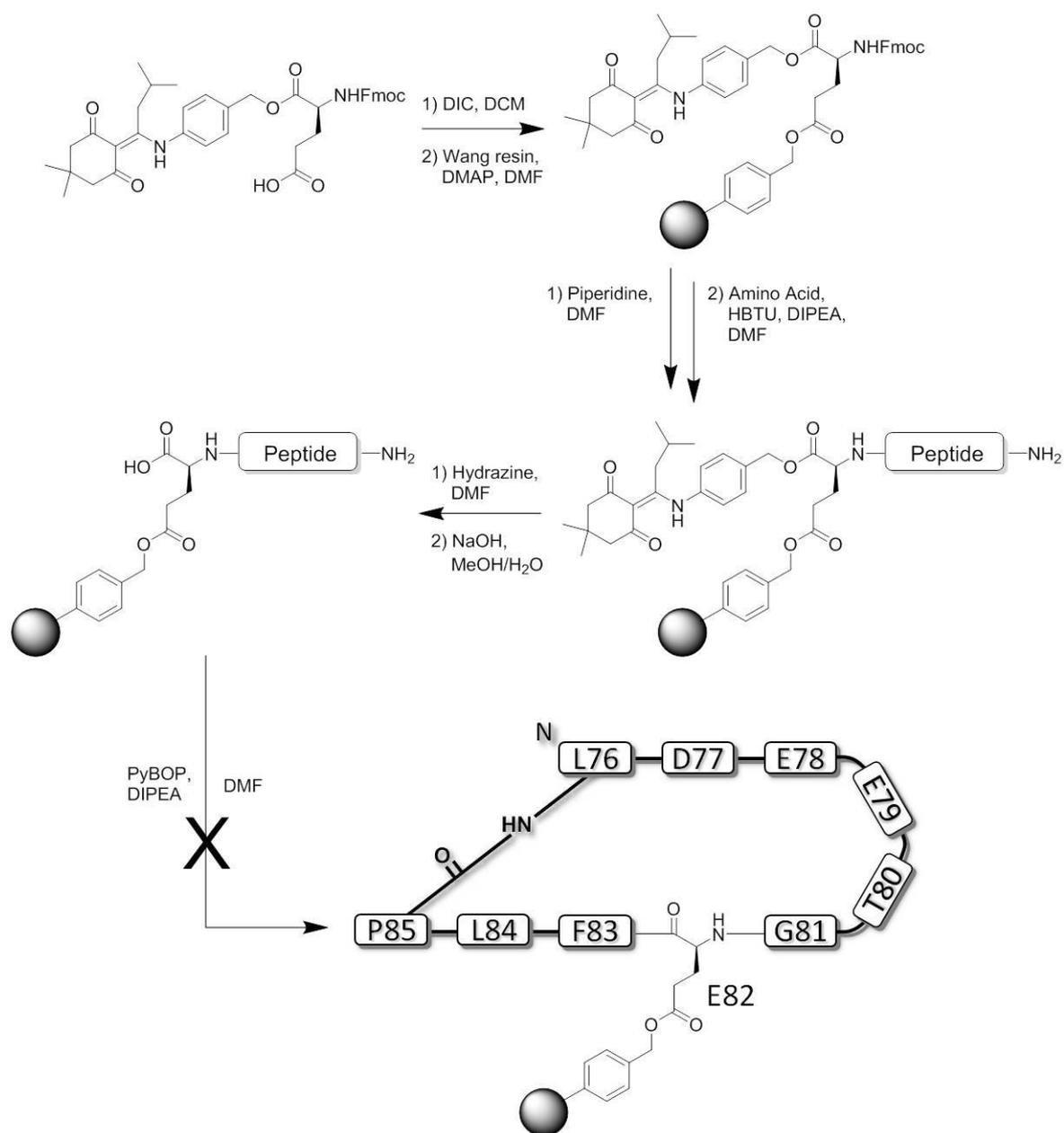
4.2.2 - Synthesis of Head-to-tail Cyclised Peptides

Initial 10 mer	$\overbrace{\text{LDEETGEFLP}}$
<i>Linear seq.</i>	<i>FLPLDEETGE</i>
Revised 10 mer (Ht10)	$\overbrace{\text{QLDEETGEFL}}$
<i>Linear seq.</i>	<i>LDEETGEFLQ</i>
9 mer (Ht9)	$\overbrace{\text{QLDEETGEF}}$
<i>Linear seq.</i>	<i>LDEETGEFQ</i>

Table 4.4: Peptide sequences for head-to-tail cyclised Nrf2 peptides, linear sequence as synthesised in italics

Unlike the disulfide bridged peptides, a head-to-tail cyclic peptide must incorporate a planar amide bond into the cyclic turn. In order for this not to impact the binding of the peptide, it was decided that a larger cycle would be a good starting point. Based on the shortest of the three

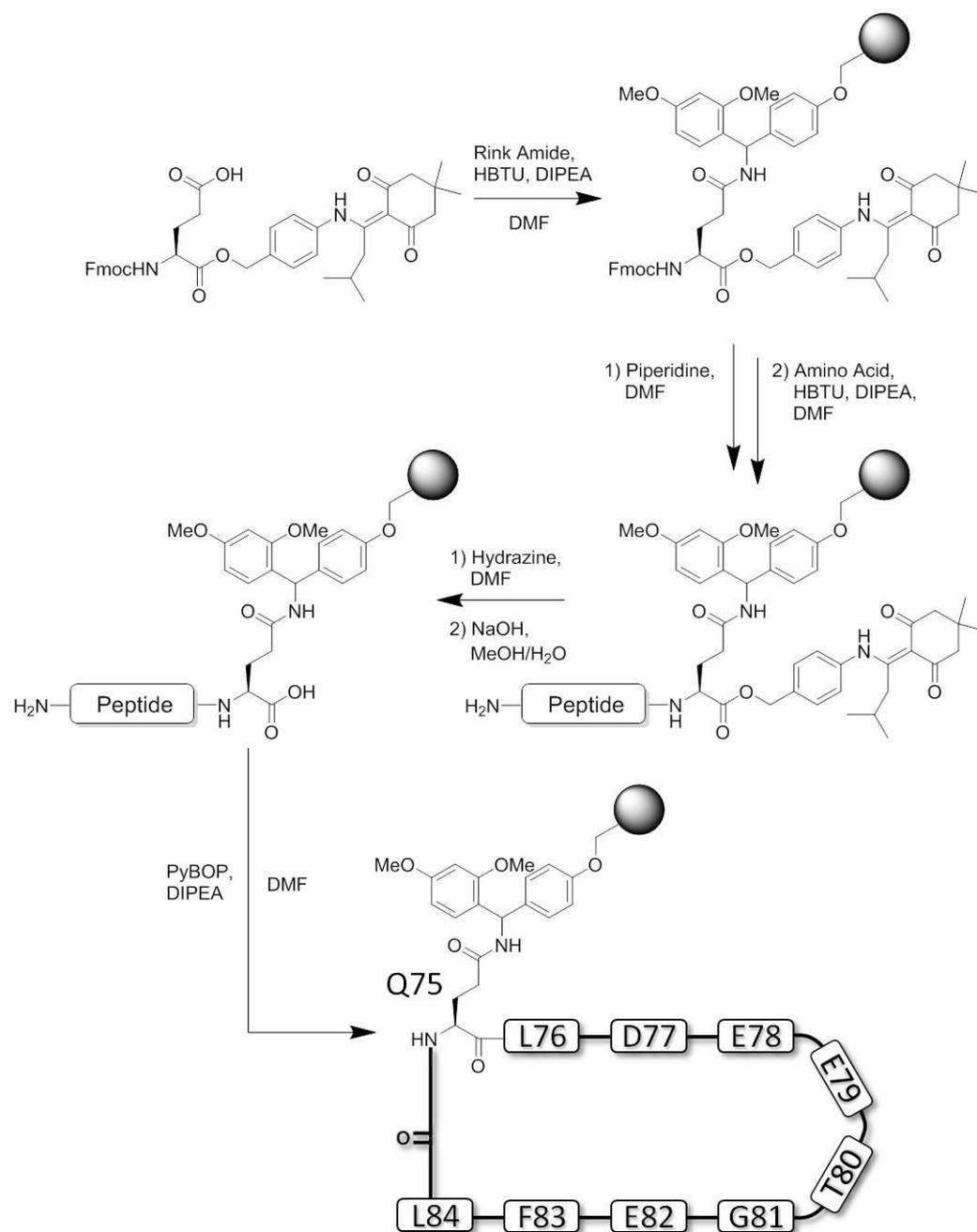
peptides identified as binders by Lo *et al.*, a linear sequence was designed that when cyclised would produce a cyclic version of the 10 mer peptide (Table 4.4). There were three possible options for placement of the C-terminus of the linear sequence, determined by the placement of Glu residues in the peptide. One option was to start at Glu79, which would break the sequence in the middle of the β -turn section. This was discounted as it was hoped the natural turn of the peptide would aid cyclisation. The remaining two residues are located at either end of the turn section, Glu82 was selected as the first residue, though Glu78 would have been equally suitable (Scheme 4.3).



Scheme 4.3: Initial synthesis of a head-to-tail cyclised peptide, resin loading, chain elongation and attempted cyclisation

Fmoc-Glu-ODMAB was immobilised on the resin via its side chain using symmetrical anhydride loading. The chain was elongated under normal conditions and following completion a small portion of resin was cleaved for analysis. Both HPLC and mass analysis showed the presence of the linear chain as the majority product, with DMAB group intact. Following DMAB removal with 2% hydrazine, the process was repeated and again the linear peptide, this time without DMAB, was observed. On-resin cyclisation was attempted using 1.5 equivalents DIC and 1 equivalent HOBt in DMF, following a literature procedure.¹⁷⁷ Coupling was carried out for 48 hours, with refreshed coupling agents after 24 hours. No significant reaction was observed by Kaiser test in this time. As a result, coupling conditions were altered. PyBOP was employed instead of DIC, equivalents were increased to 4 for both PyBOP and HOBt and 8 equivalents DIPEA were added. Following a further 48 hours, some reduction in colour was observed by Kaiser test and the peptide was cleaved from the resin to assess the extent of cyclisation. HPLC analysis of the crude product showed multiple small peaks, while analysis by MALDI showed multiple products including linear peptide but no presence of the desired product.

Due to the difficulties in the cyclisation step, an alternate peptide sequence was devised, which it was hoped would be more conducive to cyclisation (Table 4.4). In the new sequence, the β -turn was positioned in the middle of the linear chain. With a natural turning point in the peptide, it was hoped that the two ends of the chain would be in closer proximity for the cyclisation reaction. In order to place the β -turn in the centre of the sequence, none of the previously suggested Glu residues could be used for resin attachment. Aside from the high concentration of Glu in the binding motif, there are no other Glu residues in the natural peptides. By shifting the frame of the 10 mer however, from residues 76-85 to 75-84, a Gln residue is added to the N-terminus. As the final peptide has no N- or C-terminus, the chain can be broken at any point, making it possible to place the Gln residue at the C-terminus instead of the N-terminus. By synthesising this peptide using Fmoc-Glu-ODMAB at the C-terminus as before, but immobilising on Rink Amide resin instead of Wang resin, following cleavage the side chain is converted to an amide. Using this method, the β -turn motif is in the centre of the sequence and the same Fmoc-Glu-ODMAB prepared previously could be used (Scheme 4.4).

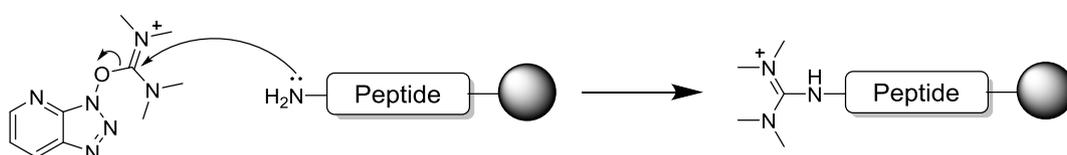


Scheme 4.4: Synthesis of the Ht10 peptide, resin loading, chain elongation and on-resin cyclisation

The linear sequence was constructed on a 0.195 mmol scale, attaching all the amino acids using standard HBTU mediated couplings. Following synthesis the resin was split into three batches to allow several attempts at cyclisation. As before, the linear peptide was the majority product both before and after DMAB removal. The initial cyclisation was based upon the protocol laid out by Meller *et al.* using DIC as the condensating agent as before, using 1.5 equivalents DIC and 1 equivalent HOBt in a minimum volume of DMF.¹⁷⁷ The reaction was monitored by Kaiser test and the reagents refreshed every 24 hours. After reacting for 4 days, the Kaiser test results had become weak (pale blue), though the reaction was still not complete. The reaction was continued

using PyBOP as a stronger coupling agent and anhydrous conditions as PyBOP is known to be less stable to hydrolysis than other coupling agents. Coupling was complete overnight using the new conditions and the peptide was cleaved from the resin. HPLC analysis showed consumption of the linear peptide and new peak corresponding to 2 M+Na by MALDI TOF. The presence of this peak was unlikely to be due to a physical head-to-tail dimerisation of two peptide chains, as the pseudo dilution effects of on-resin cyclisation do not allow chains to come into close proximity. In addition to this new peak, there were a large number of small impurities, which it was hoped could be eliminated by optimising the coupling conditions.

As use of a more potent coupling agent had driven the reaction to completion, coupling agents were considered based on their relative efficiencies. Of the reagents commonly used in SPPS, HATU is known to be one of the most effective. Its cost is prohibitive for general use in synthesis, however it has been found to be particularly effective in difficult couplings.¹⁷⁸ A second portion of resin was deprotected as before and swollen in dry DMF. The peptide was reacted overnight with 4 equivalents HATU and HOBt and 8 equivalents DIPEA in dry DMF. Encouragingly, the Kaiser test was negative and the peptide was cleaved from the resin. While HPLC and mass analysis showed consumption of the starting material, the major product had a mass corresponding to M+99 which is indicative of guanidinium capping by HATU (Scheme 4.5).¹⁷⁹ As a result, the final portion of resin was cyclised using PyBOP in the place of HATU, which resulted in fewer side products than the initial attempt, though still requiring three days to cyclise. The major peak seen by MALDI was still 2 M+Na, however. Further analysis of the mass spectrum, showed peaks which could be attributed to either adducts of the desired peptide or double ions of the dimer adducts. These species would both have identical mass spectra, making determination of the product difficult using this method. This was resolved using negative mode ESI, which showed only the M-H ion for the desired cyclic product. Following purification by reverse phase flash chromatography, only 1 mg of pure peptide was isolated. This raises questions as to the fate of the remaining peptide.



Scheme 4.5: Mechanism of N-terminal guanidinium capping by HATU

To see if rate of cyclisation and overall yield was dependent on chain length, a further peptide sequence was devised to test the effect of chain length. A nine amino acid peptide corresponding to residues 75-83 was synthesised as before (Table 4.4). This used the same residues as the 10 amino acid sequence, but without the C-terminal Leu residue. Despite hopes that this sequence

would cyclise more easily, under the same conditions as before, coupling was complete in three days and produced a similarly poor yield of pure peptide. Interestingly, as with the larger peptide, the major peak by MALDI was that of 2 M+Na, though as before analysis by ESI showed the desired product.

4.2.3 - Fluorescence Polarisation Inhibition

As with the disulfide cyclised peptides, initial assessment of the Keap1 binding affinity of the backbone cyclised peptides was determined by inhibition of fluorescence polarisation. The peptides were tested from 100 μ M to 0.01 nM in 10 fold dilutions, with a Keap1 concentration of 30 nM and 5 nM fluorescent peptide. The Ht10 peptide was found to have similar affinity to the disulfide cyclised peptides with a K_i of 107.2 nM (95% CI [82.5, 135.0]) (Table 4.5). This places it close to the Ds8 peptide in terms of potency.

Peptide	K_i (nM)	95% CI (nM)	R^2
Ht10	107.2	85.2-135.0	0.9915
Ht9	66.3	49.0-89.6	0.9853

Table 4.5: Fluorescence Polarisation inhibition by the head-to-tail cyclised peptides, n = 3, 30 nM Keap1, 5 nM F-14
FP inhibition by the Ht9 peptide shows a similar potency to the Ht10 peptide, however it is increased to 66.3 nM (95% CI [50.0, 89.6]) (Table 4.5), making it more potent than the Ds8 peptide and approaching the affinity of the Ds10 peptide.

Contrary to the pattern seen with the disulfide cyclised peptides, where increased cycle size increased potency, in this case the smaller Ht9 peptide had a higher affinity than the Ht10 peptide. The cause of this is unclear and synthesis of further peptides may help elucidate the finding, however due to the difficulties experienced with the synthesis, it was not deemed practical.

4.2.4 - HO-1 Protein Induction

As with the disulfide peptides, affinity for Keap1 assessed by the FP assay is not necessarily transferable to potency in cell based assays. To assess whether, unlike the disulfide peptides, the backbone cyclised peptides were able enter cells, HO-1 protein levels were measured by ELISA.

Prior to measuring protein levels, cytotoxicity was assessed by MTS assay, no toxicity was found for either peptide up to 300 μ M.

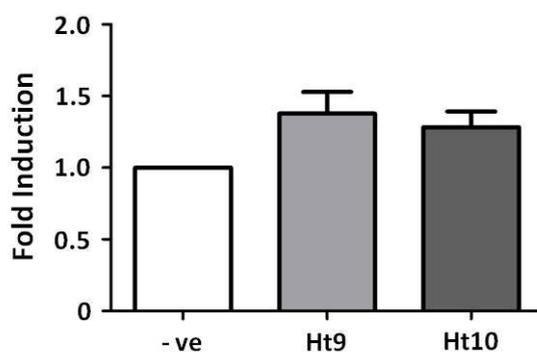


Figure 4.6: Induction of HO-1 protein by head to tail cyclised peptides, Mean \pm SEM, n = 3, 24 h, 100 μ M peptide

Due to limited quantities of peptide, only a single time point could be tested for activity. As 24 hours had shown the highest induction in tests with the TAT-14 peptide (See Chapter 2), this was the time point chosen. As with the disulfide cyclised peptides, there was no induction of HO-1 protein seen. This indicates that the peptides are unable to cross the cell membrane which considering the similarity in overall charge to the disulfide peptides was not surprising. Both the forms of cyclisation explored to this point indicate that binding affinity is increased by constraining the peptide's conformation.

4.3 - Aryl Stapled Peptides

While both the head-to-tail cyclised and disulfide bridged peptides showed promising affinity for Keap1 in the FP assay, assessment of their activity by ELISA clearly shows this is not translated into increased Nrf2 levels in whole cells. The cause of their inactivity is likely due to the multiple negatively charged side chains, preventing penetration of the cell membrane. One option to overcome this would be to attach cell penetrating sequences to the cyclic peptides, generating cyclic-tail peptides.¹⁸⁰ However, as one of the aims in generating cyclic peptides was to remove the requirement for large CPPs, it was not a desirable solution.

Having explored head-to-tail and disulfide cyclisation, the increased membrane permeability of a stapled peptide was an attractive target. While traditional metathesis mediated stapling techniques require expensive non-natural amino acids inserted into the sequence, a recent report has highlighted a method of peptide stapling which is directly applicable to the cysteine containing sequences already synthesised. Spokoyny *et al.* identified a stapling method using nucleophilic aromatic substitution (S_NAr) of perfluorinated aromatics by cysteine thiols (Scheme 4.6).¹⁸¹



Scheme 4.6: Perfluoroarylation of thiols via nucleophilic aromatic substitution

In addition to the typical improvements in binding affinity, cell permeability and proteolytic stability, they found the stapled peptides to be more chemically stable than disulfide analogs. It was also observed that even when hexafluorobenzene was used in excess, the only product was the result of 1,4 disubstitution. This is due to the steric hindrance of the 2,5-(*ortho*) C-F positions and *para* activation by the thioether following the first substitution. Stabilisation of the intermediate in the second substitution by the *para* sulfur increases the rate of the reaction, favouring disubstitution over multiple arylation. It was hoped that, as seen in olefin stapled peptides, the lipophilicity of the perfluoroaromatic linker would allow cell permeability of the peptides.

4.3.1 - Synthesis of Stapled Peptides

As this stapling strategy was amenable to the cysteine containing peptides used previously, the linear sequences were synthesised as before for the 10, 8 and 6 amino acid peptides. In addition,

two further peptides were synthesised based upon an observation by Hancock *et al.* that incorporation of a proline residue in place of Glu78 improved binding in a similar FP assay.⁹⁷ The side chain of Glu78 plays no specific role in the binding interaction of Nrf2 to Keap1 and proline is known to help induce formation of beta turns due to the reduced preference for a trans conformation. All five peptides were synthesised as described previously (Table 4.6). Following purification, the peptides were stapled using 25 equivalents 50 mM hexafluorobenzene in DMF with 25 mM TRIS base. The large excess of hexafluorobenzene, ensured the reactions were driven to completion and due to the *para* activation of the thioether following the first substitution, no disubstituted product was detected. Following reaction at room temperature for 18 hours, the solvent and excess hexafluorobenzene was removed under reduced pressure. Purification by RP-HPLC yielded four of the five expected products, which were characterised by MALDI TOF mass spec. Purity was confirmed by analytical HPLC. Interestingly, the 10 amino acid peptide produced only oxidised starting material while the proline substituted analogue reacted without issue. As a result, only four of the peptides designed could be taken on to *in vitro* testing.

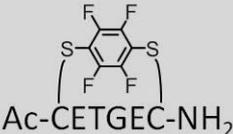
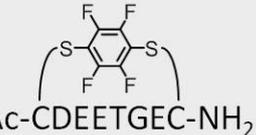
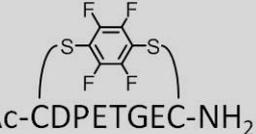
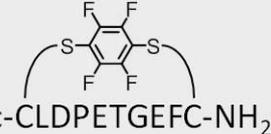
Ar6	 Ac-CETGEC-NH ₂
Ar8	 Ac-CDEETGEC-NH ₂
Ar8P	 Ac-CDPETGEC-NH ₂
Ar10P	 Ac-CLDPETGEFC-NH ₂

Table 4.6: Perfluoroaryl stapled peptide sequences

4.3.2 - Fluorescence Polarisation Inhibition

As for the other cyclic peptides, following purification the stapled peptides were assessed for Keap1 affinity by FP. As seen with the equivalent disulfide peptides, the Ar6 peptide binds very

weakly to Keap1. Due to the quantities of peptide available, it was not possible to increase the concentration of peptide in the assay in order to determine a K_i value.

Peptide	K_i (nM)	95% CI (nM)	R^2
Ar6	-	-	0.9077
Ar8	$1.4 * 10^3$	$1.1 * 10^3$ - $1.7 * 10^3$	0.9914
Ar8P	6.1	4.4-8.5	0.9827
Ar10P	61.9	46.1-83.0	0.9861

Table 4.7: Fluorescence Polarisation inhibition by the aryl stapled peptides, n = 3, 30 nM Keap1, 5 nM F-14

In comparison to the Ar6 peptide the binding of the Ar8 peptide is greatly improved, however when comparing to the equivalent Ds8 peptide, the reduction in affinity is dramatic. The K_i of 1355 nM (95% CI [1083, 1694]) is an order of magnitude greater than the disulfide. This indicates that incorporation of the fluorobenzene staple severely alters the solution conformation of the peptide, making binding to Keap1 far less favourable. As a result, the activity in cellular assays was expected to be similarly affected.

In stark contrast to the Ar8 peptide, substitution of Glu78 by proline in the Ar8P peptide not only restored binding affinity but improved it compared to all the previously tested cyclic peptides. With a K_i of 6.1 nM (95% CI [4.4, 8.5]) the Ar8P peptide has comparable affinity for Keap1 to the 16 mer peptide tested by Lo *et al.* and to the full length Nrf2 protein.^{40,42,95} This was a truly encouraging result as if the fluoroaryl moiety enabled membrane permeability, the peptide could prove to be very potent.

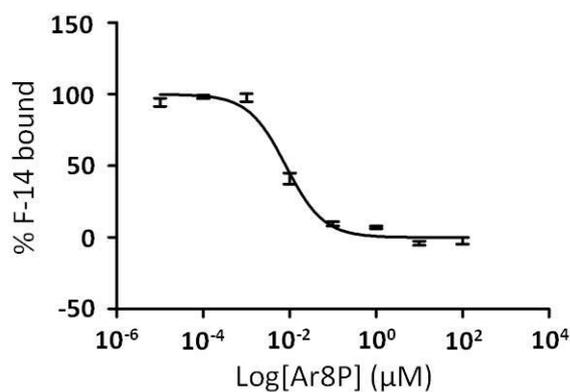


Figure 4.7: Fluorescence Polarisation inhibition by the Ar8P peptide, K_i 6.1 nM, 95% CI [4.4, 8.5], Mean \pm SEM, n = 3, 30 nM Keap1, 5 nM F-14

The Ar10P peptide does not bind as strongly as the Ar8P peptide, though still has a high affinity for Keap1, with a K_i of 61.9 nM (95% CI [46.1, 83.0]). The reasons for this are unclear, though it may be an indication that orientation of the side chains of the two cysteine residues has an impact on the position of the fluoroaryl group upon binding to Keap1. It is conceivable that in the Ar8 peptide the aromatic ring clashes with the binding pocket, while the inclusion of the proline positions the aryl away from the pocket. Upon elongation to the 10 amino acid sequence, the cysteine side chains will face in the opposite direction, once again placing the aryl ring in a less favourable position.

4.3.3 - HO-1 Protein Induction

The lipophilicity of the stapled peptides is increased in comparison to the disulfide bridged peptides, due to the fluoroaryl linker. It was anticipated that this would improve cell membrane permeability and result in HO-1 protein induction. Before assessing the effect of the peptides on HO-1 by ELISA, each of the peptides was tested for cytotoxicity by MTS assay. None of which were found to have any toxicity up to 300 μ M.

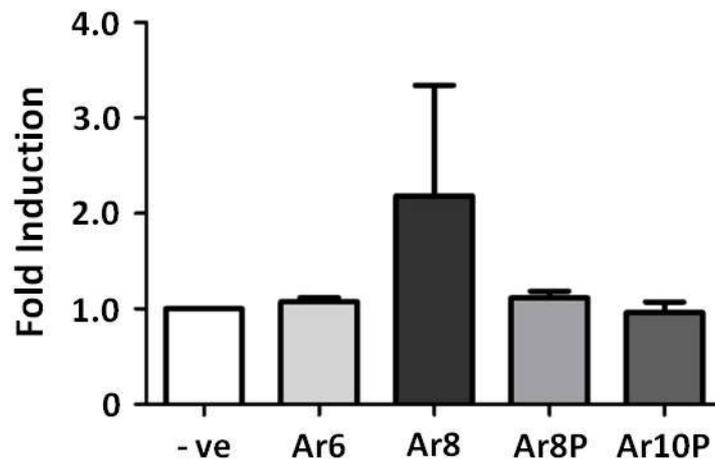


Figure 4.8: Induction of HO-1 protein by aryl stapled peptides, Mean ± SEM, n = 3, 24 h, 100 μM peptide

As can be seen from Figure 4.8, at 24 h there is no induction from the Ar6, Ar8P or Ar10P peptides. Due to a high variance between replicates, induction by the Ar8 peptide is not significant. The cause of this is unclear and time constraints prevented any further investigation. The lack of induction seen may be due to incorrect selection of time point, in which case a time course should be conducted. It may also be beneficial to use qPCR to assess these peptides as an additional screen. In spite of this, the stapled peptides and in particular the Ar8P peptide, still represent highly potent inhibitors of Keap1.

4.4 - Conclusions

Development of cyclic peptide sequences has produced inhibitors of Keap1 with significantly increased binding over their linear counterparts. The reduced size of these peptides is also important to progress towards more drug like molecules, the molecular weight of the cyclic peptides being around one third of the TAT-14 peptide. The Ds8 peptide manages an affinity of 95 nM while being below 1000 Da in weight and the Ar8P peptide, although larger at 1067 Da, has an affinity for Keap1 of just 6.1 nM.

The general pattern of binding affinity seen with the cyclic peptides indicates longer peptides bind more strongly. However, the cyclic 8 and 10 mer peptides in general show very good binding. It is only when reducing the peptides to the ETGE sequence that binding becomes weak. From the data presented here, it would appear that the DEETGE sequence, flanked by the two residues required for side chain cyclisation, is the minimum requirement for effective binding to Keap1. In that reducing the size of the peptides was one aim in producing cyclic sequences, the 8 mer peptides probably represent the best compromise of size and potency.

The head-to-tail cyclised peptides present an interesting option for stabilisation of the peptide β -turn and preventing degradation by exopeptidases. Unlike the other cyclic peptides, the head-to-tail sequences were constructed entirely on the solid phase, reducing the number of purification steps. Unfortunately the difficulty in cyclisation and resulting low yields prevented further exploration of these compounds' properties.

The aryl stapled peptides are the most promising of the cyclic peptides produced so far. While cellular assays have proved inconclusive there is reason to be optimistic. Only a single time point was assessed for HO-1 activity and while this was optimal for the CPP-Nrf2 peptides, there is no guarantee that induction by the cyclic peptides would follow the same time course. As a result it may simply be that the induction was missed. The other properties afforded by the aryl staple are also beneficial in comparison to the disulfide bridged peptides. As the fluoroaryl group bridges the two thiols, the peptides are not vulnerable to reduction, which, alongside the increased lipophilicity, should increase their bioavailability. The substitution of Glu78 with proline appears to have widely differing effects depending on the length of peptide. In the case of the Ar8P peptide, the substitution, in combination with the other benefits of aryl stapling, has resulted in a very promising lead for further development.

Production of high affinity inhibitors of a PPI can result from several differing approaches. Development of cyclic Nrf2 based peptides from the linear CPP-Nrf2 sequences tested previously

was only one possible avenue of research. Alongside the cyclic peptides, which were a rational drug design approach to developing inhibitors, a library screening approach to lead identification was employed. Using high throughput *in silico* screening, a large chemical space was explored with fewer time and financial costs than a conventional *in vitro* screen. These could then be validated using the established FP assay for rapid hit confirmation. This approach is described in the following chapter.

Chapter 5: In Silico Screening

5.0 - Introduction

5.0.1 - In Silico Screening

There are three common approaches for targeting PPIs. The first is through interface derived peptides, as shown in Chapters 2 and 4. The second is by high throughput *in vitro* screening, one of the objectives for development of the FP assay described in Chapter 3. The third exploits the advances made in both computing and elucidation of protein structures, through virtual compound screening. While *in vitro* screening still dominates the drug discovery process, virtual screening is increasing in popularity.¹⁸² It is based on a key assumption, that the function of molecules whether large proteins or small compounds, comes from their chemical and physical structure. If enough information is known about either a target or its inducers, this can be used to predict new drug leads.¹⁸³

Two methods of lead identification by computational methods are commonly used. In cases where the target may be unknown, or at least the structure is unsolved, known inducers can be used to develop a pharmacophore model.¹⁸⁴ The model identifies key features of known compounds, such as solvent accessible surface area and lipophilicity. Virtual libraries can then be screened for leads which fit these parameters and the resulting hits validated by *in vitro* assays.

If the target is known however, and the structure elucidated, either through crystallography or NMR, docking experiments are generally preferred.^{185,186} In this approach, conformational sampling of each compound in a virtual library is used to generate a series of poses for receptor-ligand interactions. Each of these complexes is then ranked using a scoring function, in order to produce hits which can be validated *in vitro*. Docking studies have been used to identify leads targeting several PPIs including MDM2 and Bcl family proteins.^{187,188}

There are two major challenges still to be solved in the field of docking calculations. The first is receptor flexibility, the classical lock and key model of ligand binding and even the more modern idea of induced fit do not take into account the dynamic nature of proteins.¹⁸⁹ Ideally every bond in both the ligand and receptor would be allowed to rotate freely in order to sample all possible conformations. This is not possible in practice however, as the number of conformations scales exponentially with the number of rotatable bonds.¹⁹⁰

The other challenge facing accurate computational techniques is assessment of binding affinity. This problem is intrinsic to the identification of hits, as once conformations have been generated,

they must be scored in a way that accurately reflects their *in vitro* potency. The way in which solvent is treated by computational models has a large effect on this. Generally in biological systems, the solvation energy of both ligand and receptor is high, which disfavours binding.¹⁹¹ This is countered by high affinity between receptor and ligand, which favours binding. The result of these two competing factors is a very small net binding affinity, which makes it difficult to calculate.

Both these problems can be tackled in a number of ways. Various algorithms exist which attempt to take the flexibility of the receptor into account, though at the expense of speed.¹⁹² A more basic approach exploits available information about binding interactions, by using crystal or NMR structures of ligand receptor complexes. It is assumed that in order to bind, the receptor must adopt a favourable conformation. If this is the case, by removing the ligand, a suitable receptor for docking studies is produced.¹⁹³ This does however prejudice results towards compounds similar to the original ligand. It has been suggested that this can be mitigated somewhat by comparison of multiple structures.¹⁹⁴ While this requires more time than assessing a single structure, it can still be faster than accounting for receptor flexibility.

With the aim of trying to increase the reliability of compound ranking, there are arguments for applying high level free energy calculations for all docking screens.¹⁹⁵ The principle being that the expense of further developing false positives outweighs the cost in computation time to score poses accurately. However, the accuracy of calculated IC_{50} values is often not as important as the order in which the docked compounds fall. This is especially true when precise affinities are determined later through *in vitro* methods. Currently available docking programs use a variety of inbuilt or external scoring functions to rank generated poses.¹⁹³ Following *in vitro* screening, it may be apparent that low level scoring is insufficient in which case higher level but slower methods can be used to rescore docked conformations.¹⁸²

With the increased interest in computational strategies for drug discovery, several public domain packages have been made available. These efforts have made docking calculations accessible to a far wider portion of the scientific community, as there is no longer the need for an expensive initial outlay before exploring computational studies. Four programs have received significant attention in recent years, DOCK, SLIDE, AutoDock and AutoDock Vina (Vina).¹⁹⁶ Despite the similarity in name, AutoDock and Vina share no components in their docking algorithms. All these programs take different approaches to the problem of conformational sampling. DOCK uses a shape-matching approach to sample alternative ligand poses.¹⁹⁷ These poses are then combined with the receptor to form a complex which can be scored using the Amber molecular force field.

Unlike most docking programs, SLIDE achieves high accuracy through intrinsic treatment of receptor flexibility.¹⁹⁸ The model assumes that multiple conformations of both ligand and receptor contribute to overall binding affinity. AutoDock divides the rigid receptor into a force field based grid, assigning electrostatic properties to each section. Poses generated by rotation of free bonds in the screened molecules can then be scored against the grid using an internal scoring function.¹⁹⁹ Vina uses stepwise mutation and local minimisation for generation of ligand conformations followed by global optimisation to calculate poses.²⁰⁰ In addition, its ability to run on multiple processors and combine multiple local minimisations to improve sampling produces an almost 60 fold increase in speed compared to AutoDock.

Vina is a recently developed docking program, which has been built with modern multiple core CPUs in mind.²⁰¹ In addition, its gradient optimisation method uses the first run as a guide for subsequent dockings, increasing the likelihood of finding the optimum conformation. As a result, it shows a significant increase in speed and increased accuracy of binding mode. In test screens using datasets which contain known ligands as well as decoys, Vina has achieved significant enrichment of high affinity ligands.²⁰² Encouragingly, rescoring the binding energies with higher level calculations did not result in generalised improvement of predictions. Vina offers a balance of speed and accuracy not found in previous programs, thanks to its ability to harness current hardware.

5.1 - *In Silico* Screening

5.1.1 - *In Silico* Screening Validation

In addition to rational drug design, molecular docking was employed in an attempt to identify potential binders of the Keap1 Kelch domain. Using this approach, it was hoped that both probes and drug leads for targeting the Keap1/Nrf2 interaction could be developed. The principle piece of software used to perform these docking experiments was Vina.²⁰¹ The program was chosen for its ability to rapidly screen thousands of molecules, its accuracy and ease of use.²⁰² In addition, it is open source and therefore freely available, with an active support community.

As the crystal structure for Keap1 is yet to be solved and the target for binding quite specific it was decided that the Kelch domain of Keap1 would be sufficient as the receptor. Previous work by Lo *et al.* has produced a crystal structure of the Kelch domain of Keap1 (Gly325-Thr609) with a bound 16 mer peptide from the Nrf2 protein.⁴⁰ The data for this structure has been deposited in the protein databank with the code 2FLU. This crystal structure provided the receptor for docking and in addition a ligand with which to test the accuracy of the calculations.

The receptor was prepared using the Swiss-Pdb Viewer²⁰³ by removing the 16 mer peptide followed by processing in AutoDock Tools (ADT) to add polar hydrogens appropriate for physiological pH and produce the receptor in the required pdbqt format. The workspace for docking was defined as a cube with sides of 30 Å to allow the largest space for docking possible while maintaining the speed of each run and the likelihood of finding the best bound conformation. The workspace was located so as to encompass the entirety of the area occupied by the peptide in the crystal structure and as much of the outer rim of the Kelch domain as possible. This also allowed for a certain amount of the central channel of the receptor to be explored. It was decided that including the sides would not be beneficial as recent work has suggested that in the complete Keap1 protein they are covered by the intervening region and binding here would be unlikely to disrupt the interaction of Nrf2.⁴⁴

Before embarking on screening it was important to validate the docking calculations with a known ligand. For this, the peptide sequence from the crystal structure was used. Due to limitations on the number of rotatable bonds that can easily be handled by Vina, two assumptions were made. Firstly, as only six of the residues present in the crystal structure form the binding sequence and the remaining residues are not in close proximity to the protein, Ala69 to Leu76, Phe83 and Leu84 were removed. These residues do however provide stabilisation of the β -turn structure of the peptide, as a result the backbone of the DEETGE peptide was fixed in the conformation found in the crystal structure. Unlike validation of docking with the full sequence, the peptide used here has the potential to rotate in the binding site if a more stable conformation can be located as the restriction of the long tail has been removed. In addition the side chain of Glu78 was disordered in the crystal data and therefore omitted. This was added back in using Swiss-Pdb Viewer. The ligand was prepared in the pdbqt format, including addition of polar hydrogens, using ADT, allowing all side chain bonds to be freely rotated.

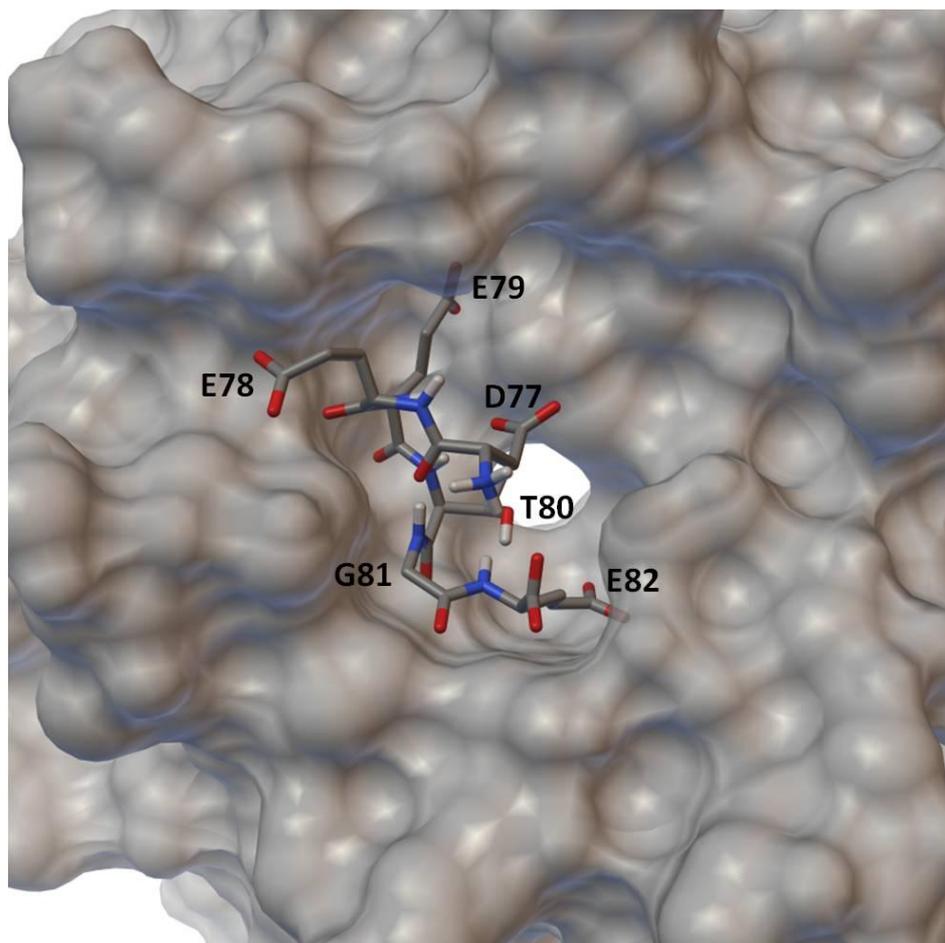


Figure 5.1: Docked conformation of the DEETGE binding sequence of Nrf2 into the Keap1 Kelch domain calculated by Vina

The resulting docking produced the conformation shown in Figure 5.1. As can be seen the peptide fits tightly into a curved pocket in the centre of the protein. The calculated binding energy of this conformation is -8.8 kcal/mol, which due to the nature of the docking calculation does not take into account the energy associated with the fixed conformation specified. However it can be assumed that any hit with a binding energy equal to or greater than this should show activity, as a greater number of free rotations can only reduce the energy of binding.

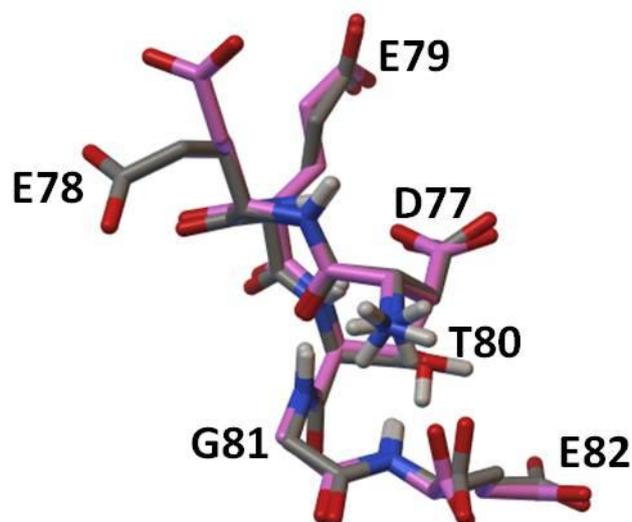


Figure 5.2: Overlay of DEETGE crystal structure (pink) and conformation calculated by Vina (grey)

In addition to the binding value calculated, overlay of the docked conformation with the crystal structure shows remarkable overlap (Figure 5.2). The only notable difference is the side chain of Glu78 which was disordered in the crystal structure and therefore to be expected. This shows that Vina is able to reproduce the binding mode found in the crystal structure, which is important for producing reliable results from screening.

5.1.2 - NCI Diversity Set II

With this initial validation completed and a benchmark in place, screening was started. Two sets of compounds were chosen for screening, the NCI Diversity Set II and the ChemBridge Building Blocks collection. The protonation states of the compounds within both these sets have been adjusted for physiological pH. The NCI Diversity Set II was chosen as a relatively small screen to test the capabilities of Vina. The aim of this set of compounds is to encompass as large a chemical space as possible using the smallest number of compounds. The result is 1880 compounds the structures of which are freely available in a number of formats including the required .pdbqt format used by Vina. As not all the structures can be successfully converted to the .pdbqt format, due to metal atoms, the screened set comprised 1540 structures. Compounds excluded contained atoms such as arsenic which would be less suitable as leads.²⁰⁰

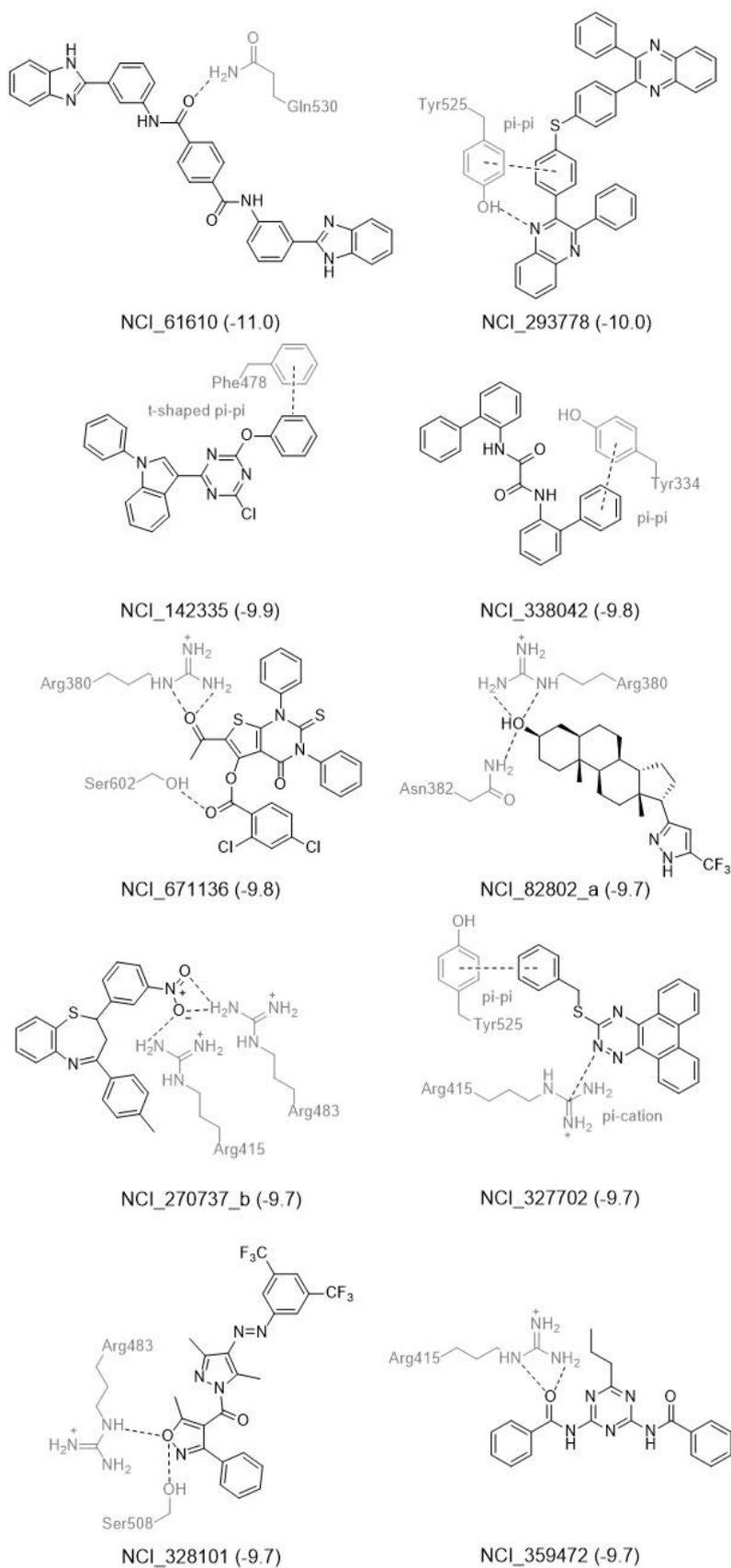


Figure 5.3: Top 10 hits identified by NCI Diversity Set II screen. NCI number, binding energy in brackets (kcal/mol) hydrogen bonding, pi-pi and pi-cation interactions shown in grey

Upon completion of this screen the top 10 hits were extracted and analysed (Figure 5.3). Hits were selected based on binding affinity, a more negative value meaning greater affinity. It was noted that all of the compounds in the top 10 had better binding energies than the test peptide and all bound within the same area of the receptor as the test peptide. Compound 61610 is a particularly interesting hit due to its binding mode. In addition to having the best binding energy, -11.0 kcal/mol, the location of its binding is unique among the top 10 (Figure 5.4). One of the benzimidazole groups is bound within the central channel, while the other is bound into a groove on the surface of the protein. The result is that the linking terephthalic acid blocks access to the peptide binding site. In addition it forms a single hydrogen bond to Gln530, which forms a backbone hydrogen bond to the test peptide. Due to the symmetry and lack of stereochemistry, synthesis of this molecule was expected to be relatively straight forward, allowing rapid preliminary validation of these results *in vitro*.

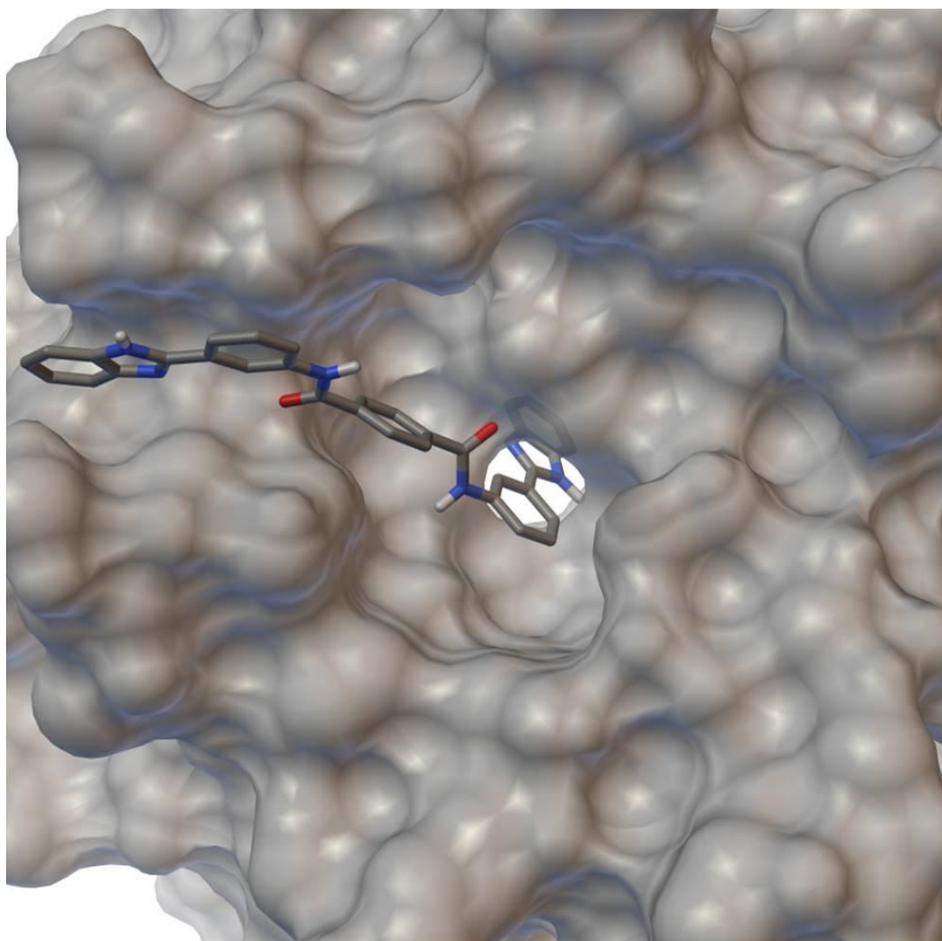


Figure 5.4: Conformation of the highest scoring hit from NCI Diversity Set II screen (NCI_61610)

The remaining hits fall into two categories, those which form pi-pi interactions with residues not involved in binding the test peptide and those that bind into one or other of the cationic pockets, which hydrogen bond to Glu79 and Glu82 of the test peptide. The distribution of affinities

between pi-pi stacking and hydrogen bonding compounds is fairly even, suggesting there may be two methods for targeting the interaction. One of the compounds, NCI_327702, interacts slightly differently with the binding pocket, instead of a hydrogen bond, it forms a pi-cation interaction via one of its nitrogens to the plane of Arg415. This may be another interaction that can be exploited further. It is important to note however, despite the high affinity and large size of these molecules none bind into both cationic pockets. This may be due to the shape of the molecules, as the binding site has a specific curve between the two pockets, though this is something that could be explored further through docking simulations. An ideal lead would interact with both cationic pockets, mimicking the key binding motif of Nrf2.

5.1.3 - ChemBridge Building Blocks

Subsequently a much larger set, the ChemBridge Building Blocks, was screened. This database, available through ZINC,²⁰⁴ comprising 11950 structures, is composed of lead-like compounds. The benefit of this is that the hits produced should have far less structural complexity than those in the previous screen, allowing greater scope for modification to improve binding and specificity. Again the top 10 hits were analysed (Figure 5.5) and all were found to bind within the same pocket as the test peptide.

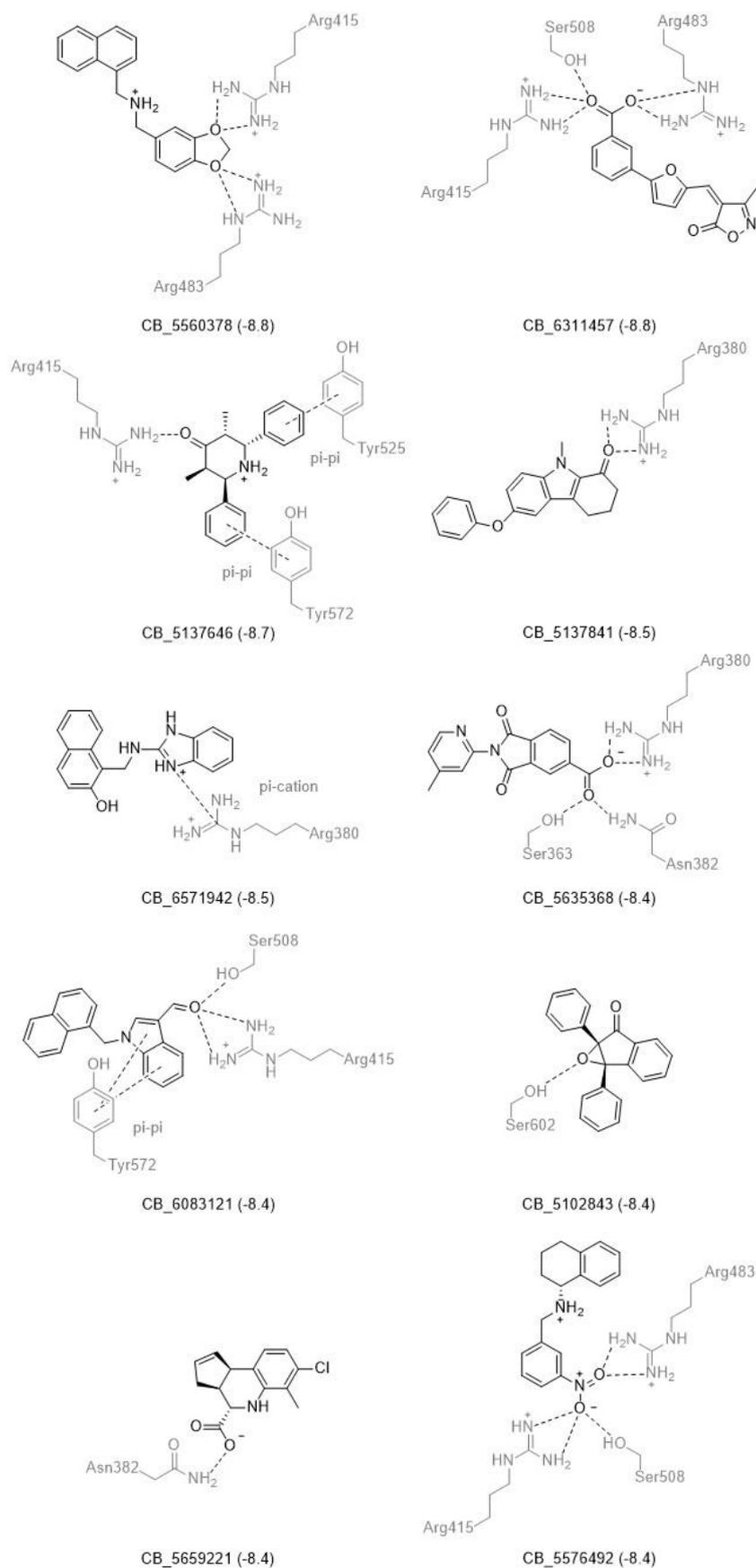


Figure 5.5: Top 10 hits identified by ChemBridge Building Blocks screen. ChemBridge number, binding energy in brackets (kcal/mol), hydrogen bonding, pi-pi and pi-cation interactions shown in grey

The best hit, CB_5560378 (Figure 5.6), only binds as strongly as the test peptide but due to its simple structure has great potential for variation. The framework for CB_5560378 can easily be accessed via reductive amination of 1-naphthaldehyde with piperonyl amine. Further variations could then be made using various aromatic aldehydes or amines to generate a library of leads.

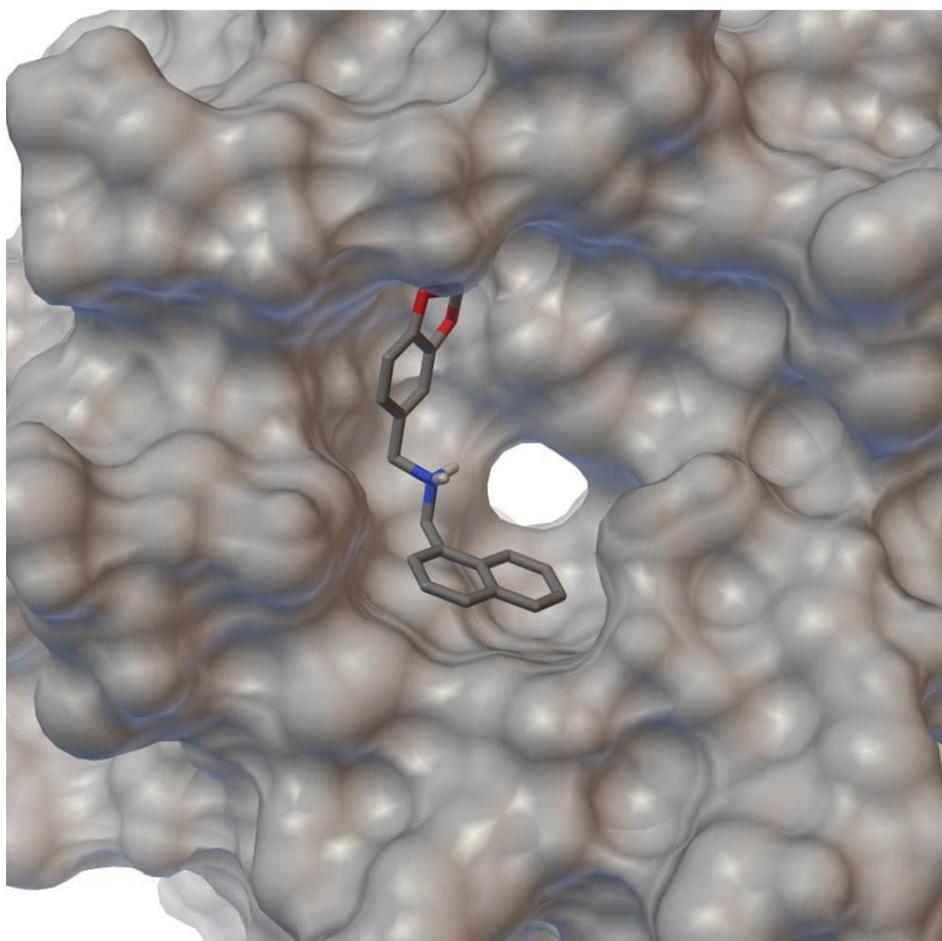
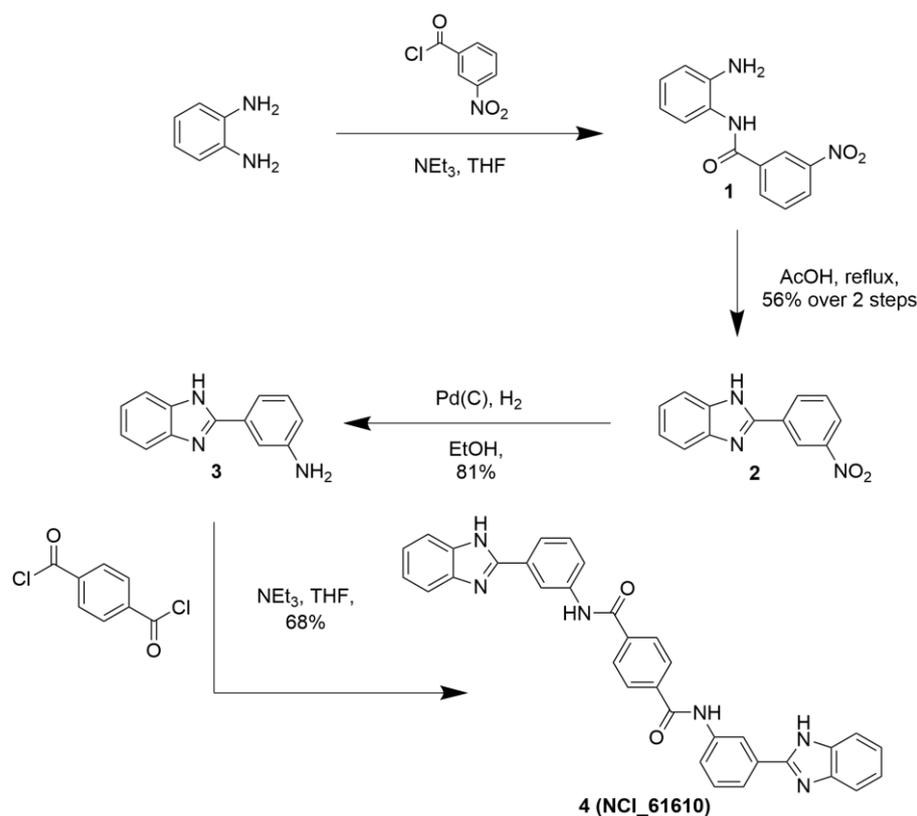


Figure 5.6: Conformation of the highest scoring hit from the ChemBridge Building Blocks screen (CB_5560378)

All the hydrogen bonds formed by the top 10 hits are to residues that the test peptide forms bonds with. Similarly to the NCI compounds, pi-pi interactions are with residues not involved in binding the test peptide, though Tyr525 is involved in interactions with both the ChemBridge and NCI compounds. Interestingly, there is another pi-cation interaction between a nitrogen in CB_6571942 and a cationic pocket arginine, however in this case it is the other binding pocket. All but one of the compounds binds into a cationic pocket, though again each compound only binds one pocket. Among these compounds, there does not seem to be a preference for one of the pockets over the other. The advantage with these compounds however, is the scope for extending their structures in order to target the second pocket, due to their simplified structures compared to the NCI set.

5.1.4 - Synthesis and *In Vitro* Screening

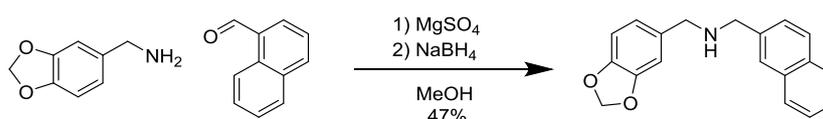


Scheme 5.1: Synthesis of NCI_61610

Having identified NCI_61610 as a potential target molecule, a synthesis was devised to afford the desired product in four steps (Scheme 5.1). Synthesis of the benzimidazole **2** was achieved in two steps by reaction of diamino benzene with a sub-stoichiometric amount of 3-nitrobenzoyl chloride. This, along with gradual addition of the acid chloride, ensured exclusive formation of the monosubstituted intermediate **1**. Refluxing this intermediate in acetic acid afforded benzimidazole **2** in an adequate 56% yield over the two steps. Reduction of the nitro substituent to generate aniline **3** was initially attempted by refluxing with iron powder in acidic conditions. This was found to be very low yielding and result in the formation of numerous side products. An alternative reduction with palladium on charcoal was found to produce the desired product in 81% yield with only minor side products. Proton NMR shows an expected shift in the nitro benzene signals upon reduction to aniline. A shift up field is characteristic of the change from an electron withdrawing nitro substituent to an electron donating aniline. Reaction of aniline **3**, with half an equivalent of terephthaloyl chloride, afforded the target compound **4** (NCI_61610). Isolation of which was complicated by the poor solubility of the product in organic solvents. Precipitation of the product from DMF by addition of water resulted in a satisfactory 68% yield. Amide formation is evident from the proton NMR, which shows the disappearance of the aniline

protons and a new signal at 10.66 ppm corresponding to the amide NH. As the molecule now has a plane of symmetry running through the terephthaloyl ring, only a single resonance is added to the aromatic region, appearing as a singlet.

Synthesis of CB_5560378 was achieved by imine formation between piperonyl amine and 1-naphthaldehyde in the presence of magnesium sulfate, followed by reduction with sodium borohydride affording the compound in a 47% yield (Scheme 5.2). Validation of these two hits was then carried out using the FP assay.



Scheme 5.2: Synthesis of CB_5560378

Both compounds were tested for their ability to disrupt the interaction at a concentration of 1000 nM using the 14 mer peptide as a positive control (Figure 5.7). At the concentration tested however, no reduction in anisotropy was observed. While higher concentrations of compound could be tested, it was decided that with no reduction in signal at 1000 nM the K_i values would be very high if they bound at all, rendering them unsuitable as leads. In addition, solubility of NCI_61610 was found to be an issue as may be expected considering the high proportion of hydrophobic groups and its polyaromatic structure which would allow for extensive pi-pi stacking.

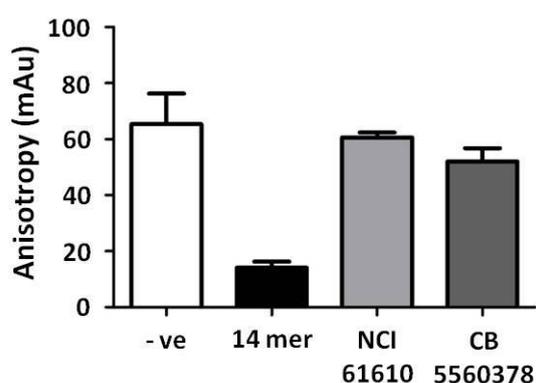


Figure 5.7: Fluorescence Polarisation inhibition screen of NCI_61610 and CB_5560378, Mean \pm SEM, n = 3, 30 nM Keap1, 5 nM F-14, 1000 nM compound

With these preliminary results indicating that neither of the predicted inhibitors would be effective, it was decided to return to the docking calculation experiments and attempt to validate the hits with other receptor structures.

5.1.5 - Cross-Receptor Screening

Crystal structures are an average of all the molecules in a crystal lattice, which can result in significant variation between structures.²⁰⁵ This is important when using the structures for drug discovery as they may not be a true representation of the protein in solution. It was decided that by screening against multiple structures and ranking hits by both frequency and affinity it would be possible to reduce the impact of variation within any one structure on lead identification. In order to do this, the other known crystal structures for the Keap1 Kelch domain were prepared for docking as before. This gave a total of eight structures to screen against.

Before screening a library of compounds against the structures, it was once again necessary to validate the docking with a known ligand. If any of the structures were altered to the point where known ligands would not bind, the results of screening could not be relied upon. The same six amino acid DEETGE sequence used in validating docking to the 2FLU structure was used to test the docking of the other seven receptor structures. The peptide was docked into each receptor using the default parameters and the resulting binding modes assessed. All but one of the docking calculations resulted in a lowest energy conformation which matches the original structure very closely. The main variation between the structures is the position of the Glu78 side chain, which is to be expected as it is missing from the original crystal structure due to disorder. The Nrf2 binding site of the majority of the crystal structures share a high degree of similarity. The exception to this is the structure of 2DYH which has a smaller central channel and smaller binding pocket in general. As a result the side chain of Glu82 is rotated out of the Arg380 pocket where it binds in its original crystal structure. In the second highest affinity mode for this docking however, the peptide assumes the same conformation as for the other structures.

As the Chembridge Building Blocks library showed the most potential for modification after lead identification, it was selected to screen against the Keap1 crystal structures. Docking calculations were carried out for the library against all eight of the receptors using identical parameters, with the exception of grid box location, which was adjusted depending on the orientation of the receptor structure. Following docking, results were processed to extract the 30 compounds with the highest binding affinity for each receptor. These 240 hits were then pooled and ranked by number of occurrences, the maximum being eight where the compound was a top 30 hit for every receptor and the minimum being one. Those with an equal rank by occurrences were then ranked by highest binding affinity for any one receptor. This produced a set of 21 compounds with high affinity for four or more receptor structures and 10 with high affinity for five or more, from a total of 107 unique structures, with binding affinities ranging from -8 to -9 Kcal/mol. 50 of the unique

compounds are only present for a single receptor structure, highlighting the difference between the reported crystal structures. The top 10 compounds, those with affinity for five or more receptor structures, were analysed for binding interactions to their highest affinity receptor (Figure 5.8).

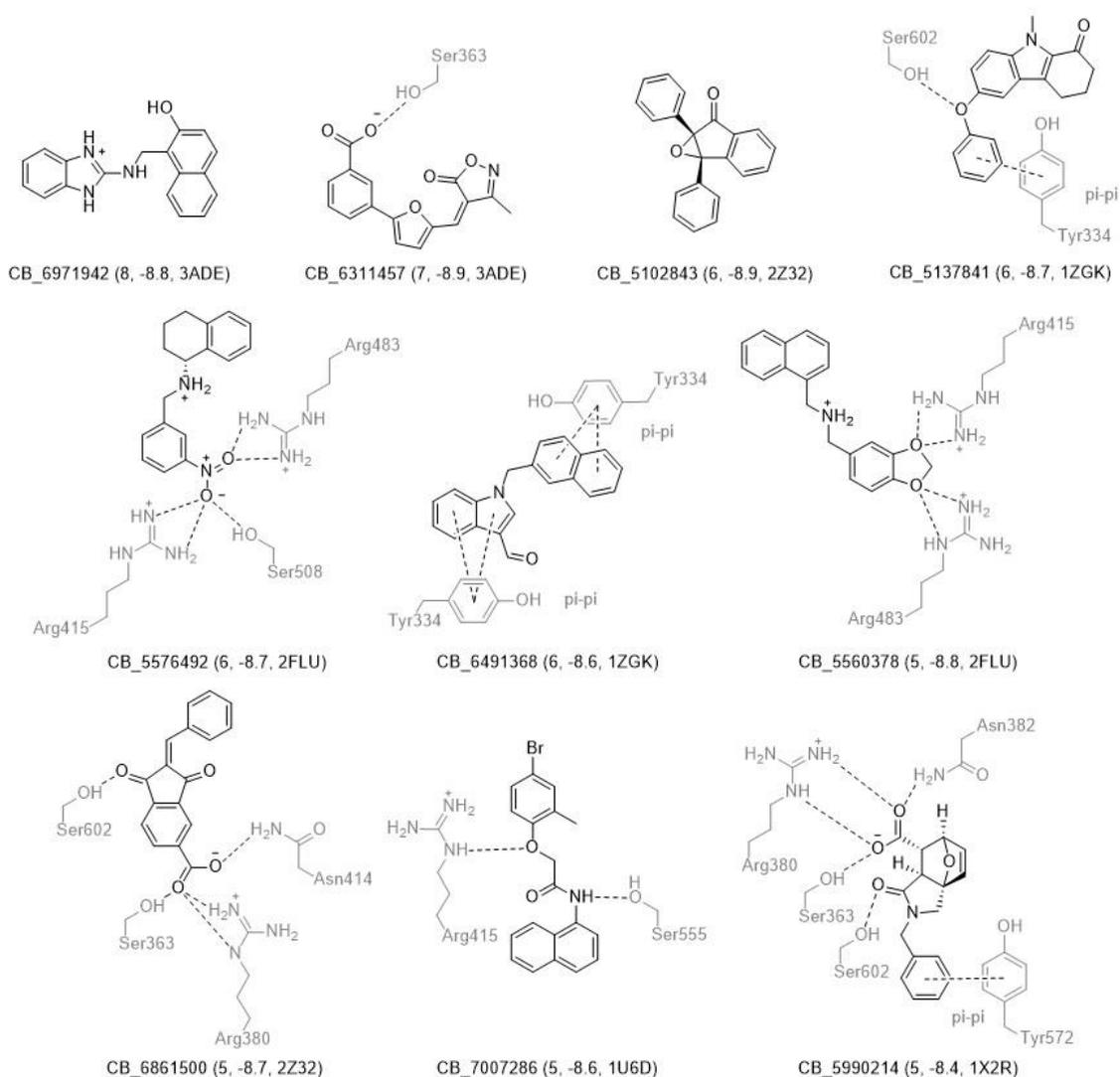


Figure 5.8: Top 10 hits from receptor screening, intermolecular interactions with highest affinity receptor in grey, ChemBridge ID. Frequency, affinity (kcal/mol), PDB ID of highest affinity receptor, in brackets

The most striking thing to note about these hits is the resemblance to those for 2FLU alone. The seven most frequent hits all appear in the top 10 for the original docking experiment, including CB_5560378, the top hit from the first screen. This indicates that despite some visual differences between the Keap1 crystal structures there is a high degree of similarity between them. It was decided that even though one of the compounds was known to not be effective *in vitro*, that the rest of the compounds may still be effective and therefore worth examining further.

The most frequent hit, CB_6571942, shows two common binding modes in its docking. With receptor structures where the central channel is open, the benzimidazole unit fits into the channel while the β -naphthol unit binds into one of the cationic pockets. Where the structure has either a very small central channel or a residue partially covering it, the two aromatic units bind into each cationic pocket. Despite the relatively strong binding, -8.8 Kcal/mol, and affinity for every receptor examined, no specific intermolecular interactions are formed between the ligand and receptor. This does however raise the possibility of adding structural modifications to introduce hydrogen bond acceptors to increase affinity. The other molecules all bind within the Nrf2 binding pocket of Keap1 and those with hydrogen bond acceptors forming specific interactions with the cationic pockets tend to have two similar sets of binding modes. One in which they form interactions with one cationic pocket and another where they bind to the other cationic pocket. In these cases it seems likely that modifications to include two hydrogen bond acceptors would increase affinity as both cationic pockets could be bound simultaneously.

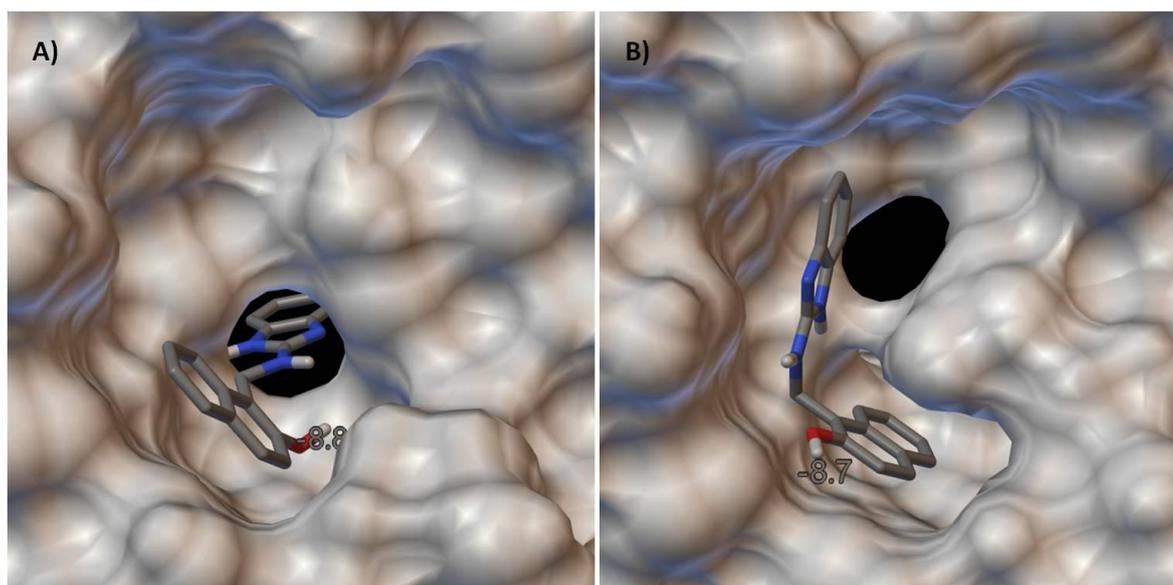


Figure 5.9: Representative calculated binding modes for CB_6571942 with Keap1 A) benzimidazole moiety binds in central channel, PDB I.D. 3ADE B) partially blocked central channel causes benzimidazole to bind into a cationic pocket, PDB I.D. 1ZGK

Of the 10 hits identified, nine were available commercially, CB_6311457 being unavailable at the time. The compounds were purchased and screened as previously in the FP assay at a concentration of 2 μ M. No significant inhibition was seen for any of the compounds and as a result investigations in this area were stopped. One issue that arose during *in vitro* testing, was the poor solubility of many of the compounds. Preparation of stock solutions in DMSO was necessary and in some cases solubility was such an issue that compounds precipitated once added to the assay wells. As a result, these compounds may indeed disrupt the interaction at high

concentrations, however their poor solubility prevents testing and renders them unsuitable as leads.

As there is such a high degree of consensus between the original ChemBridge screen and the calculations with multiple crystal structures, the fact that none of these compounds inhibits the interaction points to a fundamental problem with the docking calculations. The nature of this problem is not clear, though there is some evidence to suggest that Vina, and docking programs in general, are more suited to generation of binding modes rather than prediction of affinity.²⁰⁶ It may be possible to reanalyse these results, using an alternate molecular mechanics based binding energy calculation, in order to yield more accurate hits. With the FP assay optimised for both screening and inhibition curves, it was decided that focussing on optimisation of peptide sequences would be more productive than high throughput screening.

5.2 - Conclusions

In silico screening is emerging as an important aspect of the drug discovery process. The ability to screen large libraries without the expense of purchasing compounds is especially significant for academic research. In combination with the FP assay developed in Chapter 3, docking experiments have the potential to reduce both the time and cost required to identify new leads.

As has been shown with the work here, identification of hits using *in silico* techniques does not necessarily transfer to hits *in vitro*. It may be the case that the molecules identified would inhibit at high micromolar concentrations, however, their limited solubility prevented testing above the concentrations used. Re-examination of the results, using alternate scoring functions may be a more beneficial method to identify potent binders. Alternatively, screening more focussed libraries may produce more suitable leads.

This highlights an intrinsic problem when targeting PPIs in general. The interaction between two proteins typically comprises many weak hydrophobic contacts, spread over a large area, which combine to produce a strong overall binding. As a result, molecules predicted to bind well to these surfaces are typically large and hydrophobic as seen with the screens performed here. These factors negatively impact the physiochemical properties of the molecule, limiting their use as drug leads. This problem has to be overcome in order to develop effective inhibitors, either by insertion of hydrogen bonding functionalities where tolerated or by targeting hydrophilic "hot spots" which contribute more significantly to PPI binding. This may be effective for targeting the Nrf2/Keap1 interaction, by developing molecules which can bind into both of the cationic pockets.

Chapter 6: Conclusions and Future Work

6.1 - General Conclusions

Control of inflammation and resolution of chronic disease through manipulation of the Nrf2 pathway shows great promise. The work presented here has shown the first clear example of cellular non-covalent induction of the Nrf2/Keap1 pathway. In addition it has confirmed the validity of targeting the Nrf2/Keap1 PPI using Nrf2 derived peptides. Induction of genes regulated by Nrf2 can be triggered in a dose and time dependent manner in monocytic cells and one of these peptides, TAT-14, has been shown to reduce levels of the pro-inflammatory mediator TNF α in a model of sepsis. These peptide tools fulfil two roles, they act as a starting point for the development of smaller, more drug-like molecules and they allow detailed investigation of the Nrf2 pathway with more control than afforded by cysteine modifying inducers.

Development of an FP assay has allowed quantification of the binding affinities for each of the peptides synthesised. When considering the PPI isolated from the complex processes found in cells, the affinities of the synthesised peptides are remarkable. Single digit nanomolar inhibition by the CPP conjugated 14 mer sequences demonstrates that highly potent inhibitors can be developed from natural sequences. While this has not always correlated with data from cellular assays, it has indicated that the β -hairpin structure is of great importance to the potency of the Nrf2 based peptides. Further development of these peptides has become significantly faster thanks to the rapidity with which they can be screened in the FP assay before moving to cellular assays.

Development of disulfide cyclised peptides has shown that constraining the conformation of Nrf2 based peptides can dramatically increase binding affinity. These peptides have also identified a relationship between peptide length and potency. The 8 and 10 amino acid sequences bind strongly to the target, whereas the shorter 6 residue peptides only inhibit at high concentrations. This confirms the DEETGE sequence as the minimum required sequence for potent inhibition, from which further peptide inhibitors can be designed.

The transition from peptide inhibitors to small molecules is a lengthy process, only the beginning of which is presented here. In spite of this, the cyclic peptides produced show great promise. The key objectives of reduction in size while maintaining affinity have been met. An almost two thirds reduction in mass from the 3171 Da TAT-14 peptide to the 1067 Da Ar8P peptide while only increasing K_i from 3.6 nM to 6.1 nM is a significant development. In addition to the potent binding affinity, the properties of stapled peptides on both cell membrane permeability and proteolytic

stability will be beneficial for further work. Whether or not the aryl stapled peptides prove to have cellular activity, the groundwork has been laid for the production of small, highly potent inducers of the Nrf2/ARE pathway.

In combination, the FP assay and computational docking techniques allow for rapid screening of libraries in order to identify novel leads. The process of hit determination from *in silico* screens needs to be refined however. Either through screening of compounds selected by computation or through high throughput screening of focussed libraries, the tools are in place for the identification of small molecule inhibitors of the Nrf2/Keap1 interaction.

6.2 - Future work

Continuation of cell based assays with the aryl stapled peptides is the most pressing avenue of research. If the peptides are found to induce downstream genes, they have great potential as drug leads and should be assessed more extensively. If they are not active however, it may be possible to attach fluorescent tags in order to identify whether an inability to enter cells is responsible. If this is the case, model peptides may be of use to determine which components are reducing bioavailability.

In parallel to determining why the aryl stapled peptides are unable to induce Nrf2 dependent genes, work should commence on increasing cell membrane permeability. Initially conjugation of a cell penetrating sequence, such as TAT, would hopefully enable peptide internalisation and assessment of their properties as Nrf2 inducers. Such a peptide could incorporate the CPP sequence at either the N- or C-terminus to generate a cyclic-tail peptide. While this solution may be effective for allowing cellular activity, it should not represent a ultimate goal as one aim of the cyclic peptide work was to remove the need for a CPP sequence. Alternatively, smaller additions to the peptide to increase lipophilicity, such as conjugation of a hydrocarbon chain to one terminus, may offer a compromise in increasing molecular weight and cell membrane permeability. N-methylation of one or more residues may also benefit cell penetration, however, as the β -hairpin structure features multiple intra-molecular backbone hydrogen bonds, placement would have to be carefully considered.

Further refinement of the cyclic peptide sequence should also be undertaken. Reduction of overall charge to increase membrane permeability, such as through substitution of aspartic acid with asparagine is an option. In combination with the proline substitution already identified, this would reduce the overall charge to -2, compared to the -4 charge of the disulfide bridged peptides. Exploration of carboxylic acid isosteres such as tetrazoles may also prove beneficial for both cell membrane permeability and binding affinity. Movement away from a peptidic scaffold should be a long term goal, though this may take a considerable amount of time.

Peptidomimetic scaffolds would act as a convenient intermediate between peptide and small molecule inhibitors. Molecules analogous to the peptide synthesised previously, either CPP-Nrf2 or cyclic could be constructed using peptoid or β -peptide backbones. Both strategies benefit from increased proteolytic stability due to their non-natural structures. The N-substituted glycine peptoids offer a wide range of amide nitrogen functionalities as the necessary building blocks are widely available. The molecules produced are achiral which eliminates racemisation as an issue,

however, this may impact binding affinity. Alternatively the β -peptide scaffold inserts an additional CH_2 unit into the peptide chain producing chiral molecules with differing secondary structures to the parent molecule. Again, this may be either beneficial or detrimental for binding to the target. Either approach may open up new areas for further development to suitable drug leads.

While the CPP-Nrf2 peptides are less suitable as drug leads, they are potentially useful tools. As such, it would be beneficial to test their effects in a wide variety of cell lines. In particular, if a cell type is identified with higher resistance to octa-arginine, the R₈-14 peptide could prove to be a more potent inducer than the TAT-14 peptide, due to its increased uptake.

The results of the experiments using fluorescently tagged TAT-14 were unexpected. The rapid cell entry and decreasing fluorescence over time cannot be easily explained using normal fluorescence microscopy. Using confocal microscopy however, would allow for visualisation of sections through the cell, eliminating the issue of membrane binding. It would also allow for some degree of localisation to be assessed and may determine the cause of decreasing fluorescence over time.

Separate from continued development of peptides, identification of entirely novel Keap1 binders through high throughput screening is now possible. Computational results could be refined by implementation of an alternate scoring function. Comparison of multiple docking programs may also be beneficial to identify the shortcomings of Vina. Screening of libraries using the FP assay could be profitable, either in conjunction with *in silico* screens or alone.

Chapter 7: Experimental

7.1 - Chapter 2

7.1.1 - Peptide Synthesis

General

Unless otherwise noted all peptides were synthesised using solid phase Fmoc chemistry on a Multisynthtech Syro I automated peptide synthesiser. Couplings were performed using 4 eq protected amino acid, 4 eq O-benzotriazole-N,N,N',N'-tetramethyl-uronium-hexafluorophosphate (HBTU), 4 eq hydroxybenzotriazole (HOBt) and 8 eq N,N-diisopropylethylamine (DIPEA) in N,N-dimethylformamide (DMF). Removal of Fmoc protecting groups was achieved using piperidine (20% in DMF) for 3 x 10 min. Amino acids were dissolved in NMP at 0.69 M with the exception of arginine which was dissolved in DMF. DIPEA was prepared at 2 M in NMP and HBTU/HOBt at 0.45 M in DMF. Couplings proceeded for 25 min followed by washing with DMF and repetition of coupling. Following chain elongation on the automated synthesiser, the resin was washed with DMF (3 x 5 mL), DCM (3 x 5 mL), 1:1 DCM:MeOH (3 x 5 mL) and dried by suction.

Standard cleavage from the resin and global deprotection was performed using TFA:H₂O:TIPS 95:2.5:2.5 (5 mL) for 3 hours with constant agitation. The peptide, dissolved in the cleavage mixture was separated from the resin by filtration and the resin washed with TFA (2 x 3 mL) which was added to the filtrate. The cleavage cocktail and volatile components were removed under reduced pressure to give a gel. This product was then washed with ice cold diethyl ether (10 mL) precipitating the crude peptide as a white solid.

Analytical RP-HPLC was used to assess the purity of the product after precipitation and between purification steps. Analytical HPLC was performed on an Agilent Technologies 1200 series chromatograph with an Agilent Technologies ZORBAX Eclipse XDB-C18 (5 µm, 4.6 x 150 mm) column at 40°C and a gradient of 95:5 water:methanol with 0.05% TFA additive to 5:95 water:methanol over 15 min returning to 95:5 water:methanol over 5 min at a flow rate of 1 mL/min.

Peptides were isolated by reverse-phase chromatography using a Biotage Isolera Four providing gross purification using a 12 g C18 cartridge with a gradient of 95:5 water:methanol with a 0.05% TFA additive going to 5:95 water:methanol over 40 min with a flow rate of 20 mL/min, collecting at 214 and 254 nm. Where necessary, peptides were further purified by semi-preparative HPLC on

an Agilent Technologies 1200 series chromatograph using an Agilent Technologies ZORBAX Eclipse XDB-C18 (5 μ m, 9.4 x 250 mm) column.

TAT-10 synthesis

The peptide was synthesised on 2-chlorotrityl chloride resin (250 mg, Merck, loading: 1.01 mmol/g). Attachment of the first amino acid was carried out manually on a shaker. The resin was swollen in dry DCM, followed by addition of Fmoc-Pro-OH (340 mg, 1.01 mmol) and DIPEA (220 μ L, 1.26 mmol). The reaction proceeded for 2 hours after which the resin was capped using a mixture of DCM:MeOH:DIPEA 80:15:5 (10 mL) for 5 minutes, twice. Chain elongation and cleavage from the resin were conducted using the standard procedure. The peptide was isolated by reverse-phase chromatography using a Biotage Isolera Four using the standard method. Fractions containing pure peptide as determined by RP-HPLC and MALDI-TOF mass spec were combined, the solvent removed under reduced pressure and lyophilised from water (0.05% TFA) to produce the desired product as a white powder. TAT-10 **MS (MALDI +ve)**: calculated for: 2689.5 found: 2689.8

Initial synthesis of the TAT-14 and TAT-16 peptides

Synthesis was performed on Wang resin preloaded with the appropriate amino acid (500 mg, Merck, loading: 0.59 and 0.64 mmol/g respectively). Chain elongation and cleavage from the resin were conducted using the standard procedure. The peptides were isolated initially by reverse-phase chromatography using a Biotage Isolera Four using the standard method.

TAT-14 was further purified by semi-prep RP-HPLC using a gradient of 95:5 water:methanol to 5:95 water:methanol with 0.05% TFA additive over 45 min at a flow rate of 4 mL/min collecting at 214 nm. Fractions containing pure peptide as determined by RP-HPLC and MALDI-TOF mass spec were combined, the solvent removed under reduced pressure and lyophilised from water (0.05% TFA) to produce the desired product as a white powder. **MS (MALDI +ve)**: calculated for: 3171.7 found: 3171.5

TAT-16 was further purified by semi-prep RP-HPLC using a gradient of 95:5 water:methanol to 5:95 water:methanol with 0.05% TFA additive over 60 min at a flow rate of 4 mL/min collecting at 214 nm. Fractions containing pure peptide as determined by RP-HPLC and MALDI-TOF mass spec were combined, the solvent removed under reduced pressure and lyophilised from water (0.05% TFA) to produce the desired product as a white powder. **MS (MALDI +ve)**: calculated for: 3397.8 found: 3398.8 [M + H]

TAT-14Sc and 14 mer synthesis

Peptides were synthesised on Wang resin (125 mg, AGTC, loading: 3.8 mmol/g). Attachment of the first amino acid was carried out manually on a shaker. A symmetrical anhydride of the first amino acid was formed under an inert atmosphere at 0°C using the appropriate amino acid (10 eq, 4.788 mmol) in DCM, followed by addition of DIC (375 µL, 2.394 mmol) dropwise. The reaction was stirred for 30 min before removal of solvent under reduced pressure. The anhydride was then taken up in a minimum amount of DMF and added to the pre-swollen resin along with a catalytic amount of DMAP. The reaction proceeded for 3 hours before capping the resin using acetic anhydride (90 µL, 0.9576 mmol) and pyridine (75 µL, 0.9576 mmol) in a minimum amount of DMF for 30 min. Chain elongation and cleavage from the resin were conducted using the standard procedure. The peptides were isolated initially by reverse-phase chromatography using a Biotage Isolera Four using the standard method.

The 14 mer was further purified by semi-prep RP-HPLC using a gradient of 95:5 water:methanol to 5:95 water:methanol with 0.05% TFA additive over 30 min returning to 95:5 water:methanol over 5 min at a flow rate of 4 mL/min collecting at 214 nm. Fractions containing pure peptide as determined by RP-HPLC and MALDI-TOF mass spec were combined, the solvent removed under reduced pressure and lyophilised from water (0.05% TFA) to produce the desired product as a white powder. **MS (MALDI +ve):** calculated for: 1630.8 found: 1653.4 [M + Na]

TAT-14Sc was further purified by semi-prep RP-HPLC using a gradient of 95:5 water:methanol to 5:95 water:methanol with 0.05% TFA additive over 45 min returning to 95:5 water:methanol over 5 min at a flow rate of 4 mL/min collecting at 214 nm. Fractions containing pure peptide as determined by RP-HPLC and MALDI-TOF mass spec were combined, the solvent removed under reduced pressure and lyophilised from water (0.05% TFA) to produce the desired product as a white powder. **MS (MALDI +ve):** calculated for: 3171.7 found: 3173.2 [M + H]

Revised synthesis of TAT-14 and TAT-14Sc and synthesis of R₈-14 and R₄-14

After determination of aggregation on Wang resin, subsequent syntheses were performed using NovaSyn TGA resin (100 mg, Merck, loading 0.26 mmol/g). Attachment of the first amino acid was carried out manually on a shaker. A symmetrical anhydride of the first amino acid was formed under an inert atmosphere at 0°C using the appropriate amino acid (10 eq, 0.26 mmol) in DCM, followed by addition of DIC (20 µL, 0.13 mmol) dropwise. The reaction was stirred for 30 min before removal of solvent under reduced pressure. The anhydride was then taken up in a minimum amount of DMF and added to the pre-swollen resin along with a catalytic amount of

DMAP. The reaction proceeded for 3 hours before capping the resin using acetic anhydride (5 μ L, 0.052 mmol) and pyridine (4 μ L, 0.052 mmol) in a minimum amount of DMF for 30 min. Chain elongation and cleavage from the resin were conducted using the standard procedure. The peptides were isolated by reverse-phase chromatography using a Biotage Isolera Four using the standard method, substituting methanol for acetonitrile. Fractions containing pure peptide as determined by RP-HPLC and MALDI-TOF mass spec were combined, the solvent removed under reduced pressure and lyophilised from water (0.05% TFA) to produce the desired product as a white powder.

TAT-14 **MS (MALDI +ve)**: calculated for: 3171.7 found: 3174.4 [M + H]

TAT-14Sc **MS (MALDI +ve)**: calculated for: 3171.7 found: 3176.2 [M + H]

R₈-14 **MS (MALDI +ve)**: calculated for: 2879.6 found: 2880.8 [M + H]

R₄-14 **MS (MALDI +ve)**: calculated for: 2255.2 found: 2256.4 [M + H]

F-TAT-14 synthesis

Attachment of 5(6)-carboxyfluorescein to the N-terminus of TAT-14 was carried out on the solid phase, using TAT-14 synthesised by the NovaSyn TGA method (100 mg, 0.026 mmol). 5(6)-carboxyfluorescein (20 mg, 0.052 mmol), HBTU (20 mg, 0.052 mmol), HOBt (8 mg, 0.052 mmol) and DIPEA (18 μ L, 0.104 mmol) were preactivated in DMF before addition to preswollen peptidyl resin. The resin was shaken for 30 min, washed with DMF 4 x 5 mL and completion of coupling confirmed by Kaiser test. The resin was washed with DCM (3 x 5 mL), 1:1 DCM:MeOH (3 x 5 mL) and dried by suction. Cleavage from the resin was conducted using the standard procedure. The peptide was isolated by reverse-phase chromatography using a Biotage Isolera Four using the standard method. Fractions containing pure peptide as determined by RP-HPLC and MALDI-TOF mass spec were combined, the solvent removed under reduced pressure and lyophilised from water (0.05% TFA) to produce the desired product as a yellow powder.

7.1.2 - Cell Biology

Unless otherwise noted all reagents for cell experiments were purchased from PAA and used according to manufacturer's instructions. THP-1 cells were obtained from the European Collection of Cell Cultures.

Cell Culture

THP-1 cells were cultured in RPMI1640 medium with 10% Foetal Calf Serum (FCS), 2 mM L-glutamine and penicillin/streptomycin (P/S). Cells were grown at 37°C in an atmosphere of 5% CO₂.

MTS Cell Proliferation Assay

Cells were plated out at a concentration of 3×10^5 /mL 100 µL per well and treated with 1 µL of compound at appropriate dilution in triplicate. The cells were then incubated for 72 or 24 h before addition of 10 µL MTS per well and further incubation for 2.5 h. Absorbance of the plates was determined using a BMG Labtech Optima microplate reader (492 nm) and toxicity of compounds assessed by comparison to untreated controls.

Stimulation

Cells were plated out at a concentration of 5×10^5 /mL, 2 mL per well and incubated for 18 h to allow them to return to an unstressed state before stimulation. Cells were treated at a concentration of 1×10^6 /mL. Plates were then incubated at 37°C in an atmosphere of 5% CO₂ for the appropriate time interval.

RNA extraction and real-time PCR

Cells were prepared and treated according to the standard method. Following incubation, cells were spun down, the supernatant removed and TRI Reagent (Ambion) added followed by 100 µL 1-bromo-3-chloropropane. After centrifugation the clear colourless top layer containing RNA was transferred to clean tubes and 200 µL 2-propanol added. After vortexing, incubation for 10 min and centrifugation at 12,000 rpm for 15 min the supernatant was removed and discarded leaving a pellet containing the RNA. This was then taken up in 1 mL 70% EtOH in H₂O, followed by further centrifugation at 12,000 rpm for 10 min and subsequent removal of supernatant. The pellets were then allowed to air dry before the addition of 20 µL H₂O and freezing overnight at -80°C.

RNA concentration was quantified using a Nanodrop ND-1000 spectrophotometer and concentrations adjusted as necessary. RNA reverse transcription was achieved using a TaqMan kit (Applied Biosystems) according to the manufacturer's instructions and carried out on a PTC-100 Peltier thermal cycler using the following profile: 21°C for 10 min, 42°C for 15 min, 99°C for 5 min and 4°C for 5 min.

HO-1 and GAPDH primers were acquired from Invitrogen with the following sequences, HO-1 forward: 5'-ATGGCCTCCCTGTACCACATC-3', reverse: 5'-TGTTGCGCTCAATCTCCTCCT-3', GAPDH forward: 5'-AACAGCCTCAAGATCATCAGCA-3', reverse: 5'-TGCTAAGCAGTTGGTGGTGC-3'. mRNA expression was measured by real-time PCR using a QIAGEN Rotor-Gene Q and SYBR Green technology for 40 cycles of 95°C for 15 s and 60°C for 1 min. Each mRNA expression was normalised against GAPDH mRNA expression using the standard curve method.²⁰⁷

Western Blotting

Cells were prepared and treated according to the standard method. Following incubation, cells were spun down and the supernatant removed. SDS (Invitrogen) 100 µL was added to each sample, mixed to dissolve the pellet and then boiled for 5 min. 13.5 µL of sample was combined with 1.5 µL reducing agent (Invitrogen) and boiled for a further 5 min, 10 µL of each sample was loaded on to the gel. Separation was performed using premade Bis-Tris PAGE gels (Invitrogen) before being transferred to a nitrocellulose membrane. Antibodies used are shown in Table 7.1.

Protein	1°				2°			
	Company	Code	Dilution	Details	Company	Code	Dilution	Details
Nrf2	Abcam	AB62352	1/1000	Rabbit monoclonal	Santa Cruz	SC-2030	1/1000	Goat anti-rabbit
HO-1	R&D Systems	AF3776	1/2000	Goat polyclonal	Santa Cruz	SC-2020	1/1000	Donkey anti-goat
β-Actin	Sigma	A1978	1/100000	Mouse monoclonal	Dako	P0447	1/1000	Goat anti-mouse

Table 7.1: Antibodies used in Western blotting, manufacturer and dilution

Bands were visualised using Amersham ECL Prime according to manufacturer's instructions.

HO-1 ELISA

HO-1 protein levels were determined using an R&D Systems DuoSet IC Human Total HO-1/HMOX1 assay kit. Cells were prepared and treated according to the standard method. Following incubation, cells were spun down, the supernatant removed and washed twice with PBS. The cell pellet was then resuspended in lysis buffer (0.5% Triton X-100, Pierce protease cocktail in PBS) to a concentration of 6.66×10^6 cells/mL and kept on ice for 15 min before storage at -80°C. Plates were prepared by incubation with 100 µL of 8 µg/mL capture antibody per well over night at RT. The wells were emptied and washed with Wash Buffer (PBS + 0.05% Tween20) three times. Wells were then blocked with 300 µL Block Buffer (PBS + 1% BSA + 0.05% sodium azide) for 2 hours at

RT. Standards were prepared by serial dilution of an HO-1 protein stock (10 ng/mL) in two fold steps in IC Diluent #4 (1 mM EDTA, 0.5% Triton X-100 in PBS) to give a 7 point curve from 10 ng/mL to 0.156 ng/mL. Samples were defrosted, spun to pellet any undissolved material and diluted 2 fold in IC Diluent #4 to give an effective concentration of 3.33×10^5 cells/mL. The plate was washed as before and 100 μ L standards and samples were added using IC Diluent #4 as the zero concentration. The plate was then sealed and incubated for 2 hours at RT. The plate was washed and 100 μ L detection antibody (200 ng/mL) diluted in IC Diluent #1 (1% BSA in PBS) added, sealed and incubated for 2 hours at RT. Following another wash step, 100 μ L Streptavidin-HRP conjugate was added per well diluted as directed in IC Diluent #1 and incubated for 20 min at RT in the dark. The plate was washed and 100 μ L of a 1:1 solution of Color Reagent A (H_2O_2) and Color reagent B (tetramethylbenzidine) added to each well. The plate was then incubated for 20 min at RT in the dark. 50 μ L Stop Solution (1M H_2SO_4) was added to each well and the absorbance (450 nm) measured using a microplate reader.

The standards were analysed using a 4 parameter logistic curve fit, from which sample concentrations were calculated. Fold inductions were then determined by comparison to a non-treated control sample.

Fluorescence Microscopy

Cells were prepared according to the standard method. Cells were treated for 180, 120, 60, 30 and 5 min with 75 μ M F-TAT-14 peptide. Following incubation, cells were spun down and the supernatant removed. Cell pellets were washed 2 times with sterile PBS to remove excess fluorescent peptide. In order to remove membrane bound peptide, cells were incubated with trypsin (1 mg/mL) for 15 min before a final wash in sterile PBS. Prior to trypsin digest, half the 5 min sample was set aside as a control. The cells were resuspended to a concentration of 1×10^7 ml⁻¹ and visualised using a Leica Microsystems CMS GmbH Fluorescence Microscope and Leica Application Suite software, exciting at 490 nm and detecting at 520 nm.

7.2 - Chapter 3

7.2.1 - Peptide Synthesis

General

Unless otherwise noted all peptides were synthesised using solid phase Fmoc chemistry on a MultisynTech Syro I automated peptide synthesiser. Couplings were performed using 4 eq protected amino acid, 4 eq O-benzotriazole-N,N,N',N'-tetramethyl-uronium-hexafluorophosphate (HBTU), 4 eq hydroxybenzotriazole (HOBt) and 8 eq N,N-diisopropylethylamine (DIPEA) in N,N-dimethylformamide (DMF). Removal of Fmoc protecting groups was achieved using piperidine (20% in DMF) 3 x 10 min. Amino acids were dissolved in NMP at 0.69 M with the exception of arginine which was dissolved in DMF. DIPEA was prepared at 2 M in NMP and HBTU/HOBt at 0.45 M in DMF. Couplings proceeded for 25 min followed by washing with DMF and repetition of coupling. Following chain elongation on the automated synthesiser, the resin was washed with DMF (3 x 5 mL), DCM (3 x 5 mL), 1:1 DCM:MeOH (3 x 5 mL) and dried by suction.

F-14 Synthesis

14 mer peptide was synthesised using standard solid phase Fmoc peptide synthesis, cleaved from the resin using 95:2.5:2.5 TFA:H₂O:TIPS and purified using RP-HPLC as described previously (See Chapter 2). The fluorophore was then attached to the N-terminus of the peptide by addition of Triethylamine (0.1 μ L, 0.612 μ mol) 14 mer peptide (1 mg, 0.612 μ mol) and fluorescein isothiocyanate (0.5 mg, 1.224 μ mol) in DMSO (200 μ L). The solution was agitated for 1 hour after which the solvent was removed by lyophilisation. The crude product was purified by semi-preparative RP-HPLC using a gradient of 95:5 water:acetonitrile to 5:95 water:acetonitrile with 0.05% TFA additive over 45 min at a flow rate of 4 mL/min collecting at 214 nm. Fractions containing pure peptide as determined by RP-HPLC and MALDI-TOF mass spec were combined, the solvent removed under reduced pressure and lyophilised from water (0.05% TFA) to produce the desired product as a yellow powder. **MS (MALDI-TOF +ve):** calculated for: 2019.83 found: 2043.63 [M + Na]

TAT-16 and 16 mer synthesis

Peptides were synthesised on NovaSyn TGA resin (Merck, loading 0.26 mmol/g). Attachment of the first amino acid to 100 mg resin (0.026 mmol) was carried out manually on a shaker. A symmetrical anhydride of leucine was formed under an inert atmosphere at 0°C by dissolving Fmoc-Leu-OH (92 mg, 0.26 mmol) in DCM, followed by addition of DIC (20 μ L, 0.13 mmol)

dropwise. The reaction was stirred for 30 min before removal of solvent under reduced pressure. The anhydride was then taken up in a minimum amount of DMF and added to the pre-swollen resin along with a catalytic amount of DMAP. The reaction proceeded for 3 hours before capping the resin using acetic anhydride (5 μ L, 0.052 mmol) and pyridine (4 μ L, 0.052 mmol) in a minimum amount of DMF for 30 min. Chain elongation was achieved using standard solid phase Fmoc peptide synthesis, cleaved from the resin using 95:2.5:2.5 TFA:H₂O:TIPS and purified by reverse phase flash chromatography. Fractions containing pure peptide as determined by MALDI TOF and analytical HPLC were pooled and the solvent removed under reduced pressure. The peptides were lyophilised from water (0.05% TFA) to give the product as a white powder.

TAT-16 **MS (MALDI +ve)**: calculated for: 3397.8 found: 3400.7 [M + H]

16 mer **MS (MALDI +ve)**: calculated for: 1856.9 found: 1881.5 [M + Na]

10 mer synthesis

The peptide was synthesised on 2-chlorotriyl chloride resin (250 mg, Merck, loading: 1.01 mmol/g). Attachment of the first amino acid was carried out manually on a shaker. The resin was swollen in dry DCM, followed by addition of Fmoc-Pro-OH (340 mg, 1.01 mmol) and DIPEA (220 μ L, 1.26 mmol). The reaction proceeded for 2 hours after which the resin was capped using a mixture of DCM:MeOH:DIPEA 80:15:5 (10 mL) for 5 minutes, twice. Following chain elongation on the automated synthesiser, the resin was washed with DMF (3 x 5 mL), DCM (3 x 5 mL), 1:1 DCM:MeOH (3 x 5 mL) and dried by suction. The peptide was cleaved from the resin using a mixture of TFA:H₂O:TIPS 95:2.5:2.5 (5 mL) with agitation for 3 h. The peptide, dissolved in the cleavage mixture was separated from the resin by filtration and the resin washed with TFA (2 x 3 mL) which was added to the filtrate. The cleavage cocktail and volatile components were removed under reduced pressure to give a colourless gel. This product was then washed with ice cold diethyl ether (10 mL) resulting in a white solid. The peptide was isolated by reverse-phase chromatography using a Biotage Isolera Four using a 12 g C18 cartridge with a gradient of 95:5 water:methanol with a 0.05% TFA additive going to 5:95 water:methanol over 40 min with a flow rate of 20 mL/min, collecting at 214 nm. The peptide was further purified by semi-prep RP-HPLC using a gradient of 95:5 water:methanol to 5:95 water:methanol with 0.05% TFA additive over 15 min returning to 95:5 water:methanol over 5 min at a flow rate of 4 mL/min collecting at 214 nm. Fractions containing pure peptide as determined by RP-HPLC and MALDI-TOF mass spec were combined, the solvent removed under reduced pressure and lyophilised from water (0.05% TFA)

to produce the desired product as a white powder. 10 mer **MS (MALDI +ve)**: calculated for: 1148.5 found: 1171.3 [M + Na]

7.2.2 - Fluorescence Polarisation

General

All solutions were made using Milli-Q water. Fluorescence polarisation assays were carried out in black Costar low binding 96 well microplates. Tween20, monobasic and dibasic sodium phosphate were purchased from Sigma Aldrich. Fluorescence polarisation measurements were taken using a BMG Labtech Optima microplate reader equipped with a fluorescence polarisation optic, measuring at 490/520 nm. Data was averaged over 10 readings.

Plasmid transformation

pET15b plasmid containing the Keap1 Kelch domain gene, 325-609, was dissolved in water to a concentration of 10 ng/ μ L. 5 μ L plasmid solution was added to 50 μ L chemically competent E. Coli cells (Invitrogen BL21 Star (DE3)) and gently mixed by stirring. Plasmid and cells were then incubated on ice for 20 min. Cells were then heat shocked by placing the tube in at heat block at 42°C for 45 seconds before returning to ice for 2 min. 750 μ L sterile LB medium was added and the tube incubated at 37°C on a shaker for 1 hour. The cells were then streaked out on LB agar plates containing 100 μ g/mL ampicillin and grown in an incubator overnight at 37°C.

Protein expression

30 mL sterile LB media containing 0.1 mg/mL ampicillin was inoculated with a single colony from the plasmid transformation plates and incubated overnight at 37°C. Two flasks each containing 600 mL sterile LB medium and 0.1 mg/mL ampicillin were inoculated with 6 mL of the overnight culture and incubated at 30°C for 5 hours until the bacteria reached an O.D. of 0.6. Overexpression was induced by addition of 6 mL IPTG (100 mM) to give a final concentration of 1 mM and incubated at 30°C for 2.5 hours. Cells were then pelleted at 10000 rcf for 3 min at 4°C and stored at -80°C.

Protein Purification

Half of the cells pelleted in the previous step were resuspended in lysis buffer (50 mM NaH₂PO₄, 300 mM NaCl, 10% glycerol, 25 mM imidazole, 1 mg/mL lysozyme, pH 7.4, 20 mL) and incubated on ice for 30 min. The suspension was then sonicated on ice 10 x 15 seconds before pelleting at 10000 rcf for 30 min at 4°C. The supernatant was transferred to a clean tube and 5 g USB

PrepEase Histidine-Tagged High Yield Purification Resin added followed by incubation on a shaker for 15 min. Solution and resin was decanted into a fritted column and the supernatant eluted. The resin bed was washed with buffer under gravity (50 mM NaH₂PO₄, 300 mM NaCl, 10 % glycerol, 25 mM imidazole, pH 7.4, 3 x bed volume) and the protein eluted (50 mM NaH₂PO₄, 300 mM NaCl, 10 % glycerol, 250 mM imidazole, pH 7.4) in 8 x 2.5 mL fractions. Protein concentration and purity was assessed by SDS PAGE and fractions containing the majority of high purity protein pooled and stored at 4°C overnight. Prior to further purification by anion exchange the protein solution was diluted to 50 mM NaCl with 20 mM Tris buffer pH 8 to give a total volume of 45 mL. The 5 mL HiTrap Q HP strong anion exchange column was prepared according to manufacturer's instructions and the sample added by syringe. The column was then washed with 5 volumes 20 mM pH 8 Tris buffer and the sample eluted with increasing concentrations of NaCl in pH 8 Tris buffer from 100 mM to 1M in 100 mM steps 5 mL per step. Protein concentration and purity was assessed by SDS PAGE. A single fraction containing a high concentration of pure protein was desalted using a PD-10 desalting column according to manufacturer's instructions into buffer containing 10 mM pH 7.4 NaH₂PO₄, 10 mM β-mercapto ethanol and 10% glycerol. The protein was then concentrated by spin filtration to a final volume of 2 mL, separated into 100 μL aliquots and stored at -80°C. Protein concentration was determined by UV absorption at 280 nM giving a stock concentration of 26.7 μM.

Binding curve

Buffer solution was prepared containing 10 mM sodium phosphate at pH 7.4 and 0.1% Tween20. A serial dilution of Keap1 Kelch domain protein was prepared from 20 μM to 1 pM in buffer solution to give a 16 point dilution and kept on ice. F-14 peptide was diluted to a concentration of 5.56 nM in buffer solution and transferred to a microplate, 90 μL per well. 10 μL Keap1 dilution was then added to each well and mixed by pipetting to give a final concentration of 5 nM F-14 and a concentration range of 2 μM to 0.1 pM Keap1 in triplicate. Bubbles were removed from wells using compressed air and the plate was then incubated at room temperature for 15 min before reading. Fluorescence anisotropy readings were averaged over 10 measurements and analysed using non-linear regression with no cooperativity of binding. Values were normalised and expressed as percentage F-14 bound.

Z' test

Buffer solution was prepared containing 10 mM NaH₂PO₄ at pH 7.4 and 0.1% Tween20. A solution of 5.56 nM F-14 and 33.33 nM Keap1 was prepared in buffer. 90 μL per well of this solution was

transferred to a microplate, 96 wells in total. 14 mer peptide was diluted in buffer to a concentration of 20 μM and 10 μL added to each of 48 wells to give a final concentration of 5 nM F-14, 30 nM Keap1 and 2 μM 14 mer peptide. 10 μL buffer was added to the remaining 48 wells to give the same concentration without 14 mer peptide. Bubbles were removed from wells using compressed air and the plate was then incubated at room temperature for 15 min before reading. The plate was read 10 times and the data averaged. The data for each half was averaged and analysed using the Z' test for well to well variation where SD is the standard deviation of the anisotropy readings and mAu is the average anisotropy (Equation 2).

$$Z' = 1 - \frac{3SD_{Bound} - 3SD_{Free}}{mAu_{Bound} - mAu_{Free}}$$

Equation 3: Z' test for well-to-well variation

Compound screening

Buffer solution was prepared containing 10 mM sodium phosphate at pH 7.4 and 0.1% Tween20. A solution of 5.56 nM F-14 and 33.33 nM Keap1 was prepared in buffer and kept on ice. 90 μL per well of this solution was transferred to a microplate. Compounds were diluted in DMSO to 10 fold their final plate concentration. Compounds were transferred to the plate, 10 μL per well in triplicate and mixed by pipetting. Bubbles were removed from wells using compressed air and the plate was then incubated at room temperature for 15 min before reading 10 times. Replicates were averaged and analysed for significant inhibition by One-way ANOVA followed by Tukey's test for significance.

Inhibition curve

Buffer solution was prepared containing 10 mM sodium phosphate at pH 7.4 and 0.1% Tween20. A solution of 5.56 nM F-14 and 33.33 nM Keap1 was prepared in buffer and kept on ice. 90 μL per well of this solution was transferred to a microplate. Compounds were serially diluted in buffer in 10 fold steps, from 5 μM to 0.1 pM to give 16 data points. 10 μL compound solution was added to each well and mixed by pipetting. Bubbles were removed from wells using compressed air and the plate was then incubated at room temperature for 15 min before reading 10 times. Data was averaged for the 10 readings followed by averaging replicates. Data was analysed by non-linear regression assuming a Hill Slope of -1.0 to calculate IC_{50} values. K_i values were then calculated using the Cheng-Prusoff equation, with ligand concentration fixed at 5 nM and a K_d of 14.6 nM (Equation 4).

$$K_i = \frac{IC_{50}}{1 + \frac{[ligand]}{K_d}}$$

Equation 4: Cheng-Prusoff equation for conversion of IC_{50} to K_i for protein-ligand complexes

7.3 - Chapter 4

7.3.1 - Peptide Synthesis

Unless otherwise noted all peptides were synthesised using solid phase Fmoc chemistry on a MultisynTech Syro I automated peptide synthesiser. Couplings were performed using 4 eq protected amino acid, 4 eq O-benzotriazole-N,N',N'-tetramethyl-uronium-hexafluorophosphate (HBTU), 4 eq hydroxybenzotriazole (HOBT) and 8 eq N,N-diisopropylethylamine (DIPEA) in N,N-dimethylformamide (DMF) and with double coupling. Removal of Fmoc protecting groups was achieved using piperidine (20% in DMF) 3 x 10 min. Amino acids were dissolved in NMP at 0.69 M with the exception of arginine which was dissolved in DMF. DIPEA was prepared at 2 M in NMP and HBTU/HOBT at 0.45 M in DMF. Couplings proceeded for 25 min followed by washing with DMF and repetition of coupling.

General synthesis of N-terminally acetylated C-terminally amidated peptide linear sequences (Ds10, Ds8, Ds6, Ln10, Ln8, Ln6, Ar6, Ar8, Ar8P, Ar10P)

Peptides were synthesised on Rink Amide MBHA resin (Merck, 100 mg, loading: 0.64 mmol/g). After chain elongation on the automated synthesiser, the N-terminal amine was acetylated by addition of acetyl chloride (18 μ L, 0.26 mmol) and DIPEA (89 μ L, 0.51 mmol) in a minimum amount of DMF and agitated for 30 min. The resin was then washed with DMF (3 x 5 mL), DCM (3 x 5 mL), 1:1 DCM:MeOH (3 x 5 mL) and dried by suction. Peptides were cleaved from the resin using a mixture of TFA:H₂O:EDT:TIPS 94:2.5:2.5:1 (5 mL) with agitation for 3 h. The peptide, dissolved in the cleavage mixture was separated from the resin by filtration and the resin washed with TFA (2 x 3 mL) which was added to the filtrate. The cleavage cocktail and volatile components were removed under reduced pressure to give a colourless gel. This product was then washed with ice cold diethyl ether (10 mL) resulting in a white solid. Peptides were isolated by reverse-phase chromatography using a Biotage Isolera Four providing gross purification using a 12 g C18 cartridge with a gradient of 95:5 water:methanol with a 0.05% TFA additive going to 5:95 water:methanol over 40 min with a flow rate of 20 mL/min, collecting at 214 nm. Fractions

containing pure peptide as determined by RP-HPLC and MALDI-TOF mass spec were combined, the solvent removed under reduced pressure and lyophilised from water (0.05% TFA) to produce the desired products as white powders.

Ds10 linear **MS (MALDI +ve)**: calculated for: 1185.4 found: 1209.3 [M + Na]

Ds8 linear **MS (MALDI +ve)**: calculated for: 925.3 found: 949.4 [M + Na]

Ds6 linear **MS (MALDI +ve)**: calculated for: 681.2 found: 703.7 [M + Na]

Ln10 **MS (MALDI +ve)**: calculated for: 1153.5 found: 1178.0 [M + Na]

Ln6 **MS (MALDI +ve)**: calculated for: 649.2 found: 672.5 [M + Na]

Synthesis of N-terminally acetylated peptide linear sequence (Ds8-OH)

The peptide was synthesised manually on 2-chlorotrityl resin preloaded with H-Cys(Acm) (Merck, loading: 0.32 mmol/g). 300 mg resin (0.096 mmol) was swollen by agitating in DMF for 20 min. Chain elongation was effected by addition of the appropriate amino acid (4 eq, 0.384 mmol), HBTU (145 mg, 0.384 mmol), HOBT (59 mg, 0.384 mmol) and DIPEA (134 μ L, 0.768 mmol) in a minimum volume of DMF to the resin and agitating for 30 min. Following this the solution was removed by filtration and the resin washed with DMF (3 x 10 mL). Completion of the coupling was assessed by Kaiser test. Subsequently the N-terminal Fmoc group was removed using 20% piperidine in DMF (3 x 10 mL) for 10 min. The coupling process was then repeated for the next amino acid until the chain was completed. Following removal of the final N-terminal Fmoc protecting group the amine was acetylated using acetyl chloride (27 μ L, 0.384 mmol) and DIPEA (134 μ L, 0.768 mmol) in a minimum volume of DMF. The resin was agitated for 30 min, the solution removed by filtration and the resin washed with DMF 3 x 10 mL, DCM 3 x 10 mL and 1:1 DCM:MeOH 3 x 10 mL. The resin was then dried by compressed air. The peptide was cleaved from the resin using a mixture of TFA:H₂O:EDT:TIPS 94:2.5:2.5:1 (5 mL) with agitation for 3 h. The peptide, dissolved in the cleavage mixture was separated from the resin by filtration and the resin washed with TFA (2 x 3 mL) which was added to the filtrate. The cleavage cocktail and volatile components were removed under reduced pressure to give a colourless gel. This product was then washed with ice cold diethyl ether (10 mL) resulting in a white solid. The peptide was isolated by reverse-phase chromatography using a Biotage Isolera Four using a 12 g C18 cartridge with a gradient of 95:5 water:methanol with a 0.05% TFA additive going to 5:95 water:methanol over 40 min with a flow rate of 20 mL/min, collecting at 214 nm. Fractions containing pure peptide as determined by RP-HPLC and MALDI-TOF mass spec were combined, the solvent

removed under reduced pressure and lyophilised from water (0.05% TFA) to produce the desired product as a white powder. Ds8-OH linear **MS (MALDI +ve)**: calculated for: 1068.3 found: 1092.4 [M + Na]

Synthesis of C-terminally amidated peptide linear sequence (H₂N-Ds8)

The peptide was synthesised on Rink Amide MBHA resin (Merck, 100 mg, loading: 0.65 mmol/g). After chain elongation on the automated synthesiser, the resin was washed with DMF (3 x 5 mL), DCM (3 x 5 mL), 1:1 DCM:MeOH (3 x 5 mL) and dried by suction. The peptide was cleaved from the resin using a mixture of TFA:H₂O:EDT:TIPS 94:2.5:2.5:1 (5 mL) with agitation for 3 h. The peptide, dissolved in the cleavage mixture was separated from the resin by filtration and the resin washed with TFA (2 x 3 mL) which was added to the filtrate. The cleavage cocktail and volatile components were removed under reduced pressure to give a colourless gel. This product was then washed with ice cold diethyl ether (10 mL) resulting in a white solid. The peptide was isolated by reverse-phase chromatography using a Biotage Isolera Four using a 12 g C18 cartridge with a gradient of 95:5 water:methanol with a 0.05% TFA additive going to 5:95 water:methanol over 40 min with a flow rate of 20 mL/min, collecting at 214 nm. Fractions containing pure peptide as determined by RP-HPLC and MALDI-TOF mass spec were combined, the solvent removed under reduced pressure and lyophilised from water (0.05% TFA) to produce the desired product as a white powder. H₂N-Ds8 linear **MS (MALDI +ve)**: calculated for: 883.3 found: 907.2 [M + Na]

General procedure for formation of disulfide cyclised peptides (Ds10, Ds8, Ds6, Ds8-OH, H₂N-Ds8)

Peptides were dissolved in 80% acetic acid to a concentration of 2 mg/mL, to which solid iodine (10 eq) was added in a single portion. The solution was stirred at room temperature and the reaction monitored by RP-HPLC. In order to monitor the reaction aliquots of the mixture were taken, diluted 2 fold with water and washed with carbon tetrachloride to remove iodine. Samples were then filtered before analysis. Upon completion the reaction mixture was diluted 2 fold with water and washed with carbon tetrachloride (10 x 20 mL) until no colour was present in the carbon tetrachloride. The solvent was removed under reduced pressure and the resulting colourless gel was purified using reverse-phase chromatography on a Biotage Isolera Four. Fractions containing pure peptide as determined by RP-HPLC and MALDI-TOF mass spec were combined, the solvent removed under reduced pressure and lyophilised from water (0.05% TFA) to produce the desired products as white powders.

Ds10 **MS (MALDI +ve)**: calculated for: 1183.4 found: 1206.2 [M + Na]

Ds8 **MS (MALDI +ve)**: calculated for: 923.3 found: 946.9 [M + Na]

Ds6 **MS (MALDI +ve)**: calculated for: 679.2 found: 701.8 [M + Na]

Ds8-OH **MS (MALDI +ve)**: calculated for: 924.3 found: 947.6 [M + Na]

H₂N-Ds8 **MS (MALDI +ve)**: calculated for: 881.3 found: 905.5 [M + Na]

General synthesis of head-to-tail cyclised peptides (Ht10, Ht9)

Peptides were synthesised on Rink Amide MBHA resin (Merck, 100 mg, loading: 0.65 mmol/g). The first residue (Fmoc-Glu-ODMAB) was attached via its side chain. Following chain elongation on the automated synthesiser the resin was washed with DMF (3 x 5 mL) and the DMAB protecting group cleaved with 2% hydrazine monohydrate in DMF (3 x 10 min). The resin was washed with 1:1 MeOH:H₂O (3 x 5 mL) and complete removal of aminobenzyl alcohol was ensured by shaking overnight in a 5 mM NaOH solution of 1:1 MeOH:H₂O. The resin was washed with 1:1 MeOH:H₂O (3 x 5 mL) followed by anhydrous DMF (3 x 5 mL). Cyclisation was achieved with PyBOP (4 eq, 0.26 mmol) and DIPEA (8 eq, 0.52 mmol) in a minimum volume of anhydrous DMF. The resin was shaken at room temperature for 72 h, with fresh coupling agents every 12 h. The progress of the reaction was monitored by Kaiser test and upon completion the resin was washed with DMF (3 x 5 mL), DCM (3 x 5 mL), 1:1 DCM:MeOH (3 x 5 mL) and dried by suction. Peptides were cleaved from the resin using a mixture of TFA:H₂O:TIPS 95:2.5:2.5 (5 mL) with agitation for 3 h. The peptide, dissolved in the cleavage mixture was separated from the resin by filtration and the resin washed with TFA (2 x 3 mL) which was added to the filtrate. The cleavage cocktail and volatile components were removed under reduced pressure to give a colourless gel. This product was then washed with ice cold diethyl ether (10 mL) resulting in a white solid. Peptides were isolated by reverse-phase chromatography using a Biotage Isolera using a 12 g C18 cartridge with a gradient of 95:5 water:methanol with a 0.05% TFA additive going to 5:95 water:methanol over 40 min with a flow rate of 20 mL/min, collecting at 214 nm. Fractions containing pure peptide as determined by RP-HPLC and MALDI-TOF mass spec were combined, the solvent removed under reduced pressure and lyophilised from water (0.05% TFA) to produce the desired products as white powders.

Ht10 **MS (ESI -ve)**: calculated for: 1161.52 found: 1161.01 [M - H]

Ht9 **MS (ESI -ve)**: calculated for: 1048.43 found: 1047.93 [M - H]

General synthesis of perfluoroaryl stapled peptides (Ar6, Ar8, Ar8P, Ar10P)

Peptides were dissolved in a solution of hexafluorobenzene (25 eq, 100 mM) in DMF. An equal volume of TRIS base (50 mM) in DMF was added, the vessel agitated and then left at room temperature for 18 h. The solvent was removed under reduced pressure and the peptides purified by RP-HPLC. Fractions containing pure product as determined by MALDI-TOF mass spec were combined, the solvent removed under reduced pressure and lyophilised from water (0.05% TFA) to produce the desired products as white powders.

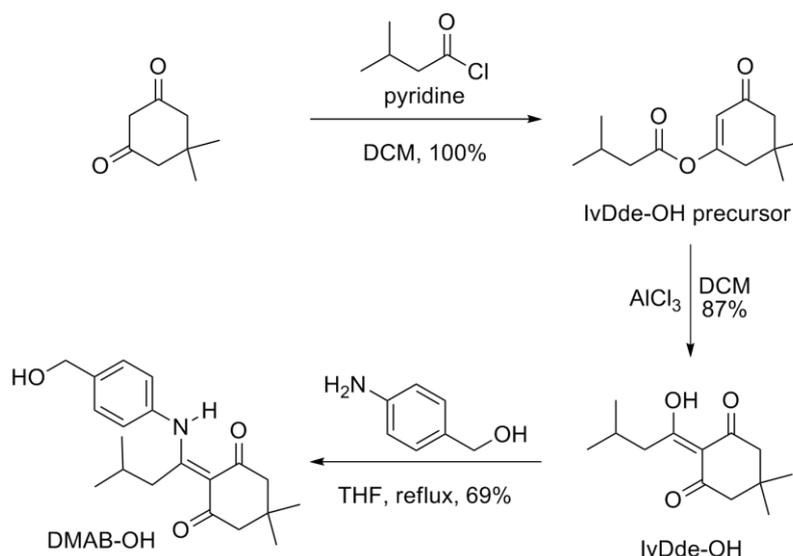
Ar6 **MS (MALDI +ve)**: calculated for: 827.19 found: 851.02 [M + Na]

Ar8 **MS (MALDI +ve)**: calculated for: 1071.26 found: 1095.25 [M + Na]

Ar8P **MS (MALDI +ve)**: calculated for: 1039.27 found: 1063.39 [M + Na]

Ar10P **MS (MALDI +ve)**: calculated for: 1299.42 found: 1324.15 [M + Na]

7.3.2 - Small Molecule Synthesis



Scheme 7.1: Synthesis of the DMAB-OH protecting group

IvDde-OH precursor¹⁷⁵

Dimedone (14 g, 100 mmol) was dissolved in DCM (150 mL) followed by addition of pyridine (9.7 mL, 120 mmol). To this solution isovaleryl chloride (14.8 mL, 120 mmol) in DCM (40 mL) was added dropwise over 10 min. The reaction was stirred at room temperature for 18h with monitoring by TLC. Upon completion the reaction was washed with aqueous HCl (2 x 50 mL, 2 M), H₂O (50 mL) and saturated sodium bicarbonate solution (50 mL) before drying over magnesium

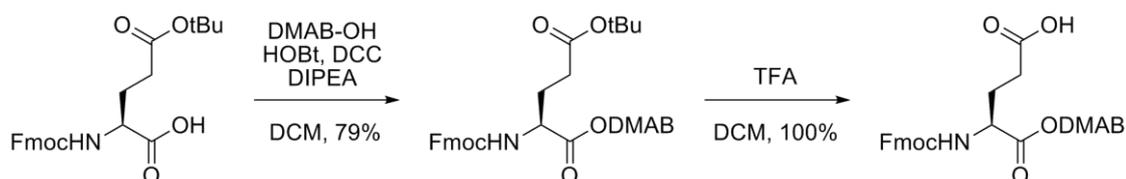
sulfate, filtration and removal of solvent under reduced pressure. The product was obtained as a yellow oil in quantitative yield and was used without further purification. $^1\text{H NMR}$ (400 MHz, CDCl_3): δ_{H} 5.89 (s, 1 H, Dde C(-O)=CH), 2.41 (s, 2H, Dde C(-O)-CH₂), 2.33 (d, J = 6.8 Hz, 2H, lv CH₂), 2.27 (s, 2H, Dde C(=O)CH₂), 2.14 (sep, J = 6.8 Hz, 1 H, lv CH), 1.11 (s, 6H, Dde 2 x CH₃), 1.00 (d, J = 6.4 Hz, 6H, lv 2 x CH₃). NMR data matches literature values.

lvDde-OH¹⁷⁵

To a suspension of AlCl_3 (16 g, 120 mmol) in dry DCM (150 mL) in an ice bath, lvDde-OH precursor (12 g, 53 mmol) in DCM (40 mL) was added via a dropping funnel over 1 h. Following addition the reaction was stirred at room temperature for 2 h until complete consumption of starting material was seen by TLC. The reaction mixture was then poured into a mixture of 37% HCl and ice (150 g) in an ice bath ensuring the temperature did not exceed 5°C. After the addition of brine (200 mL) the DCM was run off and the aqueous layer extracted with DCM (8 x 50 mL) before drying the organic layer over magnesium sulfate and removal of solvent under reduced pressure. Filtration of the crude product through a silica plug using Et_2O and subsequent removal of solvent afforded the product in 87% yield (10.5 g) as a yellow oil. $^1\text{H NMR}$ (400 MHz, CDCl_3): δ_{H} 2.91 (d, J = 6 Hz, 2H, lv CH₂), 2.53 (s, 2H Dde CH₂), 2.35 (s, 2H, Dde CH₂), 2.13 (sep, J = 6.8 Hz, 1H lv CH), 1.07 (s, 6H, Dde 2 x CH₃), 0.98 (d, J = 7.6 Hz, 6H, lv 2 x CH₃). NMR data matches literature values.

DMAB-OH¹⁶⁸

lvDde-OH (5 g, 22 mmol) in THF (50 mL) was added to a solution of 4-aminobenzyl alcohol (3.29 g, 26 mmol) in THF (100 mL). The mixture was refluxed at 90°C for 60 h followed by removal of solvent under reduced pressure. The crude mixture was purified by flash chromatography using a gradient of hexanes → ethyl acetate to afford the product as a yellow solid (69%, 5 g) along with unreacted lvDde-OH (1.25 g, 6 mmol). $^1\text{H NMR}$ (400 MHz, CDCl_3): δ_{H} 15.24 (s, 1H, NH), 7.43 (d, J = 8 Hz, 2H, 2 x CH Bn Ar), 7.12 (d, J = 8 Hz, 2H, 2 x CH Bn Ar), 4.75 (d, J = 8 Hz, 2H, Bn CH₂), 3.00 (d, J = 6 Hz, 2H, lv CH₂), 2.49 (s, 2H, Dde CH₂), 2.39 (s, 2H, Dde CH₂), 1.84 (sep, J = 6.9 Hz, 1H, lv CH), 1.08 (s, 6H, Dde 2 x CH₃), 0.77 (d, J = 6.7 Hz, 6H, lv 2 x CH₃). NMR data matches literature values.



Scheme 7.2: Synthesis of DMAB protected glutamic acid

Fmoc-Glu(^tBu)-ODMAB¹⁷⁶

DIPEA (1.15 mL, 6.72 mmol) was added to a solution of DMAB-OH (3.2 g, 9.6 mmol), Fmoc-Glu(^tBu)-OH (4.92 g 11.5 mmol) and HOBt (1.28 g, 9.6 mmol) in DCM (80 mL) followed by DCC (3.97 g, 19.2 mmol) in DCM (15 mL) dropwise over 10 min. The reaction was stirred at room temperature for 18 h after which it was filtered and the solvent removed under reduced pressure. The crude product was purified by flash chromatography using a gradient of hexanes → ethyl acetate to afford the product as a white solid (79%, 5.5 g). ¹H NMR (400 MHz, CDCl₃): δ_H 15.29 (s, 1H, NH), 7.76, 7.40, 7.30 (m, 8H, Fmoc Ar CH), 7.59 (d, J = 8 Hz, 2H, 2 x CH Bn Ar), 7.11 (d, J = 8 Hz, 2H, 2 x CH Bn Ar), 5.52 (d, J = 8 Hz, 1H, Glu NH), 5.21 (m, 2H, Bn CH₂), 4.45 (m, 1H, Glu C^αH), 4.40 (m, 2H, Fmoc CH₂), 4.21 (m, 1H, Fmoc CH), 2.99 (d, J = 4 Hz, 2H, Iv CH₂), 2.49 (s, 2H, Dde CH₂), 2.39 (s, 2H, Dde CH₂), 2.33 (m, 2H, Glu C^γH₂), 2.19, 1.98 (2m, 2H, Glu C^βH₂), 1.84 (m, 1H, Iv CH), 1.44 (s, 9H, ^tBu), 1.08 (s, 6H, Dde 2 x CH₃), 0.76 (d, J = 8 Hz, 6H, Iv 2 x CH₃). NMR data matches literature values.

Fmoc-Glu-ODMAB¹⁷⁶

Fmoc-Glu(^tBu)-ODMAB (1.6 g, 2.2 mmol) was dissolved in DCM:TFA 1:1 (20 mL) and stirred at room temperature for 1 h. The solvent was then removed under reduced pressure and the product coevaporated with MeOH (3 x 20 mL). This afforded the product as a yellow solid (86%, 1.3 g). ¹H NMR (400 MHz, CDCl₃): δ_H 15.29 (s, 1H, NH), 7.76, 7.40, 7.31 (m, 8H, Fmoc Ar CH), 7.58 (d, J = 8 Hz, 2H, 2 x CH Bn Ar), 7.12 (d, J = 8 Hz, 2H, 2 x CH Bn Ar), 5.49 (d, J = 8 Hz, 1H, Glu NH), 5.22 (s, 2H, Bn CH₂), 4.49 (m, 1H, Glu C^αH), 4.41 (m, 2H, Fmoc CH₂), 4.20 (m, 1H, Fmoc CH), 2.99 (d, J = 4 Hz, 2H, Iv CH₂), 2.50 (s, 4H, Dde 2 x CH₂), 2.43 (m, 2H, Glu C^γH₂), 2.26, 2.00 (2m, 2H, Glu C^βH₂), 1.83 (m, 1H, Iv CH), 1.09 (s, 6H, Dde 2 x CH₃), 0.76 (d, J = 8 Hz, 6H, Iv 2 x CH₃). NMR data matches literature values. **MS (ESI +ve)**: calculated for: 680.31 found: 681.32 [M + H]

7.3.3 - Fluorescence Polarisation

General

All solutions were made using Milli-Q water. Fluorescence polarisation assays were carried out in black Costar low binding 96 well microplates. Tween20, monobasic and dibasic sodium phosphate were purchased from Sigma Aldrich. Fluorescence polarisation measurements were taken using a BMG Labtech Optima microplate reader equipped with a fluorescence polarisation optic, measuring at 490/520 nm. Data was averaged over 10 readings.

Inhibition curve

Buffer solution was prepared containing 10 mM sodium phosphate at pH 7.4 and 0.1% Tween20. A solution of 5.56 nM F-14 and 33.33 nM Keap1 was prepared in buffer and kept on ice. 90 μ L per well of this solution was transferred to a microplate. Compounds were serially diluted in buffer in 10 fold steps, from 1 mM to 0.1 nM to give 8 data points. 10 μ L compound solution was added to each well and mixed by pipetting. Bubbles were removed from wells using compressed air and the plate was then incubated at room temperature for 15 min before reading 10 times. Data was averaged for the 10 readings followed by averaging replicates. Data was analysed by non-linear regression assuming a Hill Slope of -1.0 to calculate IC_{50} values. K_i values were then calculated using the Cheng-Prusoff equation, with ligand concentration fixed at 5 nM and a K_d of 14.6 nM.

7.3.4 - Cell Biology

HO-1 ELISA

HO-1 protein levels were determined using an R&D Systems DuoSet IC Human Total HO-1/HMOX1 assay kit. Cells were spun down, the supernatant removed and washed twice with PBS. The cell pellet was then resuspended in lysis buffer (0.5% Triton X-100, Pierce protease cocktail in PBS) to a concentration of 6.66×10^6 cells/mL and kept on ice for 15 min before storage at -80°C . Plates were prepared by incubation with 100 μ L of 8 μ g/mL capture antibody per well over night at RT. The wells were emptied and washed with Wash Buffer (PBS + 0.05% Tween20) three times. Wells were then blocked with 300 μ L Block Buffer (PBS + 1% BSA + 0.05% sodium azide) for 2 hours at RT. Standards were prepared by serial dilution of an HO-1 protein stock (10 ng/mL) in two fold steps in IC Diluent #4 (1 mM EDTA, 0.5% Triton X-100 in PBS) to give a 7 point curve from 10 ng/mL to 0.156 ng/mL. Samples were defrosted, spun to pellet any undissolved material and diluted 2 fold in IC Diluent #4 to give an effective concentration of 3.33×10^5 cells/mL. The plate was washed as before and 100 μ L standards and samples were added using IC Diluent #4 as the zero concentration. The plate was then sealed and incubated for 2 hours at RT. The plate was washed and 100 μ L detection antibody (200 ng/mL) diluted in IC Diluent #1 (1% BSA in PBS) added, sealed and incubated for 2 hours at RT. Following another wash step, 100 μ L Streptavidin-HRP conjugate was added per well diluted as directed in IC Diluent #1 and incubated for 20 min at RT in the dark. The plate was washed and 100 μ L of a 1:1 solution of Color Reagent A (H_2O_2) and Color reagent B (tetramethylbenzidine) added to each well. The plate was then incubated for 20 min at RT in the dark. 50 μ L Stop Solution (1M H_2SO_4) was added to each well and the absorbance (450 nm) measured using a microplate reader.

The standards were analysed using a 4 parameter logistic curve fit, from which sample concentrations were calculated. Fold inductions were then determined by comparison to a non-treated control sample.

7.4 - Chapter 5

7.4.1 - In Silico Screening

Docking studies

Data sets for NCI Diversity Set II and Chembridge Building Blocks were obtained from zinc.docking.org pre-prepared in the required .pdbqt format.²⁰⁴ Receptors were acquired from www.pdb.org.²⁰⁸ Counter ions, ligands and solvent molecules were removed using DeepView 4.0.1²⁰³ and prepared for docking by addition of polar hydrogens and calculation of gasteiger charges using AutoDock Tools 1.5.6.¹⁹⁹ The peptide used for initial validation was extracted from the crystal structure 2FLU and prepared by removing residues around the key binding motif using DeepView and polar hydrogens and gasteiger charges added using AutoDock Tools. A grid box within which molecules would be docked was defined with sides of 30 Å and centred in the Nrf2 binding pocket. Docking calculations were performed using AutoDock Vina with default parameters.²⁰¹ Initial docking studies were conducted with the protein portion of crystal structure 2FLU as the receptor.

Analysis of docking results

The top 30 hits, those with highest binding affinity as calculated using the internal free energy scoring function, were extracted from the data sets and visualised with AutoDock Tools. Hits were then assessed for binding mode in AutoDock Tools to ensure they bound within the same pocket as the Nrf2 peptide. Intermolecular hydrogen bonds, pi-pi interactions and pi-cation interactions were then calculated for each hit. The 10 hits with highest affinity and appropriate binding mode from each set were selected for further investigation.

Ligand and receptor screening

After initial studies screening against a single crystal structure of the Keap1 Kelch Domain (2FLU) the screening was extended to multiple receptors. The eight reported crystal structures available at the time (1U6D, 1X2J, 1X2R, 1ZGK, 2DYH, 2FLU, 2Z32, 3ADE) were prepared as before by removing solvent and ligands and adding polar hydrogens and gasteiger charges. The Chembridge Building Blocks library was then screened against each of these receptors using AutoDock Vina.

Analysis of docking results

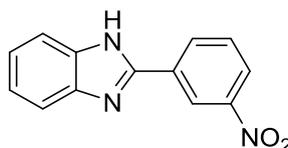
The top 30 hits as determined by highest binding affinity for each receptor were extracted and combined to give a data set of combined hits. Only the top 30 highest affinity hits were extracted

to ensure that frequently occurring but low affinity hits were excluded. These 240 hits were then sorted by frequency, those that occurred most commonly being given the highest score. Among those hits, those with equal frequency were sorted by binding affinity to give a set of the most commonly high binding ligands. The top 10 hits were assessed for appropriate binding mode and protein ligand interactions with their highest affinity receptor as before, in AutoDock Tools and then selected for further investigation.

Script for sorting top hits

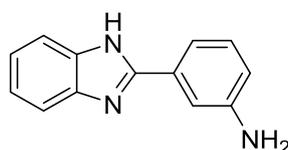
```
#!/bin/bash
# search all files ending in .pdbqt
for f in *.pdbqt
# create a variable $b which is equal to everything before the first _ in the filename
do b="{f%%_}"
# find all files which contain $b at the beginning with anything following
# then use word count to return how many times $b has been found
o=$(find . -name "$b*" | wc -l)
# first for every file containing $b extract characters 14-17 from the filename
# in this case the PDB I.D. for the receptor
# then extract the second line from the file and print the fourth word
# in this case the energy of docking
e=$(for f in $b*.pdbqt
do n=$(echo $f | cut -c 14-17)
echo $n ,
sed -n 2p $f | awk '{print $4 ","}'
done)
# put all of the pieces collected so far into a file called Sort_top.txt
echo $b, $o, $e >> Sort_top.txt
done
# sort through the file Sort_top.txt and delete any duplicate lines and save
# to Sort_top_uniq.txt
uniq Sort_top.txt > Sort_top_uniq.txt
```

7.4.2 - Small Molecule Synthesis



2-(3-nitrophenyl)-1H-1,3-benzodiazole (2)

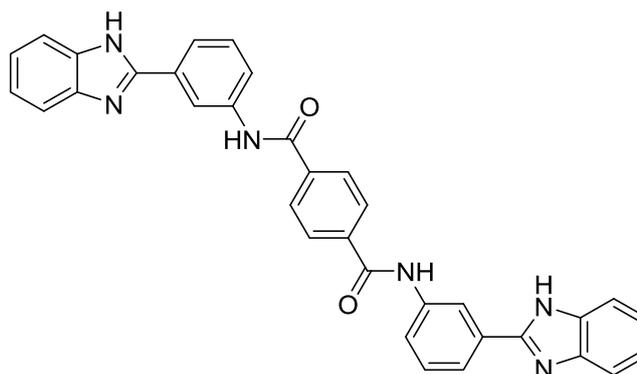
3-nitrobenzoyl chloride (2.27g, 12 mmol) in THF (20 mL) was added dropwise over 15 min to a solution of 1,2-diaminobenzene (1.62 g, 15 mmol) and triethylamine (2 mL, 15 mmol) in THF (80 mL) at 0°C. The reaction mixture was stirred at room temperature for 4 h before being filtered and the solvent removed under reduced pressure. The residue was then taken up in ethyl acetate and the organic solution washed with water and brine. Before drying over MgSO₄ the organic phase was filtered to isolate the disubstituted anilide. The solvent was then removed under reduced pressure to afford intermediate benzyl amide **1** which was used without further purification. A solution of compound **1** in acetic acid (50 mL) was refluxed at 135°C for 4 h. The solution was then allowed to cool to room temperature and poured over ice (50 g) before extracting the aqueous mixture with DCM (4 x 50 mL) and drying over MgSO₄. Removal of the solvent under reduced pressure afforded benzimidazole **2** as a beige solid (2.0 g, 56%). **MS (ESI -ve)**: calculated for: 239.1 found: 237.9 [M - H]. **¹H NMR (400 MHz, DMSO-d₆)**: δ – 13.29 (bs, 1H, NH), 8.98-8.97 (m, 1H, mABA CH), 8.59-8.56 (m, 1H, mABA CH), 8.31-8.28 (m, 1H, mABA CH), 7.84-7.80 (m, 1H, mABA CH), 7.66-7.61 (m, 2H, 2 x Bnlm CH), 7.27-7.22 (m, 2H, 2 x Bnlm CH); **¹³C NMR (100 MHz, DMSO-d₆)**: δ – 172.31 (mABA C-NO₂), 149.21 (mABA C), 148.49 (lm C), 132.61 (mABA CH), 131.78 (2 x Bnlm C), 130.84 (mABA CH), 126.32 (2 x Bnlm CH), 124.37 (mABA CH), 122.93 (2 x Bnlm CH), 120.98 (mABA CH).



3-(1H-1,3-benzodiazol-2-yl)aniline (3)

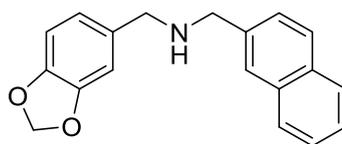
A catalytic amount of palladium on charcoal was added to a solution of benzimidazole **2** (500 mg, 2 mmol) in ethanol (10 mL) under nitrogen. Hydrogen gas was then bubbled through the mixture before stirring for 2 h under a hydrogen atmosphere. The reaction mixture was filtered, the solvent removed under reduced pressure and the residue purified by column chromatography (DCM to Ethyl Acetate) to afford aniline **3** (355 mg, 81%). **MS (ESI -ve)**: calculated for: 209.1

found: 208.0 [M - H]. ¹H NMR (400 MHz, DMSO-d₆): δ – 7.61-7.59 (m, 2H, 2 x Bnlm CH), 7.51-7.50 (m, 1H, mABA CH), 7.38-7.36 (m, 1H, mABA CH), 7.38-7.36 (m, 1H, mABA CH), 7.23-7.18 (m, 3H, 2 x Bnlm CH, mABA CH), 6.77-6.74 (m, 1H, mABA CH), 4.25 (bs, 2H, NH₂); ¹³C NMR (100 MHz, DMSO-d₆): δ – 152.60 (mABA C), 149.49 (mABA C-NH₂), 139.72 (lm C), 131.10 (2 x Bnlm C), 129.94 (mABA CH), 122.47 (2 x Bnlm CH), 116.23 (mABA CH), 115.37 (2 x Bnlm CH), 114.72 (mABA CH), 112.45 (mABA CH).



1-N,4-N-bis[3-(1H-1,3-benzodiazol-2-yl)phenyl]benzene-1,4-dicarboxamide (4) (NCI_61610)

Terephthaloyl chloride (65 mg, 0.323 mmol) in anhydrous THF (2 mL) was added dropwise under nitrogen to a solution of aniline **3** (135 mg, 0.645 mmol) in anhydrous THF (6 mL). Triethylamine (100 μL, 0.968 mmol) was added to this solution and the mixture stirred for 2 h. The solvent was then removed under reduced pressure and the residue taken up in DMF. Upon addition of water a cream precipitate formed which was filtered, washed with methanol and dried to afford NCI_61610 as a cream solid (120 mg, 34%). **MS (MALDI-TOF +ve)**: calculated for: 548.20 found: 549.12 [M + H]. ¹H NMR (400 MHz, DMSO-d₆): δ – 13.01 (bs, 2H, 2 x lm NH), 10.66 (s, 2H, 2 x Am NH), 8.76 (m, 2H, 2 x mABA CH), 8.19 (s, 4H, 4 x Trpt CH), 7.93-7.90 (m, 4H, 4 x mABA CH), 7.63-7.55 (m, 6H, 4 x Bnlm CH, 2 x mABA CH), 7.24-7.21 (m, 4H, 4 x Bnlm CH); ¹³C NMR (100 MHz, DMSO-d₆): δ – 165.53 (2 x Trpt CO), 151.61 (2 x mABA C), 139.96 (2 x lm C), 137.81 (2 x mABA C-NH), 131.06 (4 x Bnlm C), 129.85 (2 x mABA CH), 128.34 (4 x Trpt CH), 122.55 (4 x Bnlm CH), 122.37 (4 x mABA CH), 119.46 (4 x Bnlm CH, 2 x mABA CH).



(2H-1,3-benzodioxol-5-ylmethyl)(naphthalen-2-ylmethyl)amine (5) (CB_5560378)

Piperonyl amine (200 μ L, 1.6 mmol) and 1-naphthaldehyde (73 μ L, 0.535 mmol) in MeOH (2 mL) was stirred over MgSO_4 for 1 h before addition of NaBH_4 (60 mg, 1.6 mmol). The reaction mixture was then stirred at room temperature for 12 h before filtration and removal the solvent under reduced pressure. The residue was taken up in DCM (20 mL), washed with 1 M NaOH (3 x 10 mL), 1 M HCl (3 x 10 mL), dried over MgSO_4 and the solvent removed under reduced pressure. The crude product was then dissolved in diethyl ether and purified by filtration through silica and removal of the solvent under reduced pressure to afford the desired secondary amine **5** in 47% Yield. **MS (ESI +ve)**: calculated for: 291.1 found: 291.9 [M + H]. **$^1\text{H NMR}$ (400 MHz, CDCl_3)**: δ – 8.14 (m, 1H, Nap), 7.90 (m, 1H, Nap), 7.81 (m, 1H, Nap), 7.54 (m, 3H, Nap), 7.47 (m, 1H, Nap), 6.96 (m, 1H, Pip), 6.84 (m, 2H, Pip), 5.96 (s, 2H, O-CH₂-O), 4.25 (s, 2H, NH-CH₂-Nap), 3.85 (s, 2H, NH-CH₂-Pip), 1.79 ppm (bs, 1H, NH); **$^{13}\text{C NMR}$ (100 MHz, CDCl_3)**: δ – 147.8 (Ar-O), 146.6 (Ar-O), 135.9 (Nap), 134.4 (Ar-CH₂-NH), 134.0 (Ar-CH₂-NH), 131.9 (Nap), 128.8 (CH Nap), 127.8 (CH Nap), 126.2 (CH Nap), 126.1 (CH Nap), 125.7 (CH Nap), 125.5 (CH Nap), 123.8 (CH Nap), 121.4 (CH Pip), 108.9 (CH Pip), 108.1 (CH Pip), 101.0 (O-C-O), 53.6 (NH-CH₂-Ar), 50.8 (NH-CH₂-Ar).

Chapter 8: References

- (1) Ward, P. A. *Fundamentals of Inflammation*; Serhan, C. N.; Ward, P. A.; Gilroy, D. W., Eds.; 1st ed.; Cambridge University Press, 2010; pp. 1–16.
- (2) Nathan, C. *Nature* **2002**, *420*, 846–852.
- (3) Medzhitov, R. *Nature* **2008**, *454*, 428–435.
- (4) Thelen, M.; Stein, J. *Nat. Immunol.* **2008**, *9*, 953–959.
- (5) Serhan, C. N.; Savill, J. *Nat. Immunol.* **2005**, *6*, 1191–1197.
- (6) Cohen, J. *Nature* **2002**, *420*, 885–891.
- (7) DeCross, A. J.; Marshall, B. J. *Am. J. Med. Sci.* **1993**, *306*, 381–392.
- (8) Dinarello, C. A. *Cell* **2010**, *140*, 935–950.
- (9) McColl, A.; Michlewska, S.; Dransfield, I.; Rossi, A. G. *ScientificWorldJournal.* **2007**, *7*, 1165–1181.
- (10) Lundberg, I. E.; Grundtman, C.; Larsson, E.; Klareskog, L. *Best Pract. Res. Clin. Rheumatol.* **2004**, *18*, 7–19.
- (11) Whitehouse, M. W. *Inflammopharmacology* **2011**, *19*, 1–19.
- (12) Dinarello, C. A. *J. Rheumatol. Suppl.* **2005**, *74*, 40–47.
- (13) Lee, J.-M.; Li, J.; Johnson, D. a; Stein, T. D.; Kraft, A. D.; Calkins, M. J.; Jakel, R. J.; Johnson, J. A. *FASEB J.* **2005**, *19*, 1061–1066.
- (14) Bryan, H. K.; Olayanju, A.; Goldring, C. E.; Park, B. K. *Biochem. Pharmacol.* **2013**, *85*, 705–717.
- (15) Reuter, S.; Gupta, S. C.; Chaturvedi, M. M.; Aggarwal, B. B. *Free Radic. Biol. Med.* **2010**, *49*, 1603–1616.
- (16) Chen, X.-L.; Varner, S. E.; Rao, A. S.; Grey, J. Y.; Thomas, S.; Cook, C. K.; Wasserman, M. a; Medford, R. M.; Jaiswal, A. K.; Kunsch, C. *J. Biol. Chem.* **2003**, *278*, 703–711.
- (17) Adenuga, D.; Caito, S.; Yao, H.; Sundar, I. K.; Hwang, J.-W.; Chung, S.; Rahman, I. *Biochem. Biophys. Res. Commun.* **2010**, *403*, 452–456.
- (18) Wruck, C. J.; Streetz, K.; Pavic, G.; Götz, M. E.; Tohidnezhad, M.; Brandenburg, L.-O.; Varoga, D.; Eickelberg, O.; Herdegen, T.; Trautwein, C.; Cha, K.; Kan, Y. W.; Pufe, T. *J. Biol. Chem.* **2011**, *286*, 4493–4499.
- (19) Paine, A.; Eiz-Vesper, B.; Blasczyk, R.; Immenschuh, S. *Biochem. Pharmacol.* **2010**, *80*, 1895–1903.

- (20) Ryter, S.; Alam, J.; Choi, A. *Physiol. Rev.* **2006**, *86*, 583–650.
- (21) Yachie, A.; Niida, Y.; Wada, T.; Igarashi, N.; Kaneda, H.; Toma, T.; Ohta, K.; Kasahara, Y.; Koizumi, S. *J. Clin. Invest.* **1999**, *103*, 129–135.
- (22) Poss, K. D.; Tonegawa, S. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 10925–10930.
- (23) Wakabayashi, N.; Itoh, K.; Wakabayashi, J.; Motohashi, H.; Noda, S.; Takahashi, S.; Imakado, S.; Kotsuji, T.; Otsuka, F.; Roop, D. R.; Harada, T.; Engel, J. D.; Yamamoto, M. *Nat. Genet.* **2003**, *35*, 238–245.
- (24) Schäfer, M.; Farwanah, H.; Willrodt, A.-H.; Huebner, A. J.; Sandhoff, K.; Roop, D.; Hohl, D.; Bloch, W.; Werner, S. *EMBO Mol. Med.* **2012**, *4*, 364–379.
- (25) Lau, A.; Villeneuve, N. F.; Sun, Z.; Wong, P. K.; Zhang, D. D. *Pharmacol. Res.* **2008**, *58*, 262–270.
- (26) Wang, X.-J.; Sun, Z.; Villeneuve, N. F.; Zhang, S.; Zhao, F.; Li, Y.; Chen, W.; Yi, X.; Zheng, W.; Wondrak, G. T.; Wong, P. K.; Zhang, D. D. *Carcinogenesis* **2008**, *29*, 1235–1243.
- (27) Cho, J.-M.; Manandhar, S.; Lee, H.-R.; Park, H.-M.; Kwak, M.-K. *Cancer Lett.* **2008**, *260*, 96–108.
- (28) Motohashi, H.; Katsuoka, F.; Engel, J. D.; Yamamoto, M. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 6379–6384.
- (29) Nioi, P.; McMahon, M.; Itoh, K.; Yamamoto, M.; Hayes, J. D. *Biochem. J.* **2003**, *374*, 337–48.
- (30) Venugopal, R.; Jaiswal, a K. *Oncogene* **1998**, *17*, 3145–3156.
- (31) Alam, J. *J. Biol. Chem.* **1999**, *274*, 26071–26078.
- (32) Suzuki, T.; Motohashi, H.; Yamamoto, M. *Trends Pharmacol. Sci.* **2013**, *34*, 340–346.
- (33) Itoh, K.; Wakabayashi, N.; Katoh, Y.; Ishii, T.; Igarashi, K.; Engel, J. D.; Yamamoto, M. *Genes Dev.* **1999**, *13*, 76–86.
- (34) Motohashi, H.; Yamamoto, M. *Trends Mol. Med.* **2004**, *10*, 549–557.
- (35) Soltysik-Espanola, M.; Rogers, R. A.; Jiang, S.; Kim, T. A.; Gaedigk, R.; White, R. A.; Avraham, H.; Avraham, S. *Mol. Biol. Cell* **1999**, *10*, 2361–2375.
- (36) Zipper, L. M.; Mulcahy, R. T. *J. Biol. Chem.* **2002**, *277*, 36544–36552.
- (37) Kobayashi, M.; Itoh, K.; Suzuki, T.; Osanai, H.; Nishikawa, K.; Katoh, Y.; Takagi, Y.; Yamamoto, M. *Genes to Cells* **2002**, *7*, 807–820.
- (38) Li, X.; Zhang, D.; Hannink, M.; Beamer, L. J. *J. Biol. Chem.* **2004**, *279*, 54750–54758.
- (39) Padmanabhan, B.; Tong, K. I.; Ohta, T.; Nakamura, Y.; Scharlock, M.; Ohtsuji, M.; Kang, M.-I.; Kobayashi, A.; Yokoyama, S.; Yamamoto, M. *Mol. Cell* **2006**, *21*, 689–700.

- (40) Lo, S.-C.; Li, X.; Henzl, M. T.; Beamer, L. J.; Hannink, M. *EMBO J.* **2006**, *25*, 3605–3617.
- (41) Katoh, Y.; Iida, K.; Kang, M.-I.; Kobayashi, A.; Mizukami, M.; Tong, K. I.; McMahon, M.; Hayes, J. D.; Itoh, K.; Yamamoto, M. *Arch. Biochem. Biophys.* **2005**, *433*, 342–350.
- (42) Tong, K.; Katoh, Y.; Kusunoki, H. *Mol. Cell. Biol.* **2006**, *26*, 2887–2900.
- (43) Zhang, D.; Lo, S. *Mol. Cell. Biol.* **2004**, *24*, 10941–10953.
- (44) Ogura, T.; Tong, K. I.; Mio, K.; Maruyama, Y.; Kurokawa, H.; Sato, C.; Yamamoto, M. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 2842–2847.
- (45) Miseta, A.; Csutora, P. *Mol. Biol. Evol.* **2000**, *17*, 1232–1239.
- (46) Dinkova-Kostova, A. T.; Holtzclaw, W. D.; Cole, R. N.; Itoh, K.; Wakabayashi, N.; Katoh, Y.; Yamamoto, M.; Talalay, P. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 11908–11913.
- (47) Baird, L.; Llères, D.; Swift, S.; Dinkova-Kostova, A. T. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 15259–15264.
- (48) Fahey, J. W.; Talalay, P. *Food Chem. Toxicol.* **1999**, *37*, 973–979.
- (49) Hong, F.; Freeman, M. L.; Liebler, D. C. *Chem. Res. Toxicol.* **2005**, *18*, 1917–1926.
- (50) Talalay, P.; Fahey, J. *J. Nutr.* **2001**, *131*, 3027S–3033S.
- (51) Ramos-Gomez, M.; Kwak, M. K.; Dolan, P. M.; Itoh, K.; Yamamoto, M.; Talalay, P.; Kensler, T. W. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 3410–3415.
- (52) O'Dwyer, P. J.; Szarka, C. E.; Yao, K. S.; Halbherr, T. C.; Pfeiffer, G. R.; Green, F.; Gallo, J. M.; Brennan, J.; Frucht, H.; Goosenberg, E. B.; Hamilton, T. C.; Litwin, S.; Balshem, A. M.; Engstrom, P. F.; Clapper, M. L. *J. Clin. Invest.* **1996**, *98*, 1210–1217.
- (53) Szarka, C. E.; Yao, K. S.; Pfeiffer, G. R.; Balshem, A. M.; Litwin, S.; Frucht, H.; Goosenberg, E. B.; Engstrom, P. F.; Clapper, M. L.; O'Dwyer, P. J. *Cancer Detect. Prev.* **2001**, *25*, 352–361.
- (54) Jacobson, L.; Zhang, B.; Zhu, Y.; Wang, J. *Cancer Epidemiol. Biomarkers Prev.* **1997**, *6*, 257–265.
- (55) Ansher, S. S.; Dolan, P.; Bueding, E. *Hepatology* **1983**, *3*, 932–935.
- (56) Lam, S.; MacAulay, C.; Le Riche, J. C.; Dyachkova, Y.; Coldman, A.; Guillaud, M.; Hawk, E.; Christen, M.-O.; Gazdar, A. F. *J. Natl. Cancer Inst.* **2002**, *94*, 1001–1009.
- (57) Egner, P. A.; Kensler, T. W.; Prestera, T.; Talalay, P.; Libby, A. H.; Joyner, H. H.; Curphey, T. J. *Carcinogenesis* **1994**, *15*, 177–181.
- (58) Zhang, Y.; Munday, R. *Mol. Cancer Ther.* **2008**, *7*, 3470–3479.
- (59) Hong, F.; Sekhar, K. R.; Freeman, M. L.; Liebler, D. C. *J. Biol. Chem.* **2005**, *280*, 31768–31775.

- (60) Hur, W.; Sun, Z.; Jiang, T.; Mason, D. E.; Peters, E. C.; Zhang, D. D.; Luesch, H.; Schultz, P. G.; Gray, N. S. *Chem. Biol.* **2010**, *17*, 537–547.
- (61) Sekhar, K.; Spitz, D.; Harris, S. *Free Radic. Biol. Med.* **2002**, *32*, 650–662.
- (62) Sekhar, K.; Crooks, P.; Sonar, V. *Cancer Res.* **2003**, *63*, 5636–5645.
- (63) Jeong, W.-S.; Keum, Y.-S.; Chen, C.; Jain, M. R.; Shen, G.; Kim, J.-H.; Li, W.; Kong, A.-N. T. *J. Biochem. Mol. Biol.* **2005**, *38*, 167–176.
- (64) Aggarwal, B.; Ichikawa, H. *Cell Cycle* **2005**, *4*, 1201–1215.
- (65) Ernst, I. M. A.; Schuemann, C.; Wagner, A. E.; Rimbach, G. *Free Radic. Res.* **2011**, *45*, 941–949.
- (66) Balogun, E.; Hoque, M.; Gong, P.; Killeen, E.; Green, C. J.; Foresti, R.; Alam, J.; Motterlini, R. *Biochem. J.* **2003**, *371*, 887–895.
- (67) Nishinaka, T.; Ichijo, Y.; Ito, M.; Kimura, M.; Katsuyama, M.; Iwata, K.; Miura, T.; Terada, T.; Yabe-Nishimura, C. *Toxicol. Lett.* **2007**, *170*, 238–247.
- (68) Thompson, D. C.; Trush, M. A. *Toxicol. Appl. Pharmacol.* **1988**, *96*, 122–131.
- (69) Abiko, Y.; Miura, T.; Phuc, B. H.; Shinkai, Y.; Kumagai, Y. *Toxicol. Appl. Pharmacol.* **2011**, *255*, 32–39.
- (70) Na, H.-K.; Surh, Y.-J. *Biochem. Pharmacol.* **2003**, *66*, 1381–1391.
- (71) Itoh, K.; Mochizuki, M.; Ishii, Y.; Ishii, T.; Shibata, T.; Kawamoto, Y.; Kelly, V.; Sekizawa, K.; Uchida, K.; Yamamoto, M. *Mol. Cell. Biol.* **2004**, *24*, 36–45.
- (72) Hosoya, T.; Maruyama, A.; Kang, M.-I.; Kawatani, Y.; Shibata, T.; Uchida, K.; Warabi, E.; Noguchi, N.; Itoh, K.; Yamamoto, M. *J. Biol. Chem.* **2005**, *280*, 27244–27250.
- (73) Liby, K. T.; Yore, M. M.; Sporn, M. B. *Nat Rev Cancer* **2007**, *7*, 357–369.
- (74) Liby, K.; Hock, T.; Yore, M.; Suh, N. *Cancer Res.* **2005**, *65*, 4789–4798.
- (75) Dinkova-Kostova, A. T.; Liby, K. T.; Stephenson, K. K.; Holtzclaw, W. D.; Gao, X.; Suh, N.; Williams, C.; Risingsong, R.; Honda, T.; Gribble, G. W.; Sporn, M. B.; Talalay, P. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 4584–4589.
- (76) Dinkova-Kostova, A. T.; Talalay, P.; Sharkey, J.; Zhang, Y.; Holtzclaw, W. D.; Wang, X. J.; David, E.; Schiavoni, K. H.; Finlayson, S.; Mierke, D. F.; Honda, T. *J. Biol. Chem.* **2010**, *285*, 33747–33755.
- (77) De Zeeuw, D.; Akizawa, T.; Audhya, P.; Bakris, G. L.; Chin, M.; Christ-Schmidt, H.; Goldsberry, A.; Houser, M.; Krauth, M.; Heerspink, H. J. L.; McMurray, J. J.; Meyer, C. J.; Parving, H.-H.; Remuzzi, G.; Toto, R. D.; Vaziri, N. D.; Wanner, C.; Wittes, J.; Wrolstad, D.; Chertow, G. M. *N. Engl. J. Med.* **2013**, 1–12.

- (78) Ivanov, A. A.; Khuri, F. R.; Fu, H. Targeting protein-protein interactions as an anticancer strategy. *Trends Pharmacol. Sci.* **2013**, *34*, 393–400.
- (79) Wells, J. A.; McClendon, C. L. *Nature* **2007**, *450*, 1001–1009.
- (80) Sperandio, O.; Reynès, C. H.; Camproux, A.-C.; Villoutreix, B. O. *Drug Discov. Today* **2010**, *15*, 220–229.
- (81) Arkin, M. R.; Randal, M.; DeLano, W. L.; Hyde, J.; Luong, T. N.; Oslob, J. D.; Raphael, D. R.; Taylor, L.; Wang, J.; McDowell, R. S.; Wells, J. A.; Braisted, A. C. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 1603–1608.
- (82) He, M. M.; Smith, A. S.; Oslob, J. D.; Flanagan, W. M.; Braisted, A. C.; Whitty, A.; Cancilla, M. T.; Wang, J.; Lugovskoy, A. A.; Yoburn, J. C.; Fung, A. D.; Farrington, G.; Eldredge, J. K.; Day, E. S.; Cruz, L. A.; Cachero, T. G.; Miller, S. K.; Friedman, J. E.; Choong, I. C.; Cunningham, B. C. *Science* **2005**, *310*, 1022–1025.
- (83) Arkin, M. R.; Whitty, A. *Curr. Opin. Chem. Biol.* **2009**, *13*, 284–290.
- (84) Busschots, K.; De Rijck, J.; Christ, F.; Debyser, Z. *Mol. Biosyst.* **2009**, *5*, 21–31.
- (85) Rudin, C. M.; Hann, C. L.; Garon, E. B.; Ribeiro de Oliveira, M.; Bonomi, P. D.; Camidge, D. R.; Chu, Q.; Giaccone, G.; Khaira, D.; Ramalingam, S. S.; Ranson, M. R.; Dive, C.; McKeegan, E. M.; Chyla, B. J.; Dowell, B. L.; Chakravartty, A.; Nolan, C. E.; Rudersdorf, N.; Busman, T. A.; Mabry, M. H.; Krivoschik, A. P.; Humerickhouse, R. A.; Shapiro, G. I.; Gandhi, L. *Clin. Cancer Res.* **2012**, *18*, 3163–3169.
- (86) Murray, J. K.; Gellman, S. H. *Biopolymers* **2007**, *88*, 657–686.
- (87) Kussie, P. H.; Gorina, S.; Marechal, V.; Elenbaas, B.; Moreau, J.; Levine, A. J.; Pavletich, N. P. *Science* **1996**, *274*, 948–953.
- (88) Klein, C.; Vassilev, L. T. *Br. J. Cancer* **2004**, *91*, 1415–1419.
- (89) Duncan, S. J.; Grünschow, S.; Williams, D. H.; McNicholas, C.; Purewal, R.; Hajek, M.; Gerlitz, M.; Martin, S.; Wrigley, S. K.; Moore, M. *J. Am. Chem. Soc.* **2001**, *123*, 554–560.
- (90) Grasberger, B. L.; Lu, T.; Schubert, C.; Parks, D. J.; Carver, T. E.; Koblisch, H. K.; Cummings, M. D.; LaFrance, L. V.; Milkiewicz, K. L.; Calvo, R. R.; Maguire, D.; Lattanze, J.; Franks, C. F.; Zhao, S.; Ramachandren, K.; Bylebyl, G. R.; Zhang, M.; Manthey, C. L.; Petrella, E. C.; Pantoliano, M. W.; Deckman, I. C.; Spurlino, J. C.; Maroney, A. C.; Tomczuk, B. E.; Molloy, C. J.; Bone, R. F. *J. Med. Chem.* **2005**, *48*, 909–912.
- (91) Vassilev, L. T.; Vu, B. T.; Graves, B.; Carvajal, D.; Podlaski, F.; Filipovic, Z.; Kong, N.; Kammlott, U.; Lukacs, C.; Klein, C.; Fotouhi, N.; Liu, E. *Science* **2004**, *303*, 844–848.
- (92) Ding, K.; Lu, Y.; Nikolovska-Coleska, Z.; Qiu, S.; Ding, Y.; Gao, W.; Stuckey, J.; Krajewski, K.; Roller, P. P.; Tomita, Y.; Parrish, D. A.; Deschamps, J. R.; Wang, S. *J. Am. Chem. Soc.* **2005**, *127*, 10130–10131.

- (93) Zhao, J.; Wang, M.; Chen, J.; Luo, A.; Wang, X.; Wu, M.; Yin, D.; Liu, Z. *Cancer Lett.* **2002**, *183*, 69–77.
- (94) Zauli, G.; Celeghini, C.; Melloni, E.; Voltan, R.; Ongari, M.; Tiribelli, M.; di Iasio, M. G.; Lanza, F.; Secchiero, P. *Haematologica* **2012**, *97*, 1722–1730.
- (95) Eggler, A. L.; Liu, G.; Pezzuto, J. M.; van Breemen, R. B.; Mesecar, A. D. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 10070–10075.
- (96) Zhao, J.; Redell, J. B.; Moore, A. N.; Dash, P. K. *Biochem. Biophys. Res. Commun.* **2011**, *407*, 501–506.
- (97) Hancock, R.; Bertrand, H. C.; Tsujita, T.; Naz, S.; El-Bakry, A.; Laoruchupong, J.; Hayes, J. D.; Wells, G. *Free Radic. Biol. Med.* **2012**, *52*, 444–451.
- (98) Hancock, R.; Schaap, M.; Pfister, H.; Wells, G. *Org. Biomol. Chem.* **2013**, *11*, 3553–3557.
- (99) Hu, L.; Magesh, S.; Chen, L.; Wang, L.; Lewis, T. A.; Chen, Y.; Khodier, C.; Inoyama, D.; Beamer, L. J.; Emge, T. J.; Shen, J.; Kerrigan, J. E.; Kong, A.-N. T.; Dandapani, S.; Palmer, M.; Schreiber, S. L.; Munoz, B. *Bioorg. Med. Chem. Lett.* **2013**, *23*, 3039–3043.
- (100) Marcotte, D.; Zeng, W.; Hus, J.-C.; McKenzie, A.; Hession, C.; Jin, P.; Bergeron, C.; Lugovskoy, A.; Enyedy, I.; Cuervo, H.; Wang, D.; Atmanene, C.; Roecklin, D.; Vecchi, M.; Vivat, V.; Kraemer, J.; Winkler, D.; Hong, V.; Chao, J.; Lukashev, M.; Silvian, L. *Bioorg. Med. Chem.* **2013**, *21*, 4011–4019.
- (101) Sun, H.-P.; Jiang, Z.-Y.; Zhang, M.-Y.; Lu, M.-C.; Yang, T.-T.; Pan, Y.; Huang, H.-Z.; Zhang, X.-J.; You, Q. *Med. Chem. Commun.* **2014**, *5*, 93–98.
- (102) Smirnova, N. A.; Haskew-Layton, R. E.; Basso, M.; Hushpulian, D. M.; Payappilly, J. B.; Speer, R. E.; Ahn, Y.-H.; Rakhman, I.; Cole, P. a; Pinto, J. T.; Ratan, R. R.; Gazaryan, I. G. *Chem. Biol.* **2011**, *18*, 752–765.
- (103) Brasseur, R.; Divita, G. *Biochim. Biophys. Acta* **2010**, *1798*, 2177–2181.
- (104) Jones, A. T.; Sayers, E. J. *J. Control. Release* **2012**, *161*, 582–591.
- (105) Johnson, R. M.; Harrison, S. D.; Maclean, D. *Cell-Penetrating Peptides*; Langel, Ü., Ed.; Humana Press: Totowa, NJ, 2011; Vol. 683, pp. 535–551.
- (106) Frankel, A.; Pabo, C. *Cell* **1988**, *55*, 1189–1193.
- (107) Lundberg, M.; Wikström, S.; Johansson, M. *Mol. Ther.* **2003**, *8*, 143–150.
- (108) Richard, J. P.; Melikov, K.; Vives, E.; Ramos, C.; Verbeure, B.; Gait, M. J.; Chernomordik, L. V.; Lebleu, B. *J. Biol. Chem.* **2003**, *278*, 585–590.
- (109) Joliot, A.; Pernelle, C.; Deagostini-Bazin, H.; Prochiantz, A. *Proc. Natl. Acad. Sci. U. S. A.* **1991**, *88*, 1864–1868.
- (110) Vivès, E.; Brodin, P.; Lebleu, B. *J. Biol. Chem.* **1997**, *272*, 16010–16017.

- (111) Derossi, D.; Joliot, A. H.; Chassaing, G.; Prochiantz, A. *J. Biol. Chem.* **1994**, *269*, 10444–10450.
- (112) Schwarze, S. R.; Ho, A.; Vocero-Akbani, A.; Dowdy, S. F. *Science* **1999**, *285*, 1569–1572.
- (113) Wender, P. A.; Mitchell, D. J.; Pattabiraman, K.; Pelkey, E. T.; Steinman, L.; Rothbard, J. B. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 13003–13008.
- (114) Futaki, S.; Suzuki, T.; Ohashi, W.; Yagami, T.; Tanaka, S.; Ueda, K.; Sugiura, Y. *J. Biol. Chem.* **2001**, *276*, 5836–5840.
- (115) Futaki, S.; Goto, S.; Sugiura, Y. *J. Mol. Recognit.* **2003**, *16*, 260–264.
- (116) Nakase, I.; Tadokoro, A.; Kawabata, N.; Takeuchi, T.; Katoh, H.; Hiramoto, K.; Negishi, M.; Nomizu, M.; Sugiura, Y.; Futaki, S. *Biochemistry* **2007**, *46*, 492–501.
- (117) Futaki, S.; Nakase, I.; Tadokoro, A.; Takeuchi, T.; Jones, A. T. *Biochem. Soc. Trans.* **2007**, *35*, 784–787.
- (118) Richard, J. P.; Melikov, K.; Brooks, H.; Prevot, P.; Lebleu, B.; Chernomordik, L. V. *J. Biol. Chem.* **2005**, *280*, 15300–15306.
- (119) Sugita, T.; Yoshikawa, T.; Mukai, Y.; Yamanada, N.; Imai, S.; Nagano, K.; Yoshida, Y.; Shibata, H.; Yoshioka, Y.; Nakagawa, S.; Kamada, H.; Tsunoda, S.-I.; Tsutsumi, Y. *Br. J. Pharmacol.* **2008**, *153*, 1143–1152.
- (120) Sheppard, R. C. In *Fmoc Solid Phase Peptide Synthesis: A Practical Approach*; Chan, W. C.; White, P. D., Eds.; Oxford University Press, 2000; pp. 1–8.
- (121) Jones, J. H. *J. Pept. Sci.* **2007**, *13*, 363–367.
- (122) Merrifield, R. *J. Am. Chem. Soc.* **1963**, *85*, 2149–2154.
- (123) Merrifield, R. *J. Am. Chem. Soc.* **1964**, *86*, 304–305.
- (124) Carpino, L. A.; Han, G. Y. *J. Org. Chem.* **1972**, *37*, 3404–3409.
- (125) Kaiser, E.; Colescott, R. L.; Bossinger, C. D.; Cook, P. I. *Anal. Biochem.* **1970**, *34*, 595–598.
- (126) Atherton, E.; Clive, D. L. J.; Sheppard, R. C. *J. Am. Chem. Soc.* **1975**, *97*, 6584–6585.
- (127) Thaler, A.; Seebach, D.; Cardinaux, F. *Helv. Chim. Acta* **1991**, *74*, 628–643.
- (128) Fields, G. B.; Fields, C. G. *J. Am. Chem. Soc.* **1991**, *113*, 4202–4207.
- (129) Hyde, C.; Johnson, T.; Sheppard, R. C. *J. Chem. Soc. Chem. Commun.* **1992**, 1573–1575.
- (130) Quibell, M.; Johnson, T. In *Fmoc Solid Phase Peptide Synthesis: A Practical Approach*; Chan, W. C.; White, P. D., Eds.; Oxford University Press, 2000; pp. 115–135.

- (131) Nicolás, E.; Pujades, M.; Bacardit, J.; Giralt, E.; Albericio, F. *Tetrahedron Lett.* **1997**, *38*, 2317–2320.
- (132) Tsuchiya, S.; Yamabe, M.; Yamaguchi, Y.; Kobayashi, Y.; Konno, T.; Tada, K. *Int. J. Cancer* **1980**, *26*, 171–176.
- (133) Rushworth, S. A.; Ogborne, R. M.; Charalambos, C. A.; O’Connell, M. A. *Biochem Biophys Res Commun* **2006**, *341*, 1007–1016.
- (134) Inoyama, D.; Chen, Y.; Huang, X.; Beamer, L. J.; Kong, A.-N. T.; Hu, L. *J. Biomol. Screen.* **2012**, *17*, 435–447.
- (135) Rushworth, S. A.; Chen, X. L.; Mackman, N.; Ogborne, R. M.; O’Connell, M. A. *J Immunol* **2005**, *175*, 4408–4415.
- (136) Rushworth, S. A.; MacEwan, D. J.; O’Connell, M. A. *J. Immunol.* **2008**, *181*, 6730–6737.
- (137) Jones, S. W.; Christison, R.; Bundell, K.; Voyce, C. J.; Brockbank, S. M. V; Newham, P.; Lindsay, M. a *Br. J. Pharmacol.* **2005**, *145*, 1093–102.
- (138) Wegrowski, Y.; Milard, A.-L.; Kotlarz, G.; Toulmonde, E.; Maquart, F.-X.; Bernard, J. *Clin. Exp. Immunol.* **2006**, *144*, 485–493.
- (139) Langelaan, D. N.; Ngweniform, P.; Rainey, J. K. *Biochem. Cell Biol.* **2011**, *89*, 98–105.
- (140) PanVera *Fluorescence Polarization - Technical Resource Guide*; 4th ed.; Invitrogen, 2006; pp. 1–118.
- (141) Moerke, N. J. *Curr. Protoc. Chem. Biol.* **2009**, *1*, 1–15.
- (142) Rossi, A. M.; Taylor, C. W. *Nat. Protoc.* **2011**, *6*, 365–387.
- (143) Cox, J. H.; Dean, R. A.; Roberts, C. R.; Overall, C. M. *J. Biol. Chem.* **2008**, *283*, 19389–19399.
- (144) Cheng, Y.; Prusoff, W. H. *Biochem. Pharmacol.* **1973**, *22*, 3099–3108.
- (145) Moulton, H. M.; Nelson, M. H.; Hatlevig, S. A.; Reddy, M. T.; Iversen, P. L. *Bioconjug. Chem.* **2004**, *15*, 290–299.
- (146) Schust, J.; Berg, T. *Anal. Biochem.* **2004**, *330*, 114–118.
- (147) Liskamp, R.; Rijkers, D. T. S.; Bakker, S. *Modern Supramolecular Chemistry: Strategies for Macrocyclic Synthesis*; F, D.; PJ, S.; RR, T., Eds.; WILEY-VCH Verlag GmbH, 2008; pp. 1–28.
- (148) Horswill, A.; Benkovic, S. *Cell Cycle* **2005**, *4*, 552–555.
- (149) Fletcher, S.; Hamilton, A. D. *Interface J. R. Soc.* **2006**, *3*, 215–33.
- (150) Boger, D. L. *Med. Res. Rev.* **2001**, *21*, 356–381.
- (151) Lunn, R. M. *Report on Carcinogens Twelfth Edition*; 2011; p. 125.

- (152) Choc, M. G. *Int. J. Dermatol.* **1997**, *36 Suppl 1*, 1–6.
- (153) Clark, R. C.; Lee, S. Y.; Searcey, M.; Boger, D. L. *Nat. Prod. Rep.* **2009**, *26*, 465–477.
- (154) Woon, E. C. Y.; Arcieri, M.; Wilderspin, A. F.; Malkinson, J. P.; Searcey, M. *J. Org. Chem.* **2007**, *72*, 5146–5151.
- (155) Bernal, F.; Tyler, A. F.; Korsmeyer, S. J.; Walensky, L. D.; Verdine, G. L. *J. Am. Chem. Soc.* **2007**, *129*, 2456–2457.
- (156) Machida, K.; Mayer, B. J. *Biochim. Biophys. Acta* **2005**, *1747*, 1–25.
- (157) Gao, Y.; Voigt, J.; Wu, J. X.; Yang, D.; Burke, T. R. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 1889–1892.
- (158) Shi, Z.-D.; Lee, K.; Liu, H.; Zhang, M.; Roberts, L. R.; Worthy, K. M.; Fivash, M. J.; Fisher, R. J.; Yang, D.; Burke, T. R. *Biochem. Biophys. Res. Commun.* **2003**, *310*, 378–383.
- (159) Annis, I.; Hargittai, B.; Barany, G. *Methods Enzymol.* **1997**, *289*, 198–221.
- (160) Galande, A. K.; Bramlett, K. S.; Trent, J. O.; Burris, T. P.; Wittliff, J. L.; Spatola, A. F. *ChemBioChem* **2005**, *6*, 1991–1998.
- (161) Hynes, R. O. *Cell* **1992**, *69*, 11–25.
- (162) Gurrath, M.; Müller, G.; Kessler, H.; Aumailley, M.; Timpl, R. *Eur. J. Biochem.* **1992**, *210*, 911–921.
- (163) Xiong, J.-P.; Stehle, T.; Zhang, R.; Joachimiak, A.; Frech, M.; Goodman, S. L.; Arnaout, M. A. *Science* **2002**, *296*, 151–155.
- (164) Davies, J. S. *J. Pept. Sci.* **2003**, *9*, 471–501.
- (165) Claerhout, S.; Ermolat'ev, D. S.; Van der Eycken, E. V. *J. Comb. Chem.* **2008**, *10*, 580–585.
- (166) Alberico, F.; Annis, I.; Royo, M.; Barany, G. In *Fmoc Solid Phase Peptide Synthesis: A Practical Approach*; Chan, W. C.; White, P. D., Eds.; Oxford University Press, 2000; pp. 83–85.
- (167) Grieco, P.; Gitu, P. M.; Hruby, V. J. *J. Pept. Res.* **2001**, *57*, 250–256.
- (168) Chan, W. C.; Bycroft, B. W.; Evans, D. J.; White, P. D. *J. Chem. Soc. Chem. Commun.* **1995**, 2209.
- (169) Gellerman, G.; Elgavi, A.; Salitra, Y.; Kramer, M. *J. Pept. Res.* **2001**, *57*, 277–291.
- (170) Carpino, L. A. *J. Am. Chem. Soc.* **1993**, *115*, 4397–4398.
- (171) Coste, J.; Le-Nguyen, D.; Castro, B. *Tetrahedron Lett.* **1990**, *31*, 205–208.
- (172) Henchey, L. K.; Jochim, A. L.; Arora, P. S. *Curr. Opin. Chem. Biol.* **2008**, *12*, 692–697.

- (173) Blackwell, H. E.; Grubbs, R. H. *Angew. Chemie Int. Ed.* **1998**, *37*, 3281–3284.
- (174) Sun, T.-L.; Sun, Y.; Lee, C.-C.; Huang, H. *Biophys. J.* **2013**, *104*, 1923–1932.
- (175) Hillman, J. D.; Orugunty, R. S.; Smith, J. L. DIFFERENTIALLY PROTECTED ORTHOGONAL LANTHIONINE TECHNOLOGY. U.S. Patent 7,521,592.
- (176) Hossany, B. R.; Johnston, B. D.; Wen, X.; Borrelli, S.; Yuan, Y.; Johnson, M. A.; Pinto, B. M. *Carbohydr. Res.* **2009**, *344*, 1412–1427.
- (177) Mellor, S. L.; Wellings, D. A.; Fehrentz, J.-A.; Paris, M.; Martinez, J.; Ede, N. J.; Bray, A. M.; Evans, D. J.; Bloomberg, G. B. In *Fmoc Solid Phase Peptide Synthesis: A Practical Approach*; Chan, W. C.; White, P. D., Eds.; Oxford University Press, 2000; pp. 137–181.
- (178) White, P. D.; Chan, W. C. In *Fmoc Solid Phase Peptide Synthesis: A Practical Approach*; Chan, W. C.; White, P. D., Eds.; Oxford University Press, 2000; pp. 9–37.
- (179) Gausepohl, H.; Pieleles, U.; Frank, R. W. In *Peptides: Chemistry & Biology*; Smith, J. A.; Rivier, J. E., Eds.; ESCOM, 1992; pp. 523–524.
- (180) Zhu, J.; Marchant, R. E. *J. Pept. Sci.* **2008**, *14*, 690–696.
- (181) Spokoyny, A. M.; Zou, Y.; Ling, J. J.; Yu, H.; Lin, Y.-S.; Pentelute, B. L. *J. Am. Chem. Soc.* **2013**, *135*, 5946–5949.
- (182) Shoichet, B. K. *Nature* **2004**, *432*, 862–865.
- (183) Jorgensen, W. L. *Science* **2004**, *303*, 1813–1818.
- (184) Gillet, V. J.; Willett, P.; Bradshaw, J.; Green, D. V. S. *J. Chem. Inf. Model.* **1999**, *39*, 169–177.
- (185) Shoichet, B. K.; McGovern, S. L.; Wei, B.; Irwin, J. J. *Curr. Opin. Chem. Biol.* **2002**, *6*, 439–446.
- (186) Blake, J. F.; Laird, E. R. In *Annual reports in medicinal chemistry*; Doherty, A., Ed.; Elsevier Inc., 2003; Vol. Volume 38, pp. 305–314.
- (187) Bowman, A. L.; Nikolovska-Coleska, Z.; Zhong, H.; Wang, S.; Carlson, H. A. *J. Am. Chem. Soc.* **2007**, *129*, 12809–12814.
- (188) Almerico, A. M.; Tutone, M.; Lauria, A. *J. Mol. Model.* **2009**, *15*, 349–355.
- (189) Levy, Y.; Cho, S. S.; Onuchic, J. N.; Wolynes, P. G. *J. Mol. Biol.* **2005**, *346*, 1121–1145.
- (190) Wei, B. Q.; Weaver, L. H.; Ferrari, A. M.; Matthews, B. W.; Shoichet, B. K. *J. Mol. Biol.* **2004**, *337*, 1161–1182.
- (191) McGovern, S. L.; Shoichet, B. K. *J. Med. Chem.* **2003**, *46*, 1478–1483.
- (192) May, A.; Zacharias, M. *Biochim. Biophys. Acta* **2005**, *1754*, 225–231.

- (193) Gonzalez-Ruiz, D.; Gohlke, H. *Curr. Med. Chem.* **2006**, *13*, 2607–2625.
- (194) Marco, E.; Gago, F. *ChemMedChem* **2007**, *2*, 1388–1401.
- (195) Pearlman, D. A.; Charifson, P. S. *J. Med. Chem.* **2001**, *44*, 3417–3423.
- (196) Biesiada, J.; Porollo, A.; Velayutham, P.; Kouril, M.; Meller, J. *Hum. Genomics* **2011**, *5*, 497–505.
- (197) Lang, P. T.; Brozell, S. R.; Mukherjee, S.; Pettersen, E. F.; Meng, E. C.; Thomas, V.; Rizzo, R. C.; Case, D. A.; James, T. L.; Kuntz, I. D. *RNA* **2009**, *15*, 1219–1230.
- (198) Zavodszky, M. I.; Rohatgi, A.; Van Voorst, J. R.; Yan, H.; Kuhn, L. A. *J. Mol. Recognit.* **2009**, *22*, 280–292.
- (199) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. *J. Comput. Chem.* **2009**, *30*, 2785–2791.
- (200) Chang, M. W.; Ayeni, C.; Breuer, S.; Torbett, B. E. *PLoS One* **2010**, *5*, e11955.
- (201) Trott, O.; Olson, A. J. *J. Comput. Chem.* **2010**, *31*, 455–461.
- (202) Kukol, A. *Eur. J. Med. Chem.* **2011**, *46*, 4661–4664.
- (203) Guex, N.; Peitsch, M. C. *Electrophoresis* **1997**, *18*, 2714–2723.
- (204) Irwin, J. J.; Sterling, T.; Mysinger, M. M.; Bolstad, E. S.; Coleman, R. G. *J. Chem. Inf. Model.* **2012**, *52*, 1757–1768.
- (205) Acharya, K. R.; Lloyd, M. D. *Trends Pharmacol. Sci.* **2005**, *26*, 10–14.
- (206) Suenaga, A.; Okimoto, N.; Hirano, Y.; Fukui, K. *PLoS One* **2012**, *7*, e42846.
- (207) Steel, R.; Cowan, J.; Payerne, E.; O’Connell, M. A.; Searcey, M. *ACS Med. Chem. Lett.* **2012**, *3*, 407–410.
- (208) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235–242.

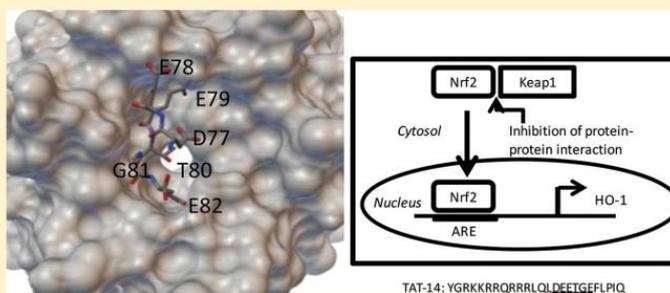
Appendix 1: Published Work

Anti-inflammatory Effect of a Cell-Penetrating Peptide Targeting the Nrf2/Keap1 Interaction

Richard Steel, Jonathan Cowan, Estelle Payerne, Maria A. O'Connell,* and Mark Searcey*

School of Pharmacy, University of East Anglia, Norwich NR4 7TJ, United Kingdom

S Supporting Information



ABSTRACT: Nuclear factor (erythroid-derived 2)-like 2 (Nrf2) is increasingly recognized as a central regulator of multiple signaling pathways in inflammation and cancer, and the ability to use chemical biological tools to investigate its biological effects is very attractive. A peptide comprising a TAT-conjugated Nrf2 sequence is shown to activate Nrf2 and its downstream target gene heme-oxygenase-1 (HO-1) in a dose-dependent manner in intact human THP-1 monocytes. Levels of Nrf2 protein peak after 3 h, whereas HO-1 mRNA and protein peak after 6 and 12 h, respectively. The peptide is also shown to inhibit the production of the pro-inflammatory cytokine TNF. The TAT-14mer constitutes a useful chemical biology tool with potential therapeutic applications.

KEYWORDS: cell-penetrating peptide, Nrf2, Keap-1, protein–protein interaction, inflammation, antioxidant response

Nuclear factor (erythroid-derived 2)-like 2 (Nrf2) is a transcription factor and member of the Cap'N'collar family of proteins.¹ Nrf2 plays a central role in the response to various types of stress by activating a myriad of cellular antioxidants,² detoxification enzymes,³ drug efflux pumps,⁴ and other cytoprotective proteins.⁵ As such, it is closely controlled in most cells by the presence of a protein called Kelch-like ECH-associated protein 1 (Keap1).⁶ Keap1 binds to Nrf2 and induces ubiquitination and proteasome-mediated degradation, thus maintaining low levels of the protein under normal conditions. In response to oxidative stimuli, the Nrf2-Keap1 interaction is perturbed, and Nrf2 translocates into the nucleus where it forms a dimer with small Maf proteins⁷ and binds to the antioxidant response element (ARE) in the regulatory regions of its target genes (Figure 1A). This results in the production of proteins that have a protective effect for the cell. These include the cellular antioxidants heme-oxygenase 1 (HO-1),⁸ proteins that regulate glutathione biosynthesis,⁹ and the phase II enzyme NAD(P)H:quinone oxidoreductase 1 (NQO1).¹⁰

Recently, Nrf2 has been recognized as an important regulator of inflammation. Nrf2 knockout mice are more susceptible to inflammation than wild-type mice, and Nrf2 exerts anti-inflammatory effects in cells *in vitro*.¹¹ This makes Nrf2 an interesting target for the development of compounds with

potential anti-inflammatory activity. To date, most effort has focused on either the triterpenoid compounds CDDO-Im and CDDO-Me or on covalent inhibitors that bind to cysteines on the surface of Keap1 that have been suggested to be important for the protein–protein interactions. CDDO-Im and CDDO-Me are synthetic analogues of a triterpene and have been shown to activate the Nrf2 pathway and induce high levels of antioxidant enzymes.¹² Both have protective effects in animal models of inflammation, with CDDO-Im demonstrating protection in an animal model of sepsis (a blood-borne infection causing acute inflammation and affecting several tissues). In addition, CDDO-Me is currently in phase III clinical trials in type II diabetics with chronic kidney disease.¹³ More recently, Gray and co-workers described the identification of a small molecule that inhibited the Nrf2/Keap1 interaction through covalent modification of Cys151 of the human Keap1 and led to activation of Nrf2.¹⁴ The authors built a structure–activity relationship for the relatively simple compound AI-1 that acts as a good lead for further development. However, covalent modification may be less desirable in a drug due to off-target effects *in vivo*. Electrophiles that interact with the

Received: February 17, 2012

Accepted: March 12, 2012

Published: March 12, 2012

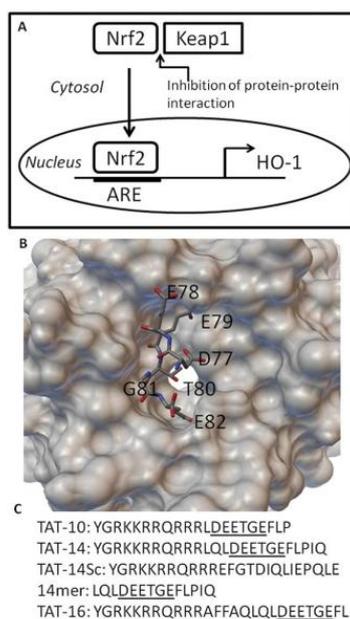


Figure 1. (A) Nrf2 is constitutively expressed but normally binds to Keap1 and is degraded. Inhibition of Nrf2/Keap1 through interaction with TAT-14 leads to nuclear translocation and activation of downstream proteins such as HO-1. (B) Structure of the conserved DEETGE sequence binding to the Kelch domain and Keap1.¹⁵ (C) The peptide sequences used in this study.

cysteines on Keap1, particularly with such a simple structure as AI-1, may well interact with the cysteines or other nucleophiles on other proteins and may also be substrates for the reaction with, for example, high levels of glutathione.

These observations suggest that an approach involving simpler compounds than CDDO-Me that interact to block Nrf2/Keap1 through noncovalent interactions are desirable and have great therapeutic potential. Hannink and co-workers have shown that short peptides based upon the Nrf2 binding region to the Kelch domain of Keap1 can bind with high affinity where the central DEETGE site is maintained.¹⁵ Three peptides of 10, 14, and 16 amino acid length were designed, and the 16mer and 14mer showed high affinity binding, while the 10mer also bound but with lower affinity. The structure of the 16mer peptide bound to the Kelch domain is shown in Figure 1B. While peptides are not ideal drug candidates, they are very useful as chemical tools, particularly when conjugated to a cell-penetrating peptide (CPP) such as the trans-activating transcriptional activator (TAT) peptide derived from HIV.¹⁶ A peptide targeted at the protein–protein interaction between the PSD95 and the NR2B subunit of the NMDA receptor has proved to be an effective intracellular regulator of the interaction, useful in stroke.¹⁷ Although it has been suggested that TAT-peptides targeting Nrf2/Keap1 do not function¹⁸ and more recently that even shorter peptides can be used to target the interaction,¹⁹ we focused on the conjugation of TAT to the Hannink peptides, for which there is structural data for binding.

Using standard Fmoc chemistry and automated synthesis, we generated three peptides with the TAT sequence located at the amine terminus of the 10-, 14-, and 16mer peptides (TAT-10, TAT-14, and TAT-16, Figure 1C). The peptides were purified

by HPLC to >90% pure and were then screened for their ability to activate HO-1 gene expression in intact THP-1 cells (a human monocytic cell line), previously validated as a model for the study of the Nrf2 pathway in human monocytes.²⁰ HO-1 is a downstream product of the activation of Nrf2. Only the TAT-14 peptide was seen to activate HO-1 expression, suggesting that the conjugation of the TAT peptide to the 10- and 16mer peptide was either not enhancing uptake into cells or was inhibiting the binding of the peptides to the Kelch domain (Figure 2A). Activation by the TAT-14 compound was significant, however, and this conjugated product was chosen for further analysis.

Investigation of the effect of the TAT-14 peptide on Nrf2 protein expression was then studied using Western blot analysis. Cells were incubated with the peptide and a version of the TAT-14 peptide scrambled in the 14mer region (TAT-14Sc), which would affect binding to the Kelch domain but the peptide would still enter the cell (see Figure 1C). Levels of Nrf2 protein reached a peak after 3 h of incubation with the TAT-14 peptide and showed no enhancement with the TAT-14Sc peptide (Figure 2B). The TAT-14 peptide had no effect on Nrf2 mRNA expression, measured by qRT-PCR (data not shown), suggesting that the increased levels of protein were due to the interaction of TAT-14 with Keap1 rather than through a stress-induced increase in basal levels of the transcription factor. This is also supported by the lack of effect of the scrambled peptide.

As noted above, the initial screen for the TAT-14 peptide involved investigation of its effect on expression of HO-1, a downstream product of the binding of Nrf2 to the ARE. To clearly define the timing of this effect and also investigate the effects on HO-1 protein, experiments were carried out using both qRT-PCR (for mRNA) and Western blotting (for protein). Effects on HO-1 mRNA by the TAT-14 peptide were compared with the nonconjugated peptide (14mer) and also the TAT-14Sc peptide. The effects of the TAT-14 peptide (75 μ M) on HO-1 mRNA expression were significant with an average 24-fold induction at the peak incubation time of 6 h (Figure 2C). The induction then gradually returned to basal levels. Neither the 14mer nor the TAT-14Sc peptides had any significant effect on the levels of HO-1 mRNA. HO-1 protein expression was similarly investigated in comparison with the TAT-14Sc peptide, and this peaked at 12 h, with the TAT-14Sc having no effect (Figure 2D).

The high levels of HO-1 expression allowed the examination of the dose response of the peptide (Figure 2E). The effect of the peptide was assessed at 7.5, 18.75, 37.5, and 75 μ M and gave a clear dose response with significant activation also occurring at the 37.5 μ M concentration.

Nrf2 deficiency results in susceptibility to inflammation in animal models, including bacteria-induced sepsis in mice, and this is due to an excessive production of pro-inflammatory cytokines including tumor necrosis factor α (TNF) in the Nrf2 knockout mice. We have previously shown that bacterial lipopolysaccharide (LPS)-induced TNF production is a valid cell model of sepsis. Monocytes play a central role in the inflammatory response during sepsis. Inhibition of Nrf2 using siRNA leads to the increased production of TNF following stimulation of THP-1 monocytes by LPS.²⁰ This suggests that conversely the release of Nrf2 from Keap1 and its resultant binding to the ARE may inhibit the production of the cytokine following a similar stimulation by LPS in these cells. In turn, this would have an anti-inflammatory effect in this model for

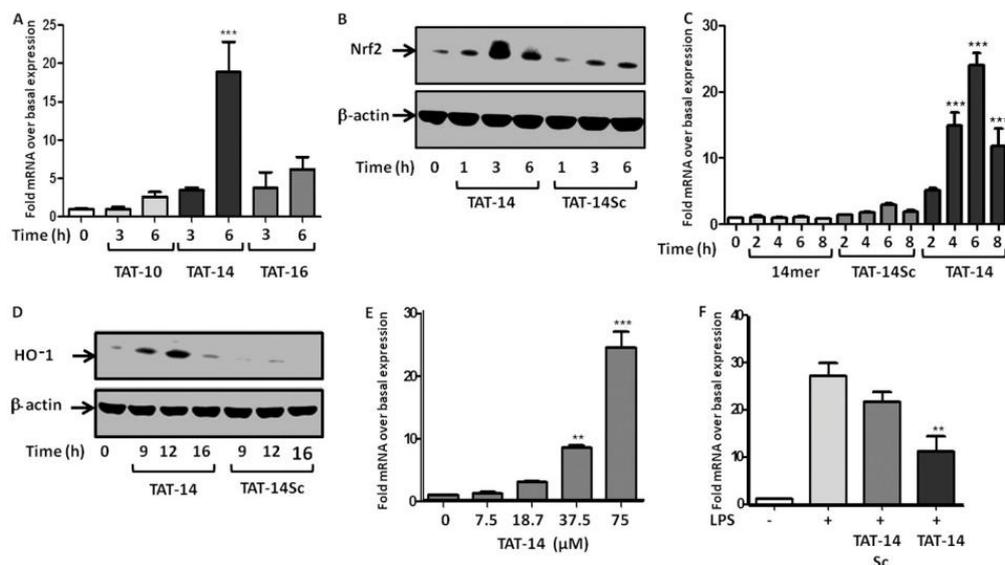


Figure 2. (A) Induction of HO-1 by the three TAT peptides. THP-1 cells were treated with the peptide for the time indicated, and levels of HO-1 mRNA were measured using qRT-PCR and normalized to the expression of the housekeeping gene GAPDH (see the Supporting Information for full experimental details). Means \pm SEMs, $n = 3$, $***p < 0.001$. (B) Western blot analysis of the induction of Nrf2 protein following incubation with the TAT-14 peptide or the TAT-14sc peptide. (C) Induction of HO-1 by the nonconjugated 14mer, the TAT-14sc, and TAT-14. THP-1 cells were treated with the peptide for the time indicated, and levels of HO-1 mRNA were analyzed using qRT-PCR and normalized to GAPDH expression. Means \pm SEMs, $n = 3$, $***p < 0.001$. (D) Western blot analysis of HO-1 protein levels following treatment with TAT-14 and TAT-14sc. Representative experiment of $n = 3$. (E) Dose dependence of HO-1 induction as assessed by qRT-PCR and normalized to GAPDH mRNA expression. Means \pm SEMs, $n = 3$, $**p < 0.01$, $***p < 0.001$. (F) TNF α induction following stimulation with LPS in the presence of either TAT-14 or TAT-14Sc. mRNA levels were analyzed using qRT-PCR and normalized to GAPDH. Means \pm SEMs, $n = 3$, $**p < 0.01$.

septic shock. THP-1 cells were rested for 24 h to generate stress-free conditions and then treated with TAT-14 or TAT-14sc for 4 h prior to stimulation with LPS for 3 h. After this, the cells were lysed and analyzed for TNF mRNA expression by qRT-PCR (Figure 2F). TAT-14 significantly suppressed LPS-induced TNF expression, with an average of 61% inhibition of the cytokine, signifying that the TAT-14 peptide has an anti-inflammatory effect.

These results demonstrate that the conjugation of a CPP to a sequence targeting the Nrf2 binding site on Keap1 generates a chemical-biological tool that is able to enter cells and inhibit the protein–protein interaction, thereby activating Nrf2. Such a construct also has some biological activity and can act as an anti-inflammatory. Importantly, this study also demonstrates that carefully designed “druglike” small molecules that target the same site on Keap1 could have a similar activity. These compounds have potential in the treatment of diseases where resolution of inflammatory processes can play a beneficial role. The progression of CDDO-Me into the clinic suggests that the Nrf2/Keap1 interaction is an important target and that chemical tools such as the TAT-14 peptide will have a role to play in understanding the biological processes involved in this pathway, as well as defining a clear therapeutic target.

■ ASSOCIATED CONTENT

● Supporting Information

Experimental procedures including the peptide synthesis, purification, and biological evaluation. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*Tel: +44-1603-592030. E-mail: m.oconnell@uea.ac.uk (M.O.). Tel: +44-1603-592026. E-mail: m.searcey@uea.ac.uk (M.S.).

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Funding

R.S. is funded through an EPSRC DTA. E.P. was funded by MRC Grant No. G0801127. J.C. is funded by UEA.

Notes

The authors declare no competing financial interest.

■ ABBREVIATIONS

Nrf2, nuclear factor (erythroid-derived 2)-like 2; Keap1, Kelch-like ECH-associated protein 1; ARE, antioxidant response element; TAT, trans-activating transcriptional activator; HO-1, heme-oxygenase 1; CPP, cell-penetrating peptide; TNF, tumor necrosis factor

■ REFERENCES

- (1) Moi, P.; Chan, K.; Asunis, I.; Cao, A.; Kan, Y. W. Isolation of NF-E2-related factor 2 (Nrf2), a NF-E2-like basic leucine zipper transcriptional activator that binds to the tandem NF-E2/AP1 repeat of the beta-globin locus control region. *Proc. Natl. Acad. Sci. U.S.A.* 1994, 91, 9926–9930.

- (2) Kobayashi, M.; Yamamoto, M. Molecular mechanisms activating the Nrf2-Keap1 pathway of antioxidant gene regulation. *Antioxid. Redox Signaling* **2005**, *7*, 385–394.
- (3) Itoh, K.; Chiba, T.; Takahashi, S.; Ishii, T.; Igarashi, K.; Katoh, Y.; Oyake, T.; Hayashi, N.; Satoh, K.; Htayama, I.; Yamamoto, Y.; Nabeshima, Y. An Nrf2/small Maf heterodimer mediates the induction of phase II detoxifying enzyme genes through antioxidant response elements. *Biochem. Biophys. Res. Commun.* **1997**, *236*, 313–322.
- (4) Hayashi, A.; Suzuki, H.; Itoh, M.; Yamamoto, Y.; Sugiyama, Y. Transcription factor Nrf2 is required for the constitutive and inducible expression of multidrug resistance-associated protein 1 in mouse embryo fibroblasts. *Biochem. Biophys. Res. Commun.* **2003**, *310*, 824–829.
- (5) Chen, X. L.; Kunsch, C. Induction of cytoprotective genes through Nrf2/antioxidant response element pathway: A new therapeutic approach for the treatment of inflammatory diseases. *Curr. Pharm. Des.* **2004**, *10*, 879–891.
- (6) Itoh, K.; Wakabayashi, N.; Katoh, Y.; Ishii, T.; Igarashi, K.; Engel, J. D.; Yamamoto, M. Keap1 represses nuclear activation of antioxidant responsive elements by Nrf2 through binding to the amino-terminal Neh2 domain. *Genes Dev.* **1999**, *13*, 76–86.
- (7) Itoh, K.; Igarashi, K.; Hayashi, N.; Nishizawa, M.; Yamamoto, M. Cloning and characterization of a novel erythroid cell-derived CNC family transcription factor heterodimerizing with the small Maf family proteins. *Mol. Cell. Biol.* **1995**, *15*, 4184–4193.
- (8) Alam, J.; Stewart, D.; Touchard, C.; Boinapally, S.; Choi, A. M.; Cook, J. L. Nrf2, a Cap'n'Collar transcription factor, regulates induction of the heme oxygenase-1 gene. *J. Biol. Chem.* **1999**, *274*, 26071–26078.
- (9) Chan, J. Y.; Kwong, M. Impaired expression of glutathione synthetic enzyme genes in mice with targeted deletion of the Nrf2 basic-leucine zipper protein. *Biochim. Biophys. Acta* **2000**, *1517*, 19–26.
- (10) Venugopal, R.; Jaiswal, A. K. Nrf1 and Nrf2 positively and c-Fos and Fra1 negatively regulate the human antioxidant response element-mediated expression of NAD(P)H:quinone oxidoreductase1 gene. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 14960–14965.
- (11) Ma, Q.; Battelli, L.; Hubbs, A. F. Multiorgan autoimmune inflammation, enhanced lymphoproliferation, and impaired homeostasis of reactive oxygen species in mice lacking the antioxidant-activated transcription factor Nrf2. *Am. J. Pathol.* **2006**, *168*, 1960–1974.
- (12) Thimmulappa, R. K.; Fuchs, R. J.; Malhotra, D.; Scollick, C.; Traore, K.; Bream, J. H.; Trush, M. A.; Liby, K. T.; Spron, M. B.; Kensler, T. W.; Biswal, S. Preclinical evaluation of targeting the Nrf2 pathway by triterpenoids (CDDO-Im and CDDO-Me) for protection from LPS-induced inflammatory response and reactive oxygen species in human peripheral blood mononuclear cells and neutrophils. *Antioxid. Redox Signaling* **2007**, *9*, 1963–1970.
- (13) Pergola, P. E.; Raskin, P.; Toto, R. D.; Meyer, C. J.; Huff, J. W.; Grossman, E. B.; Krauth, M.; Ruiz, S.; Audhya, P.; Christ-Schmidt, H.; Wittes, J.; Warnock, D. G. Bardoxolone methyl and kidney function in CKD with type 2 diabetes. *N. Engl. J. Med.* **2011**, *365*, 327–336.
- (14) Hur, W.; Sun, Z.; Jiang, T.; Mason, D. E.; Peters, E. C.; Zhang, D. D.; Luesch, H.; Schultz, P. G.; Gray, N. S. A small-molecule inducer of the antioxidant response element. *Chem. Biol.* **2010**, *17*, 537–547.
- (15) Lo, S. C.; Li, X.; Henzl, M. T.; Beamer, L. J.; Hannink, M. Structure of the Keap1:Nrf2 interface provides mechanistic insight into Nrf2 signaling. *EMBO J.* **2006**, *25*, 3605–3617.
- (16) Deshayes, S.; Morris, M. C.; Divita, G.; Heitz, F. Cell-penetrating peptides: tools for intracellular delivery of therapeutics. *Cell. Mol. Life Sci.* **2005**, *62*, 1839–1849.
- (17) Schwarze, S. R.; Ho, A.; Vocero-Akbani, A.; Dowdy, S. F. In vivo protein transduction: Delivery of a biologically active protein into the mouse. *Science* **1999**, *285*, 1569–1572.
- (18) Zhao, J.; Redell, J. B.; Moore, A. N.; Dash, P. K. A novel strategy to activate cytoprotective genes in the injured brain. *Biochem. Biophys. Res. Commun.* **2011**, *407*, 501–506.
- (19) Hancock, R.; Bertrand, H. C.; Tsujita, T.; Naz, S.; El-Bakry, A.; Laoruchupong, J.; Hayes, J. D.; Wells, G. Peptide inhibitors of the Keap1-Nrf2 protein-protein interaction. *Free Radical Biol. Med.* **2012**, *52*, 444–451.
- (20) Rushworth, S. A.; MacEwan, D. J.; O'Connell, M. A. Lipopolysaccharide-induced expression of NAD(P)H:quinone oxidoreductase 1 and heme oxygenase-1 protects against excessive inflammatory responses in human monocytes. *J. Immunol.* **2008**, *181*, 6730–6737.