Design and Synthesis of Inhibitors of the Nrf2/MafG Protein-Protein Interaction as a Research Probe.

Ellen Elizabeth Hyde

School of Chemistry, Pharmacy and Pharmacology

University of East Anglia

September 2025

Thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy of the University of East Anglia.

Declaration

This thesis has been submitted to the University of East Anglia for the degree of Doctor of Philosophy and is, to the best of my knowledge, original except where stated, referenced and acknowledged.

Ellen Hyde

Abstract

Nuclear erythroid factor 2 (Nrf2) and v-maf musculoaponeurotic fibrosarcoma oncogene family protein G (MafG) are basic leucine zipper (bZIP) transcription factors. A coiled-coil protein-protein interaction (PPI) forms between the leucine zipper domains of Nrf2 and MafG, stabilising Nrf2/DNA binding for activation of gene transcription. Nrf2 is responsible for regulating oxidative stress, controlling gene expression of antioxidant and cytoprotective activity in cells. The overexpression of Nrf2 is linked to the development of chemoresistance in non-small cell lung cancer (NSCLC). Consequently, there is a need for developing inhibitors of Nrf2 as an anti-cancer therapeutic.

In **Chapter 1** the discovery of PPI inhibitors is outlined, highlighting the techniques used for different categories of interaction. A literature review of existing human PPI inhibitors is presented, demonstrating a need to apply new approaches towards the design of coiled-coil inhibitors. This thesis presents research into novel inhibitors of the coiled-coil interaction between Nrf2 and MafG.

In the absence of an experimentally resolved structure of the Nrf2/MafG PPI, AlphaFold offers a new method for studying the Nrf2/MafG PPI. Applying this approach, **Chapter 2** reports the design and synthesis of peptides derived from the MafG leucine zipper, exploring the minimum sequence length required to afford inhibition. Using recombinantly expressed Nrf2 and MafG protein, described in **Chapter 5**, fluorescence polarisation, electrophoretic mobility shift assays and surface plasmon resonance assays were investigated, to evaluate peptide activity towards the Nrf2/MafG PPI. A lead 21-mer peptide was found capable of disrupting the PPI (IC_{50} of 36 μM) by fluorescence polarisation assay.

In **Chapter 3**, the lead peptide was applied to peptide-directed ligand design to explore small molecule fragments capable of disrupting the coiled-coil interaction. Peptide-small molecule hybrids were synthesised and screened by fluorescence polarisation assay, with two hybrid compounds capable of disrupting ternary complex formation between Nrf2/MafG/DNA (IC_{50} of 62 and 125 μM).

Inhibitory peptide and peptide-small molecule hybrids were evaluated for cellular efficacy in **Chapter 4**. A549 cells, a NSCLC cell line known to express high levels of Nrf2, were selected for the study. Inhibitors demonstrated suppression of NQO1 activity, a target gene of Nrf2. Excitingly, our lead peptide-small molecule hybrid was capable of sensitising resistant A549 cells to gemcitabine.

Contents

Abbreviations

Ac Acetyl

ARE Antioxidant recognition element

Bax BCL2 associated X protein

Bcl-2 B-cell lymphoma 2

bHLH Basic helix-loop-helix

β-ME β-mercaptoethanol

Boc tert-butyloxycarbonyl

BSA Bovine serum albumin

β-TrCP Beta-Transducin repeat containing protein

bZIP Basic Leucine Zipper

CuAAC Copper(I)-catalyzed azide alkyne cycloaddition

Da Daltons

DBF Dibenzofulvene

DIC N,N'-diisopropylcarbodiimide

DIPEA N,N-Diisopropylethylamine

DMSO Dimethyl sulfoxide

DNA Deoxyribonucleic acid

DODT 2,2'-(Ethylenedioxy)diethanethiol

DTT Dithiothreitol

EC_{50} Half maximal effective concentration

E.coli Escherichia coli

EDC 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide

EMSA Electrophoretic mobility shift assay

EP⁺ EDTA and P20 combined additives

FAD Flavin adenine dinucleotide disodium salt hydrate

FITC Fluorescein isothiocyanate

FP Fluorescence polarisation

FRET Förster resonance energy transfer

GSH Glutathione

HATU Hexafluorophosphate azabenzotriazole tetramethyl uronium

HBTU Hexafluorophosphate benzotriazole tetramethyl uronium

HEPES 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

HO-1 Heme oxygenase-1

HOBt 1-hydroxybenzotriazole hydrate

HTRF homologous time resolved FRET

HPLC High-performance liquid chromatography

IC_{50} Half maximal inhibitory concentration

IDR Intrinsically disordered region

IEC Ion exchange chromatography

IPTG Isopropyl β-D-1-thiogalactopyranoside

K_D Equilibrium dissociation constant

K_d Dissociation constant

Keap1 Kelch-like ECH associated protein 1

K_i Inhibition constant

mAbs Monoclonal antibodies

MALDI Matrix-assisted laser desorption/ionisation

MafG v-maf musculoaponeurotic fibrosarcoma oncogene family protein G

MeCN Acetonitrile

MeOH Methanol

MOPS 3-(N-morpholino)propanesulfonic acid

MRP multidrug-resistance protein

MTS 3-[4,5,dimethylthiazol-2-yl]-5-[3-carboxymethoxy-phenyl]-2-[4-sulfophenyl]- 2H-tetrazolium

NHS *N*-hydroxysuccinimide

NLS Nuclear localisation sequence

NES Nuclear exportation sequence

NMR Nuclear magnetic resonance

Nrf2 Nuclear Factor erythroid 2 related factor 2

NSCLC Non-small cell lung cancer

NTA Nitriloacetic acid

NQO1 NAD(P)H quinone dehydrogenase

OD_{600} Optical density at 600 nm

Oxyma Ethyl 2-cyano-2-(hydroxyamino)acetate

p53 Tumour protein 53

Pbf 2,2,4,6,7-Pentamethyldihydrobenzofuran-5-sulfonyl

PBS Phosphate buffered saline

PDB Protein Data Bank

PDLD Peptide-directed ligand design

PEG Polyethylene glycol

Pgp P-glycoprotein

pI Isoelectric point

PPI Protein-protein interaction

PVDF Polyvinylidene difluoride

ROS Reactive oxygen species

RPMI Roswell Park Memorial Institute 1640 Medium

SDS Sodium dodecyl sulphate

SDS-PAGE Sodium dodecyl sulphate polyacrylamide gel electrophoresis

SEC Size exclusion chromatography

siRNA small interfering ribonucleic acid

sMaf Small musculoaponeurotic fibrosarcoma

SPPS Solid-phase peptide synthesis

SPR Surface plasmon resonance

tBu tert-butyl

TF Transcription factor

TFA Trifluoroacetic acid

TFE Trifluoroethanol

TIPS Triisopropylsilane

TRIS Tris(hydroxymethyl)aminomethane

TR-FRET Time resolved- föster resonance energy transfer

In Memory of Jennifer Neal-Smith

Acknowledgements

Overall, this research has been supported by many people, but I would like to thank these individuals in particular for being so involved. I would not have wanted to do it without them.

I would like to begin by thanking Andrew for being a great supervisor, from whom I have learnt so much. I am grateful for all his advice and encouragement at work and in everyday life. I also really appreciate Andrew's ability to try and fix everything, including (but not limited to) the HPLC, my bike and the hole I drilled in my central heating system. I would like to thank all the members of the Beekman group for their support, especially Dr Suzanne van Wier and Dr Dee Hayward, who have helped shape me as a scientist and have become great friends during my PhD.

I am extremely thankful to Mark for all the thought-provoking questions, good humour and for having such a brilliant research group to work with. In particular, Dr Zoë Goddard and Dr Marco Cominetti, who were excellent at keeping the lab running and always had insightful scientific advice. I am especially grateful to have completed my PhD alongside Vicky Kamperi, her constant friendship has made my time at university incredibly enjoyable, and I am glad to have had her support both at home and in the lab.

I also had the great pleasure of working with Professor Maria O'Connell, Dr Emily Hobson and Dr Sean Tattan, who provided instrumental expertise on Nrf2, pharmacology and the cellular experiments I carried out. Thank you for all your support.

Thanks also to the dogs and cats of Norwich (Monty, Izzy, Juno, Bluey and Berry) for all the cuddles and chaos, keeping life interesting.

From home, I am grateful to my Mum and Dad for encouraging me to pursue a PhD and supporting me all the way through. I could probably write another thesis on all the ways how. I would also like to give a special thank you to my Aunt Elaine and my Grandad Rob whom I visited during my write up period, I fear you now know more about coiled coils than you ever wanted! Thank you to all the Hyde's, Neal-Smith's and Vicker's, who asked how the PhD was going… a daunting question to ask and to answer.

Finally, to Oliver and Sarah Pearson, who first brought me to visit UEA in 2017, for convincing me of what a wonderful time I would have here - of course, they were right!

This work was supported by the UKRI Biotechnology and Biological Sciences Research Council Norwich Research Park Biosciences Doctoral Training Partnership [grant number BB/T008717/1].

Chapter 1
Introduction

1.1 Protein-Protein Interactions

Protein-protein interactions (PPIs) are specific, functional contacts between proteins within an organism.[1] They underpin all biological functions in living and viral systems, influencing entire signalling cascades, supporting metabolic activity and gene expression.[1–3]

The broad application of PPIs means they encompass a huge amount of variety in the types that exist, consequently, they can be defined in a number of different ways. Quantitative characterisation of PPIs can be achieved through mathematical analysis of native surface contacts, chemical properties at the interface or biophysical parameters such as association and dissociation rates.[4–6] A more binary classification of PPIs is possible by identifying obligate/non-obligate interactions, where proteins can be found as a part of a complex (obligate) or exist independently (non-obligate). Furthermore, non-obligate proteins can produce transient or permanent interactions based on a PPI's stability.[7] Qualitative classification can be collated by protein features, including sequence motifs or individual residue properties, observing binding hotspots by sequence mutagenesis.[8,9] Alternatively, PPIs can also be defined by peptide-domain interactions, as peptide motifs are increasingly being recognised for their contributions towards an interaction.[10,11]

Despite the vast complexity of these interactions, PPIs still make compelling drug targets due to their high level of specificity and broad application towards a range of disease targets. Dysregulation of PPIs has been linked to many diseases, including cancer, neurodegeneration and immune diseases.[12–14] What makes PPIs particularly attractive in drug discovery is their significant role in biological networks, offering powerful control over disease pathogenesis, instead of targeting a singular protein such as a receptor or enzyme.[15] Described here is an overview of the current landscape in developing PPI inhibitors.

1.2 Inhibitors of Protein-Protein Interactions

PPIs are notoriously challenging to target due to their large surface area and shallow binding grooves. For instance, although more than 650,000 PPIs are predicted in the human interactome, as of 2012, it was predicted that only 0.01% had an inhibitor.[16–20] Despite the increasing popularity of PPI targeting for drug discovery (**Figure 1-1, A**), 85% of the human proteome remains undrugged, highlighting a need to improve our understanding of targeting PPIs.[21,22]

Characterisations of PPI inhibition through the contributions of buried surface area has found that larger surfaces are more accessible to peptide inhibitors whereas small molecules are better suited to smaller surface areas.[17,18,23] Whilst databases exist that categorise small molecule inhibitors of PPIs, these do not consider the contributions of peptide inhibitors towards the field.[24,25] Peptide-based PPI inhibition is typically understood by residue hot-spots, amino acids with significant contribution to the free binding energy of an interaction.[26] The approach taken here looks at the secondary structures informing PPIs, categorising them to better understand the strategies used to design and identify peptides or small molecules inhibitors.

Surveyance of the literature on "protein-protein interactions inhibitors" between 1972-2025, found 117 human PPIs (0.02%) with inhibitors (**Table 1-1**)[*], improving on the proposition that only 0.01% of human interactome is druggable. Here we provide a review of the human PPI literature, exhaustively identifying those with peptide or small molecule inhibitors, categorising the types of interactions, and highlighting the techniques used to tackle these challenging targets.

PPIs are primarily controlled by a short peptide sequence that can form secondary structures, such as α-helices or β-strands or exist as intrinsically disordered regions (IDR), such as dynamic loop structures or sequence recognition-containing key residues.[27] Categorisation of human PPIs based on structural characteristics demonstrated over 50% of protein-protein interactions reported were mediated by an α-helix and over 30% by β-strands. The remaining interactions were categorised as disordered or loop mediated interactions (<20%) (**Figure 1-1, B**). Across all secondary structure characterisations, rationally designed peptides are often the first examples of inhibitors towards PPIs. Equally, screening assays are a popular method for identifying small molecules inhibitors across all categories of PPI. (**Table 1-1**).

[*] Search performed using PubMed and Web of Science using search terms "protein protein interaction" and "inhibitor". Final access on 30/06/2025.

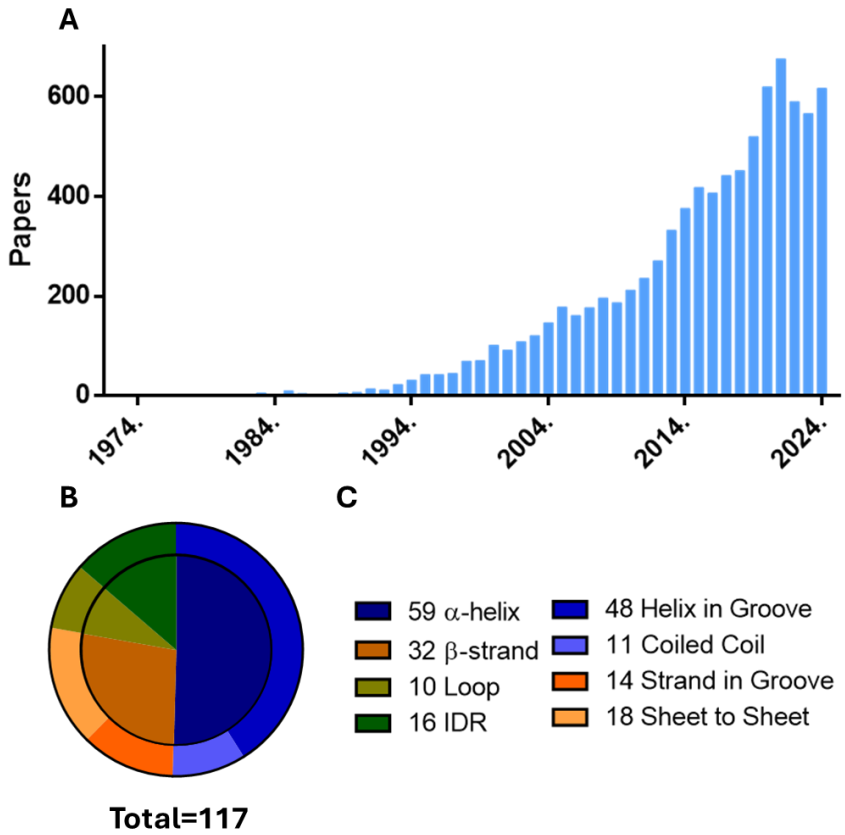


Figure 1-1 *(**A**) Search results by year: "Protein-Protein Interaction Inhibitor" (**B**) Pie chart of PPI inhibitors categorised by key secondary structure motif from "Protein-Protein Interaction" and "Inhibitor" from 1972 - June 2025. (**C**) Key for pie chart categorisation.*

1.2.1 Categorisation of interactions by secondary structure

Protein-protein interactions can be organised by the presence of interacting α-helix, β-strand or disordered secondary structures (**Figure 1-2**). The most prevalent secondary structure reported at PPI interfaces are α-helices, featured in 62% of multi-protein complexes according to data accessed from the protein data bank (PDB).[28,29] This is likely because much of our understanding of protein interactions comes from crystal structure studies. This may have introduced a bias towards the prevalence of α-helices, as they tend to crystalise more readily due to the stability of their secondary structure.[30,31]

α-helices most often form a PPI through a helix-in-groove interaction, where the surface of a partner protein forms a hydrophobic pocket (**Figure 1- 2, C**). Alternatively, they can interact with one another to form coiled-coils. These structure-to-structure interactions form a shallow hydrophobic surface, supported by neighbouring ionic interactions between two helices (**Figure 1-2, A**).

β-strands are also capable of forming features in groove (**Figure 1-2, D**) and structure to structure interactions (**Figure 1-2, B**). Analysis of PDB complexes finds β-strands contribute to 22% of binding interactions.[32] Considering our analysis of PPI secondary structure (**Figure 1-1**), these statistics appear to correlate with our PPI inhibitor categorisation (50% α-helical, 30% β-strand), highlighting the majority of PPIs known currently rely on α-helix or β-sheet binding into a groove on a partner protein surface.[33]

Dynamic proteins or intrinsically disordered protein regions (IDPRs) lack permanent secondary structures, which can also makes IDPs promiscuous in their protein binding partners.[34] However, some IDPRs can create highly specific contacts through recognition sequences of amino acids, making them an attractive druggable target (**Figure 1-2, E**).[35]

A sub-category of IDPRs considered to be disordered are loop structures. Loops can be flanked by supporting secondary structures (**Figure 1-2, F-H**) but rely on dynamic sequences to bind to the surface of a partner protein. Whilst no literature currently defines the prevalence of these structures, there are increasing reports of inhibitors disrupting these PPIs (**Figure 1-1, D**).[36–39]

Utilising this categorisation, we can review the approaches currently used to target different types of PPIs, highlighting recent advancements and the challenges that remain in PPI drug discovery.

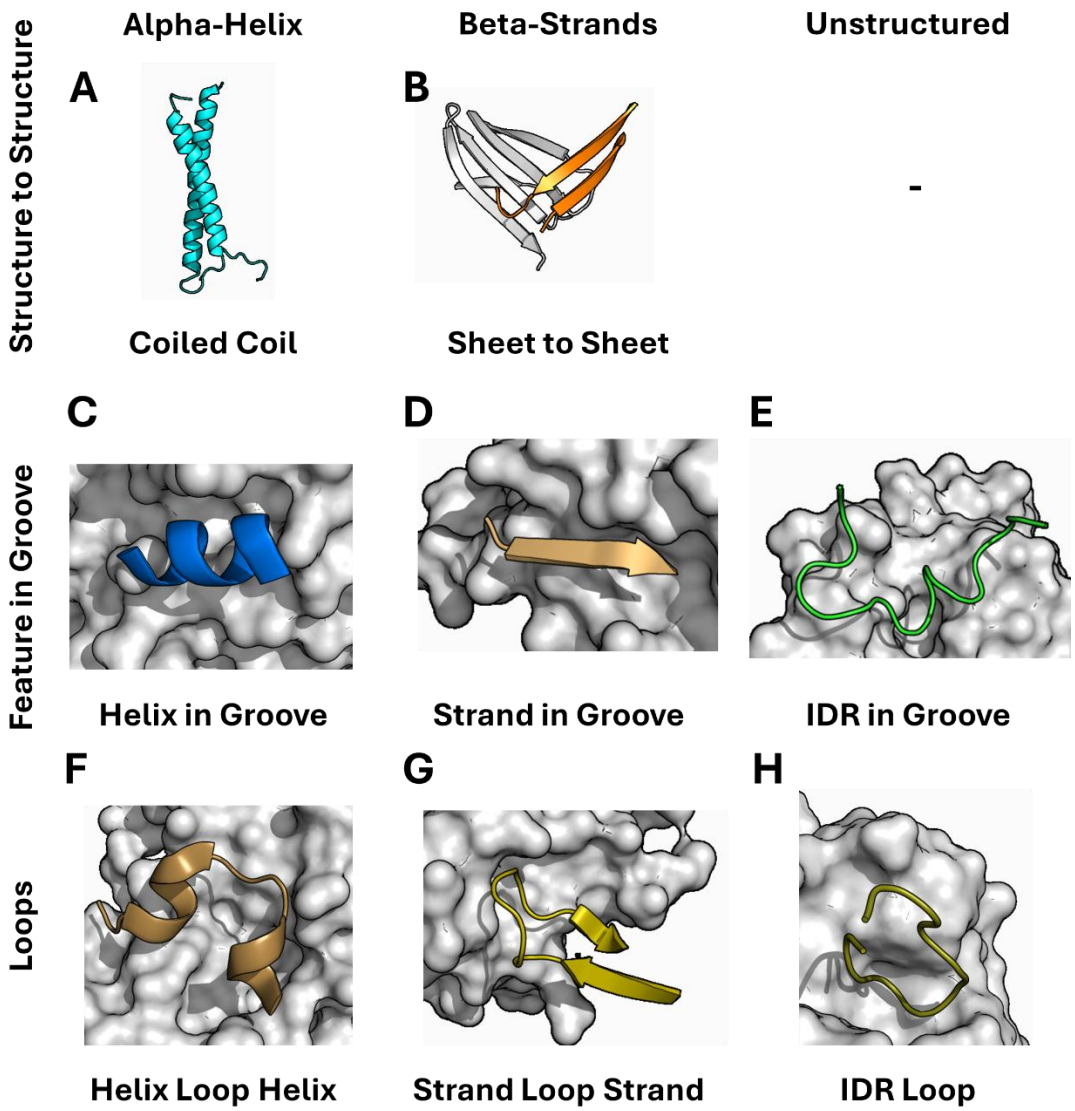


Figure 1-2: *Examples of categorisations of protein-protein interactions described in this work.* ***(A****) Coiled-coil interaction of MBD2/P66a (PDB: 2L2L).* ***(B****) Sheet to sheet interaction ICOS/ICOSL (PDB: 6X4G).* ***(C****) Helix-in-groove interaction oestrogen receptor/nuclear coactivator 2 (PDB: 1GWQ).* ***(D****) Strand-in-groove interaction GKAP/SHANK (PDB: 7A00)* ***(E****) IDR in groove interaction Rev1-CT/PolD3 (PDB: 2N1G)* ***(F****) Helix-loop-helix, (PDB: 6Q36)* ***(G****) Strand-loop-strand PRL/CNNM (PDB: 5K22)* ***(H****) IDR loop ERCC1/XPA (PDB: 2JNW).*

Table 1-1 *Human PPI with reported inhibitors, year first inhibitor of PPI was reported, with corresponding reference.*

Protein-Protein Interactions							
Peptide		**Peptidomimetic**		**Small Molecule**		**mAb**	
Helix-in-groove							
PPI	**Year**	**PPI**	**Year**	**PPI**	**Year**	**PPI**	**Year**
Bcl-2/BH3 domain[40]	1996	Gαq/PLCβ3[41]	2004	β-catenin/BCL9[42,43]	2012	DCN1/UBC12[44]	2017
14–3–3/α-synuclein[45]	1999	Smac/XIAP[46]	2007	Med25/ATF6a[47]	2013	APC/Asef[48]	2017
p53/hDM2[49]	2000	Mcl-1/BH3[50]	2010	HIFα/p300[51]	2013	NF-Y dimer[52]	2019
CBP/p300[53]	2000	AKAP/Calmodulin[54]	2010	EZH2/EED[55]	2013	RbAp48/MTA1[56]	2020
AR/TIF2[57]	2002	ppaF/Sec61[58]	2011	4E-BP/elF4E[59]	2015	PALB2/BRCA2[60]	2020
Cdk5/p25[61]	2002	PP1/RIPPO[62]	2012	TGIF1/SIN3A[63]	2016	SRSF1/SRSF3[64]	2021
Rac1/Tiam1[65]	2004	Collagen/Hsp47[66]	2017	Hsp70/Bim[67]	2021	NPAS3/ARNT[68]	2023
Skp2/CKS1[69]	2005	β-arrestin/AP-2[70]	2017	PRMT5/RioK1[71]	2021	DRP1/MiD49[72]	2023
ATG5/ATG16L[73]	2014	CSF1R/DAPK1[74]	2019	CDK9-cyclin T1[75-77]	2022	RFFL/CFTR[78]	2023
ERCC1/XPF[79]	2015	HSP90/CDC37[80]	2019	CDC42/IQGAP1[81]	2022	NLRP3/NEK7[82]	2024
XRCC4/lig4[83]	2016	NCOA4/FTH1[84]	2021	GIT/PIX[85]	2022	Smad2/3/Smad4[86]	2024
Coiled Coils							
Fos/Jun[87]	2006	EGFR JXT[88]	2014	FOXP3/FOXP[89]	2022	HOP/HSP90[90]	2024
MITF[91]	2012	HSF1[92]	2018	Nrf2/MafG[93]	2023	Myc/Max[94]	2002
Shroom/Rho[95]	2015	MBD2 /P66a[96]	2019	Beclin 1/ATG14L[97]	2020		
Strand-in-Groove							
IgG/FcRn[98]	2006	SENP/SUMO[99]	2004	GKAP/SHANK1-PDZ[100]	2006	PKCe/RACK2[101]	1997
HER2/EGFR[102]	2011	PCSK9/LDLR[103]	2013	APP/Mint2[104]	2021	BRCA2/RAD51[105]	2010
CAD/iCAD[106]	2012	SORT1/PGRN[107]	2013	BCL6/BCOR[108]	2018	α-synuclein dimer[109]	2022
Mad2/Cdc20[110]	2015	Syndecan/Syntenin[111]	2021				
Sheet to Sheet							
CD2/CD58[112]	2002	IL33/ST2[113]	2009	ICOS/ICOS-L[114]	2020	PICK1/GluA[118]	2018
uPAR/uPA[115]	1994	RANKL/RANK[116]	2010	PD-1/PD-L1[117]	2013	CCL5/HNP1[121]	2016
Grb2 SH2[119]	1996	IL17/IL17A[120]	2013	TNFa[122]	2005	PSD95/nNOS[123]	2009
RUNX1/CBFβ[124]	2015	PRL/CNNM[125]	2008	TG2/FN[126]	2014	RAS/PI3Kγ[127]	2018
Loops							
Keap1/Nrf2[128]	2006	calpain/calpastatin[129]	2009	Menin/MLL[130]	2010	GADPH/δPKC[39]	2016
ERCC1/XPA[131]	2007	MLL1/WDR5[132]	2010	YAP/TEAD[36]	2014		
p62/Keap1[133]	2016	NCS-1/Ric8a[38]	2017	β-TrCP/Nrf2[134]	2022		
Disordered Sequences							
CID/PLCγ1[135]	2018	PLK1/PDB[136]	2003	SALL4/NuRD[137]	2018	Rev1-CT/PolD3[138]	2017
α-synuclein/CHMP2B[139]	2023	PCNA/PIP-box[140]	2012	PARP1/BRCT[141]	2015	SRPK1/SRSF1[64]	2021
p47phox/p22phox[142]	2012	TRIM24/PHD[143]	2015	NHERF1/EBP50[144]	2019	SUV39H1/HP1[145]	2023
BRD9/Histone[146]	2015	DX2/HSP70[147]	2019	ZO-1/CLD[148]	2020	ELF3/MED23[149]	2023

1.2.2 Inhibitors of α-helix mediated interactions

The number of reported PPIs with helices at the interface is more than all other categories combined (**Table 1-1**).[28,29,150] Perhaps, the landscape of PPI inhibitors is skewed towards designs for α-helical targets as a result (50% of inhibitors reported, **Figure 1-1**). These largely target helix-in-groove interactions, with a smaller collection of α-helix mediated PPI inhibitors reported against coiled-coil interactions (9% of inhibitors reported, **Figure 1-1**). This may be due to coiled-coils reliance on shallow knob-in-hole interfaces distributed evenly across the α-helical surface of both protein

partners. Multiple approaches for designing peptide and small molecule inhibitors have been described, utilising knowledge of hot-spot residues across small surface areas for both coiled-coil and helical binding grooves.[151,152]

1.2.2.1 Helix-in-groove inhibitor design strategy: rational peptide design

Helix-in-groove PPIs can be categorised based on the relative energy contributions of residues across three interfacial domains of an α-helix (**Figure 1-3**).[152] Notably, it is predicted that 60% of α-helical PPIs present hot-spot residues concentrated on a singular face of a helix (**Figure 1-3, A**).[29] Peptides often go on to act as probes for screening assays or to direct small molecule inhibitor discovery towards a desired interaction site. Alternatively, small molecule inhibitors can be designed to mimic key binding residues of a helical peptide sequence.

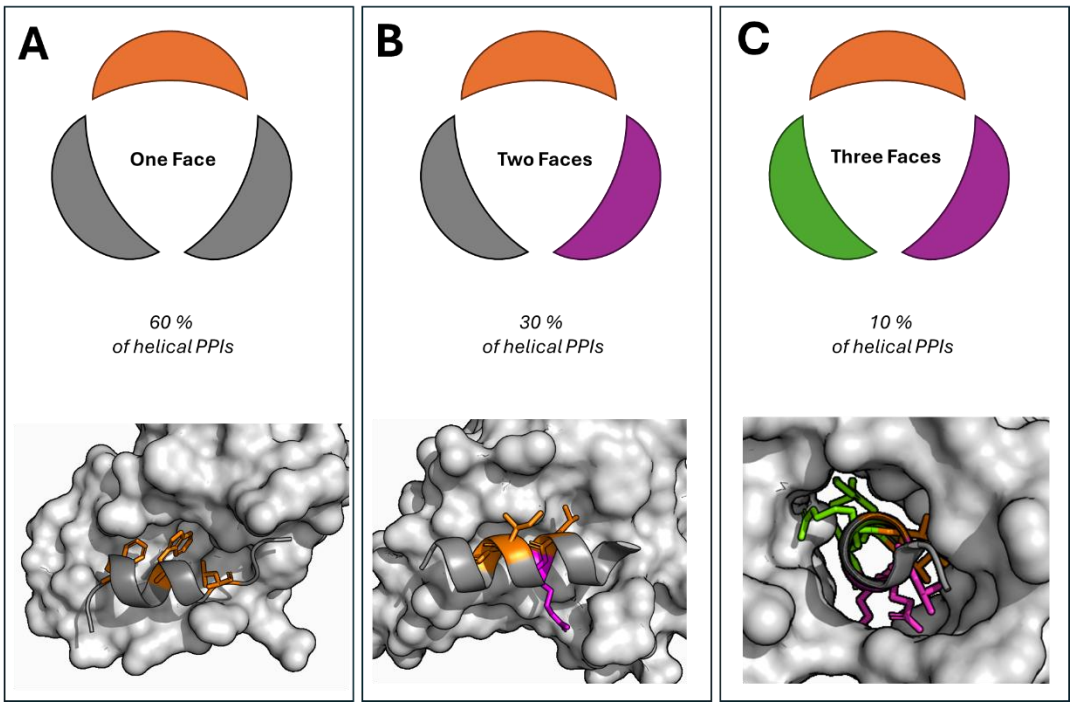


Figure 1-3 *Percentage of helix-in-groove PPIs with contributions across 1-3 faces of the helix. (**A**) Examples of p53/hdm2 (PDB: 1YCR) (**B**) two faces, Cdk2/p25 (PDB: 3O0G) (**C**) MyoA/MTIP (PDB: 4AOM).*

Structural information of a helical sequence can provide valuable information on the location of key binding pockets on a binding surface, and subsequent peptide inhibitor design. An example of this is the helix-in-groove interaction between the N-terminal transactivation domain of p53 and its negative regulators *h*DM2/X, which promotes nuclear export and degradation of p53. The α-helix of p53 engages a deep hydrophobic pocket on the *h*DM2/x via a singular face of the helix (**Figure 1-3, A**). Structural studies of p53 have revealed that this interaction is driven by three residue hotspots, Phe19,

Trp23 and Leu26 along the helix.[153,154] Peptides derived from wild-type human p53 have shown nanomolar affinity towards *h*DM2 in biophysical studies.[155]

Another well characterised example of a helix-in-groove interaction is the helical BH3 domain of Bid, Bim or Noxa binding to pro-apoptotic protein Bcl-2. The conserved BH3 helix occupies four hydrophobic pockets on a singular face of the α-helix with key isoleucine and phenylalanine residues.[156] Utilising the native sequence, high affinity peptides were synthesised, forming the basis of competition binding assays to engineer Bcl-2 specific peptides with nanomolar affinity.[157,158]

1.2.2.2 Helix-in-groove inhibitor design strategy: peptide combinatorial libraries

Surface display screening techniques or peptide arrays can help researchers screen large combinatorial peptide libraries to achieve greater affinity towards a PPI compared to wild-type peptides. Surface display technology combines genetic recombination with affinity selection to screen large varieties of peptide sequences towards a protein target.[159–161] Peptide arrays utilise immobilisation of a large number of peptide sequences on a solid support, offering a high-throughput approach to residue scanning within a peptide to gain understanding of interactions to a PPI interface. [162,163]

This has been achieved for the helix-in-groove PPI between Bim/Bcl-$_{XL}$. Researchers have been able to compare sequence hits identified from display screening in competition fluorescence polarisation assays against a wild-type BH3 helical peptide. This technique has achieved nanomolar potency for the displacement of wild-type peptide.[157,158]

Phage display methods have also improved upon the sequence affinity of p53 wild-type peptides, observing two-fold greater affinity towards *h*DM2/x proteins. Interestingly, the lead peptides retained the core hotspot residues.[164] Further investigation by residue scanning has highlighted the importance of non-contact residues in the stabilising helical conformations in both wild-type and display-derived peptides.[165] Mirror image phage display has also been used to identify D-amino acid sequences with affinity for MDM2, overcoming proteolytic susceptibility of peptides.[166]

1.2.2.3 Helix-in-groove inhibitor design strategy: helical mimetic scaffolds

Peptide inhibitors offer a scaffold for the design of small molecule inhibitors, guiding the development of peptidomimetics. Helical mimetic scaffolds have been reported for hot spot functionality across all three faces of helix-in-groove forming interactions. On a singular face of an α-helix several scaffolds have been reported, with molecular

modelling studies showing good overlap for coverage of the *i*, *i+4* and *i+7* residues within a helix.[167,168] As shown in **Figure 1-4** for inhibitors of p53/hDM2, terphenyl (**1.1**), benzamide (**1.2**) pyridyl-pyridone (**1.3**) and oxopiperazine (**1.4**) structures have demonstrated nanomolar inhibition for helix-in-groove interactions including p53/*h*DM2 (**Figure 1-4**). Helical mimetic scaffolds have also been utilised to afford inhibitors of Bcl-xL/Bak, Cdc42/Dbs and Hif1α/p300 PPI's.[29,151,169,170]

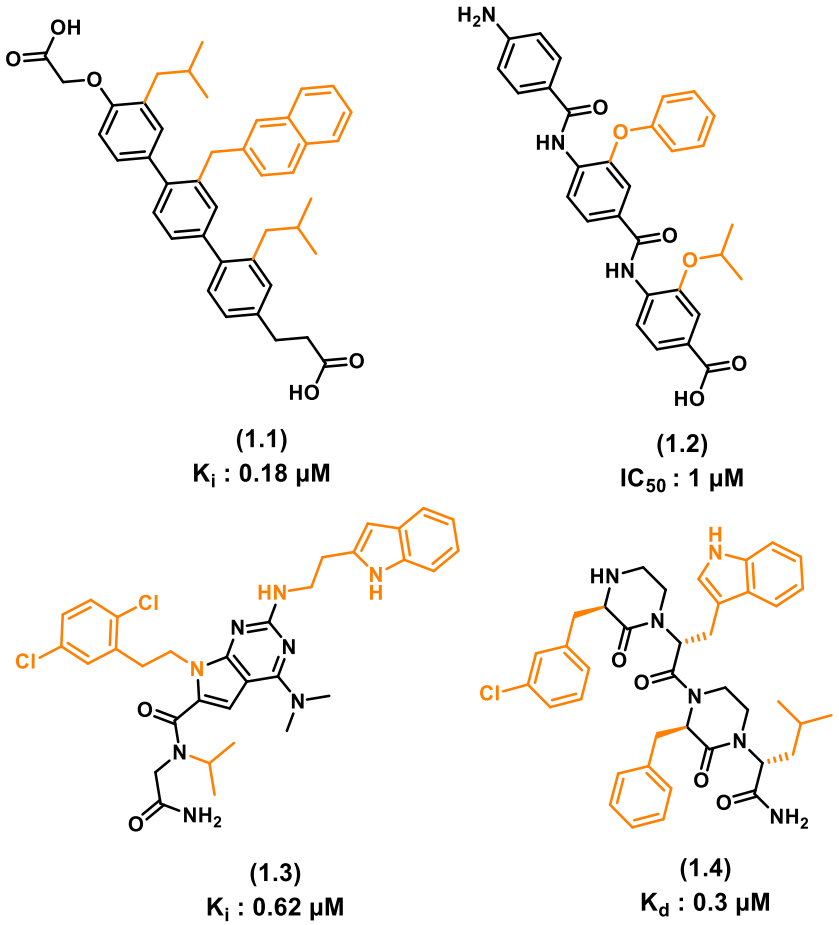


***Figure 1-4**: Helical mimetic scaffolds for inhibition of p53/hDM2 helix-in-groove interaction with nanomolar activity for the interaction, side-chain mimetics highlighted in orange.*

Most recently, oxopiperazine scaffolds have been applied for the inhibition of the PEX5-PEX14 single facing helical interaction (**Figure 1-5**). Peroxin (PEX) proteins are responsible for biogenesis of peroxisomal pathways in trypanosome parasites. PEX5 produces an amphipathic helix on a singular face with key tryptophan and phenylalanine residues across a five-residue sequence (WxxxF) (**1.5**).[171] The small interaction site makes the PPI an ideal candidate for an oxopiperazine mimetic approach, designing a small molecule inhibitor of the interaction. After a structure based design campaign, a lead compound (**1.6**) with a K_i of 27 µM was afforded for the PEX5/PEX14 interaction (**Figure 1-5, B**).[172] This demonstrates promise for helical mimetics across short surface

areas in helix-in-groove interactions without the need for excessive compound screening.

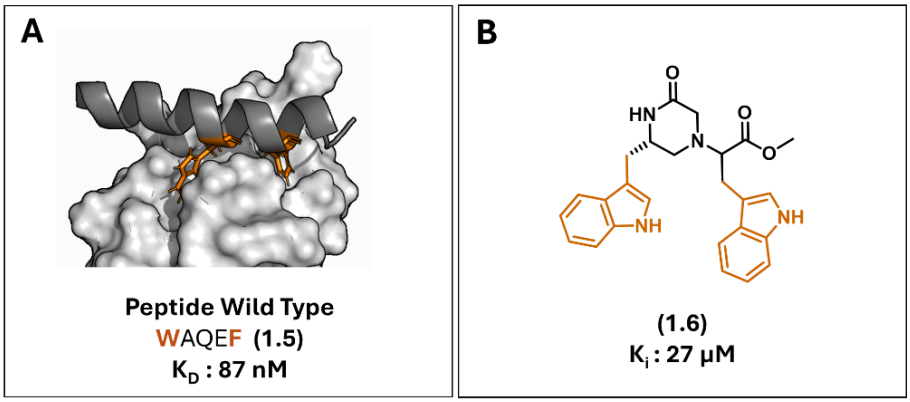


Figure 1-5 (A) *PEX5/PEX14 interaction (PDB: **2W84**) (**B**) Oxopiperazine helical mimetic (**1.6**).*

Building on this principle, other helix-in-groove PPIs have been targeted using a scaffold-based approach. Steroid receptors and their coactivators form helix-in-groove PPIs across two faces of the helix on the co-activator protein (**Figure 1-6**). The oestrogen receptor utilises three key hydrophobic points of contact with leucine residues (LxxLL) (**1.7**) which has been used to design helically constrained peptides and small molecule inhibitors utilising a pyrimidine scaffold (**1.8**) to reach the three binding pockets across the receptor surface with a K_i of 29 µM for the displacement of a peptide inhibitor containing the LxxLL motif.[173]

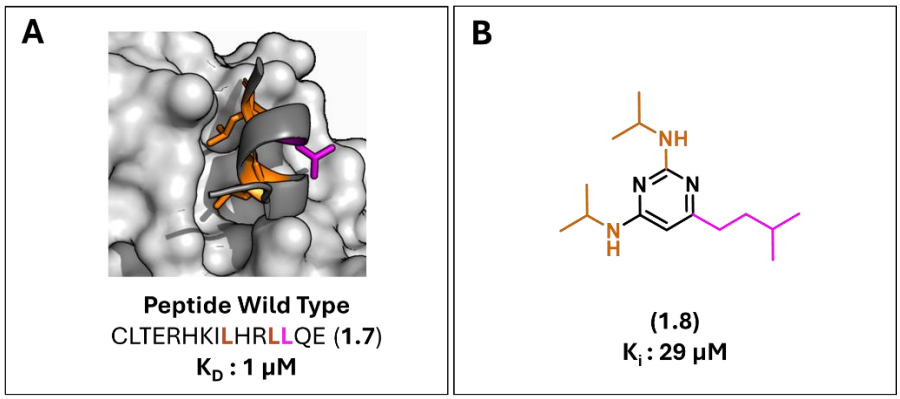


Figure 1-6: (A) *Helix-in-groove interaction of nuclear coactivator with oestrogen receptor (PDB: 1L2I). (**B**) Pyrimidine based inhibitor (**1.8**) designed off the GRIP-1 peptide (**1.7**) LxxLL motif, highlighted in orange and magenta.*

The helical interaction between MITP and the Myosin A tail in *Plasmodium falciparum* is one of the few literature examples of successful PPI inhibitors for interactions across three faces of the helix. Myosin tail interacting protein (MTIP) forms a closed complex around the Myosin A (MyoA) tail helix which stabilises the interaction using eight key charged or hydrophobic residues (**Figure 1-7**).[174] Across a 15-mer peptide sequence

(**1.9**), it was found that a helical peptide based on the MyoA tail could inhibit *P. falciparum* growth with an EC_{50} of 84 μM.[175] Using computational analysis of the hotspot residues of the 15-mer helical peptide led to the identification of a class of pyrazole-urea compounds (**1.10**) with a lead compound capable of inhibiting parasite growth (EC_{50} of 300 nM). Structural modelling demonstrated the ability of the compound to bind several of the binding pockets found by two of the helical faces of the MyoA peptide (**Figure 1-7, B**).[176]

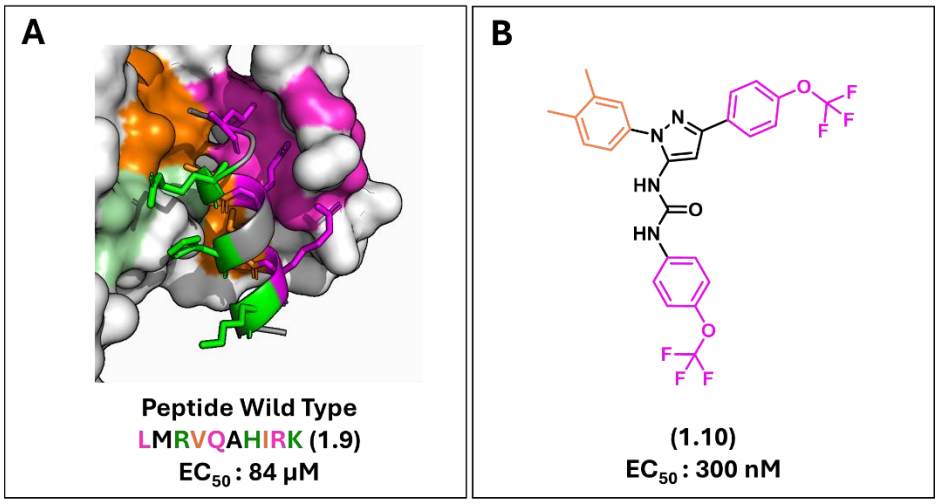


***Figure 1-7** (**A**) MyoA peptide binding to MITP protein surface (**1.9**). Three helical faces highlighted in green, orange and magenta. (PDB: 2AUC) (**B**) Pyrazole-urea compound (**1.10**) with surface interactions coloured in orange and magenta.*

1.2.2.4 Helix-in-groove inhibitor design strategy: peptide-directed ligand design

An emerging method for the identification of small molecule inhibitors of helix-in-groove interactions is peptide-directed ligand design. By utilising knowledge of hot-spot residues in a peptide sequence, peptide-small molecule hybrids are produced with improved affinity for a target protein in a PPI. This has been achieved for helix-in-groove interactions including cdk2/cyclin A, p53/*h*dm2 and Noxa/Mcl-1.[177–180] Notably, combining *in silico* screening with synthesis achieved a 50% success rate in binding assays towards the p53/*h*dm2 interaction.

A peptide inhibitor of the Noxa/Mcl-1 interaction (**Figure 1-8, A**) has been used as a scaffold (**1.11**) to investigate small molecule fragments active in displacing the wild-type helix-in-groove PPI (**Figure 1-8, D**).[178] This has led to the identification of potent small molecule inhibitors for the interaction (**1.20-21**).[178,179] This technique offers a new

strategy for drug discovery of helix-in-groove PPIs with greater hit rate† efficiency than traditional high throughput screening (HTS).[178,179]

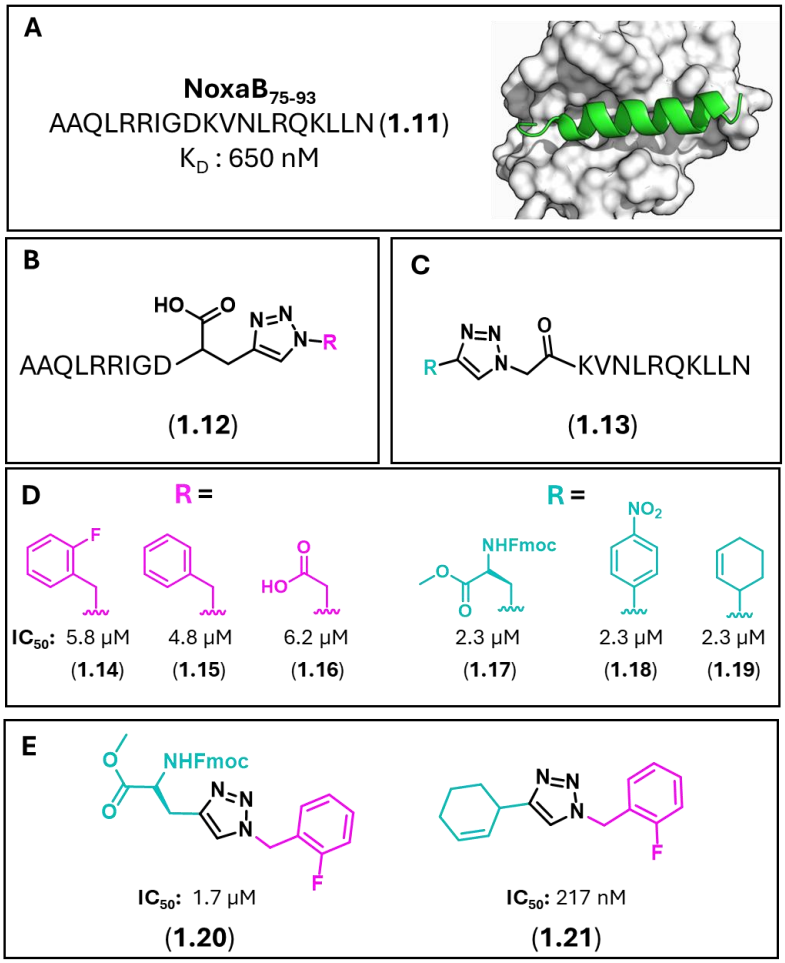


Figure 1-8 *Peptide directed-ligand design strategy (**A**) Noxa/Mcl-1 interaction (PDB: 2NLA) and NoxaB peptide (**B-C**) Preparation of triazole peptide sequences. (**D**) Peptide small molecule hybrid fragments with competitive activity against wild-type peptide. (**E**) Small molecules with improved affinity for PPI.f*

1.2.3 Inhibitors of coiled-coil mediated interactions

α-helices capable of forming helix-to-helix interactions are known as coiled-coils. These form from two or more α-helices winding into a super-helix.[181] Coiled-coils benefit from a repeated heptadic motif $(abcdefg)_n$ (**Figure 1-9**). These structures are amphipathic in nature, with hydrophobic amino acids conserved at positions *a* and *d* and polar residues at positions *e* and *g*.[182] On complementary monomers, *a* and *d* positions associate to create a hydrophobic core. Coiled-coils can exist naturally as dimers, trimers or tetramers. Coiled-coil with three or more helices are also referred to as helical bundles.

† A hit is defined as a molecule that has reproduceable target activity in a relevant screening assay. The hit rate is the percentage of active compounds out of the total number of compounds tested.

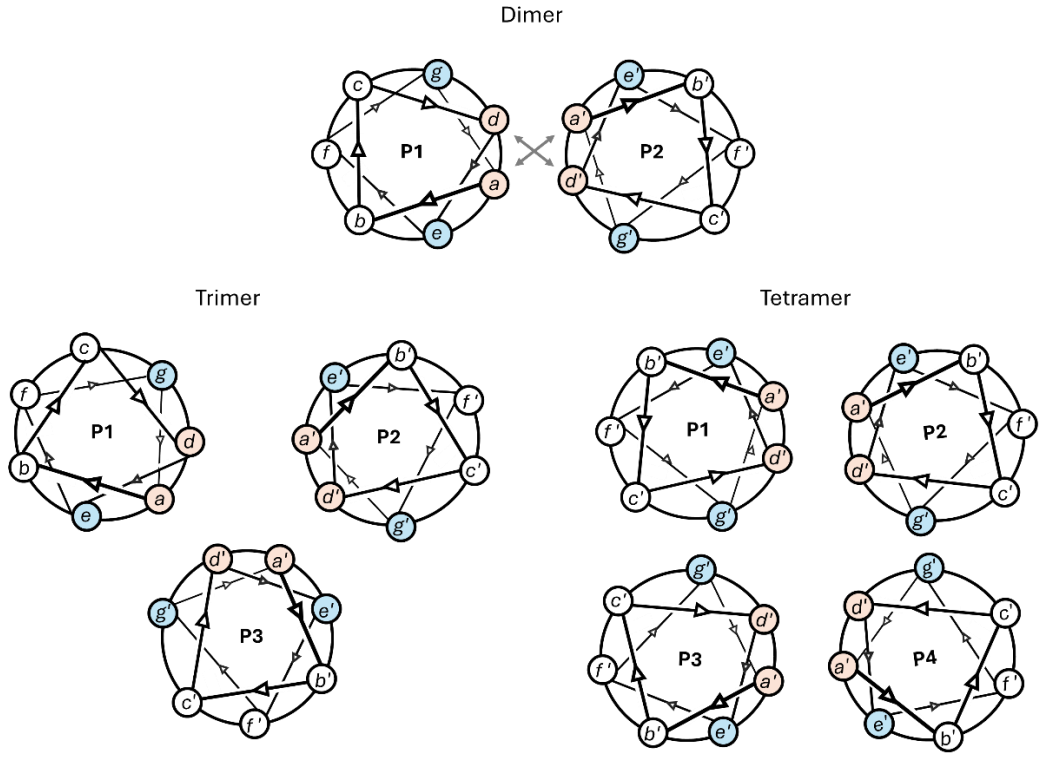


Figure 1-9: *Helical wheel diagram of residue positions on coiled-coil interacting proteins. Hydrophobic positions "a" and "d" (orange), polar positions "e" and "g" (blue).*

1.2.3.1 Coiled-coil inhibitor design strategy: rational peptide design

Coiled-coil inhibition by peptides has been reported through rational sequence design using organic constraints. Using native protein sequences peptides have been derived from dimeric coiled-coils forming transcription factor inhibitors against PPIs such as Myc-Max, Fos/Jun and GCN4.[87,183,184] Most recently, an Nrf2-derived peptide inhibitor against the coiled-coil interaction with MafG was disclosed in 2023. Based on a 16-mer sequence of the Nrf2 coiled-coil-forming helix, incorporation of a hydrocarbon staple across the external face of the helix identified one peptide with high affinity for the MafG coiled-coil (K_d of 337 nM).[93]

1.2.3.2 Coiled-coil inhibitor design strategy: combinatorial peptide libraries

Combinatorial libraries have been used to find peptide inhibitors against the Fos/Jun and microphthalmia associated transcription factor (MTIF) coiled-coil interaction.[87,91] Used in combination with protein fragment complementation assays, screening sequence libraries for improved coiled-coil interactions achieved lead peptide coiled-coil disrupters at low µM concentrations.[87,185]

1.2.3.3 Coiled-coil inhibitor design strategy: de novo peptide design

Due to the highly ordered nature of sequences found in coiled-coil interactions, it is possible to design *de novo* peptide sequences to inhibit this type of PPI.[186–188] Coiled-coil

forming peptide sequences are predicted using computational models and experimentally determined using peptide arrays.[186] For example a program called SYNZIP was used to generate peptides specific to basic leucine zipper (bZIP) and found that over 80% of peptide designs synthesised could bind to their target protein.[186] However, 40% were also capable of self-association, an added challenge in designing coiled-coil forming peptides.[189]

Most recently *de novo* peptide design was applied using a computational tool called CCbuilder, to form crosslinked helical dimers capable of selectively sequestering Myc and inhibiting PPI formation with Max. Researchers designed *de novo* peptides capable of stable coiled-coil formation in complex with Myc with nanomolar affinity, without disrupting the Max homodimer.[190]

1.2.3.4 Coiled-coil inhibitor design strategy: small molecule identification

Small molecule inhibitors of coiled-coil interactions are limited. This has largely been due to the shallow binding pockets characteristic of coiled-coils.[191] HTS has identified several small molecule inhibitors against the Myc/Max dimer.[94,192,193] Disruption of the interaction may occur through binding into the loop region in the centre of the helix-loop-helix of Myc, rather than directly binding to the Myc α-helix. *In silico* screening has helped improve the hit rate of HTS identification of small molecule inhibitors of the MDB2/p66α coiled-coil with two lead compounds demonstrating low micromolar activity (IC_{50} of 1.5-1 μM).[96] Other small molecule inhibitors are reported against larger helical bundle structures, found by HTS methods, perhaps benefitting from a deeper binding groove forming across multiple helices.[95,97]

Considering alternative routes to HTS may offer more efficient methods for disruption of this type of PPI by small molecules. The predictable heptadic assembly of coiled-coils offers guidance on key residues that could be employed to derive helical mimetic small molecules similar to those derived from helix-in-groove interactions.[194]

1.2.4 Inhibitors of β-strand mediated interactions

A β-strand is a 3-10 amino acid sequence forming a pleated backbone motif through tetrahedral bond formation around the Cα atom.[195] β-sheet structures occur from two or more β-strands connected laterally through hydrogen bond networks to form twisted flat sheets at dihedral angles of 135 °/ -135 °.[196] Most interactions mediated by β-strands can be categorised as a strand or hairpin into a binding groove or a β-sheet interaction with another β-sheet (**Figure 1-2**). Hotspot analysis of β-strand mediated PPIs has found

the majority employ residues on faces for hydrogen bonding or side chain interactions with a partner protein, often with uneven and unpredictable distribution.[32]

1.2.4.1 β-strand inhibitor design strategy: Rational peptide design modifications

Designing peptide mimetics for β-strand PPIs presents significant challenges compared to α-helices. In particular, β-strands can be prone to aggregation, driven by interstrand hydrogen bonding networks with adjacent strands. Consequently, peptide mimetics require careful design to minimise self-assembly.[197,198]

Short peptide sequences can fail to adopt secondary structure. Peptidomimetic strategies such as *N*-substitution of backbone amides can help support β-sheet formation (**Figure 1-10, B**). Most recently, *N*-methylation to constrain peptide backbones has been used to improve activity of inhibitors towards the small ubiquitin-like modifier (SUMO) interaction with the SUMO interacting motif (SIM) of RanBP2 (**Figure 1-10, A**).[199] *N*-methylation screening of a 13-mer peptide sequence (**1.21**) of the RanBP2 SIM improved the IC_{50} and K_D two-fold compared to the parent peptide in two of the twelve derivatives synthesised (**Figure 1-10, C**).[200]

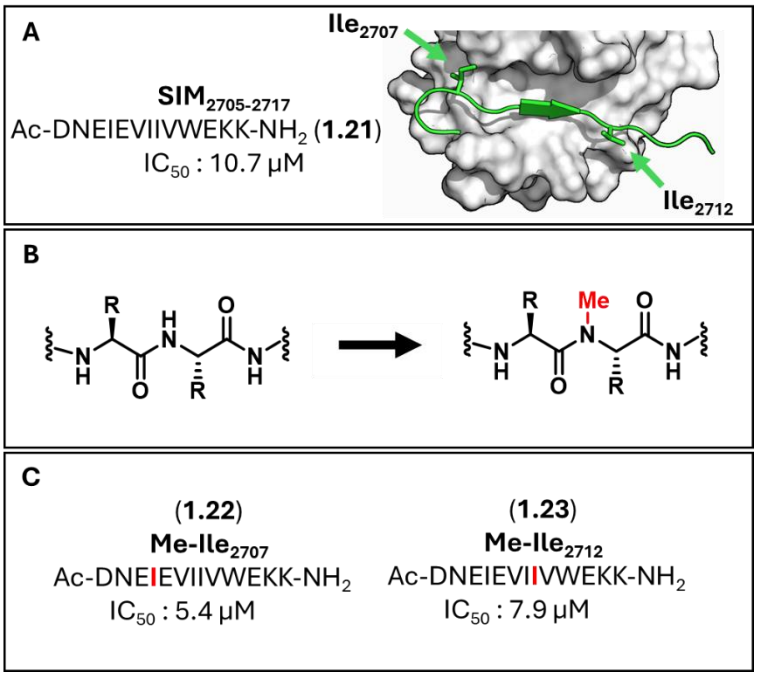


Figure 1-10: *(**A**) SIM/SUMO interaction (PDB: 2LAS) (**B**) N-methylation of the peptide backbones (**C**) Hit SIM peptides with N-methylated isoleucine residues.*

1.2.4.2 β-strand inhibitor design strategy: peptide-directed ligand design

Peptides offer an advantage over small molecules to afford β-strand inhibitors due to their ability to interact over a large flat surface.[201] Information gained from peptide inhibitors of β-sheets can direct efforts to afford small molecule inhibitors (**Figure 1-11**).

Peptide-directed ligand design has been applied to the Shank1 PDZ/GKAP PPI identifying peptide-small molecule hybrids with improved activity over the wild-type 6-mer GKAP derived β-strand peptide. Ac-EAQTRL-OH**,** (**1.23**) has a K_D of 1 μM determined by fluorescence polarisation against the Shank1 PDZ. Connecting a library of small molecule fragments to truncated 3-mer sequence (**1.24**) by arylhydrazone bond formation (**Figure 1-11, B**), led to the identification of hit compounds with low μM IC_{50} in competition with the parent peptide (**Figure 11, C**).[202]

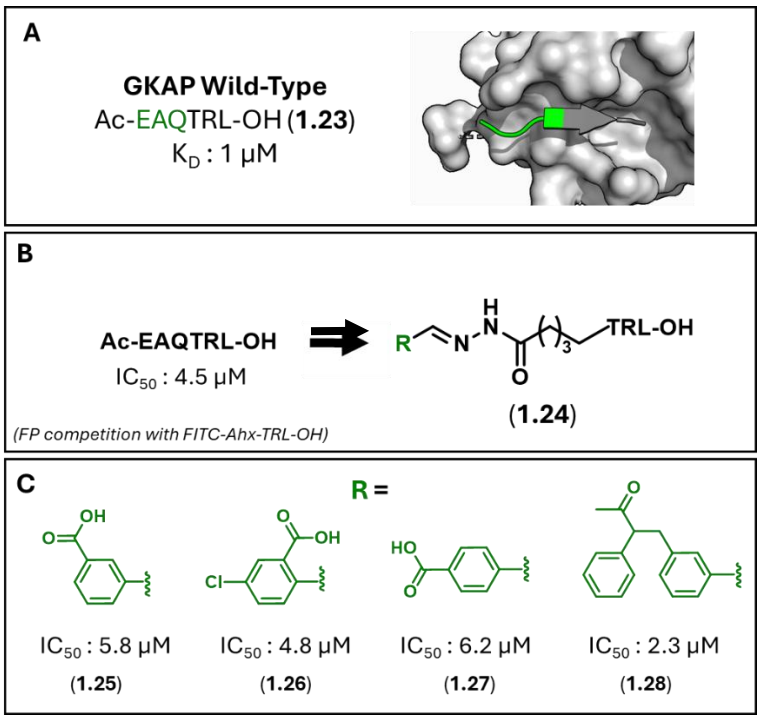


Figure 1-11*: Peptide-small molecule strategy for GKAP/SHANK1 inteaction. (**A**) GKAP C-terminal peptide (PDB: 1Q3P) (**B**) Preparation of peptide hydrazone structure (**C**) Structures of hit compounds with IC_{50} reported by competition fluorescence polarisation.*

1.2.4.3 β-strand inhibitor design strategy: Small molecule identification

Currently reported small molecule inhibitors of β-sheet interactions have been identified through HTS of large compound libraries, achieving remarkably low hit rates around 0.01%.[124,203,204] Characterised by hydrophobic binding interactions, small molecules found this way benefit from extended aromatic ring structures to capture π-π stacking interactions between the inhibitor and the target protein.[124,204–207]

In silico library screening has been applied to improve hit rates in conjunction with structure guided approaches to PPI inhibition. Interestingly, applying the same library screen towards a β-strand-in-groove interaction has been shown to produce a lower hit rate compared to an α-helix-in-groove PPI.[27] Recognition of this gap in pharmacophore

libraries designed to target β-strand interactions could guide development of tailored scaffold libraries, improving drug discovery toward alternative PPIs.

1.2.4.4 β-sheet peptide inhibitor design strategies

β-sheets form through three or more connected β-strands, twisted into parallel or anti-parallel pleats to create a large and flat surface area. Designing inhibitors of β-sheet interactions can be achieved through peptides derived from native sequences of interacting sheets. In this approach cyclic or hairpin peptides sit flat across the β-sheet surface, with every other side chain pointing towards the β-sheet surface, directly mimicking a β-sheet interaction.

The PD-1/PD-L1 interaction provides an example where β-hairpin structures have been used to induce turn structures in peptides and small molecules (**Figure 1-12**).[208–210] Where the native peptide sequence of a hairpin β-sheet of PD-L1 (**1.29**) has affinity for the PD-1 β-sheet, residue mutation from glycine to proline improved binding affinity 2-fold (**1.30**). Peptides capable of disrupting the PD-1/PD-L1 PPI have also been found by phage display techniques (**Figure 1-12, C** and **E**). Linear peptides (**1.31**) were further improved by installing an azobenzene turn unit (**1.32**) generating a light-activated β-hairpin forming peptide (IC_{50} of 79 nM), with greater activity than the parent peptide (IC_{50} of 4.6 μM).[210]

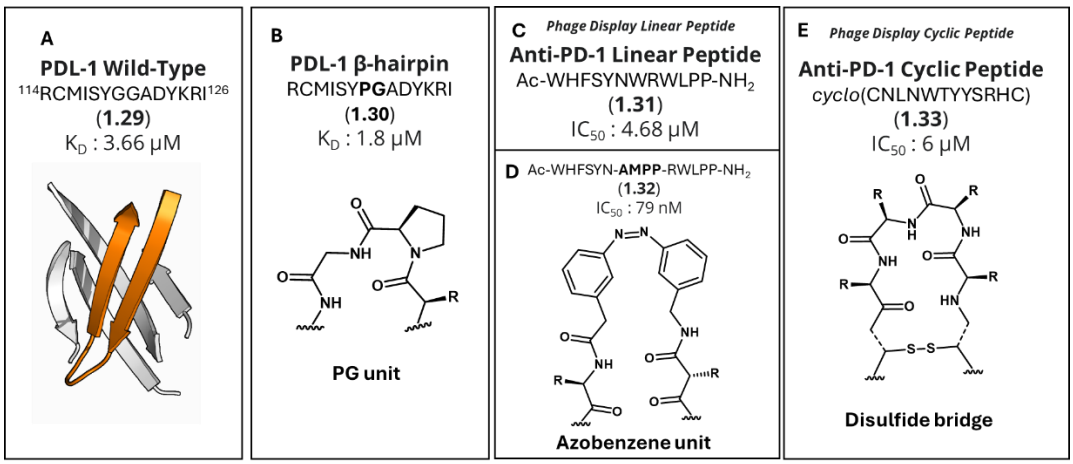


Figure 1-12 *PD-1/PDL-1 inhibitor strategies (**A**) Native PD-L1 derived peptide (orange) interacting with PD-1. (PDB: 4ZQK) (**B**) Proline-Glycine turn unit installed into native PD-L1 peptide. (**C**) Phage display derived PD-1 binding peptide (**D**) Azobenzene unit containing peptide (**E**) Cyclic phage display peptide containing disulfide bridge.*

Macrocyclisation-inducing units can be applied to peptide inhibitors to constrain β-sheet mimetics, such as the ICOS/ICOS-L interaction (**Figure 1-13, A**). Employing known key residues involved in the ICOS-L β-sheet, Tyr51, Tyr53 and Gln55, an *in silico* design strategy was used to design a macrocyclic peptide towards the ICOS β-sheet.

Introducing Pro-Gly, D-Pro-Gly turn units, disulfide bridge combinations and residue optimisation, led to a 12-mer bicyclic peptide inhibitor (**1.34**). This inhibitor shows an IC_{50} of 22.8 μM by TR-FRET against the ICOS/ICOS-L β-sheet interaction (**Figure 1-13, B**).[211]

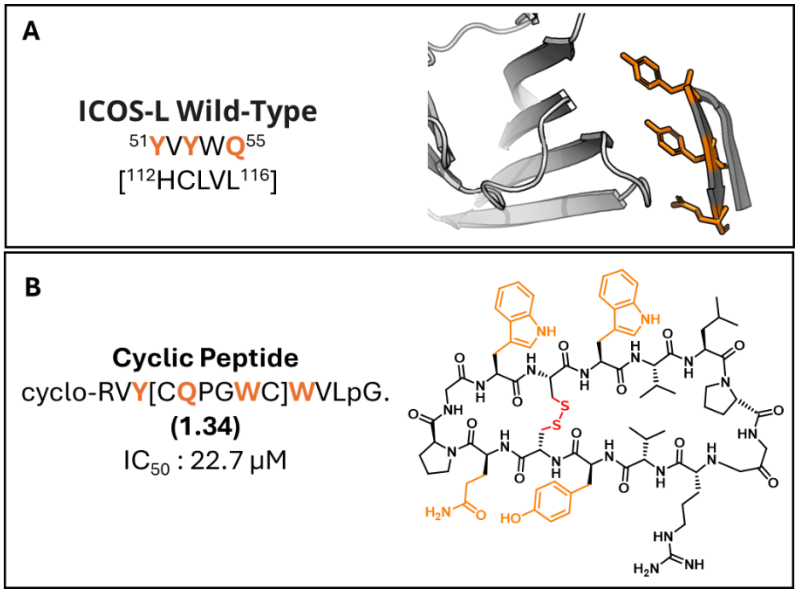


Figure 1-13 *ICOS/ICOS-L interaction* ***(A)*** *Interaction of ICOS-L β-strand, key residues highlighted in orange (PDB: 6X4G)* ***(B)*** *Cyclic peptide structure of ICOS-L derived bicyclic peptides proposed key residues highlighted in orange.*

1.2.4.5 Small molecule inhibitors of β-sheets

Small molecules have been found through HTS and fragment screening against β-sheet mediated PPIs.[212–214] However this approach remains relatively inefficient, with most inhibitors reporting micromolar inhibitory activity. Even the most potent of small molecule inhibitors found by HTS are a product of hit-rates below 2%.[215]

1.2.5 Inhibitors of dynamic protein region interactions

PPIs mediated by intrinsically disordered regions (IDR) often lack a defined secondary structure, relying on hotspot residues within a dynamic recognition sequence, or constrained loops between secondary structures that can be reproduced by macrocyclisation. Often disordered PPI structures will occupy a well characterised binding pocket, which can be exploited when designing inhibitors. Loops between fixed secondary structures are less flexible that true IDRs, but their flexibility has made inhibitor design challenging. As such, methods to constrain loop structures have been applied to peptide inhibitor design. Reports of inhibitors against dynamic PPIs has increased the most in the last decade (**Table 1-1**). Inhibitors are increasingly

incorporating structural information from peptide inhibitors to guide small molecule discovery.[135,138,140,141]

1.2.5.1 IDR inhibitor design strategies: peptide combinatorial libraries

Disordered PPIs forming from recognition sequences can be targeted by native and random peptide sequences (**Figure 1-14**). An example of this is the HIV-1 Gag p6 protein binds to host Endosomal Sorting Complex Required for Transport (ESCRT-I) Ubiquitin E2 variant (UEV) subunit through a disordered 9-mer sequence (PEPTAPPEE) (**1.35**). As the ESCRT protein is highly structured, information of the binding groove of the PPI is well characterised, despite the lack of secondary structure from the p6 binding protein (**Figure 1-14, B**).[216]

Whilst small molecules disrupting the UEV-p6 PPI have yet to be reported, peptide modifications have been explored using screening by genetically encoded libraries. Screening of the lanthipeptide library of 10^6 macrocyclic peptides was performed in a bacterial display system to confirm PPI inhibition. A lead peptide (**1.36**) demonstrated a 3-fold improvement over the parent peptide in binding to the UEV protein.[217] Interestingly the hit peptide bears no sequence similarity to the wild-type, raising questions as to the mode of PPI inhibition (**Figure 1-14, B**).

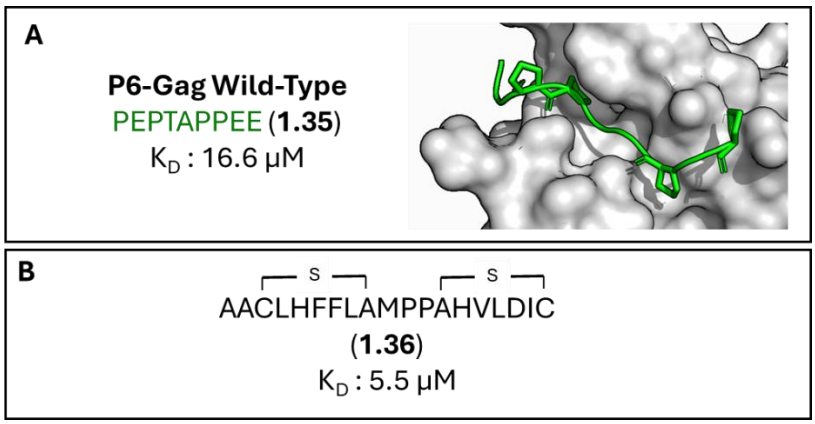


***Figure 1-14**: UEV/p6-Gag interaction (**A**) 9-mer p6 proline recognition sequence (green, **1.35**) (PDB: 3OBU) (**B**) macrocyclic peptide (**1.36**) identified through library screening with improved affinity for UEV protein.*

1.2.5.2 IDR inhibitor design strategies: peptidomimetics

Peptidomimetics can improve inhibitor activity, utilising hot-spot residues as an anchoring scaffold to explore the surrounding chemical space. This has been achieved in the IDR PPI between a phosphoprotein recognition sequence to the BRCA1 protein (**Figure 1-15, A**).[218] Utilising a small molecule microarray, peptidomimetics were generated (**1.38**), conserving a key phosphoserine anchoring to the binding pocket (**Figure 1-15, B**).[218–220] This led to a 3-fold improvement in IC_{50} against the native

phosphopeptide binding sequence (**1.37**), by retaining the key Trp, pSer and Phe moieties of the native peptide sequence (**Figure 1-15, C**).

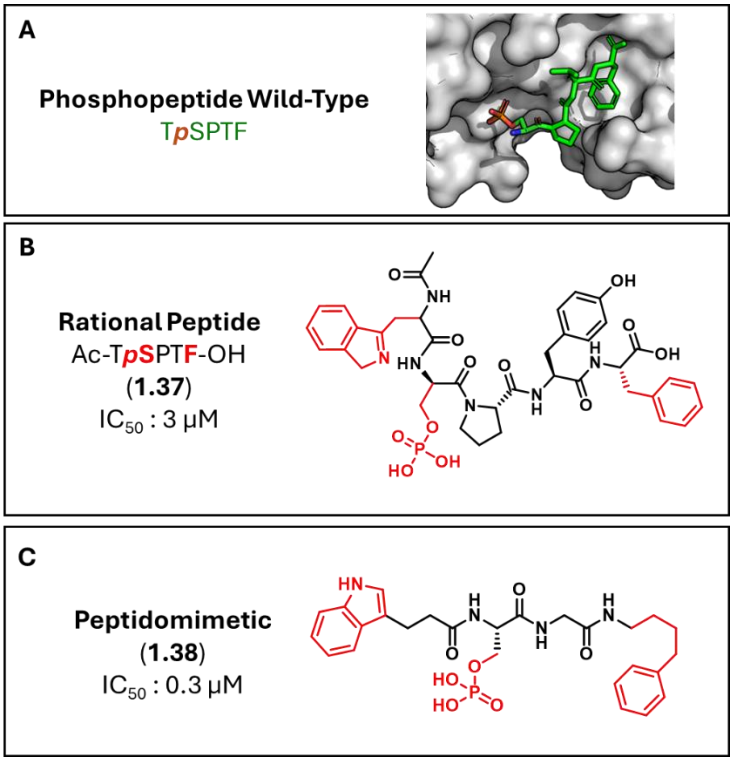


Figure 1-15: *(A) Phosphopeptide (green) recognition sequence interacting with BRCA1 (PDB: 3K0K)* ***(B)*** *Phosphopeptide derived peptide (**1.37**), hotspot residues highlighted in red. (**C**) Lead phosphopeptidomimetic (**1.38**).*

1.2.5.3 IDR inhibitor design strategies: small molecule identification

Rational approaches to IDR PPI inhibitors have yielded some highly potent small molecule inhibitors. The well-defined binding pocket found in the bromodomain and extra-terminal domain (BET) family of proteins relies on acetylated lysine (KAc) for recognition of IDR PPIs with histone proteins (**Figure 15, A**).[221,222] The first examples of bromodomain PPI inhibitors investigated thienodiazepine structures with anti-inflammatory activity. Structure-activity-relationship studies developed a small molecule (**1.40**) capable of displacing peptide binding to the bromodomain-containing protein 4 (BRD4). Modelling small molecules off the key KAc residues achieved nanomolar IC_{50} values in competition against the parent peptide (**1.39**).[222]

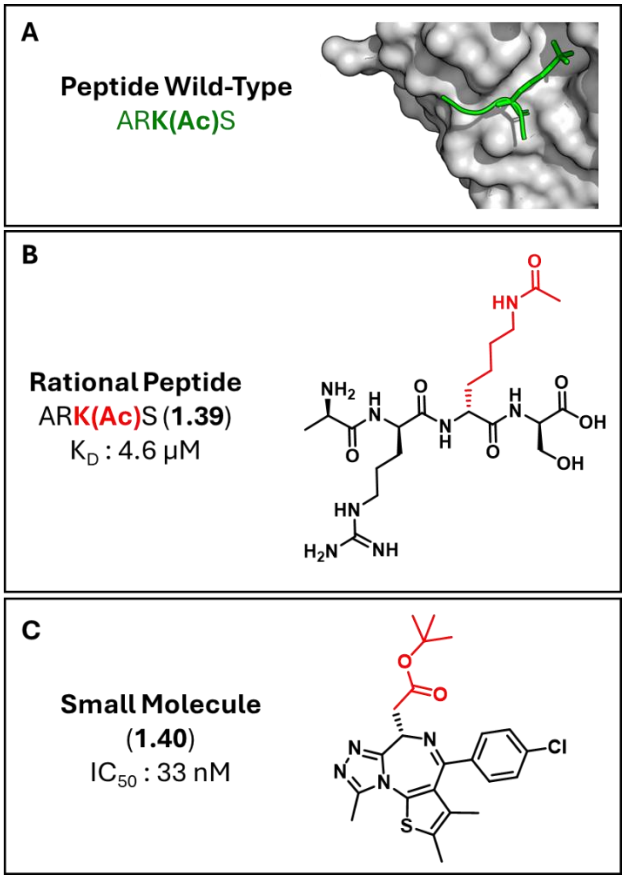


Figure 1-16 *BRD4/Histone interaction* ***(A)*** *protein H3K14ac peptide (****1.39****) responsible for acetylated histone recognition. (PDB: 3JZG), 4-mer acetylated lysine recognition site (green) (****B****) Acetylated peptide with affinity for BRD4 protein (****C****) Small molecule JQ1 (****1.40****) disrupting the BRD4/Histone complex.*

1.2.6 Inhibitors of loop mediated interactions

An emerging target for PPI inhibitors are loop structures at interaction sites, with evidence from the PDB that as many as 50% of protein complexes feature mediation by loops.[223,224] The majority of characterised loop interactions were β-turns (31%) followed by loops between α-helices (11%).[223]

1.2.6.1 Loop mediated inhibitor design strategies: rational peptide design

Loops present a unique challenge compared to α-helices and β-sheets interactions, as they lack predictable structures to manufacture standard scaffolds across a range of targets.[225] Identification of key hot-spot residues within loops has allowed for the identification of attractive interfaces such as Nrf2/Keap1 and YAP/TEAD.[223]

Macrocyclisation of wild-type peptides can improve peptide affinity in loop mediated PPIs (**Figure 1-17**). For example, inhibition of the loop unit of Nrf2 engaging the Keap1 β-propeller has been achieved using a glycine linker for head-to-tail cyclisation of the Nrf2 derived linear LDPETGEFL (**1.41**) improving the K_D from 86 nM to 18 nM in SPR, demonstrating a 4-fold increase in binding affinity (**1.42**).[226]

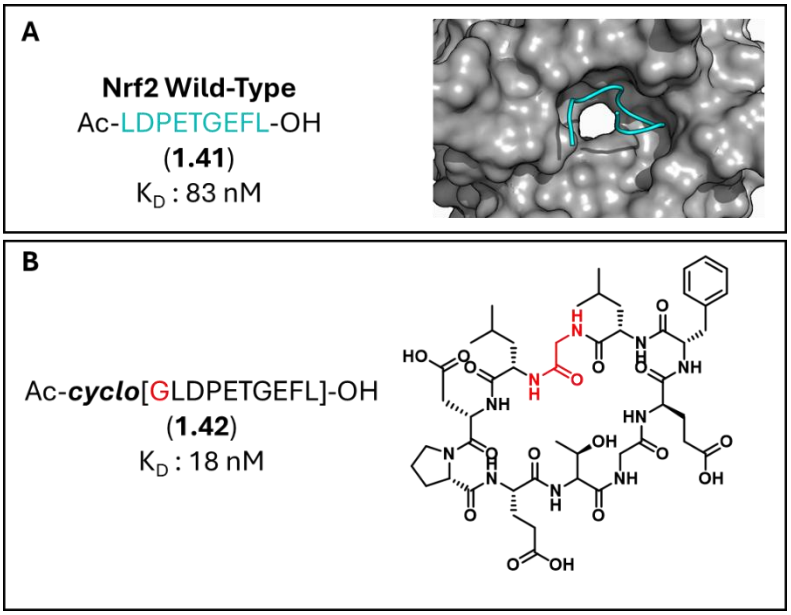


Figure 1-17: *Nrf2/Keap1 interaction* ***(A)*** *Nrf2 ETGE loop motif bound to Keap1 (PDB: 2DYH). Linear peptide (**1.41**) and cyclised Nrf2 derived peptide (**1.42**).*

Peptides incorporating non-natural amino acids have achieved potent inhibition of loop mediated interactions, such as the YAP-TEAD helix-loop-helix (**Figure 1-18, A**). Investigation by mutation studies has identified hot-spot residues and key positions for installation of turn units.[36] This led to a 15-mer peptide inhibitor (**1.44**) improving linear peptide inhibition from IC_{50} of 68 μM to 9.2 nM (**Figure 1-18, B**).[227]

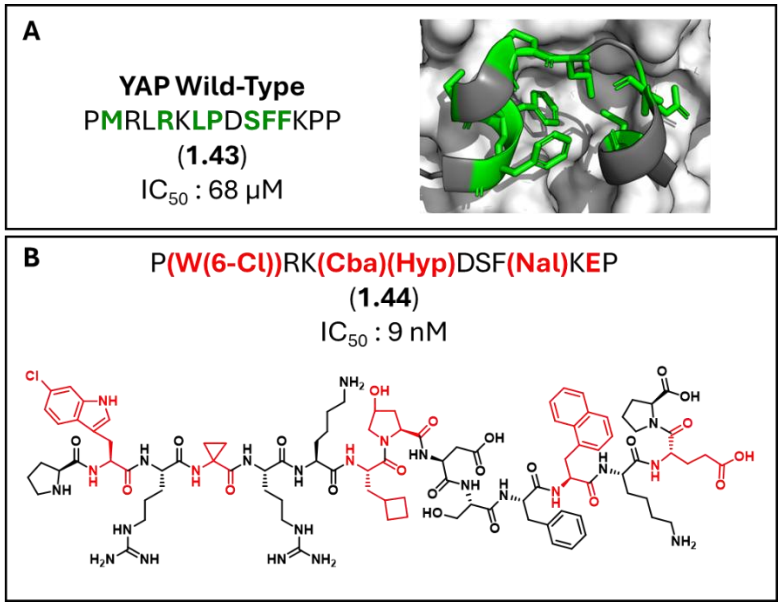


Figure 1-18: (A) *YAP/TEAD interaction (PDB:6Q36)* ***(B)*** *linear (**1.43**) and* ***(C)*** *residue mutated peptide inhibitor (**1.44**).*

1.2.6.2 Loop mediated inhibitor design strategies: small molecule identification

Following similar trajectories to the inhibitors of α-helix and β-strand mediated PPIs, screening assays have identified inhibitors of loop mediated PPIs, achieving similarly

low hit rates.[203] However, once a hit compound has been found, SAR studies can achieve greater potency towards a target PPI.

A well-documented example of this is the naphthalene containing small molecules found to inhibit Nrf2/Keap1.[228] Initially found by HTS against the wild-type peptide, several iterations explored the chemical space occupied in the hydrophobic binding pocket formed by Keap1.[229–233] Fragment-based drug discovery has also been applied towards the Nrf2/Keap1 PPI achieving 10-fold improvement in hit rate over traditional HTS.[234]

1.3 Conclusion

Drug discovery on PPIs has numerous applications for disease management, often finding molecules with high specificity for target. The literature collated here provides a snapshot into the range of success in the inhibition of different types of PPI.

α-helix mediated interactions are the most prominent secondary structure found within the proteome and are involved in over **50%** of the druggable PPIs presently investigated, predominantly describing helix-in-groove PPIs. Interestingly, coiled-coils represent an underexplored PPI within helix-mediated PPI inhibitors. Utilising tools for helical mimetics or fragment-based drug discovery to design small molecules of these PPIs could offer a new method for targeted coiled-coil interactions.

β-strands, less common in overall protein content, were featured in **30%** of PPIs with inhibitors. Strand-in-groove interactions have been described as more challenging to find hits against by HTS in comparison to helix-in-groove PPIs.[27] Alternative methods to fragment-based inhibitor discovery, such as peptide directed ligand design, may offer improved efficiency to small molecule design of β-strand PPIs.[202]

The remaining interactions with inhibitors found could be described as dynamic PPIs, including loops and IDR recognition sequences. Peptide inhibitors of β-strands and loop structures benefit from macrocyclisation techniques to stabilise the desired secondary structure, whilst also improving cell permeability and protection against degradation. Structure-based design of small molecules, guided by key PPI hotspot residues enables the development of potent inhibitors with nanomolar affinity for IDR PPIs. This rational SAR-driven approach represents a promising direction for future PPI drug discovery.

Despite advances in drug discovery, small molecule inhibitors across all categories of PPIs are still predominantly found by HTS, typically observing hits at rates lower than

1%.[235,236] Some HTS techniques use competition assays against peptide inhibitors derived from native PPI interfaces. Targeting hot-spot residues and sequences critical for binding, combined with fragment-based drug discovery, could afford greater selectivity and efficiency in small molecule design.[177–179,202]

There is an increasing abundance of structural data published depicting the secondary structures of PPIs, complemented by machine learning programs such as AlphaFold3, which can help to visualise challenging dynamic proteins in complex where crystal structure data is unavailable. Methods to identify PPI inhibitors rely on this structural data, and as such the inhibitors identified are skewed to the more stable interactions. The growth of structural biology methods such as Cryo-EM more readily allows for the characterisation of protein complexes, revealing structural details about intrinsically disordered regions, and multiple protein partners. This data, coupled with protein complex structure prediction will rapidly increase the targets for inhibitor development. There remains a challenge to develop methods more suited to less well studied interactions, particularly IDRs and coiled-coils.

1.4 Transcription Factors

1.4.1 Transcription factor categories

Transcription factors (TF) influence DNA conversion into RNA, initiating or repressing gene transcription. Characterised by specific DNA-binding domains, transcription factors interact with operator sequences at or close to promoter sequences in DNA to form the transcription initiation complexes or create transcriptional repression.[237]

There are currently over 1600 identified TFs in the human genome.[238] TFs can be general or tissue specific and are categorised based on the type of DNA-binding domain observed (**Figure 1-19**), with nearly half of all TFs being zinc fingers (ZF) or homeodomains (HD).[239,240] In the minority there are basic helix-loop-helix (bHLH), basic leucine zippers (bZIP) and nuclear hormone receptors (NHR).

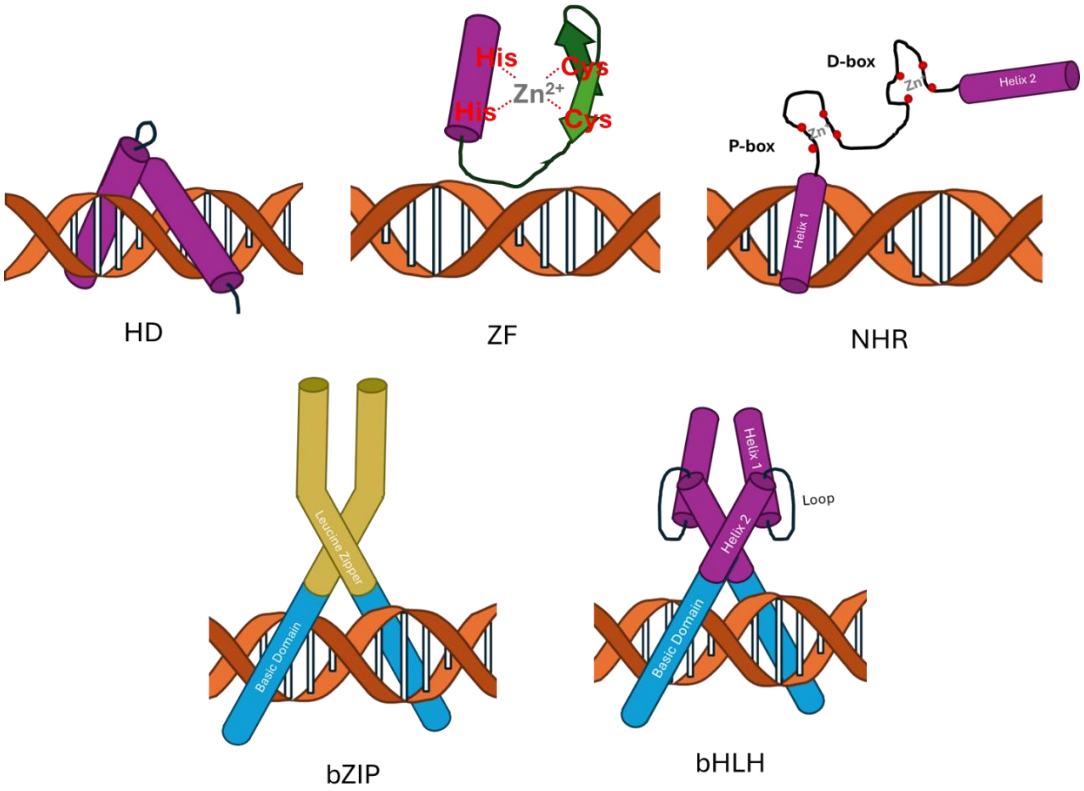


Figure 1-19 *Transcription Factor categories. Homeodomains (HD) zinc fingers (ZF), nuclear hormone receptor (NHR), basic leucine zipper (bZIP), basic helix-loop-helix (bHLH) transcription factors. Purple cylinders are α-helices. Blue cylinders are basic α-helices and yellow cylinders are leucine zippers. Green arrows are β-sheets. Amino acid residues (red) coordinating with Zn^{2+} (grey).*

Zinc Fingers (ZF) occur most prevalently in the mammalian genome, recognisable for their Zn^{2+} ion coordination with two cysteine and two histidine residues to form the 'fingers' of the DNA recognition site.[241–243] ZF transcription factors exist as monomers and do not require protein dimerisation to facilitate DNA binding, however ZF can exist as multimers for the purpose of enhancing transcription.[244] The classical ZF structure consists of an α-helix and two β-sheets within the finger structure (**Figure 1-19**).

Homeodomains (HD) TFs form a helix-turn-helix motif (**Figure 1-19**), with the second helix and turn structure directly interacting with the DNA backbone, and can act as monomers or dimers.[245] It is known that the basic region of the second helix in HD proteins confers the DNA binding specificity through a major groove DNA interaction.[246]

Nuclear hormone receptors (NHR) are ligand activated TFs, responsible for hormone induced gene activation. They contain a distinct DNA-binding domain and ligand binding dimerisation domain (**Figure 1-19**). The DNA-binding domain models three α-helices and two zinc fingers, each coordinating the four residues, identified separately as the P-box and D-box coordination sites. The P-Box, found in helix-1 informs recognition of the DNA response element through major groove binding. Helix-1 is supported by Helix 3 for

structural packing. Helices-1 and -2 exist 90° rotated from one another to form the DNA-binding domain core. The D-box, at the end of helix- 2, assists in protein dimerisation.[247]

Basic helix-loop-helix (bHLH) TFs consist of a basic domain and the helix-loop-helix motif (**Figure 1-19**).[248] bHLH is the largest family of dimerising TFs, facilitating both homo and heterodimerisation for stability during DNA binding. bHLH TFs bind to E-box DNA, a 6 nucleotide consensus sequence that bHLH proteins basic domain can recognise.[249] Protein dimerisation is essential for bHLH-DNA binding created through coiled-coil interactions between the first helix on each bHLH protein.

Basic leucine zipper (bZIP) transcription is mediated through a basic DNA-binding domain and requires TF dimerisation for stable DNA binding interactions (**Figure 1-19**). The basic domain hosts positively charged arginine and lysine residues to support major groove binding on transcription response elements.[250] bZIP dimerisation is facilitated through a coiled-coil motif in the hydrophobic leucine zipper domain to create highly specific PPIs to facilitate transcription factor activity.

1.4.2 Transcriptional regulation in health and disease

Through TFs' extensive role in cell regulation, they are highly involved in disease progression, with 20% of TFs linked to a disease phenotype in the human genome.[238] Dysregulation of cellular metabolism and growth can lead to the development of cancer.[251] TFs can also manipulate our immune system, controlling inflammatory and immune disorders.[252,253] As such, TFs have been extensively studied as drug discovery targets.[254] Transcription factors often lack well defined binding pockets for small molecules to sit in. Often TFs feature significant intrinsically disordered regions, and as a result, only a small percentage of disease relevant TFs have been targeted.[238] Key domains of transcription factors have been identified for potential sites for drug targets, demonstrated in **Figure 1-20**.

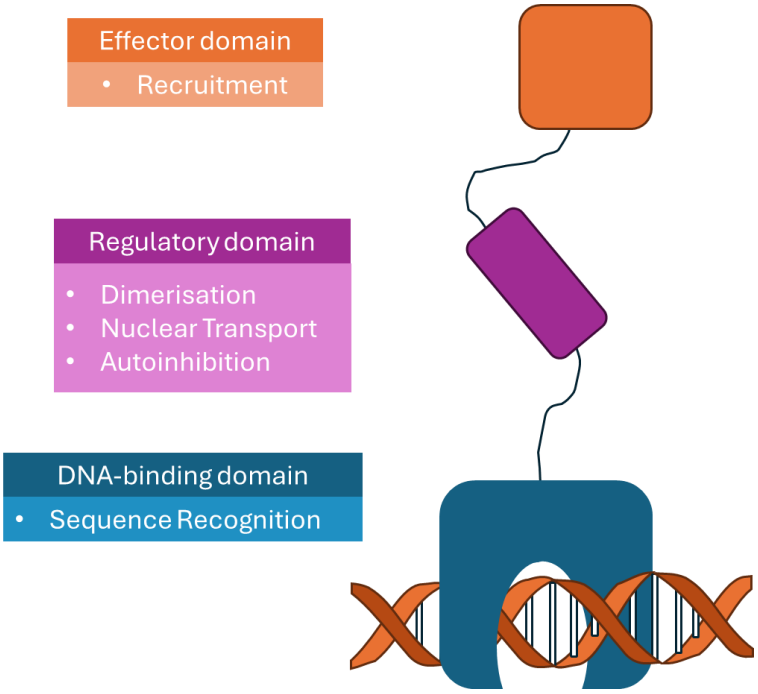


Figure 1-20 *Transcription factor domains. DNA-binding domain (blue) regulatory domain (purple) effector domain (orange). Adapted from Henry and Koehler, 2021.*[254]

Effector domains of TFs are used to inhibit activity where essential PPIs with coactivators are observed.[254] However, elucidating the structure of coactivator binding complexes is challenging as in the unbound state, they are often intrinsically disordered.[255–258]

Equally, targeting regulatory domains that are essential to bZIP and bHLH activity can pose challenges as often TF structures remain intrinsically disordered until a binding partner is present.[259] Regulatory domains are essential for controlling protein dimerisation, mediating nuclear transport and autoinhibition, making them attractive targets. Control of PPIs at regulatory domains can block downstream transcription of disease relevant genes. The design of PPI inhibitors to disrupt protein dimerisation has proven challenging due to their transient and dynamic structures.

Several inhibitors of bHLH have been reported against the Myc/Max heterodimer. As described in **Chapter 1.4.1** small molecules have been found by HTS and *in silico* docking studies, overcoming the difficulty of targeting dynamic structures. However precise binding mechanisms of these compounds remain unknown.[192,193,260–262] Protein and peptide inhibitors of Myc/Max have also been developed through design of coiled-coil binding peptides to sequester the Myc protein or to compete with TF/DNA binding by designing mini-proteins.[183,263–265] Reported inhibitors of Myc have struggled to reach the clinic due to low bioavailability.[261,264] The only Myc inhibitor to reach clinical trials to

date is Omomyc, a mini protein that competes with the native protein dimers for the DNA-binding domain to block TF activity.[264,265]

The DNA-binding domains of transcription factors are well defined, however, often observe non-specific binding interactions.[266] TFs most often bind to the major groove of DNA (**Figure 1-21**), with regulation of specificity defined by minor groove interactions.[267] Inhibitors to the TF/DNA interaction can either target the protein or DNA binding site.

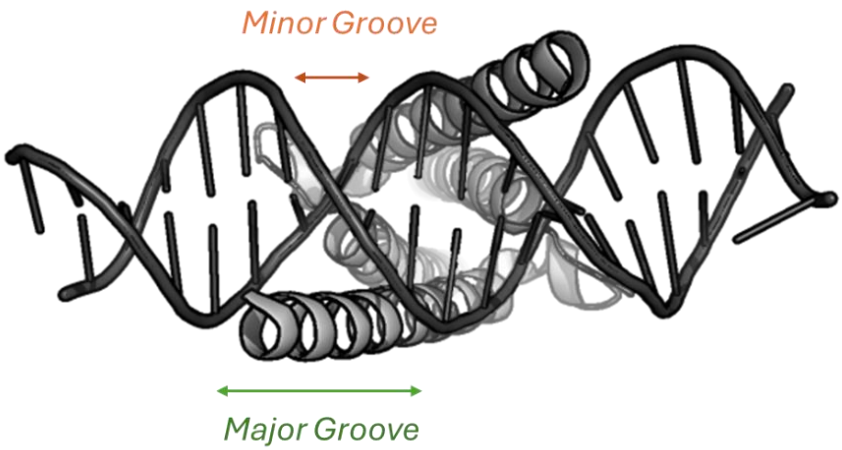


Figure 1-21 *Myc/MAX bHLH dimer bound to DNA in the major groove (PDB: 1NKP).*

Inhibition by binding to the DNA-binding domain by small molecules has been explored by HTS, but has failed to produce particularly potent compounds.[268,269] Similarly, whilst screening of small molecules against transcription factor recognition sequences has been achieved, identification of fragments with selective and potent activity towards transcription factor recognition sequences remains a challenge.[270,271,272]

The development of direct inhibitors of transcription factors to study their involvement in disease states provides a powerful research tool for drug discovery leading to the development of clinical candidates.[273] Consequently, it is imperative we continue to research transcription factors with promising interaction sites for targeted inhibitor design.

1.5 Nrf2

Nuclear factor erythroid 2-related factor 2 (Nrf2) is a basic leucine zipper (bZIP) transcription factor involved in the regulation of oxidative stress in vertebrates.[274] Its activity drives transcription of antioxidant response element (ARE) dependent genes, for the support of cells physiological response to oxidation.[275] Consequently, Nrf2 is an emerging target in inflammatory diseases for the transcription of cellular protections

against stress factors, oxidants and cytotoxic molecules.[276,277] Conversely, overactivation of Nrf2 is routinely observed in tumours, supporting cancer cell survival and promoting chemotherapy resistance.[278] Further research into the complex pathophysiology of Nrf2 will inform drug discovery of Nrf2 therapeutics across health and diseases in humans.

1.5.1 Structures of Nrf2 and MafG

Nrf2 is highly conserved across vertebrates, the protein structure can be organised into seven functional domains, Neh1-7 (**Figure 1-22, A**). Of which, structures have been experimentally resolved for Neh1 and Neh2.[279]

Cytosolic Nrf2 interacts with Kelch-like erythroid cell-derived protein (Keap1) through the Neh2 domain (**Figure 1-22, B**). Keap1 forms a homodimer through binding to two short sequences in the Neh2 domain; $_{79}$ETGE$_{82}$ and $_{29}$DLG$_{31}$.[280] The interaction between Nrf2 and Keap1 is often described as a 'hinge and latch' mechanism. The ETGE motif serves as a high affinity 'hinge' whilst the DLG motif has a lower affinity as the 'latch' (**Figure 1-22, B**). As a substrate adaptor protein for an E3 ubiquitin ligase complex, Keap1 bound Nrf2 is quickly degraded. As a result, under basal conditions, cytosolic Nrf2 has a half-life of 10-20 minutes, caused by the binding interaction with the Keap1.[281] Keap1 is cysteine rich and these thiol containing residues can be modified by reactive oxygen species to induce conformational changes that inactivate Keap1, allowing Nrf2 to translocate to the nucleus.[280]

As highlighted in **Figure 1-22, A,** Neh3-5 are transactivation domains (tAD) for coactivator complex formation during transcription. Neh6 supports β-transducin repeat-containing protein (β-TrCP) binding, which further enables ubiquitin ligase complex formation. Neh7 facilitates retinoic acid receptor-α (RARα) binding which prevents Nrf2 nuclear translocation and supports protein degradation.

Nrf2 binds to DNA through the Neh1 domain, containing the basic DNA-binding domain, and regulatory domain for TF dimerisation (**Figure 1-22, C**).[282] The DNA binding interaction is stabilised through heterodimerisation of the Nrf2 leucine zipper region with other bZIP transcription factors (**Figure 1-22, C**). Neh1 also contains a nuclear exportation sequence to promote Nrf2 migration out the cell nucleus, facilitating transient transcription factor activity (**Figure 22, C**).[283,284]

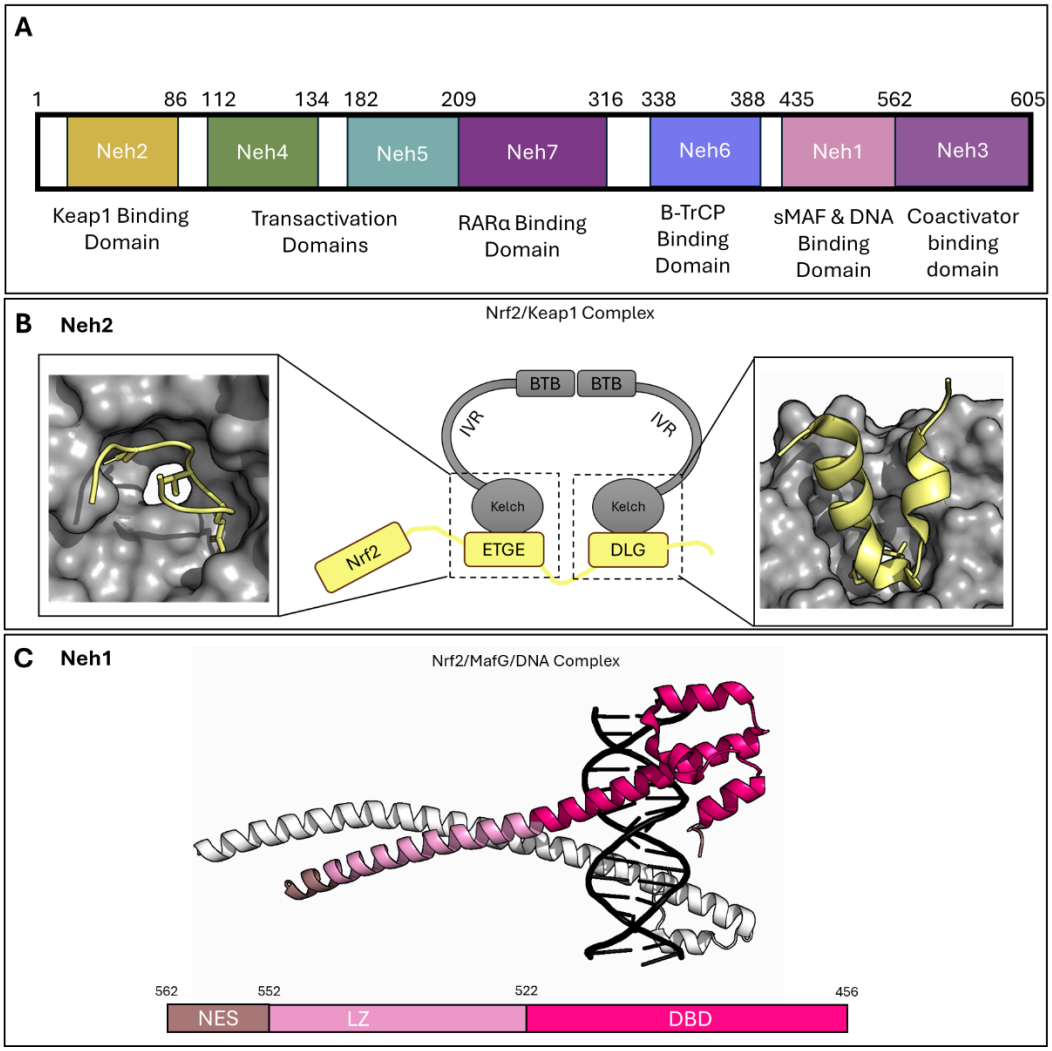


Figure 1-22 ***(A)*** *Neh1-7 domains of Nrf2 protein.* ***(B)*** *Crystal structures of Neh2 domains of Nrf2 (yellow) bound to transcriptional repressor protein Keap1 (grey) (PDB: 5FWV and 3WN7).*[279] ***(C)*** *Crystal structure of Neh1 DNA-binding domain (DBD, hot pink) leucine zipper (LZ, light pink) nuclear exportation sequence (NES, dark pink) Nrf2/MafG/DNA. (PDB: 7X5F).*[285]

The Neh1 domain towards the C-terminus of Nrf2 folds upon binding to a small musculoaponeurotic fibrosarcoma (sMaf) bZIP protein, forming a heterodimer that stabilises ARE DNA binding.[285] The sMaf protein family contains MafF, MafG and MafK, each unique bZIP transcription factors as they do not feature transactivation domains. As such, sMAF homodimers cannot initiate gene transcription, instead acting as repressors. sMaf contains a basic domain and leucine zipper domain shown in **Figure 1-23**.[116]

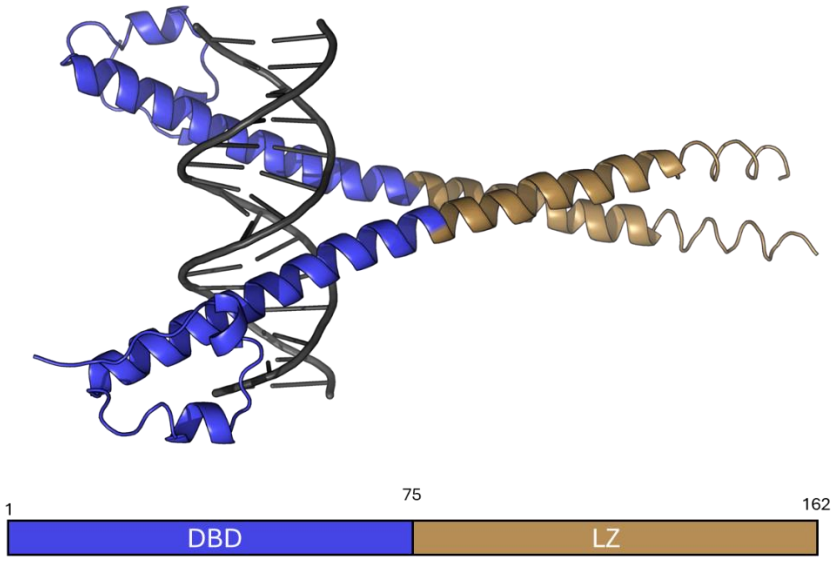


Figure 1-23 *MafG/DNA binding complex (PDB: 3A5T) DNA-binding domain (DBD) in blue, leucine zipper (LZ) in brown.*

Nrf2 is not capable of forming homodimers to stabilise DNA, instead sMaf proteins are required for heterodimerisation.[287] In sMaf knock out models, it is observed that MafG inhibition consistently creates cellular abnormalities, inflammation and tumour progression associated with aberrant Nrf2 transcriptional regulation.[288]

Nrf2 forms a parallel coiled-coil interaction with MafG, between the leucine zipper regions of the respective proteins. Regularly repeating leucine residues support the coiled-coil motif, creating a hydrophobic zipper that stabilises the heterodimer.[289] Heptadic repeats $(abcdefg)_n$ of the coiled-coil motif have been mapped to the Nrf2 and MafG leucine zippers, conveyed in the helical wheel diagram in **Figure 24, A-B**.[285,290] Key residues at positions *a* and *d* in the heptad motif create hydrophobic interactions causing protein dimerisation (**Figure 1-24, C**).

In the discovery of the Nrf2 protein, it was predicted residues 505-551 constituted the leucine zipper, theorising that Asn_{546} occupying *d* position in the heptadic repeat disrupted Nrf2 homodimer formation (**Figure 1-24, B**).[274] Consequently, MafG acts as an obligatory binding partner and critical regulator of Nrf2 gene transcription. Outside of the hydrophobic core, intermolecular salt bridges are observed between *g* positioned Asp_{528} of Nrf2 and *a* positioned Lys_{83} in MafG, as well as *a* positioned Lys_{536} of Nrf2 and *g* positioned Glu_{89} of MafG, residues in Nrf2 that are highly conserved in the bZIP sub-family.[285] This may also clarify why Nrf2 heterodimers are more preferable, as Lys_{83} and Glu_{89} equivalents do not occur in the Nrf2 leucine zipper to stabilise homodimerisation.

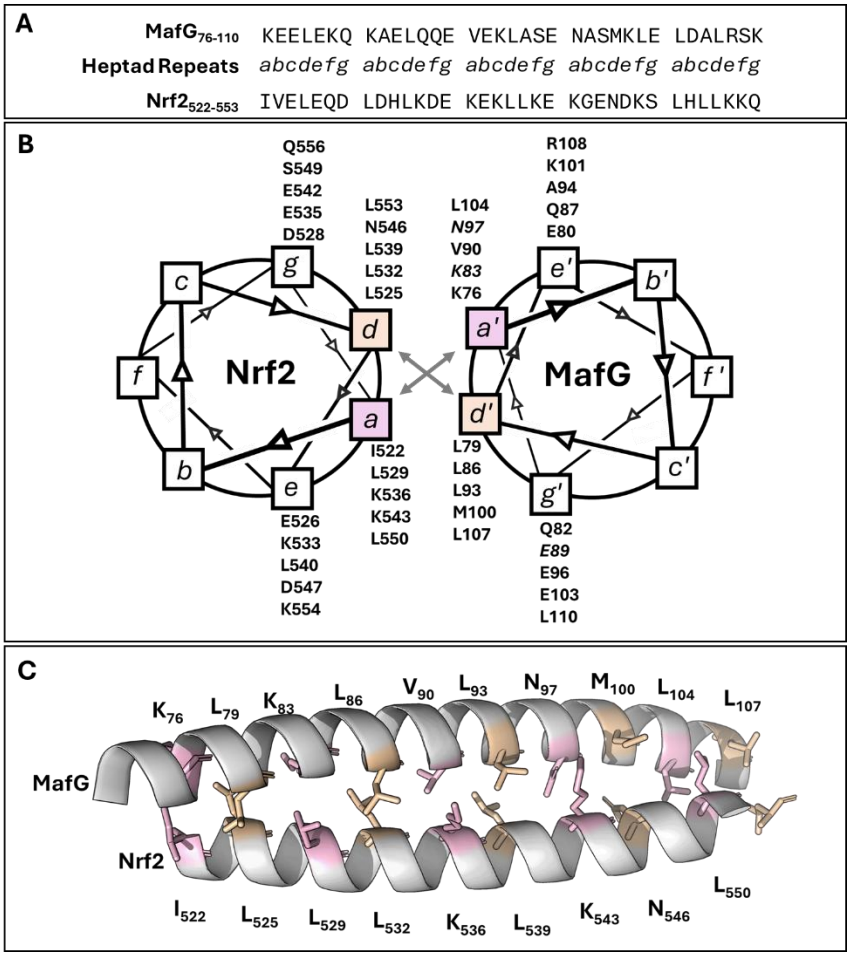


Figure 1-24 (A) *Heptadic repeats of Nrf2/MafG coiled-coil interaction. (**B**) Helical wheel diagram of the Nrf2/MafG leucine zipper domains. (**C**) PDB: 7X5F model of Nrf2/MafG leucine zipper, "a" residue (pink), "d" residue (orange).*

In MafG dimers, the bZIP domain consists of residues 46-123, residues 76-107 forming a homodimeric coiled-coil through the leucine zipper (**Figure 1-25, C**). The interaction is stablised by Leu_{79}, Leu_{86} and Val_{90} occupying *d* and interhelical electrostatic interactions such as Lys_{83} and Gln_{82} in *a* and *g'* positions respectively. Lys_{76} in *a* of the CC heptad motif and Gln_{76} of *g'* mediated by a water molecule. Asn_{97} unexpectedly encountered at an *a* position creates intermolecular hydrogen bonding, a phenomenon unique to sMaf proteins that potentially prevents their interactions with large Maf proteins (**Figure 1-25. B**).

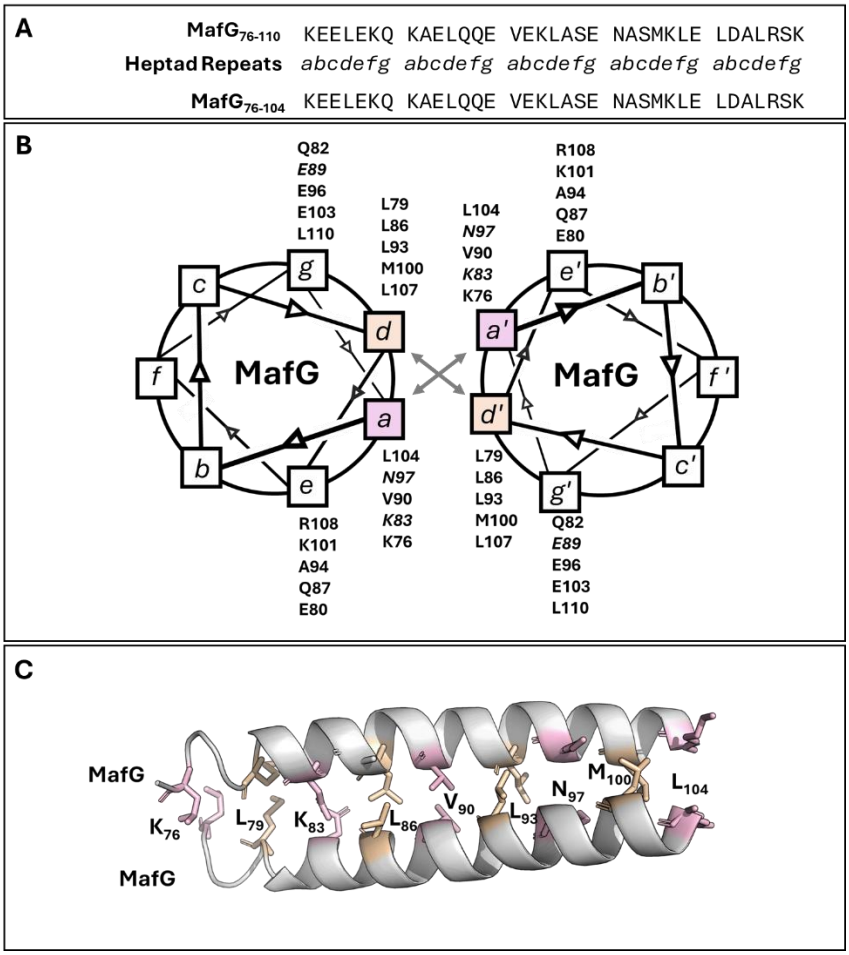


Figure 1-25 (A) *Heptadic repeats of MafG/MafG coiled-coil interaction. (**B**) Helical wheel diagram of the MafG/MafG leucine zipper domains. (**C**) PDB: 3A5T model of MafG/MafG leucine zipper, "a" residues (pink), "d" residues (orange).*

bZIP proteins recognise a core 7 base pair (bp) *5'*-TGACTCA-*3'* DNA sequence identified as the TPA-responsive element (TRE) or the 8 bp *5'*-TGACGTCA-*3'* cAMP-responsive element (CRE). The TRE and CRE consensus sequences are observed within ARE, sMAF recognition is created by GC bases on each side of the TRE/CRE core. *5'*-TGAxxxGC-*3'* sequences are observed in the promoter regions of Nrf2 target genes.[284] The leucine zipper interaction of Nrf2/MafG facilitates stabilisation of the major groove interaction of the Nrf2/MafG DNA-binding domains (**Figure 1-26**). ARE sequences are led by key basic interactions of Arg_{57}, Asn_{61} and Tyr_{64} on the MafG basic region that sits centrally for DNA recognition, specific to Maf proteins. For NF-E2 gene recognition, the Nrf2 basic region requires Arg_{515}, Asn_{507} and Ala_{510} residues for binding.

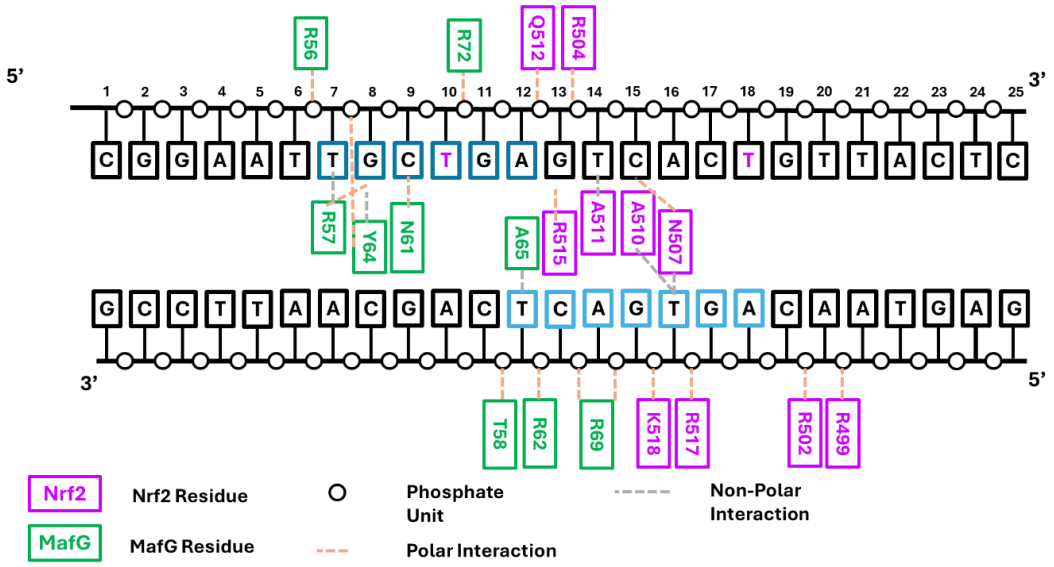


Figure 1-26 *DNA binding sites of basic region of Nrf2/MafG to DNA containing ARE sequences adapted from Sengoku et al.* [285]

ARE promoter sequences are found before genes transcribing proteins involved in inflammation and cell regulation. The NFE2L2 gene, transcribing Nrf2, also contains an ARE promoter, allowing Nrf2 to positively regulate its own transcription.[291]

1.5.2 Genes transcribed by Nrf2

Nrf2 is responsible for antioxidant production, NAD(P)H regeneration and the production of drug-metabolising enzymes.[292,293] As a key regulator of the cellular stress response, ARE-mediated gene transcription was first identified for NAD(P)H quinone dehydrogenase (NQO1) and glutathione S-transferase (GST).[294,295] Outside of inflammatory encoded gene transcription, Nrf2 also supports metabolic processes, cell autophagy, proteostasis and mitochondrial biogenesis.[296–298] This complex network that protects cellular function, summarised in **Table 1-2,** can be manipulated by cancer cells to promote cell survival functions and create resistance to toxic chemotherapeutic agents.

Table 1-2: *Downstream cytoprotective proteins transcribed as a result of Nrf2/DNA binding. Categorised by role in cancer cell survival and chemoresistance development.*

Gene Name	Function	Ref.
Drug metabolism		
NAD(P)H dehydrogenase quinone 1 (NQO1)	Detoxification enzyme reduces free radicals in cells.	262,295
Glutathione S Transferase (GST)	Cellular detoxification through GSH conjugation to electrophilic compounds.	299,300
Cytochrome P450 2A5 (CYP2A5)	Drug metabolism of drugs/toxins e.g. nicotine.	301,302
Antioxidant production		
Heme oxygenase-1 (HO-1)	Enzyme involved in heme degradation to produce antioxidants.	303,304
Thioredoxin 1 (TXN1)	Redox regulating protein involved in antioxidant production through reactive cysteines supporting thiol disulfide exchange.	305,306
Glutathione Peroxidase (GPX)	Antioxidant enzyme, uses reduced glutathione to neutralise reactive oxygen species such as H_2O_2.	307
Hypoxia-inducible factor 1-α (HIF-1-α)	Regulates cellular hypoxia.	308
Drug Efflux		
Multidrug resistance protein 1 (MRP1)	ATP-binding cassette membrane transporter for drug efflux.	309
P-glycoprotein (Pgp)	Cellular efflux of foreign molecules and xenobiotics.	310,311
Cell Survival Signalling		
Neurogenic locus notch homolog protein 1 (NOTCH1)	Initiates signalling cascade for cell survival and proliferation.	312,313
Insulin-like growth factor (IGF1)	Growth hormone for modulation of glucose transport.	314,315
Vascular endothelial growth factor C (VEGFC)	Stimulates blood vessel growth and vessel permeability.	316
B-cell lymphoma 2 (BCL-2)	Inhibition of apoptosis.	317,318
B-cell lymphoma-extra large (BCL-XL)	Binds to Bax in mitochondrial outer membrane for apoptosis inhibition.	319
Ras homolog family member A (RhoA)	Small GTPase involved in signalling for growth and cell migration.	320
Rho-associated coiled-coil kinase (ROCK)	Signalling for cellular migration and adhesions.	320
Cyclooxygenase-2 (COX-2)	Catalyses conversion of arachidonic acid to prostaglandins, for maintaining cancer stemness.	321

1.5.3 Nrf2 in disease

1.5.3.1 Nrf2 in inflammation

High levels of reactive oxygen species (ROS) activate Nrf2, increasing the transcription factors presence in nuclear and cytosolic contents in cells.[322] Nrf2 expression has been shown to protect hepatocytes against cellular damage, related to viral infections and liver disease, in connection to increased production of HO-1, GSH, CYP and COX-2 enzymes.[323–326]

Nrf2 gene transcription supports anti-inflammatory functions in cells. Expression of chemokines mediating inflammatory cells such as neutrophils and monocytes can be controlled by Nrf2.[327] In addition, Nrf2 derived cytokine production supports cell adhesion molecule production for barrier protection in endothelial cells. The overproduction of matrix metalloproteinases can also be inhibited by Nrf2, through expression of HO-1, positively influencing recovery of inflammatory bowel disease.[328,329]

Oxidative stress also plays a significant role in the pathogenesis of lung diseases.[330–332] In models of acute lung injury, Nrf2 increased levels of HO-1 and glutamate cysteine ligase (GCLM) which in turn protects alveolar epithelium from ferroptosis and oxidative stress.[333,334] Additionally, Nrf2 transcribes the aldehyde oxidase 1 enzyme, promoting epithelial barrier maintenance.[335]

Moreover, oxidative stress has been connected with the development of many neurodegenerative diseases.[336–339] Ferroptosis of the brain plays a key role in Alzheimer's disease pathogenesis, a process known to be regulated by Nrf2 transcription.[340]

Nrf2 is a central mediator of chronic kidney disease and diabetes, where HO-1 and NQO1 are downregulated in models of diabetes.[341,342] Investigating the progression of kidney disease found upregulation of inflammatory markers in Nrf2 knock-out models.[343,344] However, Nrf2 overexpression can also lead to proteinuria causing kidney injury, highlighting a need for further research into the control of Nrf2 in disease progression, which requires reliable probes to induce and inhibit Nrf2 expression.[345,346]

1.5.3.2 Nrf2 in cancer

Nrf2 can prevent tumorigenesis through a reduction in DNA damage-inducing ROS.[347,348] Mouse models with Nrf2 deficiency have an increased risk of skin cancer, and tumour development in the liver, gastrointestinal system and urinary tract.[349–352] However, Nrf2 is increasingly reported for its prominent role in the development of resistance to chemotherapy in cancer.[278,353,354] Nrf2 can become persistently activated through somatic mutations that create loss of function variants of Keap1, or gain of function Nrf2 mutations.

Overexpression of Nrf2 can become so pronounced that the term "Nrf2 addiction" has been used to describe this presentation feature in cancer physiology.[355] Nrf2-overexpression is linked with poor survival rates in lung and pancreatic cancer, known for high level of resistance to therapeutic interventions.[356,357] Additionally, Nrf2-driven

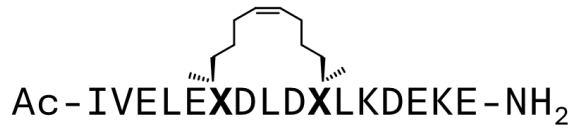
gene transcription promotes glutathione (GSH) production. This creates a favourable environment for cell proliferation, in which detoxification processes and active metabolic enzymes support fast growth of tumour cells.[358,359] Consequently, there is a need for Nrf2 inhibitors that can work in combination with current chemotherapy to increase the sensitivity of cancer cells to cytotoxic agents that Nrf2 works against.

1.5.4 Nrf2 Inhibitors

Cellular activity of Nrf2 can be modulated to varying degrees by natural products, small molecule and peptide inhibitors and activators. Over the last decade, there have been numerous peptides and small molecule reported for activation of Nrf2. Peptides are typically derived from Nrf2 domains binding to Keap1 or β-TrCP.[360–366] Small molecules have also been found to sequester Keap1 to block Nrf2 degradation.[228,367,368] Disruption of Nrf2 PPIs for inhibition of transcriptional activity is a more recent drug discovery strategy, with comparatively fewer compounds reported to date.

1.5.4.1 MafG interaction inhibitors

Inhibition of the Nrf2/MafG PPI prevents Nrf2-DNA binding, inhibiting Nrf2 transcription. In 2023, Modi et al. reported the first peptide inhibitor of the Nrf2/MafG interaction, designing a peptide that selectively binds to MafG. The 16-mer cyclic peptide (**1.45**) is a mimetic of the Nrf2 leucine zipper. The peptide features a hydrocarbon staple *i*, *i+4*, creating a constrained mimic of the native α-helix. Surface plasmon resonance found a binding affinity of 337 nM for MafG protein (**Figure 1-27**).[93] Co-treatment of the peptide with cisplatin found Nrf2 inhibition sensitised NSCLC cells by cell viability assay.[93]



(**1.45**)
K_D : 337 nM

Figure 1-27 *Sequence of Nrf2 derived stapled peptide with binding affinity for MafG.*

1.5.4.2 DNA interaction inhibitors

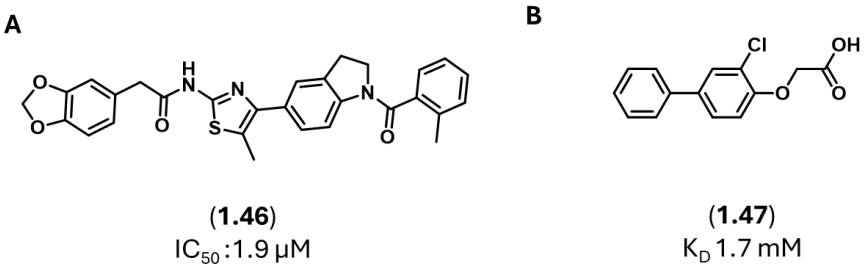
Inhibition of Nrf2/DNA binding prevents Nrf2 derived gene transcription. Small molecules (**Figure 1-28**) and peptides (**Figure 1-29**) have been identified to prevent ternary complex formation of Nrf2/MafG/ARE.

HTS of a small molecule library using a Nrf2 ARE luciferase reporter gene assay has identified inhibitors of Nrf2/ARE binding, leading to the identification of a thiazole-indoline compound with an ortho-benzoyl substitution referred to as ML385 (**1.46**).[369]

The small molecule was investigated for direct inhibition of Nrf2/MafG/ARE ternary complex using a fluorescence polarisation assay determining an IC_{50} of 1.9 μM.[369] ML385 affinity to the Neh1 domain of Nrf2 was found by biotin pull down assay against truncated forms of Nrf2, identifying the first in class Nrf2 Neh1 inhibitor.[277] Noticeably absent is an assessment of whether ML385 blocks Nrf2 activity through the protein-protein interaction or protein DNA interaction in the ternary complex formed. The study has also been questioned for broad transcription factor inhibition and the compound's mode of action.[370,371]

Despite this uncertainty, ML385 has since been assessed in models of cancer to confirm Nrf2 activated gene transcription is reduced. In models of osteosarcoma, ML385 treatment increased ROS expression.[372] Inhibition of Nrf2 to reduce chemoresistance development in cancer has been studied extensively with ML385. Models of squamous cell carcinoma observed that ML385 increased the efficacy of cisplatin.[373,374] Combination therapy of celastrol with ML385 in lung cancer has reduced cell survival.[375] Evidence has been provided that ML385 can rescue radiation treated oesophageal squamous cancers as a radio-sensitising agent.[376] Synergy has been calculated between ML385 and venetoclax for the selective treatment of acute myeloid leukaemia.[377] Despite extensive literature of ML385 as a research probe on the therapeutic benefits of Nrf2 inhibition, the compound has not entered clinical trials.

An NMR study of the Nrf2/DNA interaction has identified hit small molecule fragment with affinity for the basic DNA-binding domain of Nrf2 (**Figure 1-28, 1.47**).[370] Screening of a virtual compound library revealed hits that were experimentally investigated by NMR against Nrf2 residues 445-523, comprising the DNA-binding domain of Neh1. The study identified a binding preference for a biphenyl motif featuring a halogen substituent and acetic acid functionality on the same phenyl ring, yet observing a high K_D of 1.7 mM for Nrf2.[370]



(**1.46**)
IC_{50} :1.9 μM

(**1.47**)
K_D 1.7 mM

Figure 1-28: *Small molecules with Nrf2/ARE inhibitory activity.*

Peptides derived from the DNA-binding domain of Nrf2 have demonstrated an ability to inhibit the protein/DNA interaction, competing with the Nrf2 protein for major groove binding of the ARE.[378] A screen of Nrf2 derived sequences presented an 18-mer, *i*,*i+7* stapled peptide (**1.48**) with a K_D of 1.2 μM for ARE DNA. The peptide also showed efficient cellular uptake with evidence of nuclear localisation of the fluorescently tagged peptide. Unfortunately the peptides were not specific to ARE DNA, with similar affinity demonstrated for non-selective sequences of DNA.[378]

Achieving selective DNA binding to Nrf2/MafG ARE sites has been explored through triazole linked sequences from the Nrf2 and MafG DNA-binding domain generating a mini-protein (**1.49**) capable of potently competing with Nrf2/MafG/ARE binding with an IC_{50} of 90 nM.[379]

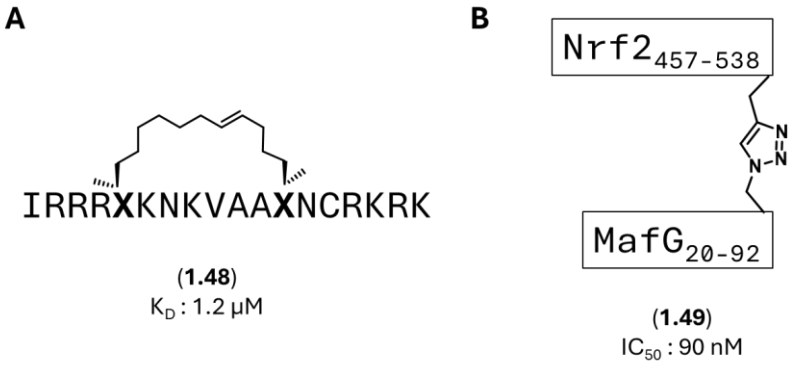


Figure 1-29 *Peptides that bind to ARE DNA sequences designed for the disruption of Nrf2/ARE binding.* ***(A)*** *Nrf2 derived stapled peptide (**1.48**) (**B**) Nrf2/MafG mini-protein (**1.49**).*

Most inhibitors of the Nrf2-DNA interaction function by competing for the ARE recognition sequence. An alternative strategy has reported the use of a DNA-linked proteolysis-targeting chimera (PROTAC) binding to the Nrf2 DNA-binding domain, promoting target degradation. Designed ARE-PROTACs achieved potent Nrf2 degradation (DC_{50} of 1.85 nM) in A549 cells, known to over-express Nrf2.[380] However, ARE sequences are regulatory domains utilised by multiple bZIP transcription factors, potentially leading to off-target protein degradation.[290,381]

1.6 Thesis Outline

This thesis investigates the discovery of novel peptide and small molecule inhibitors of the Nrf2/MafG interaction as a novel anti-cancer therapeutic, demonstrated in **Figure 1-30**. **Chapter 2** describes the synthesis and characterisation of peptide sequences mimicking the leucine zipper of MafG. From this, a linear peptide sequence has been

identified that can inhibit the formation of the Nrf2/MafG/ARE ternary complex. **Chapter 3** investigates the use of our lead peptide sequence as a scaffold for peptide-directed ligand design to identify peptide-small molecule hybrids, synthesised with Cu(I)-catalysed azide-alkyne cycloaddition (CuAAC). **Chapter 4** explores lead compounds in cellular assays to investigate inhibitor efficacy in a model of Nrf2-derived resistance in cancer. **Chapter 5** discusses the recombinant protein expression and purification of Nrf2 and MafG for biophysical assays used in **Chapters 2** and **3**.

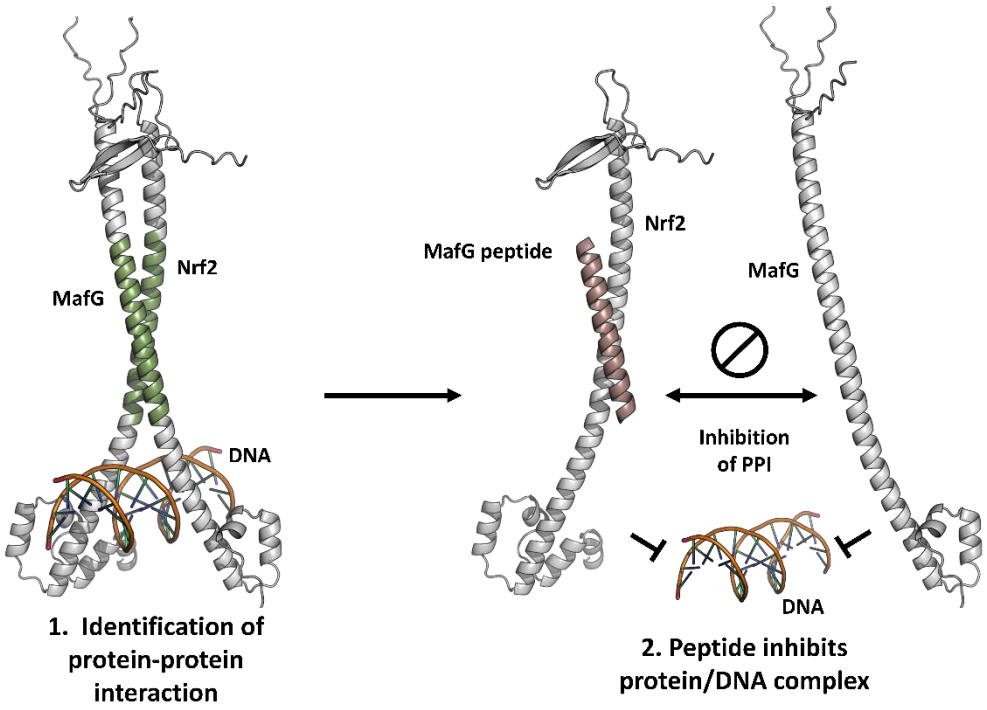


Figure 1-30 *Nrf2/MafG/NQO1 ARE homology model. In green highlights the leucine zipper coiled-coil interaction of Nrf2/MafG, from which MafG mimetic peptides are derived from (pink) to inhibit the ternary complex formation.*

1.6.1 Study aims

1. Synthesise and characterise peptides derived from the MafG leucine zipper using solid-phase peptide synthesis.
2. Evaluate MafG mimetic peptides in biophysical studies to identify inhibition of the Nrf2/MafG/ARE interaction.
3. Investigate the use of lead MafG mimetic peptide as a scaffold for peptide-directed ligand design.
4. Investigate Nrf2/MafG PPI inhibitors in cellular models of Nrf2 over-expression in cancer.
5. Report on method optimisations for the purification of recombinant Nrf2 and MafG by *E.coli* expression systems.

1.7 References

1 B. J. Mayer, *Methods Mol Biol*, 2006, **332**, 79–99.
2 A. Grigoriev, *Nucleic Acids Res*, 2001, **29**, 3513–3519.
3 P. Durek and D. Walther, *BMC Syst Biol*, 2008, **2**, 100.
4 C. H. M. Rodrigues, D. E. V. Pires, T. L. Blundell and D. B. Ascher, *Brief Bioinform*, 2022, **23**, bbac165.
5 S. Das and S. Chakrabarti, *Sci Rep*, 2021, **11**, 1761.
6 K. Dhusia and Y. Wu, *BMC Bioinformatics*, 2021, **22**, 408.
7 S. Akbarzadeh, Ö. Coşkun and B. Günçer, *J Struct Biol*, 2024, **216**, 108118.
8 H. Lu, Q. Zhou, J. He, Z. Jiang, C. Peng, R. Tong and J. Shi, *Sig Transduct Target Ther*, 2020, **5**, 1–23.
9 S. S. Raj and S. S. V. Chandra, *Protein J*, 2024, **43**, 72–83.
10 V. Neduva and R. B. Russell, *Curr Opin Biotechnol*, 2006, **17**, 465–471.
11 E. Petsalaki and R. B. Russell, *Curr Opin Biotechnol*, 2008, **19**, 344–350.
12 T. B. Thompson, P. Chaggar, E. Kuhl, A. Goriely, and Alzheimer's Disease Neuroimaging Initiative, *PLoS Comput Biol*, 2020, **16**, e1008267.
13 M. Sharifi Tabar, H. Francis, D. Yeo, C. G. Bailey and J. E. J. Rasko, *Int J Cancer*, 2022, **151**, 7–19.
14 K. M. Poluri, K. Gulati, D. K. Tripathi and N. Nagar, in *Protein-Protein Interactions: Pathophysiological and Therapeutic Aspects: Volume II*, eds K. M. Poluri, K. Gulati, D. K. Tripathi and N. Nagar, Springer Nature, Singapore, 2023, pp. 171–206.
15 C. Ruiz, M. Zitnik and J. Leskovec, *Nat Commun*, 2021, **12**, 1796.
16 K. Luck, D.-K. Kim, L. Lambourne, K. Spirohn, B. E. Begg, W. Bian, R. Brignall, T. Cafarelli, F. J. Campos-Laborie, B. Charloteaux, D. Choi, A. G. Coté, M. Daley, S. Deimling, A. Desbuleux, A. Dricot, M. Gebbia, M. F. Hardy, N. Kishore, J. J. Knapp, I. A. Kovács, I. Lemmens, M. W. Mee, J. C. Mellor, C. Pollis, C. Pons, A. D. Richardson, S. Schlabach, B. Teeking, A. Yadav, M. Babor, D. Balcha, O. Basha, C. Bowman-Colin, S.-F. Chin, S. G. Choi, C. Colabella, G. Coppin, C. D'Amata, D. D. Ridder, S. D. Rouck, M. Duran-Frigola, H. Ennajdaoui, F. Goebels, L. Goehring, A. Gopal, G. Haddad, E. Hatchi, M. Helmy, Y. Jacob, Y. Kassa, S. Landini, R. Li, N. van Lieshout, A. MacWilliams, D. Markey, J. N. Paulson, S. Rangarajan, J. Rasla, A. Rayhan, T. Rolland, A. San-Miguel, Y. Shen, D. Sheykhkarimli, G. M. Sheynkman, E. Simonovsky, M. Taşan, A. Tejeda, V. Tropepe, J.-C. Twizere, Y. Wang, R. J. Weatheritt, J. Weile, Y. Xia, X. Yang, E. Yeger-Lotem, Q. Zhong, P. Aloy, G. D. Bader, J. D. L. Rivas, S. Gaudet, T. Hao, J. Rak, J. Tavernier, D. E. Hill, M. Vidal, F. P. Roth and M. A. Calderwood, *Nature*, 2020, **580**, 402–408.
17 X. Ran and J. E. Gestwicki, *Curr Opin Chem Biol*, 2018, **44**, 75–86.
18 J. Choi, J. S. Yun, H. Song, N. H. Kim, H. S. Kim and J. I. Yook, *Sci Rep*, 2021, **11**, 13369.
19 A. D. Thompson, A. Dugan, J. E. Gestwicki and A. K. Mapp, *ACS Chem Biol*, 2012, **7**, 1311–1320.
20 A. E. Modell, S. L. Blosser and P. S. Arora, *Trends Pharmacol Sci*, 2016, **37**, 702–713.
21 Neklesa TK, Winkler JD, Crews CM., *Pharmacol Ther*, **174**, 138–144.
22 S. Pathmanathan, I. Grozavu, A. Lyakisheva and I. Stagljar, *Curr Opin Chem Biol*, 2022, **66**, 102079.
23 M. M. R. Arkin and J. A. Wells, *Nat Rev, Discov*, 2004, **3**, 301–317.
24 A. P. Higueruelo, H. Jubb and T. L. Blundell, *Database (Oxford)*, 2013, **2013**, bat039.
25 C. M. Labbé, G. Laconde, M. A. Kuenemann, B. O. Villoutreix and O. Sperandio, *Drug Discov Today*, 2013, **18**, 958–968.
26 K. S. Thorn and A. A. Bogan, *Bioinformatics*, 2001, **17**, 284–285.

27 S. Celis, F. Hobor, T. James, G. J. Bartlett, A. A. Ibarra, D. K. Shoemark, Z. Hegedüs, K. Hetherington, D. N. Woolfson, R. B. Sessions, T. A. Edwards, D. M. Andrews, A. Nelson and A. J. Wilson, *Chem Sci*, **12**, 4753–4762.
28 A. L. Jochim and P. S. Arora, *Mol BioSyst*, 2009, **5**, 924–926.
29 B. N. Bullock, A. L. Jochim and P. S. Arora, J *Am Chem Soc*, 2011, **133**, 14220–14223.
30 M. C. Deller, L. Kong and B. Rupp, *Acta Crystallogr F Struct Biol Commun*, 2016, **72**, 72–95.
31 W. N. Price, Y. Chen, S. K. Handelman, H. Neely, P. Manor, R. Karlin, R. Nair, J. Liu, M. Baran, J. Everett, S. N. Tong, F. Forouhar, S. S. Swaminathan, T. Acton, R. Xiao, J. R. Luft, A. Lauricella, G. T. DeTitta, B. Rost, G. T. Montelione and J. F. Hunt, *Nat Biotechnol*, 2009, **27**, 51–57.
32 A. M. Watkins and P. S. Arora, *ACS Chem Biol*, 2014, **9**, 1747–1754.
33 B. V. Reddy, H. A. Nagarajaram and T. L. Blundell, *Protein Sci*, 1999, **8**, 573–586.
34 V. Perovic, N. Sumonja, L. A. Marsh, S. Radovanovic, M. Vukicevic, S. G. E. Roberts and N. Veljkovic, *Sci Rep*, 2018, **8**, 10563.
35 G. Kibar and M. Vingron, *Proteins:Struct Funct Bioinf,* 2023, **91**, 980–990.
36 Z. Zhang, Z. Lin, Z. Zhou, H. C. Shen, S. F. Yan, A. V. Mayweg, Z. Xu, N. Qin, J. C. Wong, Z. Zhang, Y. Rong, D. C. Fry and T. Hu, *ACS Med Chem Lett*, 2014, **5**, 993–998.
37 D. Yasuda, M. Nakajima, A. Yuasa, R. Obata, K. Takahashi, T. Ohe, Y. Ichimura, M. Komatsu, M. Yamamoto, R. Imamura, H. Kojima, T. Okabe, T. Nagano and T. Mashino, *Bioorg Med Chem Lett*, 2016, **26**, 5956–5959.
38 A. Mansilla, A. Chaves-Sanjuan, N. E. Campillo, O. Semelidou, L. Martínez-González, L. Infantes, J. M. González-Rubio, C. Gil, S. Conde, E. M. C. Skoulakis, A. Ferrús, A. Martínez and M. J. Sánchez-Barrena, *Proc Natl Acad Sci U S A*, 2017, **114**, E999–E1008.
39 N. Qvit, A. U. Joshi, A. D. Cunningham, J. C. B. Ferreira and D. Mochly-Rosen, *J Biol Chem*, 2016, **291**, 13608–13621.
40 K. Wang, X. M. Yin, D. T. Chao, C. L. Milliman and S. J. Korsmeyer, *Genes Dev*, 1996, **10**, 2859–2869.
41 M. Taniguchi, K.-I. Suzumura, K. Nagai, T. Kawasaki, J. Takasaki, M. Sekiguchi, Y. Moritani, T. Saito, K. Hayashi, S. Fujita, S.-I. Tsukamoto and K.-I. Suzuki, *Bioorg Med Chem*, 2004, **12**, 3125–3133.
42 S. A. Kawamoto, A. Coleska, X. Ran, H. Yi, C.-Y. Yang and S. Wang, *J Med Chem*, 2012, **55**, 1137–1146.
43 M. de la Roche, T. J. Rutherford, D. Gupta, D. B. Veprintsev, B. Saxty, S. M. Freund and M. Bienz, *Nat Commun*, 2012, **3**, 680.
44 H. Zhou, J. Lu, L. Liu, D. Bernard, C.-Y. Yang, E. Fernandez-Salas, K. Chinnaswamy, S. Layton, J. Stuckey, Q. Yu, W. Zhou, Z. Pan, Y. Sun and S. Wang, *Nat Commun*, 2017, **8**, 1150.
45 B. Wang, H. Yang, Y.-C. Liu, T. Jelinek, L. Zhang, E. Ruoslahti and H. Fu, *Biochemistry*, 1999, **38**, 12499–12504.
46 Z. Gao, Y. Tian, J. Wang, Q. Yin, H. Wu, Y.-M. Li and X. Jiang, *J Biol Chem*, 2007, **282**, 30718–30727.
47 D. Sela, J. J. Conkright, L. Chen, J. Gilmore, M. P. Washburn, L. Florens, R. C. Conaway and J. W. Conaway, *J Biol Chem*, 2013, **288**, 26179–26187.
48 H. Jiang, R. Deng, X. Yang, J. Shang, S. Lu, Y. Zhao, K. Song, X. Liu, Q. Zhang, Y. Chen, Y. E. Chinn, G. Wu, J. Li, G. Chen, J. Yu and J. Zhang, *Nat Chem Biol*, 2017, **13**, 994–1001.
49 C. Garcia-Echeverria, P. Chene, M. J. J. Blommers and P. Furet, *J Med Chem*, 2000, **43**, 3205–3208.

50 M. L. Stewart, E. Fire, A. E. Keating and L. D. Walensky, *Nat Chem Biol*, 2010, **6**, 595–601.
51 S. Kushal, B. B. Lao, L. K. Henchey, R. Dubey, H. Mesallati, N. J. Traaseth, B. Z. Olenyuk and P. S. Arora, *Proc Natl Acad Sci USA*, 2013, **110**, 15602–15607.
52 S. Jeganathan, M. Wendt, S. Kiehstaller, D. Brancaccio, A. Kuepper, N. Pospiech, A. Carotenuto, E. Novellino, S. Hennig and T. N. Grossmann, *Angew Chem Int Ed Engl*, 2019, **58**, 17351–17358.
53 J. V. Frangioni, L. M. LaRiccia, L. C. Cantley and M. R. Montminy, *Nat Biotechnol*, 2000, **18**, 1080–1085.
54 I. M. Brooks and S. J. Tavalin, *J Biol Chem*, 2011, **286**, 6697–6706.
55 W. Kim, G. H. Bird, T. Neff, G. Guo, M. A. Kerenyi, L. D. Walensky and S. H. Orkin, *Nat Chem Biol*, 2013, **9**, 643–650.
56 P. 't Hart, P. Hommen, A. Noisier, A. Krzyzanowski, D. Schüler, A. T. Porfetye, M. Akbarzadeh, I. R. Vetter, H. Adihou and H. Waldmann, *Angew Chem Int Ed Engl*, 2021, **60**, 1813–1820.
57 B. He, J. T. Minges, L. W. Lee and E. M. Wilson, *J Biol Chem*, 2002, **277**, 10226–10235.
58 S. V. Maifeld, A. L. MacKinnon, J. L. Garrison, A. Sharma, E. J. Kunkel, R. S. Hegde and J. Taunton, *Chem Biol*, 2011, **18**, 1082–1088.
59 D. Peter, C. Igreja, R. Weber, L. Wohlbold, C. Weiler, L. Ebertsch, O. Weichenrieder and E. Izaurralde, *Molecular Cell*, 2015, **57**, 1074–1087.
60 J. Han, M. Yu, Y. Bai, J. Yu, F. Jin, C. Li, R. Zeng, J. Peng, A. Li, X. Song, H. Li, D. Wu and L. Li, *Cancer Cell*, 2020, **38**, 844-856.e7.
61 Y. Zheng, B.-S. Li, N. D. Amin, W. Albers and H. C. Pant, *Eur J Biochem*, 2002, **269**, 4427–4434.
62 J. Chatterjee, M. Beullens, R. Sukackaite, J. Qian, B. Lesage, D. J. Hart, M. Bollen and M. Köhn, *Angew Chem Int Ed Engl*, 2012, **51**, 10054–10059.
63 Y.-J. Kwon, B. A. Leibovitch, N. Bansal, L. Pereira, C.-Y. Chung, E. V. Ariztia, A. Zelent, E. F. Farias and S. Waxman, *Oncotarget*, 2016, **8**, 88421–88436.
64 Q. Li, C. Zeng, H. Liu, K. W. Y. Yung, C. Chen, Q. Xie, Y. Zhang, S. W. C. Wan, B. S. W. Mak, J. Xia, S. Xiong and J. C. K. Ngo, *iScience*, 2021, **24**, 102423.
65 Y. Gao, J. B. Dickerson, F. Guo, J. Zheng and Y. Zheng, *Proc Natl Acad Sci U S A*, 2004, **101**, 7618–7623.
66 S. Ito, K. Ogawa, K. Takeuchi, M. Takagi, M. Yoshida, T. Hirokawa, S. Hirayama, K. Shin-ya, I. Shimada, T. Doi, N. Goshima, T. Natsume and K. Nagata, *J Biol Chem*, 2017, **292**, 20076–20085.
67 T. Song, Y. Guo, Z. Xue, Z. Guo, Z. Wang, D. Lin, H. Zhang, H. Pan, X. Zhang, F. Yin, H. Wang, L. B. Uwituze and Z. Zhang, *Leukemia*, 2021, **35**, 2862–2874.
68 P. Li, Y. Tian, Q. Shang, C. Tang, Z. Hou, Y. Li, L. Cao, S. Xue, J. Bian, C. Luo, D. Wu, Z. Li and H. Ding, *Bioorg Chem*, 2023, **139**, 106676.
69 K. Huang and L. T. Vassilev, in *Methods in Enzymology*, Academic Press, 2005, vol. 399, pp. 717–728.
70 A. Beautrait, J. S. Paradis, B. Zimmerman, J. Giubilaro, L. Nikolajev, S. Armando, H. Kobayashi, L. Yamani, Y. Namkung, F. M. Heydenreich, E. Khoury, M. Audet, P. P. Roux, D. B. Veprintsev, S. A. Laporte and M. Bouvier, *Nat Commun*, 2017, **8**, 15054.
71 D. C. McKinney, B. J. McMillan, M. J. Ranaghan, J. A. Moroco, M. Brousseau, Z. Mullin-Bernstein, M. O'Keefe, P. McCarren, M. F. Mesleh, K. M. Mulvaney, F. Robinson, R. Singh, B. Bajrami, F. F. Wagner, R. Hilgraf, M. J. Drysdale, A. J. Campbell, A. Skepner, D. E. Timm, D. Porter, V. K. Kaushik, W. R. Sellers and A. Ianari, *J Med Chem*, 2021, **64**, 11148–11168.
72 T. Furuya, J. Lin, A. Afanaseva, L. Molz, B. Lagu and B. Ma, *ACS Med Chem Lett*, 2023, **14**, 1095–1099.

73 E. Robinson, E. Leung, A. M. Matuszek, N. Krogsgaard-Larsen, D. P. Furkert, M. A. Brimble, A. Richardson and J. Reynisson, *Med Chem Commun*, 2015, **6**, 239–246.
74 A. K. Farag, A. H. E. Hassan, H. Jeong, Y. Kwon, J. G. Choi, M. S. Oh, K. D. Park, Y. K. Kim and E. J. Roh, *Eur J Med Chem*, 2019, **162**, 161–175.
75 S.-S. Cheng, Y.-Q. Qu, J. Wu, G.-J. Yang, H. Liu, W. Wang, Q. Huang, F. Chen, G. Li, C.-Y. Wong, V. K. W. Wong, D.-L. Ma and C.-H. Leung, *Acta Pharmaceutica Sinica B*, 2022, **12**, 1390–1405.
76 J. Randjelovic, S. Eric and V. Savic, *J Mol Graph Model*, 2014, **50**, 100–112.
77 J. Randjelović, S. Erić and V. Savić, *J Mol Model*, 2013, **19**, 1711–1725.
78 S. Taniguchi, Y. Ono, Y. Doi, S. Taniguchi, Y. Matsuura, A. Iwasaki, N. Hirata, R. Fukuda, K. Inoue, M. Yamaguchi, A. Tashiro, D. Egami, S. Aoki, Y. Kondoh, K. Honda, H. Osada, H. Kumeta, T. Saio and T. Okiyoneda, *Biochem Pharmacol*, 2023, **215**, 115730.
79 E. M. McNeil, K. R. Astell, A.-M. Ritchie, S. Shave, D. R. Houston, P. Bakrania, H. M. Jones, P. Khurana, C. Wallace, T. Chapman, M. A. Wear, M. D. Walkinshaw, B. Saxty and D. W. Melton, *DNA Repair (Amst)*, 2015, **31**, 19–28.
80 L. Wang, L. Zhang, L. Li, J. Jiang, Z. Zheng, J. Shang, C. Wang, W. Chen, Q. Bao, X. Xu, Z. Jiang, J. Zhang and Q. You, *Sci Adv*, 2019, **5**, eaax2277.
81 S. Sayedyahossein, J. Smith, E. Barnaeva, Z. Li, J. Choe, M. Ronzetti, C. Dextras, X. Hu, J. Marugan, N. Southall, B. Baljinnyam, L. Thines, A. D. Tran, M. Ferrer and D. B. Sacks, *Sci Rep*, 2022, **12**, 17372.
82 X. Jin, Y. Yang, D. Liu, X. Zhou and Y. Huang, *Cell. Commun. Signal*, 2024, **22**, 565.
83 G. Menchon, O. Bombarde, M. Trivedi, A. Négrel, C. Inard, B. Giudetti, M. Baltas, A. Milon, M. Modesti, G. Czaplicki and P. Calsou, *Sci Rep*, 2016, **6**, 22878.
84 Y. Fang, X. Chen, Q. Tan, H. Zhou, J. Xu and Q. Gu, *ACS Cent Sci*, 2021, **7**, 980–989.
85 J. Gu, R.-K. Peng, C.-L. Guo, M. Zhang, J. Yang, X. Yan, Q. Zhou, H. Li, N. Wang, J. Zhu and Q. Ouyang, *Nat Commun*, 2022, **13**, 7176.
86 J. Nong, S. Shen, F. Hong, F. Xiao, L. Meng, P. Li, X. Lei and Y.-G. Chen, *Mol Biol Cell*, 2024, **35**, ar95.
87 J. M. Mason, M. A. Schmitz, K. M. Müller and K. M. Arndt, *Proc Natl Acad Sci USA*, 2006, **103**, 8989–8994.
88 J. K.-L. Sinclair, E. V. Denton and A. Schepartz, *J Am Chem Soc*, 2014, **136**, 11232–11235.
89 K. M. Hawley, R. J. Eclov, M. R. Schnorenberg, Y. Tian, R. N. Shah, A. T. Thomas-Toth, M. Fefferman, G. H. Bird, L. D. Walensky, M. V. Tirrell and J. L. LaBelle, *Proc Natl Acad Sci U S A*, 2022, **119**, e2209044119.
90 M. O. Okpara, M. C. Vaaltyn, J. L. Watson, M. Alhassan, F. Albericio, B. G. de la Torre, D. J. Clarke, C. G. L. Veale and A. L. Edkins, *ACS Infect Dis*, 2024, **10**, 3853–3867.
91 J. Speck, C. Räuber, T. Kükenshöner, C. Niemöller, K. J. Mueller, P. Schleberger, P. Dondapati, J. Hecky, K. M. Arndt and K. M. Müller, *Protein Eng Des Sel*, 2013, **26**, 225–242.
92 X. Ran, E. T. Burchfiel, B. Dong, N. J. Rettko, B. M. Dunyak, H. Shao, D. J. Thiele and J. E. Gestwicki, *Bioorg Med Chem*, 2018, **26**, 5299–5306.
93 R. Modi, N. McKee, N. Zhang, A. Alwali, S. Nelson, A. Lohar, R. Ostafe, D. D. Zhang and E. I. Parkinson, *J Med Chem*, 2023, **66**, 6184–6192.
94 T. Berg, S. B. Cohen, J. Desharnais, C. Sonderegger, D. J. Maslyar, J. Goldberg, D. L. Boger and P. K. Vogt, *Proc Natl Acad Sci U S A*, 2002, **99**, 3830–3835.
95 H. M. Dickson, A. Wilbur, A. A. Reinke, M. A. Young and A. B. Vojtek, *BMC Neurosci*, 2015, **16**, 34.

96 M. Y. Kim, I. Na, J. S. Kim, S. H. Son, S. Choi, S. E. Lee, J.-H. Kim, K. Jang, G. Alterovitz, Y. Chen, A. van der Vaart, H.-S. Won, V. N. Uversky and C. G. Kim, *Sci Adv*, 2019, **5**, eaav9810.
97 I. Pavlinov, M. Salkovski and L. N. Aldrich, *J. Am. Chem Soc*, 2020, **142**, 8174–8182.
98 S. B. Petkova, S. Akilesh, T. J. Sproule, G. J. Christianson, H. Al Khabbaz, A. C. Brown, L. G. Presta, Y. G. Meng and D. C. Roopenian, *Int Immunol*, 2006, **18**, 1759–1769.
99 J. Hemelaar, A. Borodovsky, B. M. Kessler, D. Reverter, J. Cook, N. Kolli, T. Gan-Erdene, K. D. Wilkinson, G. Gill, C. D. Lima, H. L. Ploegh and H. Ovaa, *Mol Cell Biol*, 2004, **24**, 84–95.
100 M. C. Hammond, B. Z. Harris, W. A. Lim and P. A. Bartlett, *Chem Biol*, 2006, **13**, 1247–1251.
101 M. O. Gray, J. S. Karliner and D. Mochly-Rosen, *J Biol Chem*, 1997, **272**, 30945–30951.
102 A. Lesner, A. Legowska, M. Wysocka and K. Rolka, *Curr. Pharm. Des*, **17**, 4308–4317.
103 Y. Zhang, C. Eigenbrot, L. Zhou, S. Shia, W. Li, C. Quan, J. Tom, P. Moran, P. Di Lello, N. J. Skelton, M. Kong-Beltran, A. Peterson and D. Kirchhofer, *J Biol Chem*, 2014, **289**, 942–955.
104 C. R. O. Bartling, T. M. T. Jensen, S. M. Henry, A. L. Colliander, V. Sereikaite, M. Wenzler, P. Jain, H. M. Maric, K. Harpsøe, S. W. Pedersen, L. S. Clemmensen, L. M. Haugaard-Kedström, D. E. Gloriam, A. Ho and K. Strømgaard, *J Am Chem Soc*, 2021, **143**, 891–901.
105 J. Nomme, A. Renodon-Cornière, Y. Asanomi, K. Sakaguchi, A. Z. Stasiak, A. Stasiak, B. Norden, V. Tran and M. Takahashi, *J Med Chem*, 2010, **53**, 5782–5791.
106 D. Kutscher, A. Pingoud, A. Jeltsch and G. Meiss, *The FEBS Journal*, 2012, **279**, 2917–2928.
107 W. C. Lee, S. Almeida, M. Prudencio, T. R. Caulfield, Y.-J. Zhang, W. M. Tay, P. O. Bauer, J. Chew, H. Sasaguri, K. R. Jansen-West, T. F. Gendron, C. T. Stetler, N. Finch, I. R. Mackenzie, R. Rademakers, F.-B. Gao and L. Petrucelli, *Hum Mol Genet*, 2014, **23**, 1467–1478.
108 T. Sameshima, T. Yamamoto, O. Sano, S. Sogabe, S. Igaki, K. Sakamoto, K. Ida, M. Gotou, Y. Imaeda, J. Sakamoto and I. Miyahisa, *Biochemistry*, 2018, **57**, 1369–1379.
109 S. Chia, Z. Faidon Brotzakis, R. I. Horne, A. Possenti, B. Mannini, R. Cataldi, M. Nowinska, R. Staats, S. Linse, T. P. J. Knowles, J. Habchi and M. Vendruscolo, *Mol Pharm*, 2023, **20**, 183–193.
110 J. Kastl, J. Braun, A. Prestel, H. M. Möller, T. Huhn and T. U. Mayer, *ACS Chem Biol*, 2015, **10**, 1661–1666.
111 M. Garcia, L. Hoffer, R. Leblanc, F. Benmansour, M. Feracci, C. Derviaux, A. L. Egea-Jimenez, P. Roche, P. Zimmermann, X. Morelli and K. Barral, *Eur J Med Chem*, 2021, **223**, 113601.
112 A. J. da Silva, M. Brickelmaier, G. R. Majeau, Z. Li, L. Su, Y.-M. Hsu and P. S. Hochman, *J Immunol*, 2002, **168**, 4462–4471.
113 T. T. Mai, P. G. Nguyen, M.-T. Le, T.-D. Tran, P. N. H. Huynh, D.-T. T. Trinh, Q.-T. Nguyen and K.-M. Thai, *Mol Divers*, 2022, **26**, 2659–2678.
114 J. C. Chavez, F. M. Foss, B. M. William, J. E. Brammer, S. M. Smith, A. Prica, J. M. Zain, J. M. Tuscano, M. Glenn, N. Mehta-Shah, B. X. Wang, S. Zantinge, L. Wang, L. Zhang, A. Boutrin, W. Zhao, L. Cheng, N. Standifer, G. Carlesso, J. Moscow and L. Siu, *Blood*, 2020, **136**, 5–6.
115 H. Kobayashi, J. Gotoh, M. Fujie, H. Shinohara, N. Moniwa and T. Terao, *Int J Cancer*, 1994, **57**, 727–733.
116 H. M. Ta, G. T. T. Nguyen, H. M. Jin, J. Choi, H. Park, N. Kim, H.-Y. Hwang and K. K. Kim, *Proc Natl Acad Sci U S A*, 2010, **107**, 20281–20286.

117 World Intellectual Property Organization, WO2013144704A1, 2013.
118 D. J. Marcotte, J. Hus, C. C. Banos, C. Wildes, R. Arduini, C. Bergeron, C. A. Hession, D. P. Baker, E. Lin, K. M. Guckian, A. W. Dunah and L. F. Silvian, *Protein Sci*, 2018, **27**, 672–680.
119 J. Rahuel, B. Gay, D. Erdmann, A. Strauss, C. Garcia-Echeverría, P. Furet, G. Caravatti, H. Fretz, J. Schoepfer and M. G. Grütter, *Nat Struct Biol*, 1996, **3**, 586–589.
120 World Intellectual Property Organization, WO2014066726A2, 2014.
121 K. Wichapong, J.-E. Alard, A. Ortega-Gomez, C. Weber, T. M. Hackeng, O. Soehnlein and G. A. F. Nicolaes, *J Med Chem*, 2016, **59**, 4289–4301.
122 M. M. He, A. S. Smith, J. D. Oslob, W. M. Flanagan, A. C. Braisted, A. Whitty, M. T. Cancilla, J. Wang, A. A. Lugovskoy, J. C. Yoburn, A. D. Fung, G. Farrington, J. K. Eldredge, E. S. Day, L. A. Cruz, T. G. Cachero, S. K. Miller, J. E. Friedman, I. C. Choong and B. C. Cunningham, *Science*, 2005, **310**, 1022–1025.
123 S. Florio, C. Loh, S. Huang, A. Iwamaye, K. Kitto, K. Fowler, J. Treiberg, J. Hayflick, J. Walker, C. Fairbanks and Y. Lai, *Br J Pharmacol*, 2009, **158**, 494–506.
124 A. Illendula, J. A. Pulikkan, H. Zong, J. Grembecka, L. Xue, S. Sen, Y. Zhou, A. Boulton, A. Kuntimaddi, Y. Gao, R. A. Rajewski, M. L. Guzman, L. H. Castilla and J. H. Bushweller, *J Biol Chem*, 2015, **347**, 779–784.
125 F. Cai, Y. Huang, M. Wang, M. Sun, Y. Zhao and M. Hattori, *Sci Rep*, 2020, **10**, 12879.
126 B. Yakubov, L. Chen, A. M. Belkin, S. Zhang, B. Chelladurai, Z.-Y. Zhang and D. Matei, *PLoS One*, 2014, **9**, e89285.
127 C. E. Quevedo, A. Cruz-Migoni, N. Bery, A. Miller, T. Tanaka, D. Petch, C. J. R. Bataille, L. Y. W. Lee, P. S. Fallon, H. Tulmin, M. T. Ehebauer, N. Fernandez-Fuentes, A. J. Russell, S. B. Carr, S. E. V. Phillips and T. H. Rabbitts, *Nat Commun*, 2018, **9**, 3169.
128 S.-C. Lo, X. Li, M. T. Henzl, L. J. Beamer and M. Hannink, *EMBO J*, 2006, **25**, 3605–3617.
129 A. D. Abell, M. A. Jones, J. M. Coxon, J. D. Morton, S. G. Aitken, S. B. McNabb, H. Y.-Y. Lee, J. M. Mehrtens, N. A. Alexander, B. G. Stuart, A. T. Neffe and R. Bickerstaffe, *Angew Chem Int Ed*, 2009, **48**, 1455–1458.
130 J. Grembecka, A. M. Belcher, T. Hartley and T. Cierpicki, *J Biol Chem*, 2010, **285**, 40690–40698.
131 O. V. Tsodikov, D. Ivanov, B. Orelli, L. Staresincic, I. Shoshani, R. Oberman, O. D. Schärer, G. Wagner and T. Ellenberger, *EMBO J*, 2007, **26**, 4768–4776.
132 H. Karatas, E. C. Townsend, D. Bernard, Y. Dou and S. Wang, *J Med Chem*, 2010, **53**, 5179–5185.
133 T. Saito, Y. Ichimura, K. Taguchi, T. Suzuki, T. Mizushima, K. Takagi, Y. Hirose, M. Nagahashi, T. Iso, T. Fukutomi, M. Ohishi, K. Endo, T. Uemura, Y. Nishito, S. Okuda, M. Obata, T. Kouno, R. Imamura, Y. Tada, R. Obata, D. Yasuda, K. Takahashi, T. Fujimura, J. Pi, M.-S. Lee, T. Ueno, T. Ohe, T. Mashino, T. Wakai, H. Kojima, T. Okabe, T. Nagano, H. Motohashi, S. Waguri, T. Soga, M. Yamamoto, K. Tanaka and M. Komatsu, *Nat Commun*, 2016, **7**, 12030.
134 R. Fernández-Ginés, J. A. Encinar, J. D. Hayes, B. Oliva, M. I. Rodríguez-Franco, A. I. Rojo and A. Cuadrado, *Redox Biology*, 2022, **55**, 102396.
135 A. Poissonnier, J.-P. Guégan, H. T. Nguyen, D. Best, N. Levoin, G. Kozlov, K. Gehring, R. Pineau, F. Jouan, L. Morere, S. Martin, M. Thomas, E. Lazaro, I. Douchet, T. Ducret, P. van de Weghe, P. Blanco, M. Jean, P. Vacher and P. Legembre, *Nat Chem Biol*, 2018, **14**, 1079–1089.
136 A. E. H. Elia, L. C. Cantley and M. B. Yaffe, *Science*, 2003, **299**, 1228–1231.
137 B. H. Liu, C. Jobichen, C. S. B. Chia, T. H. M. Chan, J. P. Tang, T. X. Y. Chung, J. Li, A. Poulsen, A. W. Hung, X. Koh-Stenta, Y. S. Tan, C. S. Verma, H. K. Tan, C.-S. Wu, F.

Li, J. Hill, J. Joy, H. Yang, L. Chai, J. Sivaraman and D. G. Tenen, *Proc Natl Acad Sci U S A*, 2018, **115**, E7119–E7128.
138 V. Sail, A. A. Rizzo, N. Chatterjee, R. C. Dash, Z. Ozen, G. C. Walker, D. M. Korzhnev and M. K. Hadden, *ACS Chem Biol*, 2017, **12**, 1903–1912.
139 S. Nim, D. M. O'Hara, C. Corbi-Verge, A. Perez-Riba, K. Fujisawa, M. Kapadia, H. Chau, F. Albanese, G. Pawar, M. L. De Snoo, S. G. Ngana, J. Kim, O. M. A. El-Agnaf, E. Rennella, L. E. Kay, S. K. Kalia, L. V. Kalia and P. M. Kim, *Nat Commun*, 2023, **14**, 2150.
140 C. Punchihewa, A. Inoue, A. Hishiki, Y. Fujikawa, M. Connelly, B. Evison, Y. Shao, R. Heath, I. Kuraoka, P. Rodrigues, H. Hashimoto, M. Kawanishi, M. Sato, T. Yagi and N. Fujii, *J Biol Chem*, 2012, **287**, 14289–14300.
141 Z. Na, B. Peng, S. Ng, S. Pan, J.-S. Lee, H.-M. Shen and S. Q. Yao, *Angew Chem Int Ed*, 2015, **54**, 2515–2519.
142 S. M. Ø. Solbak, J. Zang, D. Narayanan, L. J. Høj, S. Bucciarelli, C. Softley, S. Meier, A. E. Langkilde, C. H. Gotfredsen, M. Sattler and A. Bach, *J Med Chem*, 2020, **63**, 1156–1177.
143 J. Bennett, O. Fedorov, C. Tallant, O. Monteiro, J. Meier, V. Gamble, P. Savitsky, G. A. Nunez-Alonso, B. Haendler, C. Rogers, P. E. Brennan, S. Müller and S. Knapp, *J Med Chem*, 2016, **59**, 1642–1647.
144 A. Coluccia, G. La Regina, V. Naccarato, M. Nalli, V. Orlando, S. Biagioni, M. L. De Angelis, M. Baiocchi, C. Gautier, S. Gianni, F. Di Pastena, L. Di Magno, G. Canettieri, A. M. L. Coluccia and R. Silvestri, *ACS Med Chem Lett*, 2019, **10**, 499–503.
145 L. Han, J. B. Lee, E. W. Indermaur and A. J. Keung, *Biochem J*, 2023, **480**, 421–432.
146 P. G. K. Clark, L. C. C. Vieira, C. Tallant, O. Fedorov, D. C. Singleton, C. M. Rogers, O. P. Monteiro, J. M. Bennett, R. Baronio, S. Müller, D. L. Daniels, J. Méndez, S. Knapp, P. E. Brennan and D. J. Dixon, *Angew Chem Int Ed*, 2015, **54**, 6217–6221.
147 S. Lim, H. Y. Cho, D. G. Kim, Y. Roh, S.-Y. Son, A. U. Mushtaq, M. Kim, D. Bhattarai, A. Sivaraman, Y. Lee, J. Lee, W. S. Yang, H. K. Kim, M. H. Kim, K. Lee, Y. H. Jeon and S. Kim, *Nat Chem Biol*, 2020, **16**, 31–41.
148 M. Hisada, M. Hiranuma, M. Nakashima, N. Goda, T. Tenno and H. Hiroaki, *Eur J Pharmacol*, 2020, **887**, 173436.
149 S.-Y. Hwang, S. Park, H. Jo, S. Hee Seo, K.-H. Jeon, S. Kim, A.-R. Jung, C. Song, M. Ahn, S. Yeon Kwak, H.-J. Lee, M. Uesugi, Y. Na and Y. Kwon, *J Adv Res*, 2023, **47**, 173–187.
150 M. Raj, B. N. Bullock and P. S. Arora, *Bioorg Med Chem*, 2013, **21**, 4051–4057.
151 B. B. Lao, K. Drew, D. A. Guarracino, T. F. Brewer, D. W. Heindel, R. Bonneau and P. S. Arora, *J Am Chem Soc*, 2014, **136**, 7877–7888.
152 H. I. Merritt, N. Sawyer and P. S. Arora, *Pept Sci (Hoboken)*, 2020, **112**, e24145.
153 P. H. Kussie, S. Gorina, V. Marechal, B. Elenbaas, J. Moreau, A. J. Levine and N. P. Pavletich, *Science*, 1996, **274**, 948–953.
154 S. Baek, P. S. Kutchukian, G. L. Verdine, R. Huber, T. A. Holak, K. W. Lee and G. M. Popowicz, *J Am Chem Soc*, 2012, **134**, 103–106.
155 O. Schon, A. Friedler, M. Bycroft, S. M. V. Freund and A. R. Fersht, *J Mol Biol*, 2002, **323**, 491–501.
156 S. Shangary and D. E. Johnson, *Biochemistry*, 2002, **41**, 9485–9495.
157 E. F. Lee, A. Fedorova, K. Zobel, M. J. Boyle, H. Yang, M. A. Perugini, P. M. Colman, D. C. S. Huang, K. Deshayes and W. D. Fairlie, *J Biol Chem*, 2009, **284**, 31315–31326.
158 S. Dutta, S. Gullá, T. S. Chen, E. Fire, R. A. Grant and A. E. Keating, *J Mol Biol*, 2010, **398**, 747–762.
159 T. Maculins, J. Garcia-Pardo, A. Skenderovic, J. Gebel, M. Putyrski, A. Vorobyov, P. Busse, G. Varga, M. Kuzikov, A. Zaliani, S. Rahighi, V. Schaeffer, M. J. Parnham, S. S.

Sidhu, A. Ernst, V. Dötsch, M. Akutsu and I. Dikic, *Cell Chem Biol*, 2020, **27**, 1441-1451.e7.
160 G. N. Sundell and Y. Ivarsson, *Biomed Res Int*, 2014, **2014**, 176172.
161 E. T. Boder and K. D. Wittrup, *Nat Biotechnol*, 1997, **15**, 553–557.
162 D. P. Hernandez and G. Dittmar, *Anal Bioanal Chem*, 2021, **413**, 5561–5566.
163 D. Iaculli and S. Ballet, *J Pept Sci*, 2025, **31**, e70029.
164 M. Pazgier, M. Liu, G. Zou, W. Yuan, C. Li, C. Li, J. Li, J. Monbo, D. Zella, S. G. Tarasov and W. Lu, *Proc Natl Acad Sci U S A*, 2009, **106**, 4665–4670.
165 C. Li, M. Pazgier, C. Li, W. Yuan, M. Liu, G. Wei, W.-Y. Lu and W. Lu, *J Mol Biol*, 2010, **398**, 200–213.
166 M. Liu, M. Pazgier, C. Li, W. Yuan, C. Li and W. Lu, *Angew Chem Int Ed Engl*, 2010, **49**, 3649–3652.
167 P. Tošovská and P. S. Arora, *Org Lett*, 2010, **12**, 1588–1591.
168 S. Marimganti, M. N. Cheemala and J.-M. Ahn, *Org Lett*, 2009, **11**, 4418–4421.
169 C. Sheng, G. Dong, Z. Miao, W. Zhang and W. Wang, *Chem Soc Rev*, 2015, **44**, 8238–8259.
170 I. Saraogi and A. D. Hamilton, *Biochem Soc Trans*, 2008, **36**, 1414–1417.
171 C. Neufeld, F. V. Filipp, B. Simon, A. Neuhaus, N. Schüller, C. David, H. Kooshapur, T. Madl, R. Erdmann, W. Schliebs, M. Wilmanns and M. Sattler, *EMBO J*, 2009, **28**, 745–754.
172 M. Marciniak, P. Mróz, V. Napolitano, V. C. Kalel, R. Fino, E. Pykacz, W. Schliebs, O. Plettenburg, R. Erdmann, M. Sattler, G. M. Popowicz and M. Dawidowski, *Eur J Med Chem*, 2023, **258**, 115587.
173 A. L. Rodriguez, A. Tamrazi, M. L. Collins and J. A. Katzenellenbogen, *J Med Chem*, 2004, **47**, 600–611.
174 C. H. Douse, J. L. Green, P. S. Salgado, P. J. Simpson, J. C. Thomas, G. Langsley, A. A. Holder, E. W. Tate and E. Cota, *J Biol Chem*, 2012, **287**, 36968–36977.
175 J. Bosch, S. Turley, T. M. Daly, S. M. Bogh, M. L. Villasmil, C. Roach, N. Zhou, J. M. Morrisey, A. B. Vaidya, L. W. Bergman and W. G. J. Hol, *Proc Natl Acad Sci U S A*, 2006, **103**, 4852–4857.
176 S. Kortagere, W. J. Welsh, J. M. Morrisey, T. Daly, I. Ejigiri, P. Sinnis, A. B. Vaidya and L. W. Bergman, *J Chem Inf Model*, 2010, **50**, 840–849.
177 M. J. I. Andrews, G. Kontopidis, C. McInnes, A. Plater, L. Innes, A. Cowan, P. Jewsbury and P. M. Fischer, *Chembiochem*, 2006, **7**, 1909–1915.
178 A. M. Beekman, M. A. O'Connell and L. A. Howell, *Angew Chem Int Ed*, 2017, **56**, 10446–10450.
179 A. M. Beekman, M. M. D. Cominetti, S. J. Walpole, S. Prabhu, M. A. O'Connell, J. Angulo and M. Searcey, *Chem Sci*, 2019, **10**, 4502–4508.
180 L. A. Howell and A. M. Beekman, *RSC Chem Biol*, 2021, **2**, 215–219.
181 F. H. C. Crick, *Nature*, 1952, **170**, 882–883.
182 L. Truebestein and T. A. Leonard, *Bioessays*, 2016, **38**, 903.
183 E. M. Jouaux, K. Schmidtkunz, K. M. Müller and K. M. Arndt, *J Pept Sci*, 2008, **14**, 1022–1031.
184 H. Domingues, D. Cregut, W. Sebald, H. Oschkinat and L. Serrano, *Nat Struct Biol*, 1999, **6**, 652–656.
185 J. M. Mason, K. M. Müller and K. M. Arndt, *Biochemistry*, 2007, **46**, 4804–4814.
186 K. E. Thompson, C. J. Bashor, W. A. Lim and A. E. Keating, *ACS Synth Biol*, 2012, **1**, 118–129.
187 J. M. Fletcher, K. A. Horner, G. J. Bartlett, G. G. Rhys, A. J. Wilson and D. N. Woolfson, *Chemical Science*, 2018, **9**, 7656–7665.
188 A. W. Reinke, R. A. Grant and A. E. Keating, *J Am Chem Soc*, 2010, **132**, 6025–6031.

189 G. Grigoryan, A. W. Reinke and A. E. Keating, *Nature*, 2009, **458**, 859–864.
190 C. W. Wood and D. N. Woolfson, *Protein Sci*, 2018, **27**, 103–111.
191 L. Lu, F. Yu, L. Cai, A. K. Debnath and S. Jiang, *Curr Top Med Chem*, 2016, **16**, 1074–1090.
192 J. R. Hart, A. L. Garner, J. Yu, Y. Ito, M. Sun, L. Ueno, J.-K. Rhee, M. M. Baksh, E. Stefan, M. Hartl, K. Bister, P. K. Vogt and K. D. Janda, *Proc Natl Acad Sci U S A*, 2014, **111**, 12556–12561.
193 J. Chauhan, H. Wang, J. L. Yap, P. E. Sabato, A. Hu, E. V. Prochownik and S. Fletcher, *ChemMedChem*, 2014, **9**, 2274–2285.
194 I. Pavlinov, M. Salkovski and L. N. Aldrich, *Autophagy*, 2020, **16**, 1547–1549.
195 O. Khakshoor and J. S. Nowick, *Curr Opin Chem Biol*, 2008, **12**, 722–729.
196 D. Eisenberg, *Proc Natl Acad Sci U S A*, 2003, **100**, 11207–11210.
197 G. Bellesia and J.-E. Shea, *J Chem Phys*, 2009, **130**, 145103.
198 A. C. Tsolis, N. C. Papandreou, V. A. Iconomidou and S. J. Hamodrakas, *PLoS One*, 2013, **8**, e54175.
199 A. T. Namanja, Y.-J. Li, Y. Su, S. Wong, J. Lu, L. T. Colson, C. Wu, S. S. C. Li and Y. Chen, *J Biol Chem*, 2012, **287**, 3231–3240.
200 E. E. Cawood, E. Baker, T. A. Edwards, D. N. Woolfson, T. K. Karamanos and A. J. Wilson, *Chem Sci*, 2024, **15**, 10237–10245.
201 L. Wang, N. Wang, W. Zhang, X. Cheng, Z. Yan, G. Shao, X. Wang, R. Wang and C. Fu, *Sig Transduct Target Ther*, 2022, **7**, 1–27.
202 Z. Hegedüs, F. Hóbor, D. K. Shoemark, S. Celis, L.-Y. Lian, C. H. Trinh, R. B. Sessions, T. A. Edwards and A. J. Wilson, *Chem Sci*, 2021, **12**, 2286–2293.
203 Qiuyue Zhang, Yan ,Ling, Zhang ,Yuxuan, Zhang ,Lixiao, Yu ,Jia, You ,Qidong, and Lei and Wang, *Future Med Chem*, 2024, **16**, 125–138.
204 J. Kastl, J. Braun, A. Prestel, H. M. Möller, T. Huhn and T. U. Mayer, *ACS Chem Biol*, 2015, **10**, 1661–1666.
205 V. E. Albrow, E. L. Ponder, D. Fasci, M. Békés, E. Deu, G. S. Salvesen and M. Bogyo, *Chem Biol*, 2011, **18**, 722–732.
206 C. W. Robinson, C. S. Rye, N. E. A. Chessum and K. Jones, *Org Biomol Chem*, 2015, **13**, 7402–7407.
207 J. Zang, F. Peters, Y. Cambet, E. Cifuentes-Pagano, M. M. S. Hissabu, C. M. Dustin, L. H. Svensson, M. M. Olesen, M. F. L. Poulsen, S. Jacobsen, P. S. Tuelung, D. Narayanan, A. E. Langkilde, M. Gajhede, P. J. Pagano, V. Jaquet, F. Vilhardt and A. Bach, *J Med Chem*, 2023, **66**, 14963–15005.
208 W.-J. Jeong, J. Bu, Y. Han, A. J. Drelich, A. Nair, P. Král and S. Hong, *J Am Chem Soc*, 2020, **142**, 1832–1837.
209 M. Bojko, K. Węgrzyn, E. Sikorska, M. Kocikowski, M. Parys, C. Battin, P. Steinberger, M. M. Kogut, M. Winnicki, A. K. Sieradzan, M. Spodzieja and S. Rodziewicz-Motowidło, *Bioorg Chem*, 2022, **128**, 106047.
210 D. Hayward, Z. R. Goddard, M. M. D. Cominetti, M. Searcey and A. M. Beekman, *Chem Commun*, 2024, **60**, 8228–8231.
211 L. Calvo-Barreiro, M. Secor, J. Damjanovic, S. A. Abdel-Rahman, Y.-S. Lin and M. Gabr, *Chem Biol Drug Des*, 2025, **105**, e70117.
212 S. Bothe, P. Hänzelmann, S. Böhler, J. Kehrein, M. Zehe, C. Wiedemann, U. A. Hellmich, R. Brenk, H. Schindelin and C. Sotriffer, *Commun Chem*, 2022, **5**, 1–15.
213 Z. M. Oo, A. Illendula, J. Grembecka, C. Schmidt, Y. Zhou, V. Esain, W. Kwan, I. Frost, T. E. North, R. A. Rajewski, N. A. Speck and J. H. Bushweller, *Leuk Lymphoma*, 2018, **59**, 2188–2200.

214 H. Mouhsine, H. Guillemain, G. Moreau, N. Fourati, C. Zerrouki, B. Baron, L. Desallais, P. Gizzi, N. Ben Nasr, J. Perrier, R. Ratsimandresey, J.-L. Spadoni, H. Do, P. England, M. Montes and J.-F. Zagury, *Sci Rep*, 2017, **7**, 3424.
215 S. A. Abdel-Rahman and M. T. Gabr, *SLAS Discovery*, 2025, **34**, 100240.
216 Y. J. Im, L. Kuo, X. Ren, P. V. Burgos, X. Z. Zhao, F. Liu, T. R. Burke, J. S. Bonifacino, E. O. Freed and J. H. Hurley, *Structure*, 2010, **18**, 1536–1547.
217 X. Yang, K. R. Lennard, C. He, M. C. Walker, A. T. Ball, C. Doigneaux, A. Tavassoli and W. A. van der Donk, *Nat Chem Biol*, 2018, **14**, 375–380.
218 Z. Na, S. Pan, M. Uttamchandani and S. Q. Yao, in *Small Molecule Microarrays: Methods and Protocols*, eds M. Uttamchandani and S. Q. Yao, Springer, New York, NY, 2017, pp. 139–156.
219 Y. M. Foong, J. Fu, S. Q. Yao and M. Uttamchandani, *Curr Opin Chem Biol*, 2012, **16**, 234–242.
220 C. H. S. Lu, H. Sun, F. B. Abu Bakar, M. Uttamchandani, W. Zhou, Y.-C. Liou and S. Q. Yao, *Angew Chem Int Ed*, 2008, **47**, 7438–7441.
221 C. Chung and J. Witherington, *SLAS Discovery*, 2011, **16**, 1170–1185.
222 P. Filippakopoulos, J. Qi, S. Picaud, Y. Shen, W. B. Smith, O. Fedorov, E. M. Morse, T. Keates, T. T. Hickman, I. Felletar, M. Philpott, S. Munro, M. R. McKeown, Y. Wang, A. L. Christie, N. West, M. J. Cameron, B. Schwartz, T. D. Heightman, N. La Thangue, C. A. French, O. Wiest, A. L. Kung, S. Knapp and J. E. Bradner, *Nature*, 2010, **468**, 1067–1073.
223 J. Gavenonis, B. A. Sheneman, T. R. Siegert, M. R. Eshelman and J. A. Kritzer, *Nat Chem Biol*, 2014, **10**, 716–722.
224 J. A. Wells and C. L. McClendon, *Nature*, 2007, **450**, 1001–1009.
225 L. Lombardi, L. A. Granger, R. J. Shattock and D. R. Williams, *J Org Chem*, 2025, **90**, 1467–1477.
226 M.-C. Lu, Q. Jiao, T. Liu, S.-J. Tan, H.-S. Zhou, Q.-D. You and Z.-Y. Jiang, *Eur J Med Chem*, 2018, **143**, 1578–1589.
227 P. Furet, B. Salem, Y. Mesrouze, T. Schmelzle, I. Lewis, J. Kallen and P. Chène, *Bioorg Med Chem Lett*, 2019, **29**, 2316–2319.
228 D. Marcotte, W. Zeng, J.-C. Hus, A. McKenzie, C. Hession, P. Jin, C. Bergeron, A. Lugovskoy, I. Enyedy, H. Cuervo, D. Wang, C. Atmanene, D. Roecklin, M. Vecchi, V. Vivat, J. Kraemer, D. Winkler, V. Hong, J. Chao, M. Lukashev and L. Silvian, *Bioorg Med Chem*, 2013, **21**, 4011–4019.
229 S. Kaseda, Y. Sannomiya, J. Horizono, J. Kuwazuru, M. A. Suico, S. Ogi, R. Sasaki, H. Sunamoto, H. Fukiya, H. Nishiyama, M. Kamura, S. Niinou, Y. Koyama, F. Nara, T. Shuto, K. Onuma and H. Kai, *Kidney360*, 2021, **3**, 687–699.
230 M.-C. Lu, X. Zhang, F. Wu, S.-J. Tan, J. Zhao, Q.-D. You and Z.-Y. Jiang, *J Med Chem*, 2019, **62**, 6796–6813.
231 D. A. Abed, A. R. Ali, S. Lee, M.-U. Nguyen, M. P. Verzi and L. Hu, *Eur J Med Chem*, 2023, **252**, 115302.
232 S. Lee, D. A. Abed, M.-U. Nguyen, M. P. Verzi and L. Hu, *Eur. J Med Chem*, 2022, **237**, 114380.
233 X. Zhang, K. Cui, X. Wang, Y. Tong, C. Liu, Y. Zhu, Q. You, Z. Jiang and X. Guo, *Antioxidants*, 2023, **12**, 1062.
234 D. Narayanan, K. T. Tran, J. S. Pallesen, S. M. Ø. Solbak, Y. Qin, E. Mukminova, M. Luchini, K. O. Vasilyeva, D. González Chichón, G. Goutsiou, C. Poulsen, N. Haapanen, G. M. Popowicz, M. Sattler, D. Olagnier, M. Gajhede and A. Bach, *J Med Chem*, 2022, **65**, 14481–14526.
235 G. H. S. Dreiman, M. Bictash, P. V. Fish, L. Griffin and F. Svensson, *SLAS Discovery*, 2021, **26**, 257–262.

236 T. Zhu, S. Cao, P.-C. Su, R. Patel, D. Shah, H. B. Chokshi, R. Szukala, M. E. Johnson and K. E. Hevener, *J Med Chem*, 2013, **56**, 6560–6572.
237 D. L. Fulton, S. Sundararajan, G. Badis, T. R. Hughes, W. W. Wasserman, J. C. Roach and R. Sladek, *Genome Biol*, 2009, **10**, 1–14.
238 S. A. Lambert, A. Jolma, L. F. Campitelli, P. K. Das, Y. Yin, M. Albu, X. Chen, J. Taipale, T. R. Hughes and M. T. Weirauch, *Cell*, 2018, **172**, 650–665.
239 S. V. Razin, V. V. Borunova, O. G. Maksimenko and O. L. Kantidze, *Biochem (Mosc)*, 2012, **77**, 217–226.
240 A. de Mendoza, A. Sebé-Pedrós, M. S. Šestak, M. Matejcic, G. Torruella, T. Domazet-Loso and I. Ruiz-Trillo, *Proc Natl Acad Sci U S A*, 2013, **110**, E4858-4866.
241 X. Zhang, R. M. Blumenthal and X. Cheng, *Curr Opin Struct Biol*, 2024, **87**, 102836.
242 A. Klug, *Annu Rev Biochem*, 2010, **79**, 213–231.
243 J. Miller, A. D. McLachlan and A. Klug, *EMBO J*, 1985, **4**, 1609–1614.
244 J. P. Mackay and M. Crossley, *Trends Biochem Sci*, 1998, **23**, 1–4.
245 W. J. Gehring, M. Affolter and T. Bürglin, *Annu Rev Biochem*, 1994, **63**, 487–526.
246 W. J. Gehring, Y. Q. Qian, M. Billeter, K. Furukubo-Tokunaga, A. F. Schier, D. Resendez-Perez, M. Affolter, G. Otting and K. Wüthrich, *Cell*, 1994, **78**, 211–223.
247 A. Aranda and A. Pascual, *Physiological Reviews*, 2001, **81**, 1269–1304.
248 C. Murre, P. S. McCaw, H. Vaessin, M. Caudy, L. Y. Jan, Y. N. Jan, C. V. Cabrera, J. N. Buskin, S. D. Hauschka and A. B. Lassar, *Cell*, 1989, **58**, 537–544.
249 S. Jones, *Genome Biol*, 2004, **5**, 1–6.
250 K. Jindrich and B. M. Degnan, *BMC Evolutionary Biology*, 2016, **16**, 28.
251 A. Hasan, N. A. Khan, S. Uddin, A. Q. Khan and M. Steinhoff, *Semin Cancer Biol*, 2024, **98**, 31–50.
252 A. U. Ahmed, B. R. G. Williams and G. E. Hannigan, *Biomolecules*, 2015, **5**, 3087–3111.
253 C. Liu, K. Omilusik, C. Toma, N. S. Kurd, J. T. Chang, A. W. Goldrath and W. Wang, *PLoS Comput Biol*, 2022, **18**, e1010116.
254 M. J. Henley and A. N. Koehler, *Nat Rev Drug Discov*, 2021, **20**, 669–688.
255 E. Viziteu, C. Grandmougin, H. Goldschmidt, A. Seckinger, D. Hose, B. Klein and J. Moreaux, *Br J Cancer*, 2016, **114**, 519–523.
256 H. J. Dyson and P. E. Wright, *J Biol Chem*, 2016, **291**, 6714–6722.
257 T. Zor, R. N. De Guzman, H. J. Dyson and P. E. Wright, *J Mol Biol*, 2004, **337**, 521–534.
258 S. A. Dames, M. Martinez-Yamout, R. N. De Guzman, H. J. Dyson and P. E. Wright, *Proc Natl Acad Sci U S A*, 2002, **99**, 5271–5276.
259 A. Lobley, M. B. Swindells, C. A. Orengo and D. T. Jones, *PLOS Computational Biology*, 2007, **3**, e162.
260 J. H. Bushweller, *Nature Reviews Cancer 2019 19:11*, 2019, **19**, 611–624.
261 J. R. Whitfield and L. Soucek, *J Cell Biol*, 2021, **220**, e202103090.
262 P. Nioi, M. McMahon, K. Itoh, M. Yamamoto and J. D. Hayes, *Biochem J*, 2003, **374**, 337–348.
263 T. Nguyen, S. H. Hong and P. Arora, *J Am Chem Soc*, 2025, **147**, 13296–13302.
264 D. Massó-Vallés and L. Soucek, *Cells*, 2020, **9**, 883.
265 L. Soucek, M. Helmer-Citterich, A. Sacco, R. Jucker, G. Cesareni and S. Nasi, *Oncogene*, 1998, **17**, 2463–2472.
266 I. Goldstein and G. L. Hager, *WIRES Syst Biol Med*, 2018, **10**, e1390.
267 R. Schleif, *Science*, 1988, **241**, 1182–1187.
268 K. Fujita and N. Nonomura, *World J Mens Health*, 2019, **37**, 288–295.
269 K. Dalal, M. Roshan-Moniri, A. Sharma, H. Li, F. Ban, M. Hessein, M. Hsing, K. Singh, E. LeBlanc, S. Dehm, E. S. Tomlinson Guns, A. Cherkasov and P. S. Rennie, *J Biol Chem*, 2014, **289**, 26417–26429.

270 J. P. Andrews, R. S. Joshi, M. P. Pereira, T. Oh, A. F. Haddad, K. M. Pereira, R. C. Osorio, K. C. Donohue, Z. Peeran, S. Sudhir, S. Jain, A. Beniwal, A. S. Chopra, N. S. Sandhu, T. Tihan, L. Blevins and M. K. Aghi, *World Neurosurg*, 2021, **151**, e185–e191.
271 P. Peixoto, Y. Liu, S. Depauw, M.-P. Hildebrand, D. W. Boykin, C. Bailly, W. D. Wilson and M.-H. David-Cordonnier, *Nucleic Acids Res*, 2008, **36**, 3341–3353.
272 M. Munde, A. Kumar, R. Nhili, S. Depauw, M.-H. David-Cordonnier, M. A. Ismail, C. E. Stephens, A. A. Farahat, A. Batista-Parra, D. W. Boykin and W. D. Wilson, *J Mol Biol*, 2010, **402**, 847–864.
273 A. Chen and A. N. Koehler, *Trends Mol Med*, 2020, **26**, 508–518.
274 P. Moi, K. Chan, I. Asunis, A. Cao and Y. W. Kan, *Proc Natl Acad Sci U S A*, 1994, **91**, 9926–9930.
275 F. He, X. Ru and T. Wen, *Int. J. Mol. Sci*, 2020, **21**, 4777.
276 F. Pouremamali, A. Pouremamali, M. Dadashpour, N. Soozangar and F. Jeddi, *Cell. Commun. Signal*, 2022, **20**, 100.
277 Natalia Robledinos-Antón, Raquel Fernández-Ginés, Gina Manda, and Antonio Cuadrado, *Oxid Med Cell Longev*, **14**, 9372182.
278 X.-J. Wang, Z. Sun, N. F. Villeneuve, S. Zhang, F. Zhao, Y. Li, W. Chen, X. Yi, W. Zheng, G. T. Wondrak, P. K. Wong and D. D. Zhang, *Carcinogenesis*, 2008, **29**, 1235–1243.
279 S. Saha, B. Buttari, E. Panieri, E. Profumo and L. Saso, *Molecules*, 2020, **25**, 5474.
280 K. I. Tong, A. Kobayashi, F. Katsuoka and M. Yamamoto, *Biol Chem*, 2006, **387**, 1311–1320.
281 K. Itoh, N. Wakabayashi, Y. Katoh, T. Ishii, T. O'Connor and M. Yamamoto, *Genes to Cells*, 2003, **8**, 379–391.
282 W. H. Landschulz, P. F. Johnson and S. L. McKnight, *Science*, 1988, **240**, 1759–1764.
283 A. Kopacz, A. I. Rojo, C. Patibandla, D. Lastra-Martínez, A. Piechota-Polanczyk, D. Kloska, A. Jozkowicz, C. Sutherland, A. Cuadrado and A. Grochot-Przeczek, *Free Radic Biol Med*, 2022, **192**, 37–49.
284 C. Tonelli, I. I. C. Chio and D. A. Tuveson, *Antioxid Redox Signal*, 2018, **29**, 1727–1745.
285 T. Sengoku, M. Shiina, K. Suzuki, K. Hamada, K. Sato, A. Uchiyama, S. Kobayashi, A. Oguni, H. Itaya, K. Kasahara, H. Moriwaki, C. Watanabe, T. Honma, C. Okada, S. Baba, T. Ohta, H. Motohashi, M. Yamamoto and K. Ogata, *Nucleic Acids Res*, 2022, **50**, 12543–12557.
286 W. Li, S. Yu, T. Liu, J. H. Kim, V. Blank, H. Li and A. N. T. Kong, *Biochim Biophys Acta Mol Cell Res*, 2008, **1783**, 1847–1856.
287 C. Vinson, M. Myakishev, A. Acharya, A. A. Mir, J. R. Moll and M. Bonovich, *Mol Cell Biol*, 2002, **22**, 6321–6335.
288 F. Katsuoka and M. Yamamoto, *Gene*, 2016, **586**, 197–205.
289 W. Li and A.-N. Kong, *Mol Carcinog*, 2009, **48**, 91–104.
290 H. Kurokawa, H. Motohashi, S. Sueno, M. Kimura, H. Takagawa, Y. Kanno, M. Yamamoto and T. Tanaka, *Mol Cell Biol*, 2009, **29**, 6232–6244.
291 M.-K. Kwak, K. Itoh, M. Yamamoto and T. W. Kensler, *Mol Cell Biol*, 2002, **22**, 2883–2892.
292 J. M. Maher, M. Z. Dieter, L. M. Aleksunes, A. L. Slitt, G. Guo, Y. Tanaka, G. L. Scheffer, J. Y. Chan, J. E. Manautou, Y. Chen, T. P. Dalton, M. Yamamoto and C. D. Klaassen, *Hepatology*, 2007, **46**, 1597–1610.
293 M. S. Yates, Q. T. Tran, P. M. Dolan, W. O. Osburn, S. Shin, C. C. McCulloch, J. B. Silkworth, K. Taguchi, M. Yamamoto, C. R. Williams, K. T. Liby, M. B. Sporn, T. R. Sutter and T. W. Kensler, *Carcinogenesis*, 2009, **30**, 1024–1031.
294 R. S. Friling, S. Bergelson and V. Daniel, *Proc Natl Acad Sci U S A*, 1992, **89**, 668–672.

295 T. H. Rushmore, M. R. Morton and C. B. Pickett, *J Biol Chem*, 1991, **266**, 11632–11639.
296 M. Pajares, A. Cuadrado and A. I. Rojo, *Redox Biol*, 2017, **11**, 543–553.
297 A. Luchkova, A. Mata and S. Cadenas, *FEBS Letters*, 2024, **598**, 2092–2105.
298 Y. Hu, Y. Luo and Y. Zheng, *Antioxidants (Basel)*, 2022, **11**, 1747.
299 J. D. Hayes, S. A. Chanas, C. J. Henderson, M. McMahon, C. Sun, G. J. Moffat, C. R. Wolf and M. Yamamoto, *Biochem Soc Trans*, 2000, **28**, 33–41.
300 L. E. Tebay, H. Robertson, S. T. Durant, S. R. Vitale, T. M. Penning, A. T. Dinkova-Kostova and J. D. Hayes, *Free Radic Biol Med*, 2015, **88**, 108–146.
301 A. Abu-Bakar, J. Hakkola, R. Juvonen, M. Rahnasto-Rilla, H. Raunio and M. A. Lang, *Curr Drug Metab*, 2013, **14**, 137–150.
302 T. Ashino, H. Ohkubo-Morita, M. Yamamoto, T. Yoshida and S. Numazawa, *Redox Biology*, 2014, **2**, 284–288.
303 D. Martin, A. I. Rojo, M. Salinas, R. Diaz, G. Gallardo, J. Alam, C. M. R. de Galarreta and A. Cuadrado, *J Biol Chem*, 2004, **279**, 8919–8929.
304 K. D. Poss and S. Tonegawa, *Proc Natl Acad Sci U S A*, 1997, **94**, 10925–10930.
305 J.-Y. Im, K.-W. Lee, J.-M. Woo, E. Junn and M. M. Mouradian, *Hum Mol Genet*, 2012, **21**, 3013–3024.
306 C. M. Clements, R. S. McNally, B. J. Conti, T. W. Mak and J. P.-Y. Ting, *Proc Natl Acad Sci U S A*, 2006, **103**, 15091–15096.
307 Y. Mitsuishi, K. Taguchi, Y. Kawatani, T. Shibata, T. Nukiwa, H. Aburatani, M. Yamamoto and H. Motohashi, *Cancer Cell*, 2012, **22**, 66–79.
308 V. Malec, O. R. Gottschald, S. Li, F. Rose, W. Seeger and J. Hänze, *Free Radic Biol Med*, 2010, **48**, 1626–1635.
309 L. Ji, H. Li, P. Gao, G. Shang, D. D. Zhang, N. Zhang and T. Jiang, *PLoS One*, 2013, **8**, e63404.
310 A.-A. Zimta, D. Cenariu, A. Irimie, L. Magdo, S. M. Nabavi, A. G. Atanasov and I. Berindan-Neagoe, *Cancers (Basel)*, 2019, **11**, 1755.
311 X. Wang, C. R. Campos, J. C. Peart, L. K. Smith, J. L. Boni, R. E. Cannon and D. S. Miller, *J Neurosci*, 2014, **34**, 8585–8593.
312 N. Wakabayashi, J. J. Skoko, D. V. Chartoumpekis, S. Kimura, S. L. Slocum, K. Noda, D. L. Palliyaguru, M. Fujimuro, P. A. Boley, Y. Tanaka, N. Shigemura, S. Biswal, M. Yamamoto and T. W. Kensler, *Mol Cell Biol*, 2014, **34**, 653–663.
313 Y. Chen, J. Pang, L. Ye, Z. Zhang, S. Lin, N. Lin, T. H. Lee and H. Liu, *Biomed Pharmacother*, 2023, **166**, 115383.
314 H.-B. Wang, C.-J. Zhou, S. Song, P. Chen, W. Xu, B. Liu, K. Zhu, W. Yu, H. Wu, H. Wang, S. Lin, J. Guo and C. Qin, *Pathol Res Pract*, 2011, **207**, 169–173.
315 S. Riis, J. B. Murray and R. O'Connor, *Cells*, 2020, **9**, 147.
316 N. Kweider, A. Fragoulis, C. Rosen, U. Pecks, W. Rath, T. Pufe and C. J. Wruck, *J Biol Chem*, 2011, **286**, 42863–42872.
317 S. K. Niture and A. K. Jaiswal, *J Biol Chem*, 2012, **287**, 9873–9886.
318 J. M. Hardwick and L. Soane, *Cold Spring Harb Perspect Biol*, 2013, **5**, a008722.
319 S. K. Niture and A. K. Jaiswal, *Free Radic Biol Med*, 2013, **57**, 119–131.
320 C. Zhang, H.-J. Wang, Q.-C. Bao, L. Wang, T.-K. Guo, W.-L. Chen, L.-L. Xu, H.-S. Zhou, J.-L. Bian, Y.-R. Yang, H.-P. Sun, X.-L. Xu and Q.-D. You, *Oncotarget*, 2016, **7**, 73593–73606.
321 S. Saito, H. Ozawa, Y. Imanishi, M. Sekimizu, Y. Watanabe, F. Ito, Y. Ikari, N. Nakahara, K. Kameyama and K. Ogawa, *Oncol Lett*, 2021, **22**, 533.
322 B. R. Imhoff and J. M. Hansen, *Biochem J*, 2009, **424**, 491–500.
323 W.-H. Hou, L. Rossi, Y. Shan, J.-Y. Zheng, R. W. Lambrecht and H. L. Bonkovsky, *World Journal of Gastroenterology*, 2009, **15**, 4499–4510.

324 J. Zhou, Q. Zheng and Z. Chen, *Front Cell Dev Biol*, 2022, **10**, 826204.
325 M. Fuertes-Agudo, M. Luque-Tévar, C. Cucarella, P. Martín-Sanz and M. Casado, *Antioxidants*, 2023, **12**, 1491.
326 J. Weerachayaphorn, S.-Y. Cai, C. J. Soroka and J. L. Boyer, *Hepatology*, 2009, **50**, 1588–1596.
327 D. G. Helou, B. Noël, F. Gaudin, H. Groux, Z. El Ali, M. Pallardy, S. Chollet-Martin and S. Kerdine-Römer, *J Immunol*, 2019, **202**, 2189–2194.
328 A. Chaiprasongsuk, J. Lohakul, K. Soontrapa, S. Sampattavanich, P. Akarasereenont and U. Panich, *J Pharmacol Exp Ther*, 2017, **360**, 388–398.
329 S. Peng, L. Shen, X. Yu, L. Zhang, K. Xu, Y. Xia, L. Zha, J. Wu and H. Luo, *Front Immunol*, 2023, **14**, 1200111.
330 Y. Qiu, B. Wan, G. Liu, Y. Wu, D. Chen, M. Lu, J. Chen, R. Yu, D. Chen and Q. Pang, *Respir Res*, 2020, **21**, 1–16.
331 T. E. Sussan, S. Gajghate, S. Chatterjee, P. Mandke, S. McCormick, K. Sudini, S. Kumar, P. N. Breysse, G. B. Diette, V. K. Sidhaye and S. Biswal, *Am J Physiol Lung Cell Mol Physiol*, 2015, **309**, L27–L36.
332 A. Boutten, D. Goven, E. Artaud-Macari, J. Boczkowski and M. Bonay, *Trends Mol Med*, 2011, **17**, 363–371.
333 H. Yang, H. Lv, H. Li, X. Ci and L. Peng, *Cell Commun Signal*, 2019, **17**, 1–15.
334 H. Dong, Z. Qiang, D. Chai, J. Peng, Y. Xia, R. Hu and H. Jiang, *Aging*, 2020, **12**, 12943–12959.
335 Y. Shintani, S. Maruoka, Y. Gon, D. Koyama, A. Yoshida, Y. Kozu, K. Kuroda, I. Takeshita, E. Tsuboi, K. Soda and S. Hashimoto, *Allergology International*, 2015, **64**, S54–S63.
336 D. Moretti, S. Tambone, M. Cerretani, P. Fezzardi, A. Missineo, L.-T. Sherman, I. Munoz-Sajuan, S. Harper, C. Dominquez, R. Pacifici, L. Tomei, L. Park and A. Bresciani, *Free Radic Biol Med*, 2021, **162**, 243–254.
337 S. Petrillo, T. Schirinzi, G. Di Lazzaro, J. D'Amico, V. L. Colona, E. Bertini, M. Pierantozzi, L. Mari, N. B. Mercuri, F. Piemonte and A. Pisani, *Mov Disord*, 2020, **35**, 180–184.
338 M. T. Fischer, R. Sharma, J. L. Lim, L. Haider, J. M. Frischer, J. Drexhage, D. Mahad, M. Bradl, J. van Horssen and H. Lassmann, *Brain*, 2012, **135**, 886–899.
339 Y. Xiao, C. Karam, J. Yi, L. Zhang, X. Li, D. Yoon, H. Wang, K. Dhakal, P. Ramlow, T. Yu, Z. Mo, J. Ma and J. Zhou, *Pharmacol Res*, 2018, **138**, 25–36.
340 S. Pagnotta, A. Tramutola, E. Barone, F. Di Domenico, V. Pittalà, L. Salerno, V. Folgiero, M. Caforio, F. Locatelli, S. Petrini, D. A. Butterfield and M. Perluigi, *Free Radic Biol Med*, 2022, **183**, 1–13.
341 G. Shang, X. Tang, P. Gao, F. Guo, H. Liu, Z. Zhao, Q. Chen, T. Jiang, N. Zhang and H. Li, *J Nutr Biochem*, 2015, **26**, 596–606.
342 S. Zhao, A. Ghosh, C.-S. Lo, I. Chenier, J. W. Scholey, J. G. Filep, J. R. Ingelfinger, S.-L. Zhang and J. S. D. Chan, *Endocrinology*, 2018, **159**, 836–852.
343 A. K. Aranda-Rivera, A. Cruz-Gregorio, J. Pedraza-Chaverri and A. Scholze, *Antioxidants (Basel)*, 2022, **11**, 1112.
344 S. D. Chung, T. Y. Lai, C. T. Chien and H. J. Yu, *PLoS One*, 2012, **7**, e47299.
345 B. M. Rush, C. D. Bondi, S. D. Stocker, K. M. Barry, S. A. Small, J. Ong, S. Jobbagy, D. B. Stolz, S. I. Bastacky, D. V. Chartoumpekis, T. W. Kensler and R. J. Tan, *Kidney Int*, 2021, **99**, 102–116.
346 P. Rossing, G. A. Block, M. P. Chin, A. Goldsberry, H. J. L. Heerspink, P. A. McCullough, C. J. Meyer, D. Packham, P. E. Pergola, B. Spinowitz, S. M. Sprague, D. G. Warnock and G. M. Chertow, *Kidney Int*, 2019, **96**, 1030–1036.

347 D. A. Frohlich, M. T. McCabe, R. S. Arnold and M. L. Day, *Oncogene*, 2008, **27**, 4353–4362.
348 N. M. Reddy, S. R. Kleeberger, J. H. Bream, P. G. Fallon, T. W. Kensler, M. Yamamoto and S. P. Reddy, *Oncogene*, 2008, **27**, 5821–5832.
349 D. J. Long, R. L. Waikel, X. J. Wang, L. Perlaky, D. R. Roop and A. K. Jaiswal, *Cancer Res*, 2000, **60**, 5913–5915.
350 M. Ramos-Gomez, M. K. Kwak, P. M. Dolan, K. Itoh, M. Yamamoto, P. Talalay and T. W. Kensler, *Proc Natl Acad Sci U S A*, 2001, **98**, 3410–3415.
351 Y. Kitamura, T. Umemura, K. Kanki, Y. Kodama, S. Kitamoto, K. Saito, K. Itoh, M. Yamamoto, T. Masegi, A. Nishikawa and M. Hirose, *Cancer Sci*, 2007, **98**, 19–24.
352 K. Iida, K. Itoh, J. M. Maher, Y. Kumagai, R. Oyasu, Y. Mori, T. Shimazui, H. Akaza and M. Yamamoto, *Carcinogenesis*, 2007, **28**, 2398–2403.
353 R. Grossman and Z. Ram, *World Neurosurg*, 2013, **80**, 284–286.
354 W. S, L. H and B. Y, *Cancer medicine*, 2019, **8**, 2252–2267.
355 H. Kitamura and H. Motohashi, *Cancer Sci*, 2018, **109**, 900–911.
356 L. M. Solis, C. Behrens, W. Dong, M. Suraokar, N. C. Ozburn, C. A. Moran, A. H. Corvalan, S. Biswal, S. G. Swisher, B. N. Bekele, J. D. Minna, D. J. Stewart and I. I. Wistuba, *Clin Cancer Res*, 2010, **16**, 3743–3753.
357 A. Lister, T. Nedjadi, N. R. Kitteringham, F. Campbell, E. Costello, B. Lloyd, I. M. Copple, S. Williams, A. Owen, J. P. Neoptolemos, C. E. Goldring and B. K. Park, *Mol Cancer*, 2011, **10**, 1–13.
358 K. C. Wu, J. Y. Cui and C. D. Klaassen, *Toxicol Sci*, 2011, **123**, 590–600.
359 Y. Mitsuishi, K. Taguchi, Y. Kawatani, T. Shibata, T. Nukiwa, H. Aburatani, M. Yamamoto and H. Motohashi, *Cancer Cell*, 2012, **22**, 66–79.
360 B. Padmanabhan, K. I. Tong, T. Ohta, Y. Nakamura, M. Scharlock, M. Ohtsuji, M.-I. Kang, A. Kobayashi, S. Yokoyama and M. Yamamoto, *Mol Cell*, 2006, **21**, 689–700.
361 Y. Chen, D. Inoyama, A.-N. T. Kong, L. J. Beamer and L. Hu, *Chem Biol Drug Des*, 2011, **78**, 1014–1021.
362 R. Steel, J. Cowan, E. Payerne, M. A. O'Connell and M. Searcey, *ACS Med Chem Lett*, 2012, **3**, 407–410.
363 D. Inoyama, Y. Chen, X. Huang, L. J. Beamer, A.-N. T. Kong and L. Hu, *J Biomol Screen*, 2012, **17**, 435–447.
364 C.-H. Leung, J.-T. Zhang, G.-J. Yang, H. Liu, Q.-B. Han and D.-L. Ma, *Int J Mol Sci*, 2019, **20**, 4445.
365 G. Wu, G. Xu, B. A. Schulman, P. D. Jeffrey, J. W. Harper and N. P. Pavletich, *Mol Cell*, 2003, **11**, 1445–1456.
366 J. Pons, N. Evrard-Todeschi, G. Bertho, J. Gharbi-Benarous, V. Tanchou, R. Benarous and J.-P. Girault, *Biochemistry*, 2008, **47**, 14–29.
367 S. Unni, P. Deshmukh, G. Krishnappa, P. Kommu and B. Padmanabhan, *The FEBS Journal*, 2021, **288**, 1599–1613.
368 G.-Z. Liu, L.-B. Fang, P. Hjelmström and X.-G. Gao, *Mult Scler*, 2007, **13**, 149–155.
369 A. Singh, S. Venkannagari, K. H. Oh, Y.-Q. Zhang, J. M. Rohde, L. Liu, S. Nimmagadda, K. Sudini, K. R. Brimacombe, S. Gajghate, J. Ma, A. Wang, X. Xu, S. A. Shahane, M. Xia, J. Woo, G. A. Mensah, Z. Wang, M. Ferrer, E. Gabrielson, Z. Li, F. Rastinejad, M. Shen, M. B. Boxer and S. Biswal, *ACS Chem Biol*, 2016, **11**, 3214–3225.
370 S. Brüschweiler, J. E. Fuchs, G. Bader, D. B. McConnell, R. Konrat and M. Mayer, *ChemMedChem*, 2021, **16**, 3576–3587.
371 M. Dodson, M. R. de la Vega, A. B. Cholanians, C. J. Schmidlin, E. Chapman and D. D. Zhang, *Annu Rev Pharmacol Toxicol*, 2019, **59**, 555–575.

372 Z. Tang, L. Zhao, Z. Yang, Z. Liu, J. Gu, B. Bai, J. Liu, J. Xu and H. Yang, *Int J Nanomedicine*, 2018, **13**, 2907–2919.
373 E.-J. Jeong, J. J. Choi, S. Y. Lee and Y. S. Kim, *Int J Mol Sci*, 2024, **25**, 7011.
374 L. Ji, N. Moghal, X. Zou, Y. Fang, S. Hu, Y. Wang and M. S. Tsao, *Cancer Med*, 2023, **12**, 5688–5702.
375 C. Xu, Y. Chen, Z. Zhou, Y. Yan, W. Fu, P. Zou and D. Ni, *ACS Omega*, 2024, **9**, 43697–43705.
376 L. Yan, H. Hu, L. Feng, Z. Li, C. Zheng, J. Zhang, X. Yin and B. Li, *Med Oncol*, 2024, **41**, 309.
377 X. Yu, Y. Wang, J. Tan, Y. Li, P. Yang, X. Liu, J. Lai, Y. Zhang, L. Cai, Y. Gu, L. Xu and Y. Li, *Cell Death Discov*, 2024, **10**, 1–10.
378 B. Wiedemann, D. Kamps, L. Depta, J. rn Weisner, J. Cvetreznik, S. TomassiID, S. Gentz, J.-E. Hoffmann, M. P. Mü ller, O. Koch, L. Dehmelt and D. RauhID, *PLOS ONE*, 2022, **17**, e0267651.
379 V. Simov, M. D. Altman, E. Bianchi, S. DelRizzo, E. N. DiNunzio, G. Feng, P. Goldenblatt, R. Ingenito, S. A. Johnson, M. S. Mansueto, T. Mayhood, J. D. Mortison, V. Serebrov, C. Sondey, V. Sriraman, T. J. Tucker, A. Walji, H. Wan, Y. Yue, A. Stoeck and E. F. DiMauro, *Eur J Med Chem*, 2021, **224**, 113686.
380 J. Ji, S. Ma, Y. Zhu, J. Zhao, Y. Tong, Q. You and Z. Jiang, *J Med Chem*, 2023, **66**, 6070–6081.
381 S. Dhakshinamoorthy, A. K. Jain, D. A. Bloom and A. K. Jaiswal, *J Biol Chem*, 2005, **280**, 16891–16900.

Chapter 2
Design, Synthesis and Biophysical Evaluation of Peptides to Inhibit Nrf2/MafG

2.1 Introduction

2.1.1 Identification of coiled-coil targeting peptides

Peptides are short chains of amino acids conjugated by amide bonds, containing at least two, and up to fifty, residues with longer sequences considered proteins.[1,2] Peptides offer an advantage over small molecules for the disruption of protein-protein interactions (PPIs) as they are capable of engaging large surface areas to effectively compete with native binding partners. As discussed in **Chapter 1**, peptides can offer a valuable starting point for identifying key residues and binding regions of proteins for several types of PPI.

Coiled-coil interactions have been explored extensively through protein and peptide sequence mutations to understand the drivers of interacting α-helices. Over the past 25 years, a selection of coiled-coil targeting peptides have been developed to afford inhibition of disease-relevant coiled-coil interactions.[3–7] Whilst none have progressed towards the clinic, these peptides provide valuable information on disease mechanisms and have generated understanding of coiled-coil inhibition.

Development of peptides to control coiled-coil interactions is challenging due to the long and shallow surface of the interaction. Coiled-coils require high target complementarity to afford selective PPI formation. Peptides to control these interactions must be able to form a more stable coiled-coil interaction than the parent proteins.

2.1.1.1 Random peptide libraries

One of the first methods used to identify coiled-coil binding peptides was the protein complementation assays (PCA). Used to identify novel peptide sequences against coiled-coil interactions in microphthalmia-associated transcription factor (MTIF) and Fos/Jun transcription factors.[3,4] PCA utilises survival selection in modified bacterial cells through a reporter enzyme where PPI formation rescues the activity of the enzyme, conjugated to the proteins of interest.[8–10]

A lead peptide inhibiting the homodimeric coiled-coil PPI of MITF was identified through library screening by PCA performed in bacterial cells, allowing for selection of sequences that were also stably folded in a cellular environment.[11] Through this method, a highly stable peptide with strong target affinity was identified. As measured by a thermal melt assay, the coiled-coil stability increased the thermal melt (ΔT_m) value by +13.4 °C compared to the MITF homodimer.[11]

PCA was also used to identify peptide antagonists of c-Jun and c-Fos, in combination with phage display libraries, yielding a peptide with higher stability for c-Fos than native c-Jun withstanding a ΔT_m of + 37 °C.[5] A key problem in binding heterodimeric coiled-coils is that peptide inhibiting the PPI may have affinity for either binding partner. By combining PCA with semi-rational library design, the peptide FosW was investigated by thermal melting assays to improve stability from the wild-type PPI ΔT_m by + 47 °C against c-Jun and + 28 °C binding to c-Fos.[4]

To identify peptides with selectivity for only one of the proteins in a coiled-coil dimer, PCA techniques can be improved to select for a singular interaction surface, referred to as a competitive and negative design initiative (CANDI).[12] This has led to the discovery of $JunW_{CANDI}$, which maintained a ΔT_m of + 28 °C for c-Fos but only + 7 °C for c-Jun.[12] Despite these improvements, retaining strong binding affinity of peptide inhibitors for the c-Jun/c-Fos PPI without competing with other coiled-coil transcription factors in the (Activator Protein) AP-1 family remains an ongoing challenge.[6]

2.1.1.2 De novo peptide design

The design of *de novo* peptide sequences for inhibiting coiled-coil PPIs has been developed through computational sequence generation. Combined with a solid support peptide array of sequences, binding events of sequences to known bZIP regions of transcription factors were screened.[13] This led to the design and assessment of binding partners for each of the 20 bZIP families that could be applied to future coiled-coil peptide inhibitor design.[6]

2.1.1.3 Rational peptide design

Native sequence mimetics of coiled-coil proteins have produced peptide inhibitors against the HIV gp41 protein and epidermal growth factor receptor (EGFR) dimers as well as in helix-loop-helix and basic leucine zipper transcription factors.[14,15]

Peptides based on the C-terminal helix of the HIV gp41 viral entry protein can form coiled-coil interactions with the N-terminal trimeric coiled-coil.[15] Investigation by thermal melt assay found the native peptide sequence produced a T_m of 66 °C for the N-terminal helix of gp41. Residue mutation by alanine scanning the sequence found consistent decrease in T_m stability, with the greatest decrease (ΔT_m -29 °C) through modification of the C-terminal tryptophan residue. As described in **Chapter 1.2.3,** core binding residues of coiled-coils are highly predictable, following a heptadic binding motif to inform key residues of the binding interaction, which perhaps diminishes the

need to perform residue scanning. This was further confirmed by the wild-type C-terminal peptide, which retained the highest activity by viral entry assays (2.1 nM IC_{50}) compared to mutants.[3]

Determining peptide-protein stability by melting temperatures are not indicative of inhibitory activity for a coiled-coil interaction. For example, a 31-mer peptide sequence derived from the coiled-coil region of the BZLF1 transcription factor were found to be capable of reducing protein-DNA complex formation by 76% at 10 nM by mobility shift assay, despite a relatively low ΔT_m value of 25 °C to the BZLF1 transcription factor.[7,16,17]

Biophysical assays, such as fluorescence polarisation, can be used to identify peptide inhibitors with applications for high throughput screening of novel peptide or small molecule inhibitors of a coiled-coil. Inhibition of trimeric coiled-coil forming Heat Shock Factor 1 (HSF1) transcription factor was achieved by evaluated fourteen HSF1 derived peptides. The peptides were assessed by fluorescence polarisation assay measuring the formation of the transcription factor complex with DNA. Analysis at 400 μM led to the identification of a 16-mer peptide that disrupted complex formation by 91%.[18] The lead peptide was then taken forward to design a peptide based fluorescent probe for the interaction, which demonstrated affinity for the coiled-coil region of the protein (13 μM).[18]

Peptides derived from proteins can also be investigated by *in vitro* cellular assays to determine inhibitory activity of a target pathway mediated by a coiled-coil. EGFR signalling is facilitated by a coiled-coil interaction at the juxtamembrane domain (JXD) of the intracellular kinase. The JXD has been investigated for peptide inhibition by cellular assessment.[19] A 16-mer peptide derived from the coiled-coil forming JXD was initially conjugated to cell permeable peptide sequence TAT to produce a reduction in cell viability (EC_{50} 12.6 μM) corresponding to EGFR activity inhibition. Peptide stapling can also be used on native peptide sequences to improve peptide folding towards an α-helix.[20] As the EGFR peptide alone had no inhibitory activity, hydrocarbon stapling was investigated to improve native sequence permeability, creating a constrained peptide with a 10-fold improvement in potency over the TAT conjugated variant.[14,19]

2.1.1.4 Peptide inhibitors of the Nrf2/MafG coiled-coil

The Nrf2 leucine zipper has also been used to derive a stapled peptide to disrupt Nrf2/MafG coiled-coil formation. As Nrf2 is intrinsically disordered in the absence of a binding partner, linear peptides contained no helical secondary structure by circular

dichroism. Hydrocarbon stapling was employed to improve peptide helicity and resulted in a 16-mer peptide that bound to MafG and significantly decreased the Nrf2/DNA binding response by luciferase reporter assay in HEK-293 cells.[20] Deriving peptide sequences from the Nrf2 leucine zipper benefits from the principle that Nrf2 leucine zippers cannot self-assemble, preventing the peptide from forming coiled-coil interaction with itself. However, choosing to design peptides selective for MafG must also consider that the peptide may cause inhibition of MafG to other transcription factors, reducing selectivity for the Nrf2 transcription pathway.[21]

Overall peptides have provided a valuable tool for understanding coiled-coil interactions. They are beginning to be applied to disrupt these interactions in a therapeutic setting, and as a tool for designing probes for small molecule screening in drug discovery for this challenging type of interaction.

Rational peptide design offers an accessible approach to delivering peptide inhibitors of coiled-coil interactions and computational methods are emerging for the enhancement of sequence selectivity design towards more potent peptide binding partners of leucine zipper proteins. Despite their crucial role in mediating bHLH and bZIP transcription factor activity, examples of successful coiled-coil inhibition are limited. Continued research into the disruption of therapeutically relevant coiled-coils is needed to support drug discovery into the interaction.

2.2 Chapter Aims

This chapter describes the design and synthesis of peptides derived from the MafG leucine zipper, to afford an inhibitor of the Nrf2/MafG protein-protein interaction. In the absence of a crystal structure at the beginning of this research, a homology model of the Nrf2/MafG complex bound to DNA was generated using AlphaFold2 to identify MafG residues involved in coiled-coil formation. Following the report of a crystal structure for the Nrf2/MafG/DNA complex in 2022, the accuracy of AlphaFold models was examined.

Utilising knowledge of coiled-coil theory, sequence truncation was explored to identify an optimal peptide sequence for disrupting the coiled-coil interaction of Nrf2/MafG. Beyond peptide synthesis, this chapter explores the development of biophysical assays to investigate Nrf2 inhibition, under the hypothesis that peptides derived from the MafG leucine zipper can bind to Nrf2 and disrupt the Nrf2/MafG PPI.

2.3 Results and Discussion

2.3.1 Investigation of the Nrf2/MafG coiled-coil by structural modelling

In recent years, a number of machine learning programs have been developed to predict three-dimensional protein structures and multi-protein complexes.[22–25] Of these tools, AlphaFold in particular has become widely employed in structural biology research, benefitting from a large training data set of experimentally determined structures from the Protein Data Bank (PDB).

AlphaFold benefits from statistical estimations of each models confidence in structural predictions, with evidence of near experimental accuracy being reported in some models.[26,27] AlphaFold uses a predicted local distance difference test (pLDDT) to estimate a confidence metric at a residue level on a scale of 0-100 (**Table 2-1**) with >90 indicating high confidence through to < 50 suggesting low confidence in the model produced, the latter often associated with intrinsically disordered protein regions.

Table 2-1 *Summary of pLDDT scoring towards a models estimated confidence.*[28]

Model Confidence	pLDDT Score	Comment
Very High	**> 90**	Suitable for atomistic experiments.
Confident	**70-90**	Reliable backbone placement, variable side chain positioning.
Low	**50-70**	Use model with caution.
Very Low	**<50**	Unreliable and poor interpretation of the protein structure

Upon the release of AlphaFold3, it is now possible to measure two intrinsic model accuracy estimates of how likely the overall folding of a protein is correct.[29] These include a predicted template modelling (pTM) and inter-chain pTM (ipTM) score. ipTM scoring is particularly useful for predicting multi-protein complex folding to evaluate the confidence in a protein-protein interaction. [30] A value of < 0.6 is assumed to be incorrect, 0.6-0.8 as likely to have sections of accurate folding, however will also contain errors, and a value of 0.8 predicts that the model is highly accurate in the prediction.[31]

Comparing experimentally resolved models to machine learning predictions can be achieved through comparison of the root mean squared deviation (RMSD) a metric used to quantify the difference between protein structures relative to the atomic positioning of each residue. A lower RMSD value correlates to higher structural similarity of structural models, measured in Angstroms (Å). A value of less than 2 Å would indicate that a predicted model is a close match to an experimentally determined structure.[32]

The absense of rationally designed Nrf2 inhibitors exists largely due to the lack of structural information on the interaction.[33] At the beginning of this research, a crystal structure of the full Neh1 domain in Nrf2 had not been solved experimentally. As such, AlphaFold2 was used to prepare a homology model of the protein-protein interaction between Nrf2 and MafG. AlphaFold uses previously deposited resources in the PDB to help inform the predictions of how a protein will interact.[34] Confidence in the models produced of Nrf2/MafG can be found in the pre-existing crystal structures of other leucine zipper proteins, including the MafG homodimeric ternary structure (PDB: 3A5T), perhaps explaining the "very high" confidence scoring of the Nrf2/MafG leucine zipper, and ARE DNA interaction (**Figure 2-1**).[35–40]

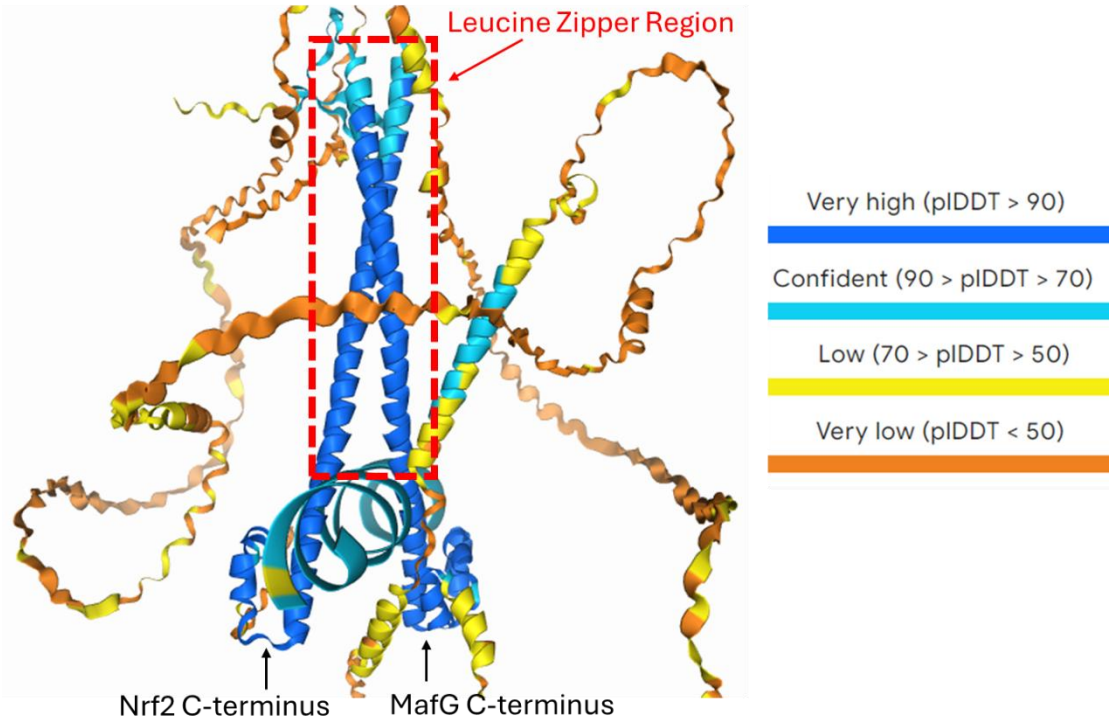


Figure 2-1 *AlphaFold3 Prediction model of full length $Nrf2_{1-605}$, $MafG_{1-162}$ and 25 base pair ARE DNA. Coloured based on model confidence very high (blue) confident (cyan) low (yellow) very low (orange).*

The subsequent report of a crystal structure of the Nrf2/MafG/ARE complex and the release of AlphaFold3 allows for a comparison to be made on the accuracy of the structural predictions.[29,41] The models demonstrate similarity in the backbone positions for interacting residues of the DNA binding domain with variability increasing up into the leucine zipper (**Figure 2-2**).

To quantify accuracy of the AlphaFold models of Nrf2 Neh1 in complex with MafG and DNA, RMSD was calculated using PyMOL against the published crystal structure (PDB:

7X5F). RMSD calculations compared the overlapping structures α-carbon (Cα) atoms after alignment. The RMSD score calculated compared to the crystal structure for Alphafold2 would suggest the model was similar to the crystal structure but was not accurate enough to confidently use the model for atomistic level experiments. (**Table 2-2**). The AlphaFold3 model brings the RMSD score closer to experimentally resolved structure of the protein-protein interaction. Whilst calculating the RSMD score from Cα does not provide information on the position of residue side chains, for the purpose of confirming general residue positioning within a coiled-coil, AlphaFold provides an appealing tool for sequence identification for PPI inhibition. Despite the advancements of machine learning programs, experimentally resolved structures remain the gold standard as they provide the most reliable information on the molecular interactions between proteins.

Table 2-2 *RMSD scores of Alphafold models in alignment with the crystal structure of bZIP protein domains of Nrf2/MafG (PDB:7X5F).*

Prediction Model	RMSD Score
AlphaFold2 (2021)	2.55 Å
AlphaFold3 (2024)	1.14 Å

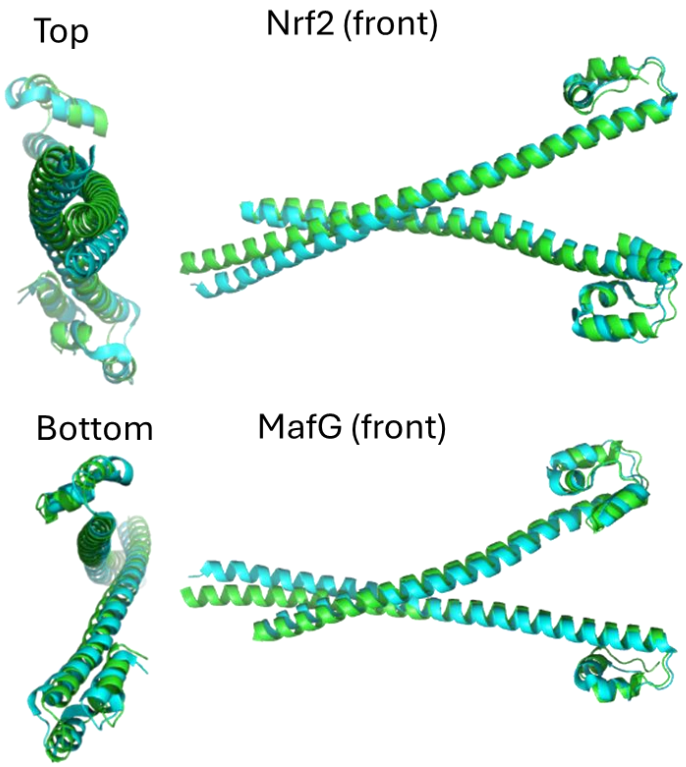


Figure 2-2 *Structural alignment of Nrf2/MafG AlphaFold3 model (cyan) against experimental (PDB: 7X5F) (green).*

2.3.2 MafG peptide synthesis

The coiled-coil interaction between Nrf2 and MafG stabilises the protein-DNA interaction with Nrf2 promoter sequences in DNA.[41] As discussed in **Chapter 1.5.1**, the

leucine zipper of $MafG_{76\text{-}110}$, forms a coiled-coil interaction with $Nrf2_{522\text{-}562}$. Notably, $_{76}$KEELEKQKAELQQEVEKLASENASMKLE$_{104}$ of MafG interacts with Nrf2 without masking a nuclear exportation sequence $_{553}$LKKQLSTYL$_{562}$ (**Figure 2-3**).[42–44] It is anticipated that peptides designed to compete with MafG for Nrf2 coiled-coil formation should leave the exportation sequence exposed to promote Nrf2 translocation into the cytoplasm. As such, peptides derived from $MafG_{76\text{-}104}$ were prepared by solid phase peptide synthesis (SPPS).

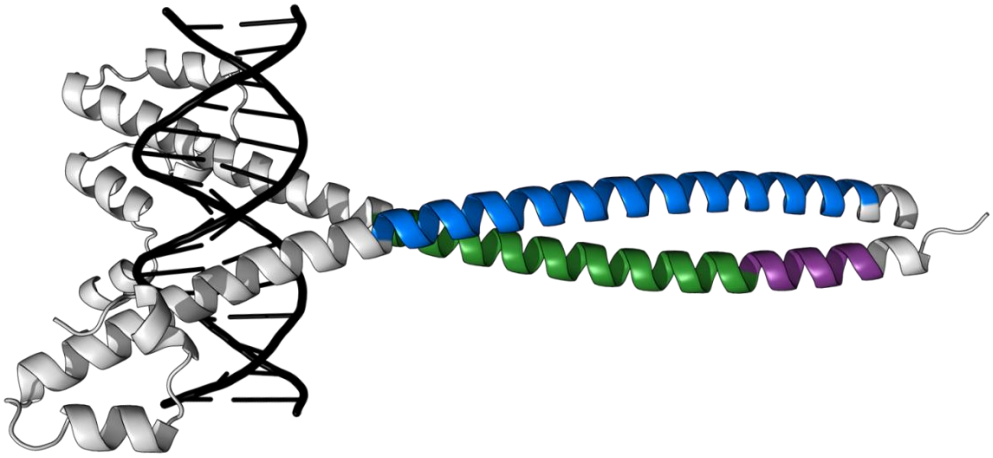


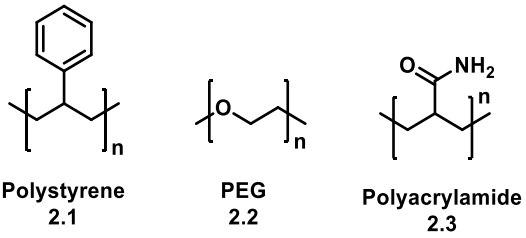
Figure 2-3 *Nrf2/MafG/ARE homology model generated using AlphaFold2. MafG leucine zipper (Blue) Nrf2 leucine zipper (green) Nrf2 nuclear localisation sequence (purple).*

2.3.2.1 Solid Phase Peptide Synthesis Resins

Peptide synthesis was first realised through a solution phase chemical synthesis of glycine dipeptide structures in the early 1900s.[45,46] Whilst synthesis can now be achieved through a range of biological and chemical techniques, the most commonly applied method is solid phase peptide synthesis (SPPS).[47–49] SPPS can be automated and reagents can be used in excess to push coupling reactions to completion. The use of a solid support significantly reduces purification requirements, allowing for reaction cycles to take place in a singular vessel.[50] Increasingly the procedure is becoming more efficient with less waste through innovative methods in industry.[51]

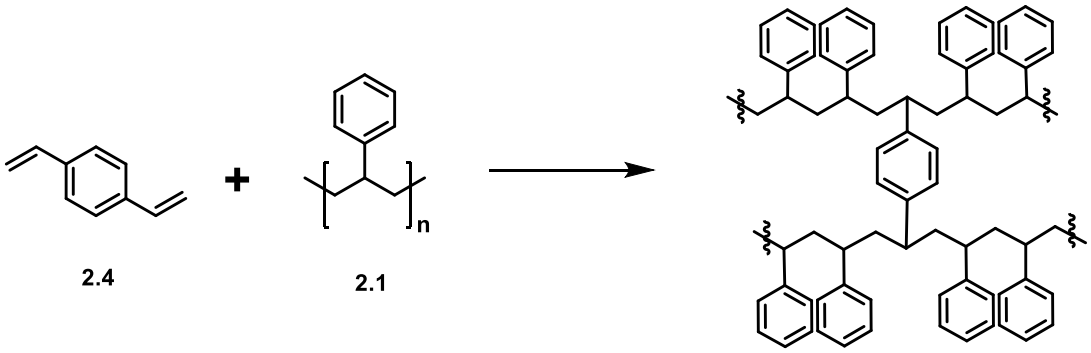
Synthesis on the solid phase utilises a polymer matrix as the anchor for building a chain of amino acids (**Figure 2-4**). Common polymers used include polystyrene (**2.1**), poly-ethylene glycol (PEG) (**2.2**), polyacrylamide (**2.3**) or can be co-polymer in composition (**Figure 2-5**). Crosslinking between chains of a polymer (**Figure 2-5**) is achieved using the

addition of 1-2% divinylbenzene (DVB) (**2.4**) to create a porous but insoluble resin support (**2.5**).



Polystyrene 2.1 | PEG 2.2 | Polyacrylamide 2.3

Figure 2-4 *Polymers used in resin supports.*



2.4 + 2.1 → 2.5

Figure 2-5 *DVB crosslinking.*

Resin for SPPS is functionalised with a handle to anchor the synthesis of a peptide to the resin surface (**Figure 2-6**). As peptide segments of proteins natively observe a carboxamide at the C-terminus, it is desirable for the peptide to end in -$CONH_2$ for stability against degradation.[52] Rink amide (**2.6**) functionalised resin was the first to support carboxamide formation at the N-terminus, utilising a trialkoxybenzylhydryl linker.[53,54] Sieber **(2.7)** and Ramage **(2.8)** linker resins have also been used for the production of a carboxamide (-$CONH_2$) at the C-terminus of the peptide and offer new variability for improving synthesis conditions through linker solubility and flexibility.

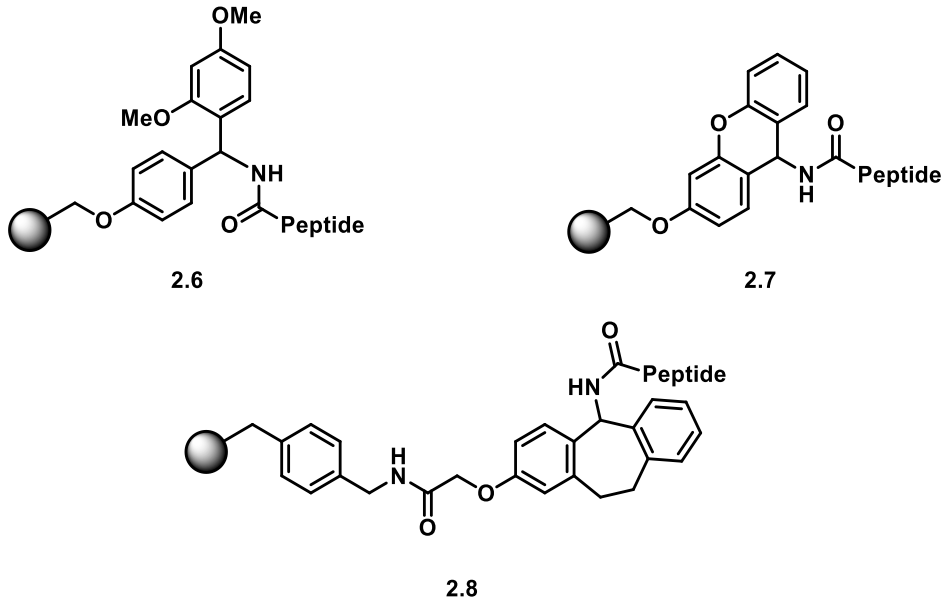


Figure 2-6 *Examples of resin linkers for solid phase peptide synthesis. Rink Amide (**2.6**) Seiber (**2.7**) and Ramage (**2.8**).*

2.3.2.2 Amino Acid Protecting Groups for Solid Phase Peptide Synthesis

Solid supports can be used for acid or base labile synthesis, dependent on the protecting groups used on the α-carbon of the amino acid units (**Figure 2-7**). Acid labile synthesis uses tert-butyloxycarbonyl (Boc) (**2.11**) protection of the terminal amine functional group, whilst base labile uses 9-fluorenylmethyloxycarbonyl (Fmoc) (**2.9)**. Fmoc synthesis is most used, as it avoids the use of harsh hydrofluoric acid for the final cleavage of the peptide from the resin. Fmoc-SPPS offers an attractive alternative to solution phase synthesis as reactions can be purified more efficiently through resin washing between reaction steps.[55]

Fmoc-SPPS utilises acid-labile protecting groups on reactive amino acid side chains to prevent side product formation during the synthesis. Tert-butyl (**2.10**) is useful in the protection of hydroxy and carboxy functional groups, trityl (**2.12**) can be used for the protection of the thiol group in cysteine and amines in histidine residues, and 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl (Pbf, **2.13**) is used for the protection of arginine**.**

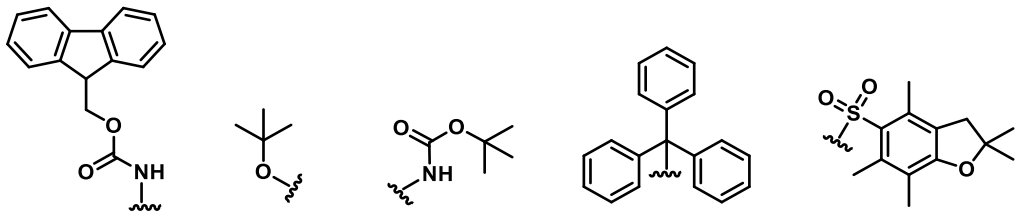
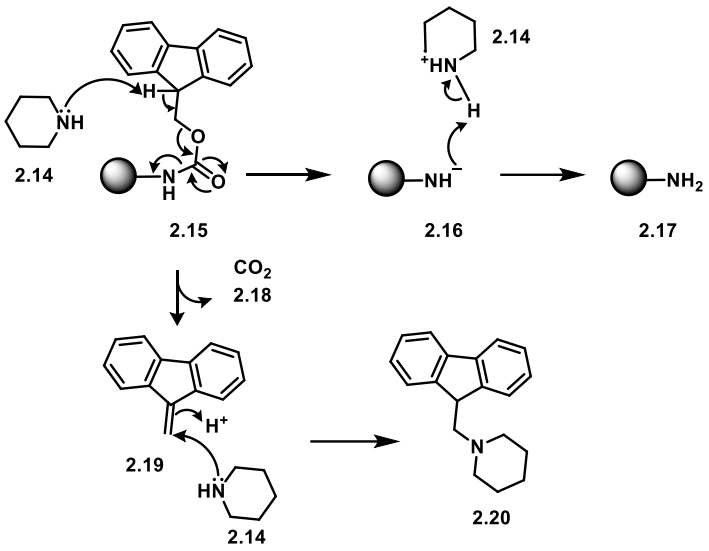


Figure 2-7 *SPPS protecting groups, Fmoc (**2.9**) for backbone amine protection,* t*Bu (**2.10**), Boc (**2.11**), Trt (**2.12**) and Pbf (**2.13**) protecting groups.*

Protecting groups are important for amino acids containing carboxamides as these residues have low solubility and are prone to aggregation on resin, through hydrogen bond bridges between parallel chains during synthesis.[56] Acylation and racemisation of amino acids can also be minimised through protection of histidine and arginine residues. Un-protected asparagine can form cyano-alanine during coupling reactions.[57] Without Trt protection of glutamine, Fmoc deprotection can lead to pyroglutamate formation.[58] Additionally some coupling reagents can react with un-protected glutamine or asparagine residues to form byproducts, which prevents further chain growth.[59]

2.3.2.3 Deprotection of amino acids in Fmoc solid phase peptide synthesis

Removal of the Fmoc protecting group to reveal a free amine is required to begin synthesis by Fmoc SPPS (**Scheme 2-1**). Fmoc deprotection is achieved under basic conditions, commonly using a cyclic secondary amine such as piperidine (**2.14**) deprotonating the acidic β-carbon (**2.15**) of the fluorenyl group.[60] Removal of this proton lead to a β-elimination reaction to produce a free amine (**2.17**) and CO_2 (**2.18**). A highly reactive dibenzofulvene intermediate (**2.19**) is also formed, scavenged by piperidine (**2.14**) allowing the formation of a more stable adduct (**2.20**).[61]



Scheme 2-1 *Basic deprotection of an Fmoc protected amino acid unit (**2.15**).*

Following deprotection, amide coupling can be performed using an excess of amino acid, an activating reagent and a base. Commonly used coupling activators in SPPS (**Figure 2-8)** include, hexafluorophosphate benzotriazole tetramethyl uronium (HBTU) (**2.23**), hydroxy benzotriazole (HOBt) (**2.22**) and Oxyma pure (**2.24**), along with basic agents N,N-diisopropylethylamine (DIPEA) (**2.25**) or diisopropylcarbodiimide (DIC) (**2.21**).

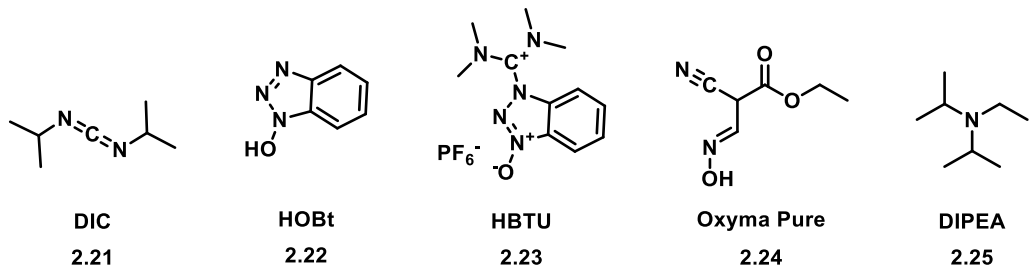
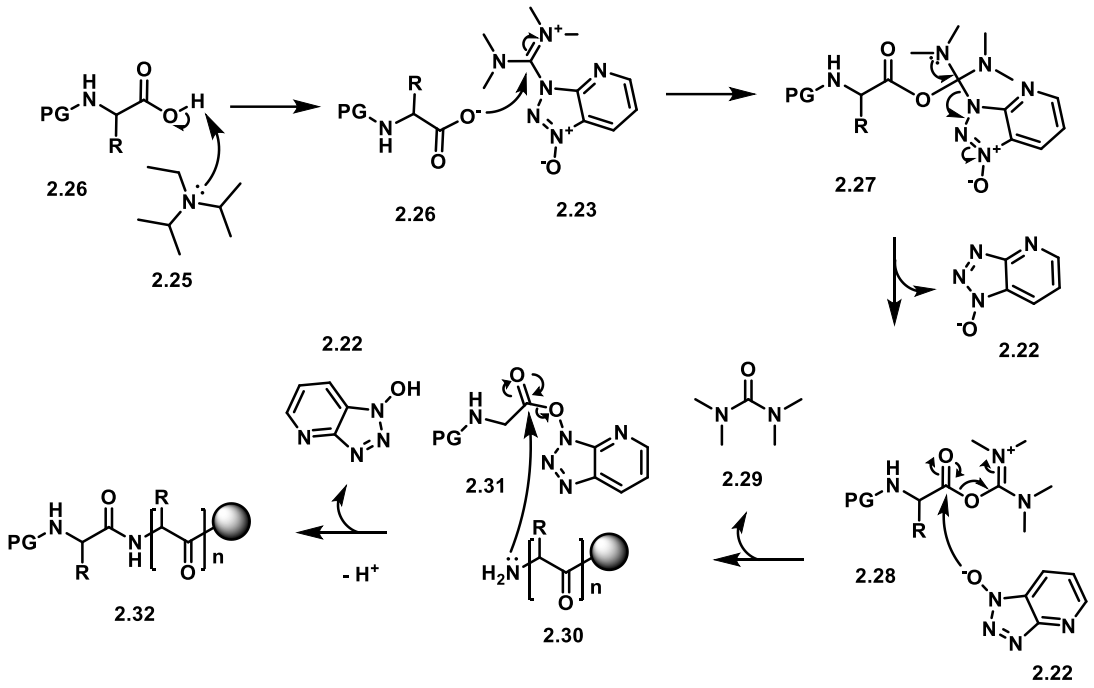


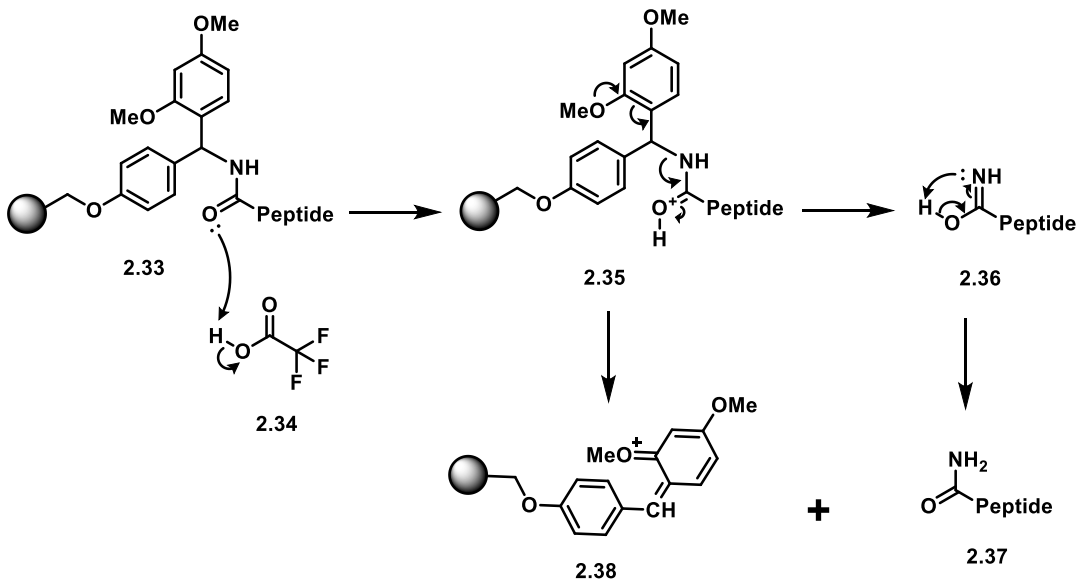
Figure 2-8 *Coupling reagents commonly used in solid phase peptide synthesis.*

Phosphonium and uronium based reagents, such as HBTU are popular activation reagents for peptide synthesis due to their stability and rapid efficiency in coupling reactions with minimal racemisation, which is further suppressed by the addition of HOBt.[62] As shown in **Scheme 2-2,** HBTU/HOBt coupling requires a base, such as DIPEA (**2.25**), to deprotonate the carboxylic acid on the incoming amino acid (**2.26**). The deprotonated unit then reacts with HBTU (**2.23**) to form a labile *O*-acylisourea (**2.28**) and HOBt (**2.22**). HOBt then reacts to form an activated ester (**2.31**) which undergoes nucleophilic attack by the free amine of the resin bound peptide (**2.30**) to form an amide bond (**2.32**) whilst reforming HOBt (**2.22**).



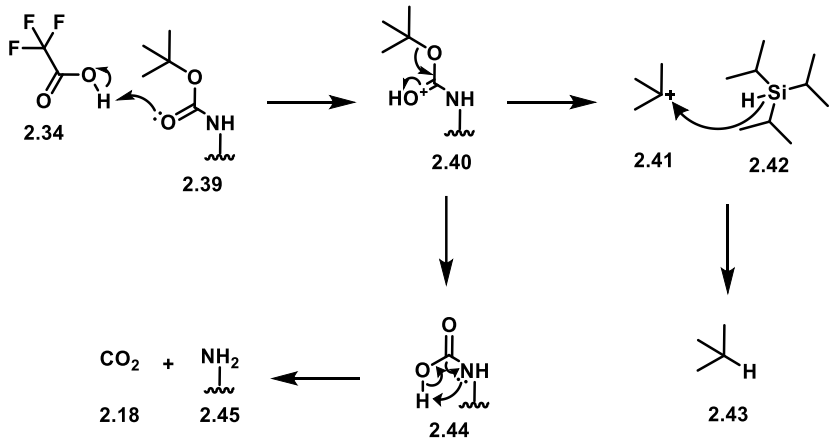
Scheme 2-2 *Coupling mechanism of HBTU (**2.23**), HOBt (**2.22**) and DIPEA (**2.25**).*

SPPS subsequently proceeds through a series of deprotection and coupling reactions until the desired peptide sequence is complete. Finally, cleavage can be achieved with acid to remove the peptide chain from resin and cleave the acid-labile protecting groups. Trifluoracetic acid (TFA) is used in Fmoc SPPS to cleave the peptide from the acid labile linker at the C-terminus of the sequence. TFA cleavage is an efficient process that also removes acid labile side-chain protecting groups (**Scheme 2-3**). The amide functional group connecting the peptide to the rink amide functionalised resin (**2.33**) is protonated by TFA (**2.34**), followed by an electron cascade (**2.35**) from the methoxy group of the rink amide allowing for amide cleavage from the resin (**2.36**).



Scheme 2-3 *Fmoc SPPS C-terminal cleavage from solid support by TFA* (**2.34**).

Scavengers are added to the cleavage reaction to prevent unwanted side reactions with the cleaved peptide (**Scheme 2-4**). After TFA cleavage (**2.39)** water or triisopropylsilane (TIPS) (**2.55**) can scavenge t-butyl cations (**2.41)**. TIPS is a hydride donor capable of reducing side products during the removal of protecting groups (**2.40**).[58,63] 2,2'-(Ethylenedioxy)diethanethiol (DODT) is also a required scavenger in the presence of thiol containing residues to prevent reactive carbonium or sulfonium ion intermediates from reacting with the peptide.[64]

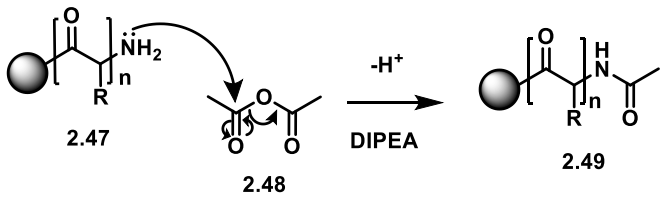


Scheme 2-4 Cleavage of Boc protecting group and subsequent reduction by TIPS (**2.42**).

2.3.2.4 Synthesis of 28-mer MafG peptide

Fmoc SPPS was used as described to synthesise the 28-mer MafG derived peptide, Ac-KEELEKQKAELQQEVEKLASENASMKLE-NH_2, **(2.46)**, using NovaPEG Rink amide resin (0.6 mmol/g loading) a PEG based resin with a mesh* of 35-100, at room temperature. Whilst it is no longer commercially available, NovaPEG was an affordable C-terminal amide forming resin benefiting from high hydrophilicity for the synthesis of hydrophobic sequences.[65,66]

Following sequence synthesis, the peptide was acetylated to afford greater peptide stability by reducing N-terminal degradation by peptidases.[67,68] Acetylation can also support the formation of helical secondary structure in peptide and proteins.[69,70] Selective acylation (**Scheme 2-5)** at the N terminus (**2.47**) can be achieved through the addition of DIPEA (**2.25**) to provide the rapid addition of acetyl group through acetic anhydride **(2.48)**.



Scheme 2-5 *Acetylation reaction of N-terminal peptide (**2.47**) with acetic anhydride (**2.48**).*

After cleavage of the 28-mer peptide, **(2.46),** from resin, the crude mixture was analysed by reverse phase HPLC. (**Figure 2-9,** blue trace) This revealed poor resolution from the synthesis with broad peaks co-eluting multiple impurities that proved impossible to remove. It was considered that the high content of hydrophobic residues within the sequence may be causing sequence aggregation.

2.3.2.5 Resin Screening

Reducing the resin loading capacity can improve synthesis by lowering the number of peptide chains extending off a singular resin bead, increasing the distance between growing peptide chains and reducing aggregation.[71] Consequently a resin screen was trialled to assess resin influence on sequence solubility and aggregation. TentaGel S RAM and Protide LL utilise a PS-PEG crosslinked backbone that supports high resin swelling and improved reagent diffusion rates within the resin but have different loading capacities TentaGel S RAM (0.24 mmol/g loading, 100-200 mesh) and Protide LL (0.15 mmol/g loading, 50-100 mesh).[72] All resins were trialled with automated parallel

* Mesh is a measure of particle size inversely related to the particle's diameter in μm.

synthesis, acetylation and cleavage followed by analysis via RP-HPLC (**Figure 2-9**). HPLC traces revealed limited improvements in purity, as broad heterogenous peaks were eluted, representing poor candidates for further purification. Protide LL (**Figure 2-9**, red trace) demonstrated the greatest improvement in crude purity, however, no mass was found by matrix associated laser desorption ionisation time of flight mass spectroscopy (MALDI-TOF MS). MALDI analysis of the crude peptide synthesised on Tentagel S RAM found the expected mass, suggesting the peptide may appear in the broad peak between 10-11 minutes. Isolation of the peptide proved impossible to separate from impurities (**Figure 2-9**, green trace**).**

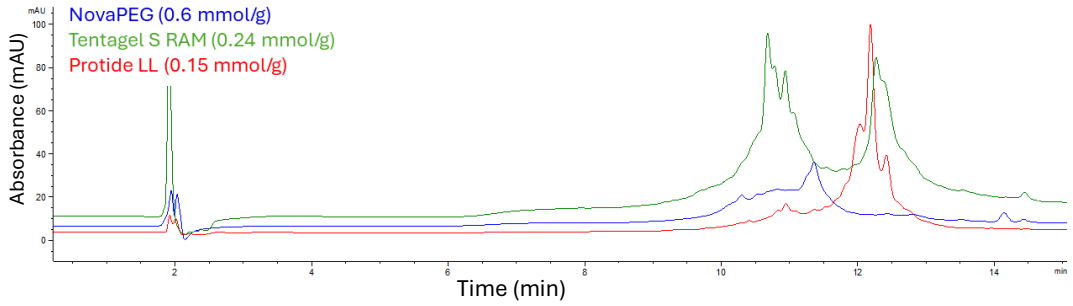
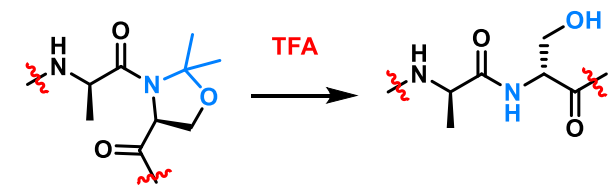


Figure 2-9 *HPLC Trace (214 nm, 5-95% in MeCN in H_2O + 0.05% TFA over 15 min),* ***2.46*** *synthesised using NovaPEG Rink Amide (Blue) Protide LL (Red) Tentagel S RAM (Green).*

With reduced resin loading proving insufficient to purify the peptide, internal sequence modifications were explored. The sequence was considered a good candidate for the use of a pseudoproline moiety to reduce inter-chain aggregation.[73]

2.3.2.6 Pseudoproline incorporation

The addition of pseudoproline dipeptides within SPPS can induce turns within the amide backbone creating kinks in linear chains, reducing the likelihood of interstrand interactions. First reported by Mutter et al, the use of dipeptides converts proline-like structures **(2.50)** into a serine **(2.51)**, threonine or cysteine containing sequences upon TFA cleavage **(Scheme 2-8)**.[74] Pseudoprolines prevent β-sheet formation of hydrophobic sequences during chain elongation by constraining the amide backbone, reducing self-association and improving peptide solvation on resin.[75]



2.50

Ala-Ser[PSI(Me,Me)Pro]

2.51

Ala-Ser

Scheme 2-6 *Pseudoproline dipeptide Ala-Ser release in the presence of TFA.*

The coupling following the pseudoproline unit is hindered by the secondary amine and the adjacent dimethyl unit, consequently a pre-built a Ala-Ser dipeptide pseudoproline unit was used.[76] There are two positions within the sequence that Ala-Ser pseudoprolines could be substituted. Reports suggest it is most effective to introduce pseudoprolines at regular intervals throughout the sequence.[77] As there are two Ala-Ser motifs were within two residues of one another, it was decided to introduce a singular pseudoproline in the sequence at positions 9-10, as aggregation is most likely to occur after the sixth residue.[78] Addition of the pseudoproline moiety produced multiple broad peaks on analytical HPLC (**Figure 2-10,** red trace) and purification of any individual major peak was not achievable. Perhaps indicative of the oxazolidine moiety producing steric hindrance, preventing efficient coupling.[79] Consequently, the use of pseudoproline to improve aggregation was not explored further.

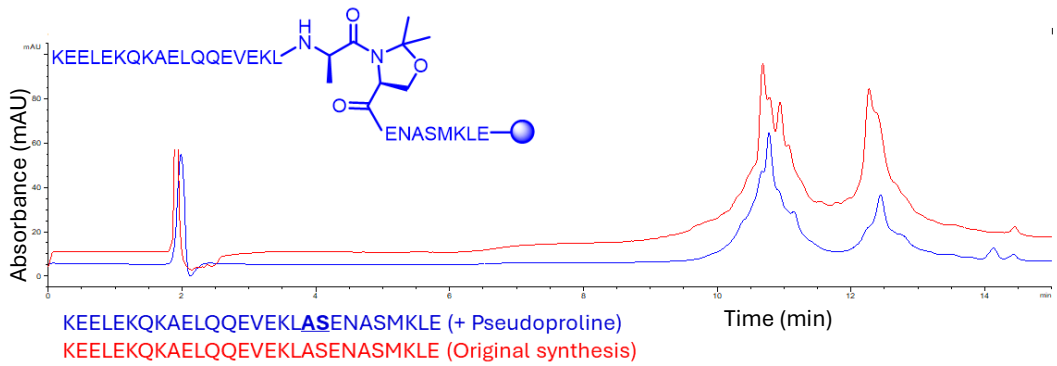


Figure 2-10 *HPLC traces of crude* ***2.46*** *using an Ala-Ser pseudoproline (**blue**) against linear synthesis on Tentagel S RAM (**Red**). Traces are offset by 10% in absorbance from one another.*

2.3.2.7 Acetyl Capping

Aggregation during peptide synthesis is driven by inter or intra-molecular hydrogen bonding, which can reduce the accessibility of the terminal amine for peptide coupling. Incomplete coupling reactions can produce peptide impurities with similar physicochemical properties to the desired product, resulting in their co-elution with the target peptide by RP-HPLC. As a result, aggregation can make separation of impurities from the desired product challenging and low yielding. Consequently, acetyl capping of unreacted N-terminal amines was introduced between coupling reactions to limit incorrect sequence assembly (**Figure 2-11**).[80,81]

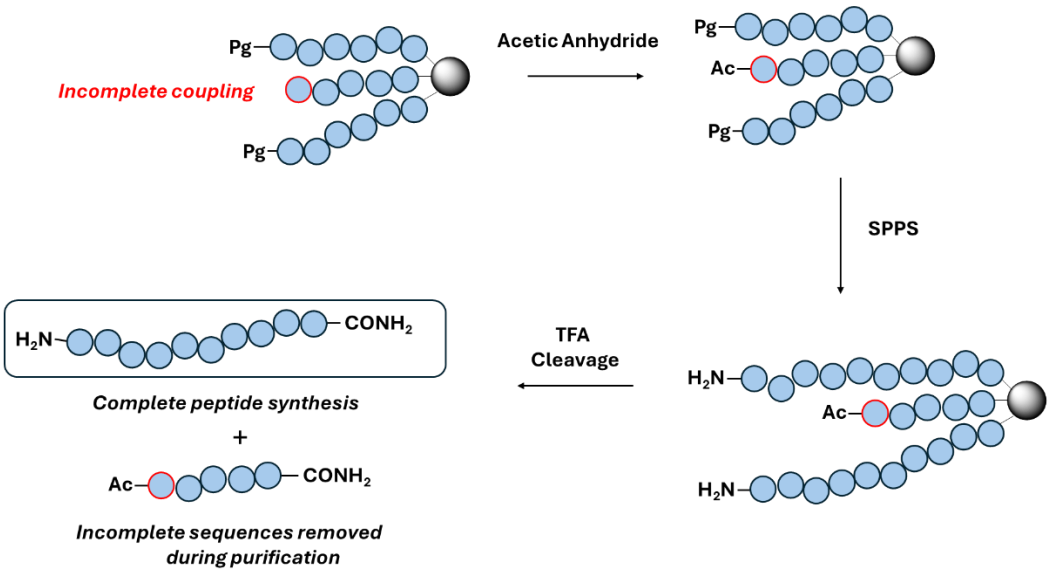


Figure 2-11 *Process of acetyl capping to reduce the synthesis of incorrect peptide sequences.*

Synthesis of **2.46** was investigated using microwave assisted peptide synthesis, employing temperatures of 75-90 °C for coupling reactions with DIC and Oxyma pure and deprotection steps with piperidine.[82] With each amino acid cycle taking under 5 minutes to complete, microwave supported reactions can decrease aggregation as short coupling times help reduce intermolecular interactions.[83] Reducing the time spent on synthesis can reduce potential degradation during synthesis.[84] Microwave peptide synthesis of **2.46** appeared to produce fewer impurities within the synthesis, yet retained large shoulders either side of the major product peak (***Figure 2-12***). Additionally, the expected mass could not be found by MALDI-TOF.

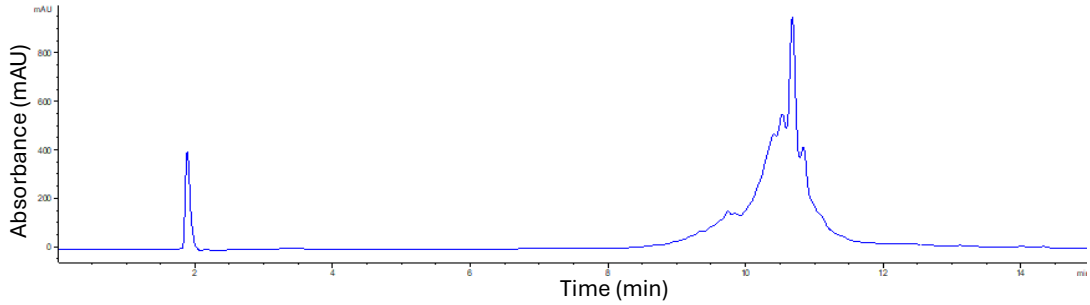


Figure 2-12 *crude HPLC trace of acetyl capping method on* ***2.46****.*

2.3.2.8 Chaotropic Salts

Chaotropic salts can be used to destabilise aggregation of the sequences by disrupting hydrogen bond formation.[85] Synthesis of **2.46** was trialled by the addition of 0.4 M lithium chloride in the dimethylformamide (DMF) used in all stages of the automated SPPS (**Figure 2-13**). After peptide acetylation, the resin was washed exhaustively in dichloromethane and methanol to remove any residual DMF and lithium chloride.

Following cleavage, crude analysis by RP-HPLC revealed a major peak, separated from other impurities, which allowed for purification by preparative HPLC in 5-95% acetonitrile over 15 minutes. The corresponding mass of **2.46** was observed at the expected *m/z* of 3300.7 $[M+H]^+$.

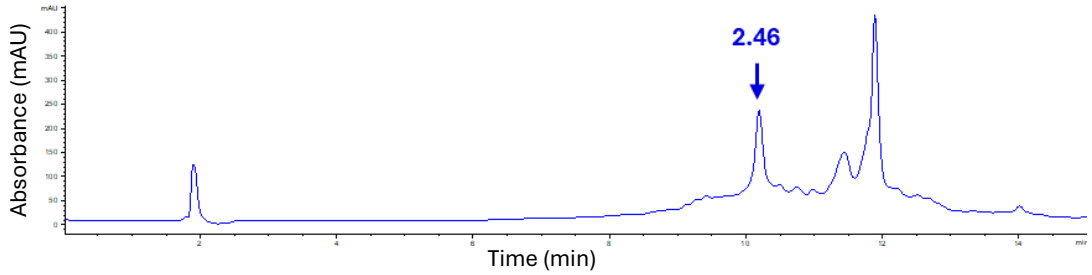


Figure 2-13 *HPLC traces of crude* ***2.46*** *using 0.4 M LiCl in DMF as a solvent for SPPS.*

2.3.2.9 Synthesis of 14 and 21-mer MafG peptides

Utilising the coiled-coil theory discussed in **Chapter 1.2.3** sequence truncation could reveal an efficient sequence length for Nrf2 inhibition. In order to explore this, shorter peptides were designed in which heptadic motifs were conserved to retain helicity, essential for the coiled-coil interaction being imitated (**Figure 2-14**).

$MafG_{76-104}$ KEELEKQ KAELQQE VEKLASE NASMKLE
abcdefg abcdefg abcdefg abcdefg

MafG Peptides

2.46: Ac-KEELEKQKAELQQEVEKLASENASMKLE-NH_2
2.52: Ac-KAELQQEVEKLASENASMKLE-NH_2
2.53: Ac-KEELEKQKAELQQEVEKLASE-NH_2
2.54: Ac-VEKLASENASMKLE-NH_2
2.55: Ac-KAELQQEVEKLASE-NH_2
2.56: Ac-KEELEKQKAELQQE-NH_2

Figure 2-14 *Design of truncated MafG leucine zipper derived peptides.*

The synthesis of 14-mer and 21-mer peptides was carried out, as summarised in **Table 2-3**. Whilst **2.56**, **2.55** and **2.53** were achieved using microwave assisted peptide synthesis, **2.54** and **2.52** required the addition of 0.4 M LiCl and were made at room temperature using Tentagel S RAM.

The synthesis of **2.52** was later revisited to optimise the yields for further assessment of the peptide. From the perspective of purification, the best separation of **2.52** from impurities was found on Protide LL resin, the lower loading potential of the resin presumably prevents excessive sequence interactions that can produce incomplete

coupling. Demonstrating the careful combination of chaotropic salts and low resin loading can improve crude purity of SPPS.

Ongoing challenges in achieving high yields of sequences containing methionine were encountered, with methionine sulfoxide conversion commonly observed. **Figure 2-15** highlights two peaks after reverse phase purification. Oxidation of **2.52** caused a shift in the retention time of the peptide. Semi-preparative purification achieved separation of the two peaks, with the product at 10.6 min observing an *m/z* corresponding to $[M+17]^+$ (**2.57**) compared to **2.52** at 10.7 min.

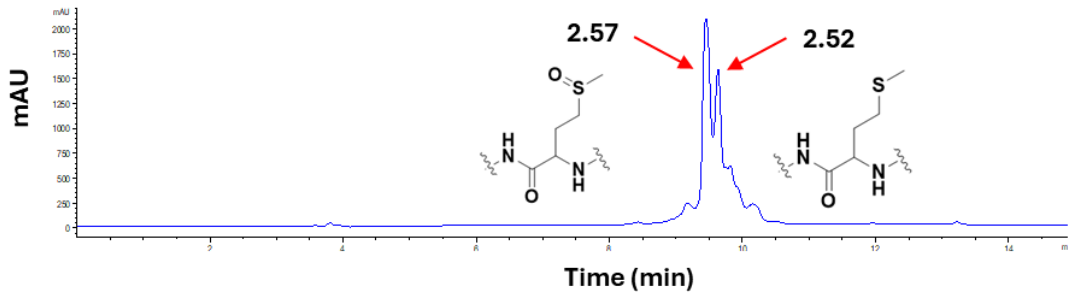


Figure 2-15 *Crude analytical HPLC traces of* ***2.52*** *at 214 nm, demonstrating closely related peaks produced by methionine oxidation to* ***2.57****.*

Methionine can be substituted for norleucine to provide similar structural properties without risk of producing a sulfoxide side product. Synthesis of this peptide (**2.58**) was achieved with 0.4 M LiCl additive in the DMF throughout the automated SPPS. The expected *m/z* of 2398 $[M+H]^+$ was observed.

Table 2-3 *Peptides synthesised based on heptadic motifs in the MafG leucine zipper.*

#	Sequence	HPLC T_R (min)	MALDI $[M+H]^+$	Yield %
2.56	Ac-KEELEKQKAELQQE-NH_2	8.9	1771.1	18.9
2.55	Ac-KAELQQEVEKLASE-NH_2	10.9	1642.2	17.1
2.54	Ac-VEKLASENASMKLE-NH_2	10.8	1591.2	9.2
2.53	Ac-KEELEKQKAELQQEVEKLASE-NH_2	5.6	2527.0	10.0
2.52	Ac-KAELQQEVEKLASENASMKLE-NH_2	10.7	2416.2	4.1
2.57	Ac-KAELQQEVEKLASENASM**(O)**KLE-NH_2	10.6	2432.9	2.0
2.58	Ac-KAELQQEVEKLASENAS(**Nle)**KLE-NH_2	8.2	2398.6	2.5
2.46	Ac-KEELEKQKAELQQEVEKLASENASMKLE-NH_2	10.6	3300.7	2.3

2.3.3 Secondary structure characterisation of MafG peptides

As discussed in **Chapter 1**, secondary structure is important for the formation of protein-protein interactions. Sequences derived from the MafG leucine zipper require an α-helical secondary structure to form a coiled-coil interaction. It is necessary to study the secondary structure of the peptide sequences, to determine if they convey the

desired secondary structure to support binding. Circular dichroism (CD) is an absorption spectroscopy technique measuring the absorption of circularly polarised light in a sample in the far-UV region of proteins, (260-180 nm). A piezoelectric transducer fused with quartz, referred to as a photoelastic modulator, induces an optical rotation of linearly polarised light (**Figure 2-16**). Due to the chiral nature of the peptide backbone, the difference between left and right circularly polarised light can be measured by detecting changes in absorption of the circularly polarised light.[86,87]

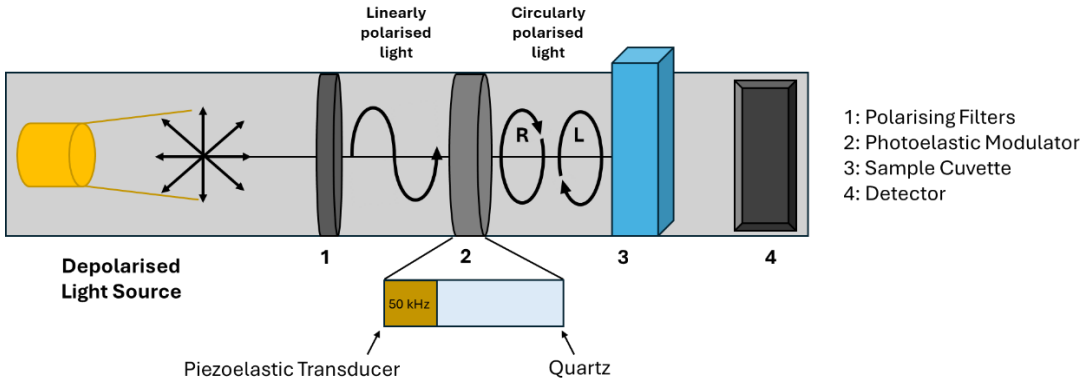


Figure 2-16 *Process of polarising light in circular dichroism.*

Conformations of β-sheets, α-helices, helical turns and even disordered structures such as random coils can be understood spectroscopically (**Figure 2-17**). α-helices in particular adopt a characteristic negative ellipsis at 222 and 208 nm and positive ellipticity at 193 nm (**Figure 2-17,** blue trace).[88] CD spectra can subsequently be used to predict the percentage of helical content within a peptide sequence using reference-based deconvolution methodologies such as CDSSTR, CONTIN or SELCON3.[89–91]

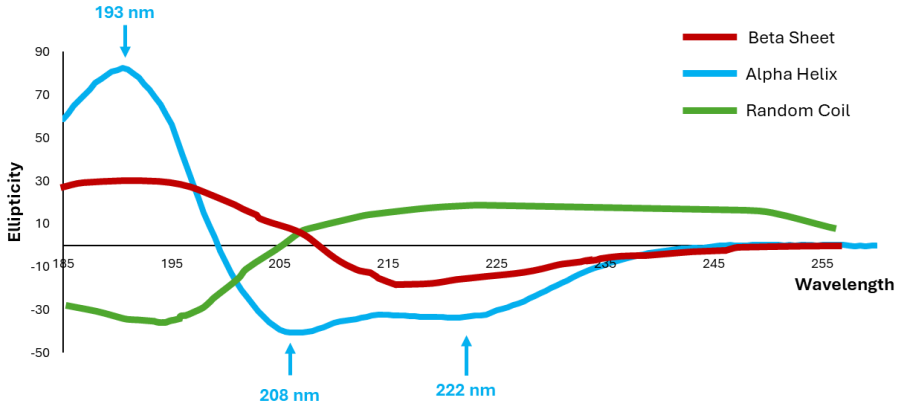


Figure 2-17 *Example spectrum recorded by circular dichroism demonstrating β-sheet (red), α-helix (cyan) and random coil (green).*

The peptides synthesised were anticipated to produce an α-helical secondary structure as generated by the Nrf2-MafG coiled-coil interaction.[41] CD was used to investigate the secondary structure over wavelengths 180-260 nm. Each peptide assessed produced

an α-helical spectrum with a maximum at 190-193 nm and a recognisable double minima at 208 and 222 nm.[92] Measurements were performed in 10 mM PBS at pH 7.4 with 50% Trifluoroethanol (TFE) as a cosolvent for stabilising secondary structure.[89,93] Peptides were measured at a concentration of 0.2 mg/mL. The results were then plotted with wavelength on the x-axis and ellipticity on the y-axis to visually observe α-helical forming spectra in all six peptides prepared (**Figure 2-18)**.

The results were then analysed in DichroWeb using the CDSSTR algorithm to quantify the helical content of the peptides, (**Table 2-4**).[91] DichroWeb utilises a data bank of reference CD spectra to compare the quality of input data. The estimated proportion of α-helical secondary structural elements exceeded 40% in each sample. The greatest α-helicity was found in the sequences of **2.53** (81%) and **2.52** (71%), both of which are 21 amino acids in length. The third greatest helicity came from **2.56** (60%) with 14 amino acids. Comparing 14 to 21 amino acids sequences, the helicity appears to increase with chain length, however the helical content is lower (48%) for the 28-mer **2.46**. Increasing the length of the peptide may create unfavourable entropy reducing the propensity for an α-helix.[94,95] The organisation of amino acids in a sequence of the MafG leucine zipper lends itself to a helical secondary structure. This supports the suggestion that these peptides could be capable of binding to Nrf2 through a coiled-coil interaction as they can independently form the α-helical structure necessary.

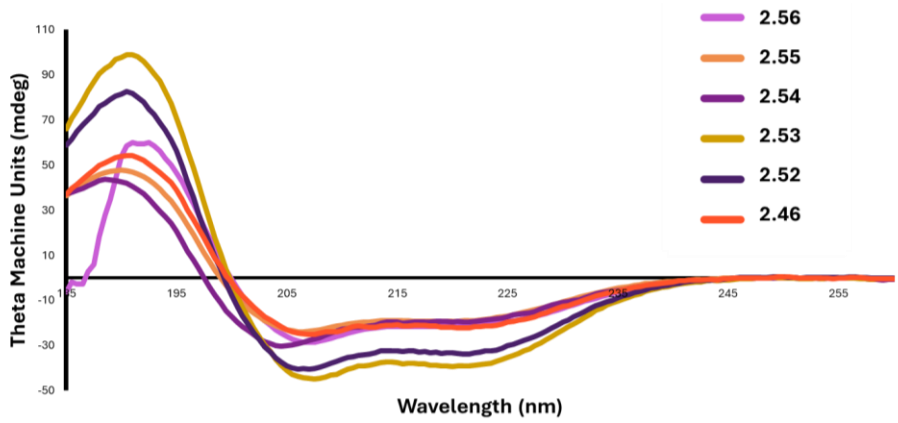


Figure 2-18 *Circular dichroism spectra of* ***2.46 and 2.52-2.56*** *at 0.2 mg/mL in 50/50 10 mM phosphate buffer at pH 7.4/TFE. Represented with mdeg (y) against wavelength (x).*

Table 2-4 *Experimentally determined helicity of* ***2.46 and 2.52-2.56*** *by circular dichroism, analysed by DichroWeb CDSSTR method.*

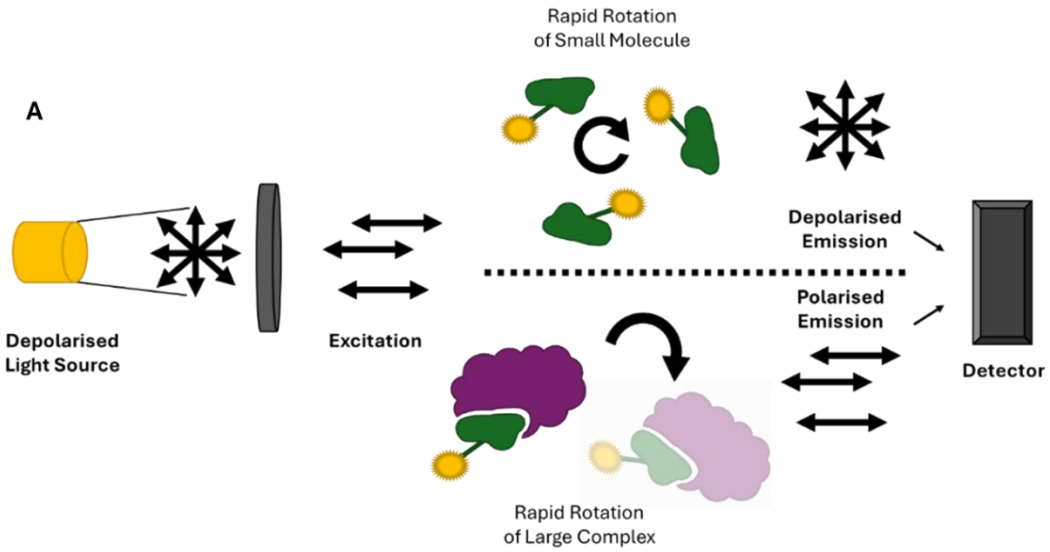
Peptide #	*Sequence*	*Helicity %*
2.56	Ac-KEELEKQKAELQQE-NH_2	60
2.55	Ac-KAELQQEVEKLASE-NH_2	45
2.54	Ac-VEKLASENASMKLE-NH_2	46
2.53	Ac-KEELEKQKAELQQEVEKLASE-NH_2	81
2.52	Ac-KAELQQEVEKLASENASMKLE-NH_2	71
2.46	Ac-KEELEKQKAELQQEVEKLASENASMKLE-NH_2	48

2.3.4 Investigation of inhibition by fluorescence polarisation

Disruption of two interacting biomolecules can be observed by fluorescence polarisation (FP). FP uses the inverse relationship between polarisation and fluorophore tumbling rate to quantify a binding interaction. Upon excitation of a fluorophore with polarised light, the resulting fluorescence produced is depolarised due to the fast rotation of the fluorophore in isolation (**Figure 2-19, A)**. In the presence of a binding partner, the rotation of the fluorophore is slowed resulting in the emission of polarised light, translating to an increased signal strength.[96] FP is reported as a ratio of emission intensity between vertical and horizontally polarised light (**Equation 2-1, Figure 2-19, B)**

The interaction between an analyte and a binding target can be measured by attaching a fluorescent tag to the smaller component within the system. The most commonly used fluorophore in FP is fluorescein (**Figure 2-19, C**), a conjugated system of aromatic rings that supports absorption at 498 nm and re-emission at 517 nm.[97] An example of fluorescein labelling for Nrf2/PPI inhibition has been reported using the H-DEETGEL-OH peptide inhibitor of an Nrf2/Keap1 binding. Affinity of the Nrf2 derived peptide was determined through a fluoroscein tagged derivative, measuring the polarisation signal of the peptide in the presence of the Keap1 protein.[98] Upon quantification of a binding curve for a ligand to a mid-point of approximately 65-80% bound can be used to prepare competition assays by FP to investigate the inhibitory activity of other compounds.[99]

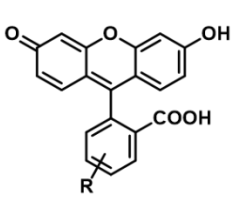
FP makes an excellent candidate for developing HTS assays for screening large numbers of compounds for an interaction, as the measurable sensitivity is quite high allowing for low quantities of reagents to be used making FP cost efficient. FP can also measure binding events between protein and DNA through a fluorescent tag attached to an oligonucleotide strand. Interestingly for DNA-analyte interactions, the fluorophore can be attached to the heavier DNA molecule, and changes in polarisation can still be measured.[100]



B *Equation: 2-1*

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

C



(2.59)

Figure 2-19 (A) Schematic of *fluorescence polarisation assay* ***(B) Equation 2-1*** *Fluorescence Polarisation equation, P; polarisation, is the difference between emission intensity (I) of the parallel (∥) and perpendicular (⊥) light divided by the total emission intensity* ($I_{\parallel} + I_{\perp}$) . ***(C)*** *Fluorescein labelling structure.*

In the absence of a sMaf binding partner, Nrf2 exists as a partially disordered protein.[101] It has been proposed that the presence of MafG and DNA may be necessary for the Neh1 domain of Nrf2 to achieve an α-helix during the dynamic formation of the ternary complex.[41,102] Guided by a previously reported methodology, fluorescence polarisation was used to measure protein binding to ARE DNA.[20,102] This required the recombinant expression of Nrf2 and MafG which is discussed in **Chapter 5.** A fluorophore labelled DNA probe was prepared to measure inhibition of DNA binding (**Table 2-5**). A 25 base pair DNA sequence was chosen containing the ARE consensus TRE sequence (5'-TGA(C/G)TCA-3').[41,100] Oligonucleotides strands were annealed by heating to 95 °C and slowly cooling overnight to room temperature for use as a fluorescent probe in the assay.

Table 2-5 *Oligonucleotide sequences procured for assay development.*

#	*Oligonucleotide Sequence*
2.60	5’ [FAM]CGGAATTGCTGAGTCACTGTTACTC 3’
2.61	3’ GCCTTAACGACTCAGTGACAATGAG 5’

Using a minimum concentration of fluorescein labelled DNA (10 nM), fluorescence polarisation was used to measure binding affinity of MafG and Nrf2 to DNA. PBS at pH

7.4 was used with 0.01% Tween-20 and 10 mM DTT to prevent non-specific interactions. The ideal conditions for testing the synthesised peptides within the assay was at 80% of B_{MAX} of the ternary complex formation with DNA, where B_{MAX} represents the concentration at which maximum binding capacity is acheived.[96] MafG was titrated against 10 nM of fluorescein labelled DNA to determine the B_{MAX} of the MafG/DNA interaction at approximately 50 nM of protein. This concentration was subsequently used to titrate Nrf2 which achieved B_{MAX} at approximately 62.5 nM of protein (**Figure 2-20**). Consequently, these concentrations were used to run the fluorescence polarisation assay against the MafG peptides synthesised.

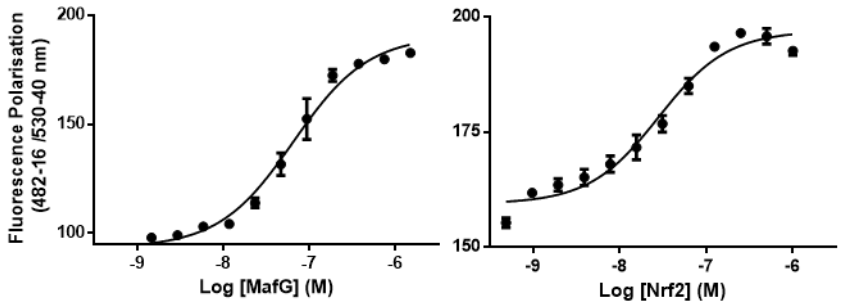


Figure 2-20 *Fluorescence polarisation of MafG (**right**) from 1.5 µM 1:2 dilutions with 10 nM of DNA. (**left**) Nrf2 from 1 µM in the presence of 50 nM MafG and 10 nM of DNA.*

2.46 and 2.52-2.56 were incubated within the established FP conditions at a concentration of 125 µM and read by CLARIOstar plate reader after 1 hour at room temperature and 16 hours at 4 °C. For both time points, the plate was incubated in darkness to prevent quenching of the fluorescein labelled DNA. The results were normalised by subtracting the signal of the fluorescein labelled DNA (negative control), expressed as a percentage of the signal produced by the Nrf2/MafG/ARE complex (positive control) (**Figure 2-21**). Excitingly **2.52,** a 21-mer sequence, produced a 30% reduction in Nrf2/MafG ternary complex at 125 µM. After 16 hours, **2.46** produced a 10% reduction in Nrf2/MafG ternary complex but demonstrated high variability in our assay which made further analysis challenging. The 14-mer sequences, **2.54-43** had no inhibitory activity against the Nrf2/MafG or MafG/MafG ternary complexes with ARE DNA at 125 µM. Suggesting sequences of this length are not competitive with the binding interaction of the full-length protein.

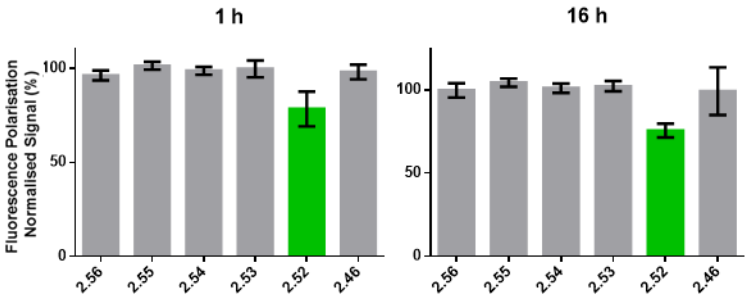


Figure 2-21 *Inhibition of Nrf2/MafG/ARE complex. Normalised fluorescence polarisation of* ***2.46 and 2.52-2.56*** *at 125 μM recorded at 1 hr and 16 hrs incubation.* ***2.52*** *highlighted in green demonstrating >25% inhibition.*

2.52 was investigated further to determine an IC_{50} against the ternary complex, from a concentration of 250 μM less than 50 % inhibition was achieved, consequently IC_{50} instead refers to 50 % of the change of activity observed. IC_{50} of 36.6 μM (95% CI of 19.0 – 70.5 μM), shown in **Figure 2-22**.

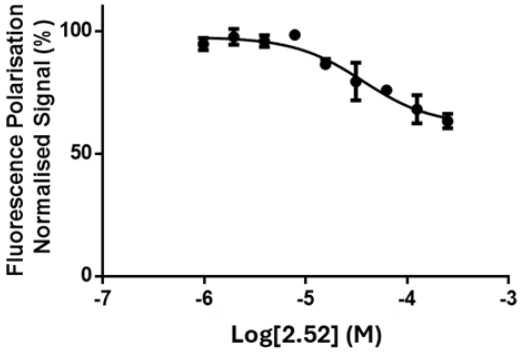


Figure 2-22 *Inhibition of Nrf2/MafG/ARE complex. Fluorescence polarisation of* ***2.52*** *at 16 h incubation against 62.5 nM Nrf2, 50 nM MafG and 10 nM ARE DNA.*

To maintain dynamic regulation of transcription the Nrf2/MafG interaction benefits from being relatively weak, this is supported by the weak inhibition observed by **2.52**. Additionally, the complexity of an assay of four components, where homo and heterodimeric interactions can be observed simultaneously, giving reason as to why 100 % inhibition is not observed. The peptides were screened against the MafG homodimer, it was found that **2.52** also reduced the binding of the homodimeric complex to DNA (**Figure 2-23**). The inhibition of MafG did not appear as strong as Nrf2 for **2.52** but a comparable IC_{50} was produced to compare the strength of the inhibitory activity at 21.3 μM (95% CI – 8.3-54.2 μM) (**Figure 2-23**). To confirm peptides disrupted the heterodimeric complex a complementary gel electrophoresis assay was established, discussed in **Chapter 2.3.5.**

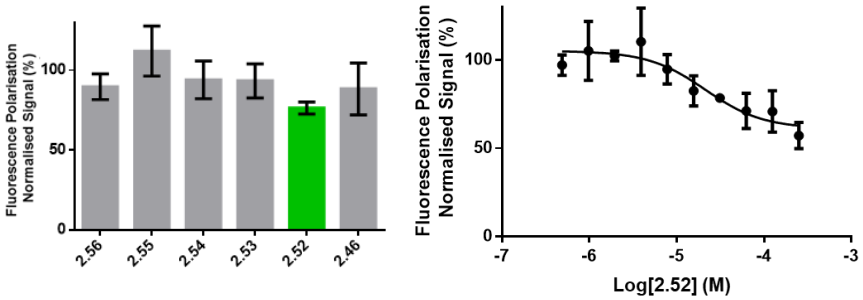


Figure 2-23 *Inhibition of MafG/MafG complex,* ***2.46 and 2.52-2.56*** *screened at 125 μM against 50 nM MafG and 10 nM ARE DNA (left).* ***2.52*** *titration from 250 μM 1:2 dilutions against 50 nM MafG and 10 nM DNA (right).*

2.3.4.1 ***Oxidation of M100 abolishes peptide activity.***

Interrogation of the impact of methionine oxidation within the **2.52** sequence using the fluorescence polarisation assay found that oxidation resulted in a substantial loss in inhibition at 125 μM (**Figure 2-24**). **2.57** reduced DNA binding by 10% whilst **2.52** retained inhibition of 25%.

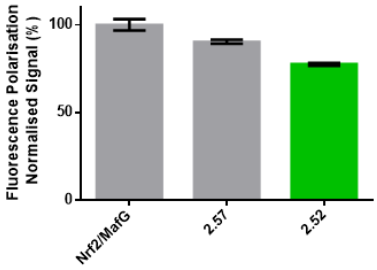


Figure 2-24 *Fluorescence Polarisation of* ***2.52*** *and* ***2.57*** *at 125 μM against 62.5 nM Nrf2, 50 nM MafG and 10 nM ARE DNA.*

Substitution of the methionine residue to norleucine (**2.58**) produced similar inhibition to **2.52** against the Nrf2 heterodimeric complex (**Figure 2-25**). **2.58** reduced DNA binding signal in both MafG (22%) and MafG/Nrf2 (24%) experiments. Titration of **2.58** against the Nrf2/MafG complex produced an IC_{50} of 130 μM (95% CI – 98.94 -170.80 μM). **2.58** produces complete inhibition of both complexes at 250 μM.

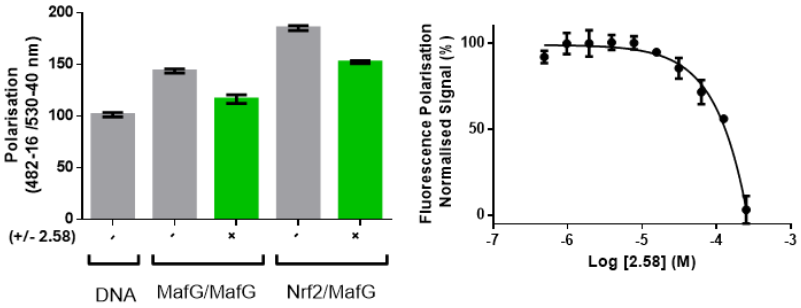


Figure 2-25 *(Left) fluorescence polarisation of* ***2.58*** *at 125 μM treated against Nrf2/MafG and MafG/MafG dimeric complexes binding to ARE DNA. (Right) titration against Nrf2/MafG from 250 μM.*

Whilst it was promising to observe inhibition of the Nrf2/MafG PPI and subsequent binding to ARE DNA, it was still uncertain if the peptide was selectively inhibiting the Nrf2/MafG interaction, or if it also inhibited MafG/MafG interactions. To investigate this question an electrophoretic mobility shift assay was explored.

2.3.5 Investigation of inhibition by electrophoretic mobility shift assay

Electrophoretic mobility shift assays (EMSA) allow observation of protein-DNA interactions, exploiting the principle that DNA bound in complex to proteins reduces DNA migration during gel electrophoresis compared to unbound DNA, based on size (**Figure 2-26**). Using fluorescently-labelled DNA, a gel can be visualised by a desired fluorophore filter in a gel imaging system. In the context of transcription factors, EMSAs have been used to measure DNA interactions of a protein to the specific DNA recognition sequences to determine binding affinity.[40,41,100,103] However, competition based EMSAs to study inhibitors activity to block DNA binding is an under-explored approach. Electrophoretic mobility shift assays have previously been used to investigate the homo and heterodimeric complexes of the Nrf2/MafG transcription factors, informing the development of an EMSA assay to investigate inhibitory activity of the peptides.[40,100]

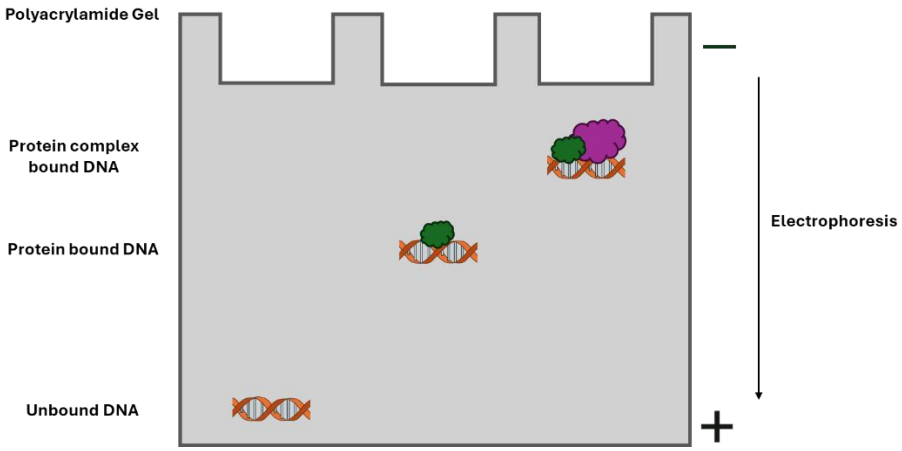


Figure 2-26 *Transport of biomolecules through gel electrophoresis in an EMSA.*

To investigate the complex formation with the fluorescent ARE DNA probe and the subsequent screening of peptide inhibitors of the Nrf2/MafG ternary complex, 250 nM of fluorophore labelled ARE DNA was incubated with MafG or Nrf2/MafG. The samples were loaded onto a 20% TBE polyacrylamide gel and subjected to 200 V for 135 minutes at 4 °C to monitor the changes in MafG/MafG/DNA, MafG/Nrf2/DNA and free DNA movement on the gel.

As seen in **Figure 2-27,** multiple shifted bands are observed in lane 3. The Nrf2/MafG complex causes slower probe migration because of formation of a heavier heterodimeric complex. This also creates a stronger intensity of shift compared to the MafG dimeric complex within the same lane.

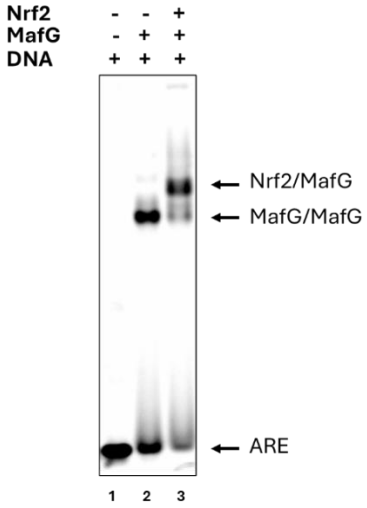


Figure 2-27 *20% TBE gel of 250 nM of fluorophore labelled ARE DNA incubated with either MafG or Nrf2/MafG (ARE:MafG:Nrf2 ratio of 1:1.5:9) in PBS, pH 7.4 with 10 mM DTT and 0.01% Tween-20.*

To directly complement the fluorescence polarisation assay, after reading, the contents of each well was directly loaded on a gel and used in the EMSA assay. Using this approach, visualisation of the MafG/MafG/ARE complex was difficult (**Figure 2-28**) as bands were comparatively weaker than the signal observed by fluorescence polarisation (Figure 2-20). But observation of changes in intensity to the free DNA probe were readily visible. Titration of Nrf2 from 1 μM – 62.5 nM saw complete loss of the free DNA probe between 1 and 0.5 μM of Nrf2 and an increase in intensity between 0.25 μM and 31 nM Nrf2, comparable to the free DNA probe in the far-right lane.

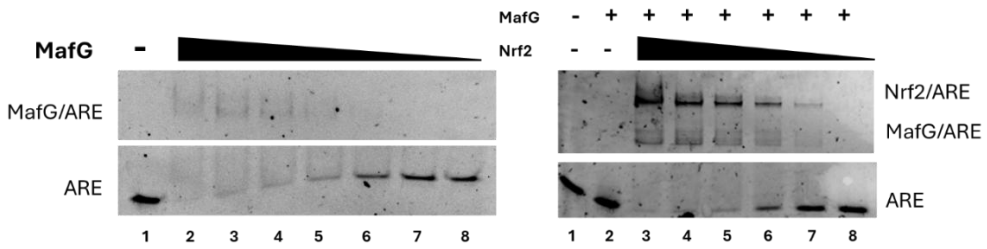


Figure 2-28 *(Left) EMSA of MafG from 0.5 μM 1:2 dilutions (lanes 2-8) incubated* ***(right)*** *EMSA of Nrf2 against 50 nM MafG 1:2 dilution of Nrf2 from 1 μM (lanes 3-8) Both gels had samples prepared in PBS with 0.1% Tween-20 and 10 mM DTT at 4 ℃ incubated with 10 nM ARE DNA.*

2.46 and 2.52-2.56 were explored at 125 μM against 10 nM ARE DNA, 50 nM MafG and 62.5 nM Nrf2 (**Figure 2-29**). The gel produced distinct bands of the Nrf2 complex and free DNA probe, allowing for comparison to be made of each peptide. In the presence of **2.52**

and **2.46,** the disappearance of the Nrf2/MafG complex band is observed, complemented by an increase in the free DNA migration band to similar intensity of free probe migration in the MafG/MafG dimer (**Figure 2-29, Lane 2**). This suggests **2.52** and **2.46** are interrupting the Nrf2/MafG/DNA complex, perhaps with specificity towards Nrf2/MafG disruption.

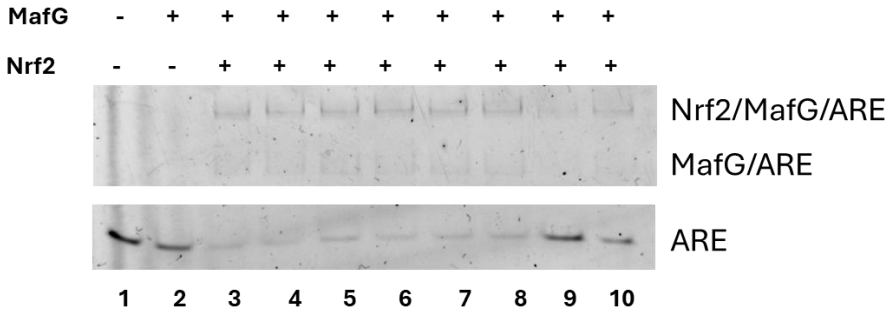


Figure 2-29 *EMSA of* ***2.46 and 2.52-2.56*** *against 62.5 nM Nrf2, 50 nM MafG and 10 nM fluorescein ARE DNA. Lane 4 ML385, Lane 5 (****2.56****), Lane 6 (****2.55****), Lane 7 (****2.54****), Lane 8 (****2.53****), Lane 9 (****2.52****), Lane 10 (****2.46****), at 125 μM.*

This approach demonstrates that EMSA's can be used as a complementary assay to confirm qualitatively confirm ternary complex inhibition by inhibitors of the interaction between Nrf2 and MafG.

2.3.6 Surface plasmon resonance of Nrf2

2.3.6.1 Introduction

To determine if MafG derived peptides had affinity for the Nrf2 an investigation using surface plasmon resonance (SPR) was conducted. SPR is a biophysical technique that can be used to measure the binding interaction between an analyte and an immobilised ligand. In the context of the Nrf2/MafG PPI, SPR has previously been used to immobilise a biotin labelled Nrf2 derived stapled peptide, measuring the K_d of $MafG_{1\text{-}162}$.[19] However, SPR has not been reported for immobilisation of full-length Nrf2 for measuring binding interactions with analytes. This may offer a method for screening Nrf2/MafG PPI inhibitors. Described here is an investigation into the use of covalently captured Nrf2 to measure binding affinity of the peptides synthesised.

SPR monitors molecular interactions in real time, by shining polarised light across a gold conductor, producing excitation of plasmons on the surface. Immobilisation of a ligand to the gold surface allows ligand binding events to be measured, as a change in the angle of resonance at which the light is reflected is caused by analyte binding.[46] The shift in the angle of incidence can be monitored through an association and dissociation of the analyte in order to perform kinetics or affinity calculations. As shown in **Figure 2-30,** SPR

plots a sensogram of response units (RU) over time, to determine whether a binding interaction has occurred between an analyte and the ligand bound to the chip surface.

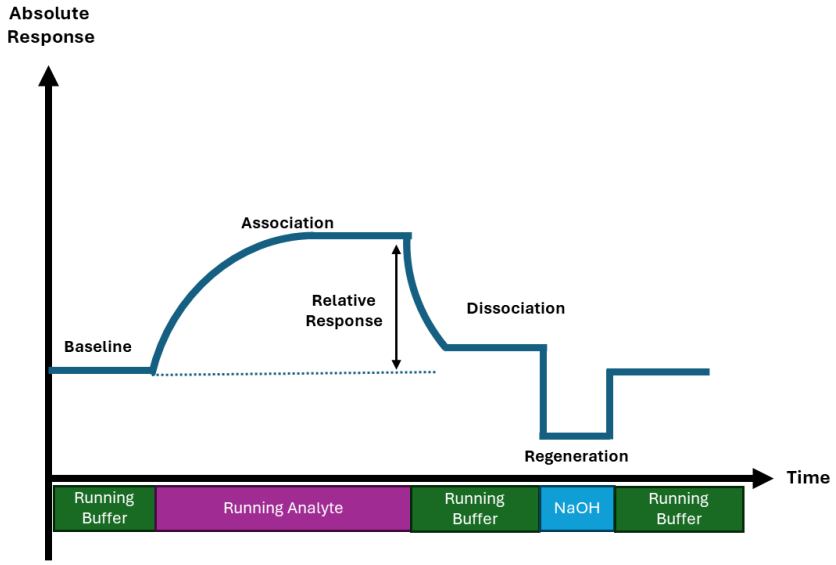


Figure 2-30 *Example SPR sensogram of association and dissociation of an analyte across a ligand immobilised surface.*

Immobilisation of a ligand to a gold surface can be achieved through a variety of methods, catering to the requirements of the interaction being measured. Ligands can be immobilised transiently through affinity capture or permanently through covalent bond formation. Covalent ligand immobilisation is led by the available reactive functional groups on the desired ligand, such as thiol, aldehyde or amine groups, allowing coupling with ease to a dextran surface.

2.3.6.2 Immobilisation of Nrf2

The immobilisation of Nrf2 to the chip surface was achieved by covalent amide coupling across a carboxymethylated dextran chip surface. Firstly, buffer pH scouting was performed to determine the optimum pH for protein pre-concentration, to localise protein at the dextran surface for efficient coupling. A positive charge of the protein is required for efficient pre-concentration in proximity to the negatively charged dextran surface. A pH that is below the isoelectric point (pI) of the ligand, is typically necessary to produce the desired charges as seen in **Figure 2-31.**

Nrf2 was diluted into a buffer of a low ionic strength (10 mM sodium acetate) to identify the optimum pH for immobilisation. The isoelectric point (pI) of Nrf2 is 4.7, consequently a buffer below this pH would be expected to produce an ionisation state of the protein to achieve effective preconcentration on the chip surface.[34] The dextran surface becomes positively charged at pH 3 (**Figure 2-31**), preventing pre-concentration across

the chip, making a higher pH buffer more favourable for ligand immobilisation.[47] 10 mM sodium acetate buffer was used to at pH 4-5.5. **Figure 2-32** establishes, in line with the pI of Nrf2, that buffer at pH 4 (highlighted in green) was most suitable for the preconcentration of ligand across the chip surface.

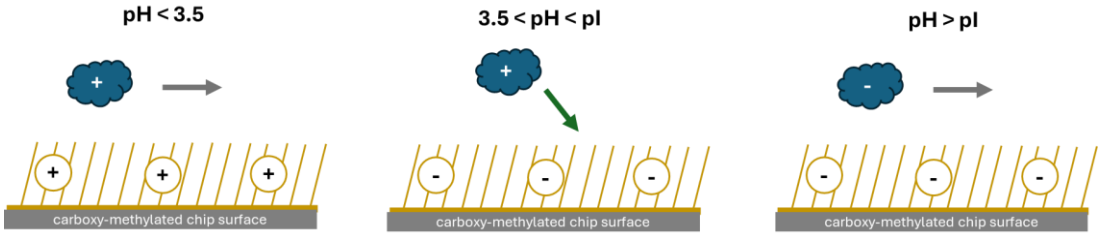


Figure 2-31 *pH scouting charge interactions between the surface of a chip and the desired ligand for immobilisation.*

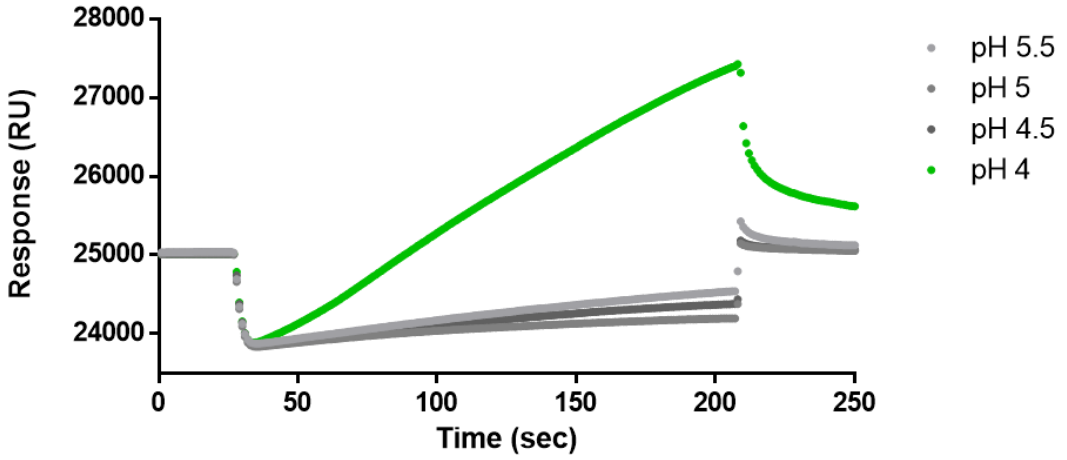
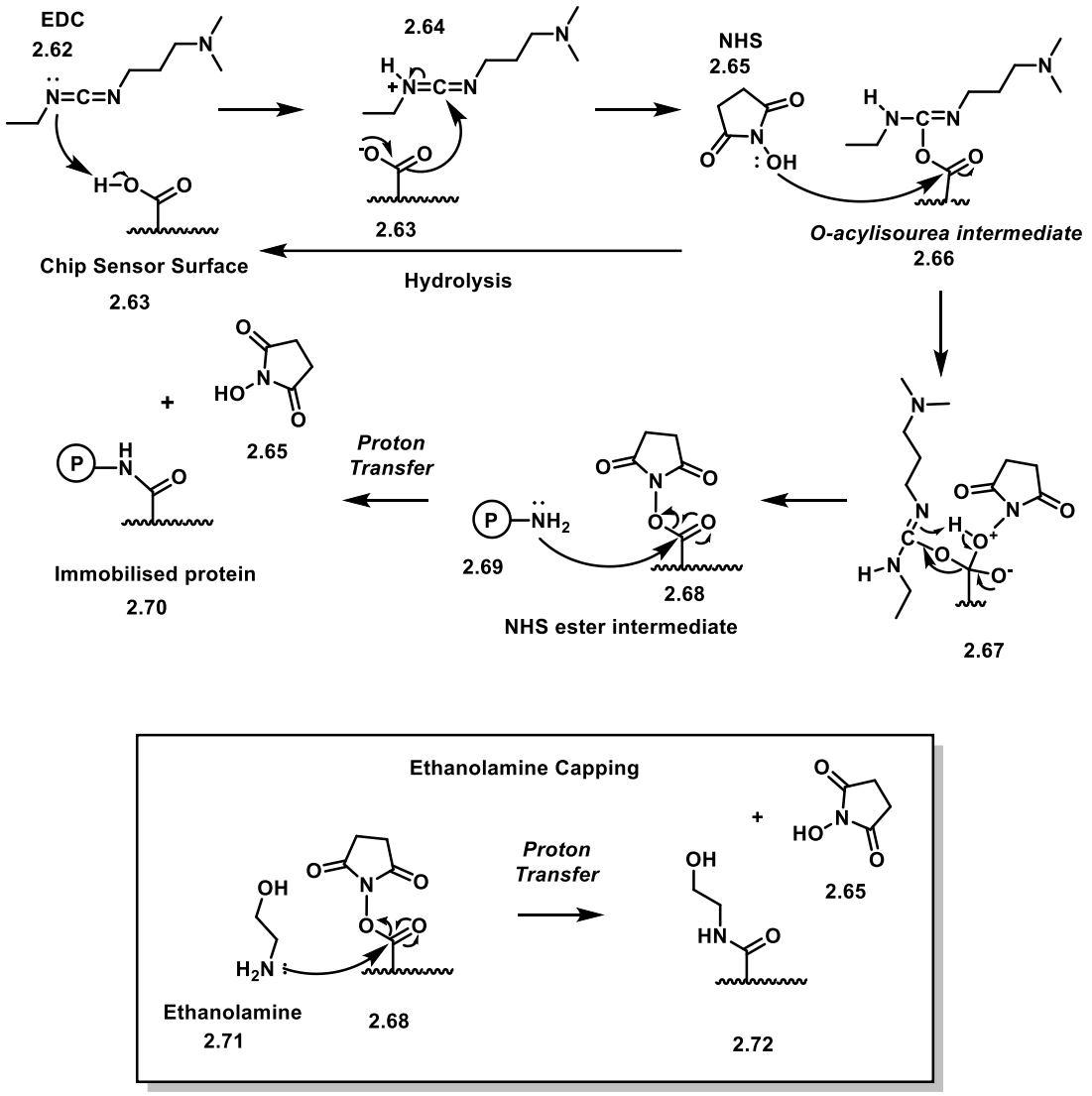


Figure 2-32 *SPR pH scouting of Nrf2 in sodium acetate buffer at pH 5.5, 5.0, 4.5 and 4.0.*

For immobilisation of recombinant protein, amide coupling is often chosen due to the relative abundance of free primary amines within a protein, without compromising relative access to a proposed binding site (**Scheme 2-7**).[48] Amide coupling is achieved through the surface activation of the carboxylic acids present by ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) (**2.62**) which reacts with the carboxyl group (**2.63**) to form an highly reactive O-acylisourea intermediate (**2.66**). This intermediate is then stabilised by the addition of *N*-hydroxysuccinimide (NHS) (**2.65**) to form an NHS ester (**2.68**), offering an electrophilic site for a proteins (**2.69**) free amine group to attack, leading to the formation of an amide bond for protein immobilisation (**2.70**). Preparation of the SPR chip is then completed by capping unreacted carboxylated dextran (**2.68**) with ethanolamine (**2.71**) to limit nonspecific binding.



Scheme 2-7 *Coupling of protein using EDC (**2.62**) and NHS (**2.65**) and ethanolamine (**2.71**) capping of unreacted positions.*

The desired change in response units for an immobilised ligand can be determined prior to immobilisation to complement a desired analyte response if binding is observed (**Equation 2-2**). Immobilisation of Nrf2 was prepared with a desired range for peptides **2.46** and **2.52-2.56**. Based on a recommended maximum response within range of 100 RU for an analyte, protein immobilisation was calculated using **Equation 2-2,** determining an immobilisation range between 5,000 RU would achieve quantitative signal strength.[46]

$$RU_L = \frac{R_{max} \times MW_L}{\eta \times MW_A}$$

Equation 2-2 *Surface Density Capacity**:** Rmax; the maximum response observed by the specified analyte. MW_A; molecular weight of the analyte. MW_L ; molecular weight of the ligand. RU_L ; response unit change produced by ligand immobilisation. η ; stoichiometry of analyte binding sites on the ligand.*

After activation of the chip surface with EDC and NHS, a target immobilisation of 5,000 RU for Nrf2 was set, flowing the protein over the chip surface at 10 µL/min for approximately 10 minutes (**Figure 2-33**). Ethanolamine was then used to cap unreacted carboxyl sites to limit non-specific interactions of the analyte with the chip surface. The chip surface baseline rose from 24800 RU to 29440 RU, indicative of a successful immobilisation of 4640 RU of Nrf2 (**Figure 2-33).** Based on this immobilisation level, theoretical Rmax values were calculated to determine the maximum change in RU from 1:1 binding of peptides **2.46 and 2.52-2.56** to Nrf2**,** summarised in **Table 2-6.**

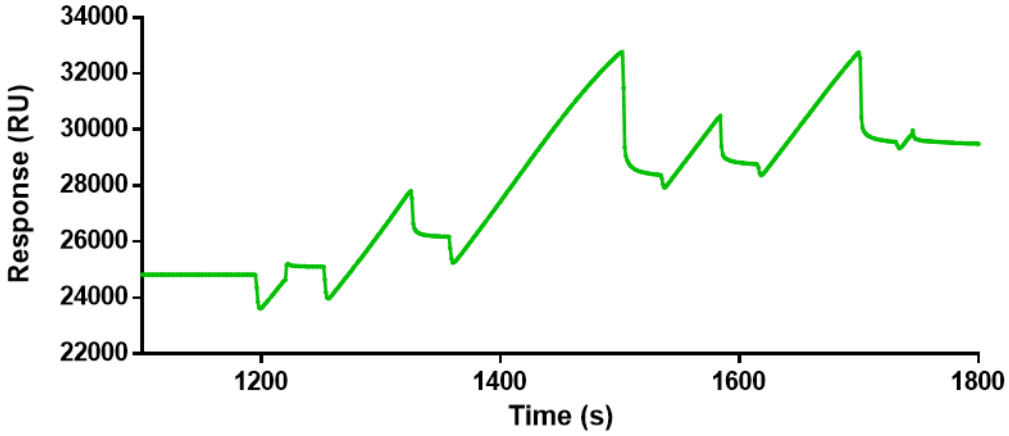


Figure 2-33 *Immobilisation of Nrf2 to CM5 chip surface*

Table 2-6 *Peptides and proposed Rmax at an RU_L of 5000.*

#	Sequence	MW	Rmax
2.56	Ac-KEELEKQKAELQQE-NH_2	1771	79
2.55	Ac-KAELQQEVEKLASE-NH_2	1643	73
2.54	Ac-VEKLASENASMKLE-NH_2	1590	71
2.53	Ac-KEELEKQKAELQQEVEKLASE-NH_2	2528	113
2.52	Ac-KAELQQEVEKLASENASMKLE-NH_2	2417	107
2.46	Ac-KEELEKQKAELQQEVEKLASENASMKLE-NH_2	3301	147

2.3.6.3 Buffer trials

Buffer composition can greatly impact the interaction of an analyte with a ligand through alterations to pH, ionic strength and additives.[49] Traditionally, Tris (TBS), HEPES (HBS) or phosphate (PBS) buffered saline can be used in SPR however a buffer screen can be performed to determine the optimum conditions for kinetic experiments to take place. As each buffer listed varies in pKa at 20 °C, the temperature experiments were run at. Each buffer was supplemented with ethylenediaminetetraacetic acid (EDTA) a chelating

agent, and the surfactant P20 to prevent aggregation and non-specific interactions, with addition of EDTA and P20, denoted as EP+.[50] Analytes and the ligand may demonstrate variable solubility in each buffer, influencing their ability to interact.[49] Due to the limited quantity of peptide available, an initial buffer screen, a compound known to bind to Nrf2, ML385, was used to investigate the quality of binding profiles to the immobilised protein (**Figure 2-34**).[51]

Key characteristics of SPR kinetic experiments aim to produce an efficient regeneration profile after sample injection. All three buffers tested produced fast association of ML385 to reach a complete association with the immobilised Nrf2 protein. In TBS-EP+ buffer, after initial rapid association, ML385 continues to produce a slower increase in association **(Figure 2-34, A)**. This suggests the tris-based buffer is less preferable for measuring binding kinetics. Additionally, the dissociation phase in TBS-EP+, measured over 300 seconds, achieved only a partial dissociation from Nrf2. Sharp changes to the RU signal are referred to as a bulk shift, caused by a mismatch in the refractive index between the analyte solution and running buffer. Additionally, the calculated K_D fell outside of the range of concentrations used (K_D ~ 307 µM). PBS-EP+ (**Figure 2-34, B**) produced a considerable bulk shift in response units at the beginning and end of the association cycle, despite best efforts to produce accurate buffer matching, potentially signifying unfavourable interactions in this buffer composition. HBS-EP+ produced a comparably smooth, fast association profile with a fast dissociation profile and complete dissociation of the compound was achieved after the end of injection, whilst also remaining in range of the theoretical Rmax for ML385, assuming 1:1 binding stoichiometry (M_r: 511, Rmax ~22 RU) (**Figure 2-34, A**). Comparison of K_D to existing literature was not possible as it has not previously been reported, however, the calculated K_D of ML385 in HBS-EP+ 132 µM felt most reasonable compared to the concentration range tested.

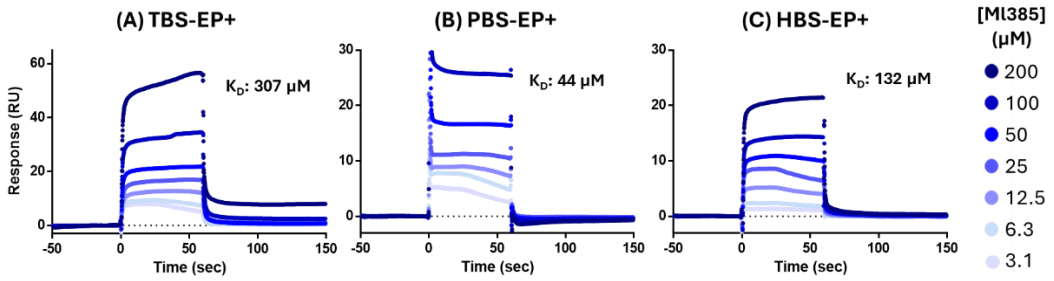


Figure 2-34 *ML385 from 200 µM 1:2 dilutions in TBS-EP+ (A), PBS-EP+ (B) or HBS-EP+ (C) injected across the surface of the chip over 60 seconds at 30 µL/min.*

2.46 and 2.52-2.56 were subsequently ran in HBS-EP+ buffer with a 60 second injection, followed by a 300 second dissociation period to determine if the peptides could interact with Nrf2 (**Figure 2-35**)**.** Across all six peptides, a negative sensogram profile was produced, (**2.55,** is shown as an example sensogram, **Figure 2-35, A**) with a negative signal during the association phase, 0-60 s. All peptides also displayed an increased interaction during the dissociation phase of the experiment, 60-300 s. It was considered that the peptides were producing a binding interaction with the reference surface of the SPR chip; a second channel exposed to the same immobilisation reagents in the absence of Nrf2 protein, capped with ethanolamine.

The use of additives to the running buffer can help to prevent non-specific binding of the analyte. Bovine serum albumin (BSA) and dextran are commonly used.[52–54] To investigate the benefits of these additives, experiments were trialled using **2.55** to compare the effect on the binding kinetics with Nrf2 (**Figure 2-35, B and C**). The use of BSA still resulted in a negative sensogram profile of **2.55** at lower concentrations. Dextran at 1 mg/mL in the running buffer appeared to improve the association profile counteracting the non-specific binding of the carboxylated dextran (**Figure 2-35, C**).

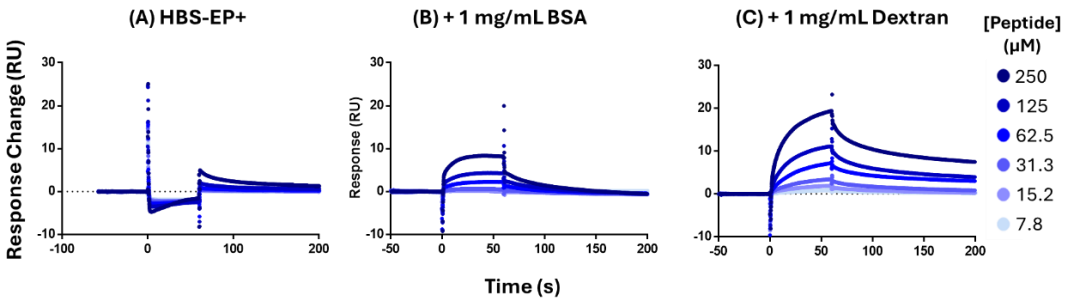


Figure 2-35: *Impact of additives.* ***2.55*** *from 250 μM 1:2 dilutions against immobilised Nrf2 in HBS-EP+, 5% DMSO. pH 7.4, 60 second injections (A) comparison with (B) 1 mg/mL dextran and* ***(C)*** *1 mg/mL BSA.*

Changes to the pH of the running buffer were subsequently trialled to determine if a pH closer to the pI of Nrf2 (4.7) improved performance, as optimised pH can improve the stability of analyte binding kinetics and further reduce non-specific binding interactions.[46] Ultimately pH 7.4 was chosen for the running buffer, as the profile produced by **2.55** has reduced bulk shift and supported responses at a higher Rmax (**Figure 2-36**).

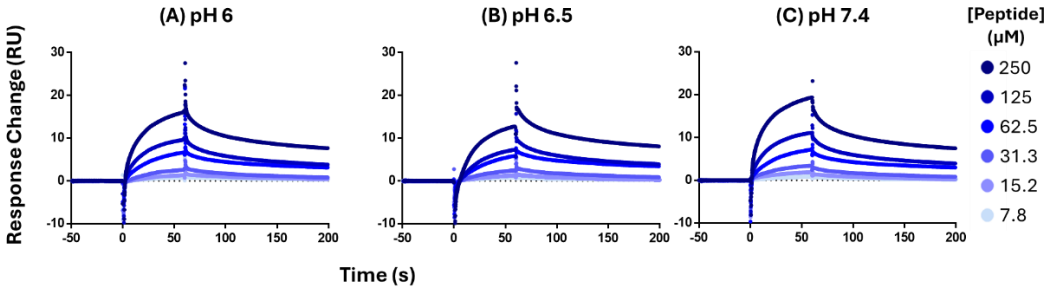


Figure 2-36 *Impact of pH on Nrf2 SPR. Sensogram results of* ***2.55*** *in HBS-EP+ 1 mg/mL dextran 5% DMSO at pH 6 (A), 6.5 (B) or 7 (C).*

With running buffer conditions established, **2.46 and 2.52-2.56** were tested again in HBS-EP+ with 1 mg/mL dextran at pH 7.4 to minimise non-specific interactions (**Figure 2-37**). The dilution series was limited to a minimum concentration of 7.8 μM as samples below this concentration produced no change in response units. For **2.46 and 2.52-2.56** the buffer conditions improved the sensogram profiles produced by adding dextran to the running buffer.

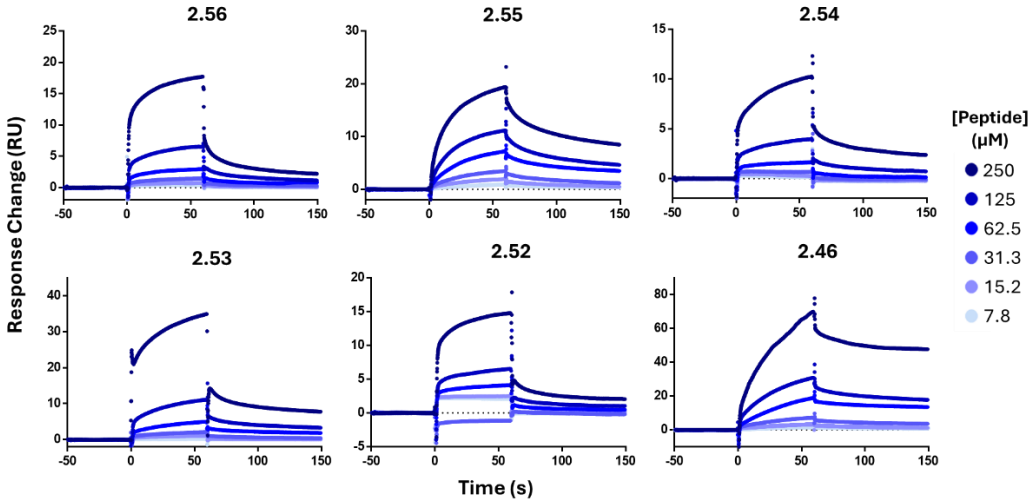


Figure 2-37 *Peptide SPR against Nrf2. SPR spectra recorded for* ***2.46 and 2.52-2.56*** *in HBS-EP+* pH *6 with 1 mg/mL dextran at 5% DMSO. Injection of peptides over 60 seconds 30 μL/min at 250-7.8 μM as 1:2 dilutions.*

Despite improved responses with optimised conditions, the RU recorded failed to plateau at high concentrations, making it difficult to determine the K_D for the binding interaction with Nrf2 (**Figure 2-38**). Additionally, none of the six peptides investigated reached the theoretical Rmax at the concentrations used. Therefore, higher concentrations of analyte would be required to accurately calculate K_D, but this was not feasible due to peptide solubility. Immobilising a greater amount of Nrf2 onto an SPR chip may overcome these issues. However further experiments observed erosion of the baseline RU, indicative that the stability of Nrf2 protein over extended periods may not be compatible with SPR to determine binding affinity. The disordered nature of the Nrf2

protein in isolation may also prevent the correct folding required for a coiled-coil interaction to form and be measured in this way.[15] Our investigation by SPR would conclude that all six peptides demonstrate equally weak affinity for the Nrf2 protein. This contradicts the competition fluorescence polarisation and EMSA data collected in **Chapter 2.3.5,** particularly towards **2.52**. Consequently, limitations to the SPR assay may have contributed to an underestimation in quantifying the peptide-protein binding interaction.

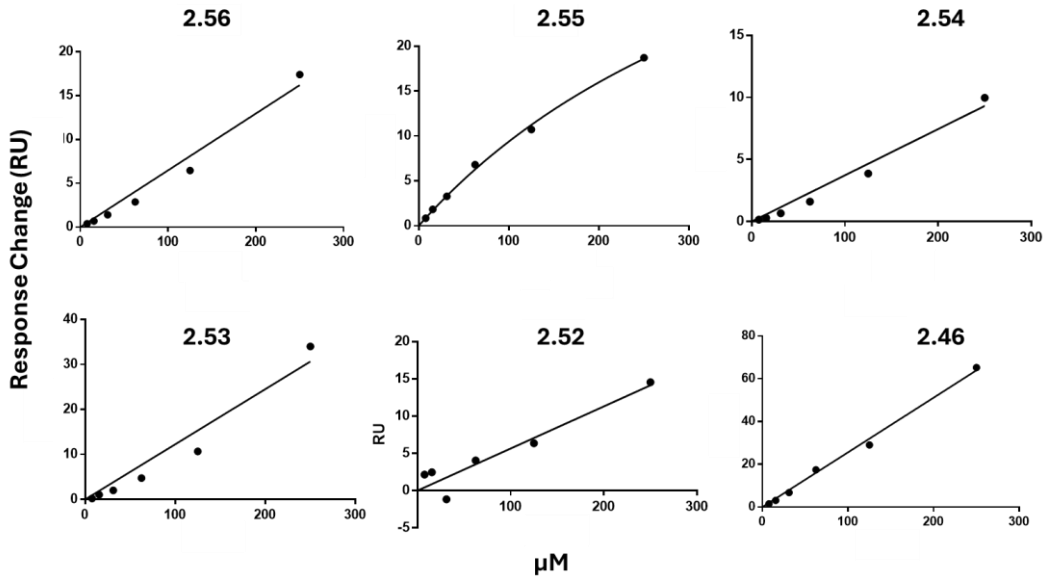


Figure 2-38: *Relationship between response units (RU) and concentration (μM) for* ***2.46 and 2.52-2.56*** *from 200 μM 1:2 dilutions.*

2.4 Conclusion

This chapter presents the synthesis and biophysical characterisation of novel peptides based on the leucine zipper of MafG. Using an AlphaFold prediction model and existing literature on the coiled-coil interaction of Nrf2 and MafG proteins, six peptides were designed (**2.46 and 2.52-2.56**, **Figure 2-14**). Derived from $MafG_{76\text{-}104}$, Peptides were synthesised at 14, 21 or 28 amino acids in length. Improvements to the synthesis of **2.54, 2.52** and **2.46** using 0.4 M LiCl in DMF as the solvent throughout the synthesis enabled purification with quantifiable yields of these peptides.

To compete with MafG to disrupt coiled-coil PPI formation, the peptides needed to form an α-helical secondary structure. Circular dichroism experiments demonstrated **2.46** and **2.52-2.56** all produced an α-helical secondary structure, with the highest percentage helicity produced by **2.53** (81%) and **2.52** (71%) (**Figure 2-18**). A fluorescence polarisation assay was established to measure the inhibition of the ternary complex of

the Nrf2/MafG interaction with DNA containing an ARE promoter sequence. **2.52** was found capable of inhibiting the complex formation with an IC_{50} of 36 μM **(Figure 2-22).** Nrf2/MafG complex inhibition by **2.52** and **2.46** was qualitatively confirmed through an EMSA, demonstrating a decrease in protein-DNA complexes and an increase in free DNA probe migration across the gel (**Figure 2-29**).

The use of SPR was explored by immobilising Nrf2 to the surface of a CM5 chip as a method for measuring binding affinity of peptides predicted to interact with the protein. SPR has not previously been used to characterise Nrf2 binding interactions, making this study an interesting first report on Nrf2 in SPR. Investigation of MafG mimicking peptides, **2.46** and **2.52-43,** binding profiles to Nrf2 reported weak binding responses (**Figure 2-38**), perhaps supporting the weak inhibitory activity observed by FP and EMSA. Additionally, a maximum response for the peptide-protein interaction was not established, where higher concentrations of peptide could not be tested due to solubility issues. The inherent challenges of studying transient protein-protein interactions may offer an explanation as to why only weak interactions were observed.[55] Nrf2 undergoes disorder-to-order transitions in the presence of its binding partner MafG.[15] Covalently constraining Nrf2 via amide coupling may further disrupt the formation of the helical secondary structure required for PPI formation with MafG derived peptides. Using Nrf2 in solution may improve the proteins' ability to fold correctly.

Additionally, the requirement for buffer to be at pH 4 for immobilisation of Nrf2 is not ideal for protein stability. Based on the results found in **Chapter 2,** a future SPR study could immobilise **2.52** to the chip surface, to investigate if a K_D of the peptide-protein interaction could be calculated inversely. Future studies could also explore alternative immobilisation strategies to minimise the impact of immobilisation induced secondary structure changes influencing the formation of the desired PPI.

Future work will look to improve the quality of the EMSA's produced to improve the signal strength. The use of minimal fluorophore concentrations posed challenges, high background noise in the gel produced artifacts unrelated to the interactions being investigated. fluorescein labelling has previously been reported to observe a low noise to signal ratio in gel visualisation.[104] Consequently, we would hope to explore the use of alternative fluorophore labelled DNA to improve the image quality of the experiments.

Investigation into the impact of methionine oxidation on activity of **2.52** against the Nrf2 protein found that methionine sulfoxide containing sequences lost inhibitory activity against Nrf2/MafG. This led to the synthesis of **2.58,** substituting methionine for norleucine. Assessment by fluorescence polarisation determined **2.58** had a similar inhibitory activity against the Nrf2 heterodimer (**Figure 2-25**) but also an increase in inhibition of the MafG homodimeric interaction.

It is possible that peptides derived from MafG $_{76}$KEELEKQKAELQQEVEKLASE$_{94}$ have a preferential ability to self-assemble as coiled-coils, which could be confirmed by further circular dichroism experiments designed to measure coiled-coil formation. **2.54, 2.52** and **2.46** contain the sequence $_{96}$NASMKLE$_{104}$, which may have a reduced preference for homodimerisation due to the polar Asn$_{97}$ residue occupying an "*a"* position (**Figure 2-39**) in the heptadic motif of MafG. This substitution may disrupt the hydrophobic core of homodimeric coiled-coil assembly, enhancing the peptides' ability to compete with native MafG. Future work may explore residue mutation of ionic positions (*e* and *g*) in the MafG peptides to further dissuade self-assembly.

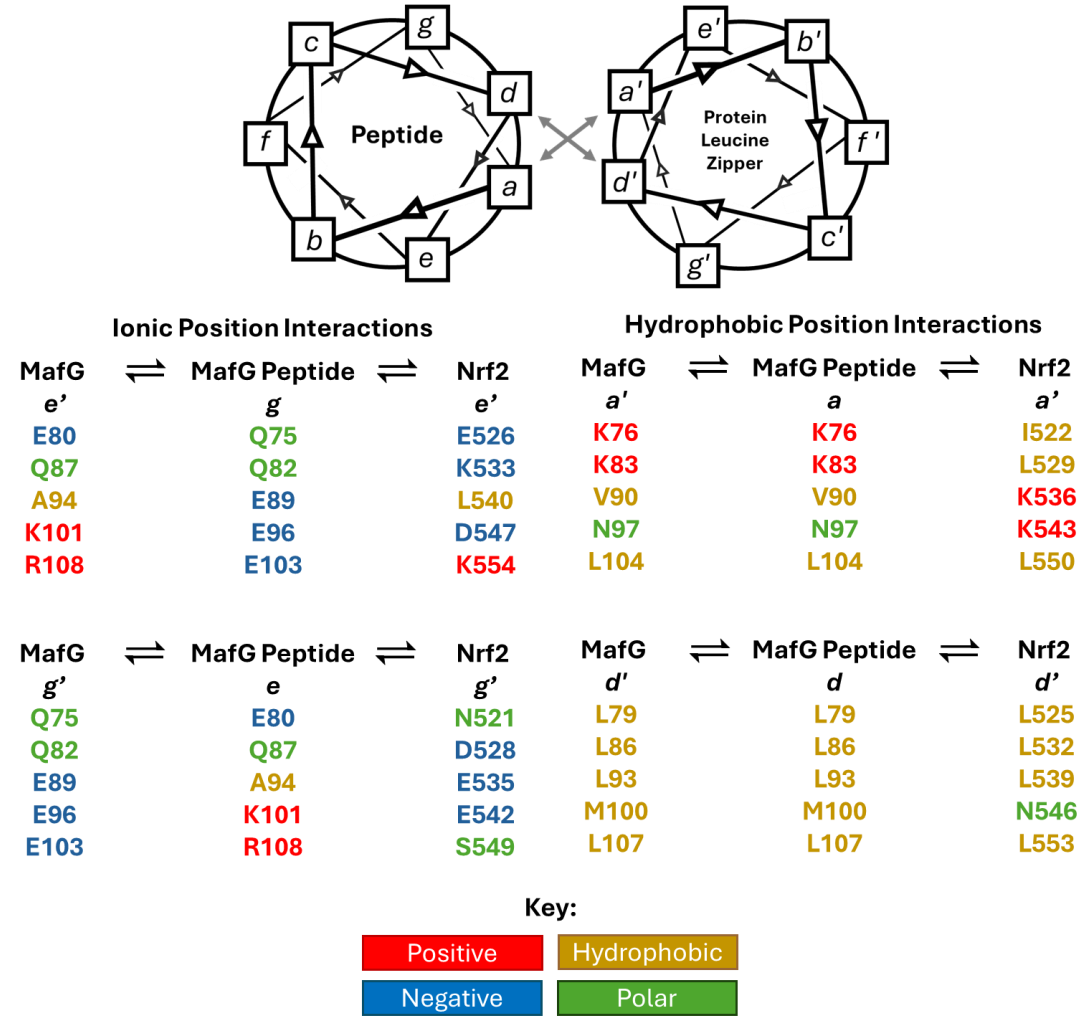


Figure 2-39: *Organisation of coiled-coil interacting residues between MafG peptide and MafG or Nrf2.*

Future work would include exploring the **2.52** structure further to improve selective inhibition of Nrf2/MafG heterodimer through residue modification within the coiled-coil to improve the hydrophobic and ionic properties of the helix to afford more potent inhibitors. Further synthesis optimisation could also be carried out, as whilst purification of lead **2.52** was achievable, the production of high yielding synthesis remains a challenge. Development of an assay to investigate peptide selectivity towards the Nrf2 or MafG leucine zipper would also provide an interesting tool to further optimise peptide inhibitors of the PPI.

Overall, the optimisation of a fluorescence polarisation and electrophoretic mobility shift assays has led to the identification of peptide inhibitors capable of disrupting the Nrf2/MafG/ARE ternary complex. This will enable the development of further inhibitors of this interaction. **2.52** and **2.58** are well placed for evaluation of Nrf2 inhibition in cell-based assays in models of Nrf2 over-expression in cancer.

2.5 Experimental

2.5.1 Chemicals

All chemicals used were reagent grade. Fmoc-Amino Acids were purchased from Sigma-Aldrich, AGTC Bioproducts, or Fluorochem. Coupling and Deprotection Reagents, HBTU, DIPEA and Oxyma Pure from Fluorochem; DIC and Piperidine from Sigma Aldrich and HOBt from AGTC Bioproducts. TentaGel® S RAM resin was purchased from Rapp Polymere. Rink Amide ProTide™ Resin (LL) was purchased from CEM. Analytical and HPLC grade solvents were purchased from Sigma Aldrich and assumed to conform to specification. Peptide synthesis grade solvents, N,N-Dimethylformamide (DMF) and 1-Methyl-2-pyrrolidone (NMP) were purchased from AGTC Bioproducts. Trifluoroacetic Acid (TFA) was purchased from Fluorochem.

2.5.2 Solid-phase peptide synthesis

2.54, 2.52 and **2.46** were synthesised using a Syro I automated peptide synthesiser. 100 mg of Tentagel S RAM resin (0.26 mmol/g) was swollen in DMF containing 0.4 M LiCl (3 mL) for 20 min with agitation. DMF was then drained and Fmoc-deprotection was performed using a solution of 20% piperidine in DMF (2 mL) for 10 min with agitation, the vessel was then drained and the step repeated. Wash step was then performed by rinsing the vessel with 0.4 M LiCl in DMF (2 mL) and agitating for 30 s, this step was repeated 3 times. Fmoc-amino acid (4 eq) in NMP was coupled using HOBt (4 eq), HBTU (4 eq) in DMF and DIPEA (8 eq) in NMP with the addition of 1 mL DMF with 0.4 M LiCl for 20 minutes with agitation, this step was then repeated. After the coupling reaction was completed, the established wash step was then repeated. The process of deprotection and coupling was carried out until the desired amino acid sequence was completed.

2.56, 2.55 and 2.53 were synthesised using a Liberty Lite automated microwave peptide synthesiser, using Tentagel S RAM resin (0.26 mmol/g) was manually swollen in DMF for 10 min before addition to the reaction vessel. Fmoc deprotection steps were carried out in 20% piperidine in DMF with microwave heating of 90 °C for 1 min. Wash steps were performed by draining the reaction vessel and washing with DMF (3 x 3 mL). Coupling reactions of all amino acids was achieved using Oxyma Pure (5 eq), DIC (5 eq) and Fmoc-protected amino acids (5 eq) at 90 °C for 5 min. The process of deprotection and coupling was carried out until the desired amino acid sequence was completed.

Upon the final deprotection all peptides were acetylated in acetic anhydride (4 eq) and DIPEA (4 eq) in DMF with agitation for 45 minutes. The final wash step rinsed the vessel with DMF (3 x 5 mL), CH_2Cl_2 (3 x 5 mL) and MeOH (3 x 5 mL) to exhaustively remove remaining DMF.

Cleavage of peptides from resin support was achieved using 10 mL TFA with addition of reaction scavenger's H_2O, DODT and TIPS at 2.5% each. Cleavage solution was added to a reaction vessel and agitated for 3 h. The cleavage solution was then collected and concentrated in vacuo, followed by precipitation in diethyl ether. Peptides were purified using RP-HPLC and analysed by analytical HPLC and MALDI-TOF.

2.5.3 Analytical techniques

2.5.3.1 RP-HPLC

An Agilent 1200 with an Agilent eclipse XDB-C18 column, 4.6 x 150 mm, 5 µm and a flow rate of 1 mL/min was used for performing Analytical RP-HPLC. Solvent system of solvent A = H_2O 95% : MeCN 5% : TFA 0.05% Solvent B: H_2O 5% : MeCN 95% : TFA 0.05% Gradient of Solvent B from 5-95% over 15 minutes.

Preparative RP-HPLC was performed using an Agilent 1200 with an Agilent eclipse XDB-C18 column 21.2 x 150 mm (5 µm) flow rate of 20 mL/min. Solvent A: H_2O 95% : MeCN 5% : TFA 0.05% Solvent B: H_2O 5% : MeCN 95% : TFA 0.05% Gradient of Solvent B from 5-95% over 15 minutes.

2.46 Ac-KEELEKQKAELQQEVEKLASENASMKLE-NH_2, RP-HPLC (RT = 10.6 min)

2.52 Ac-KAELQQEVEKLASENASMKLE-NH_2, RP-HPLC (RT = 10.7 min)

2.58 Ac-KAELQQEVEKLASENAS(Nle)KLE-NH_2, RP-HPLC (RT = 8.2 min)

2.57 Ac-KAELQQEVEKLASENASM(O)KLE-NH_2, RP-HPLC (RT = 10.6 min)

2.53 Ac-KEELEKQKAELQQEVEKLASE-NH_2, RP-HPLC (RT = 5.6 min)

2.54 Ac-VEKLASENASMKLE-NH_2, RP-HPLC (RT = 10.8 min)

2.55 Ac-KAELQQEVEKLASE-NH_2, RP-HPLC (RT = 10.9 min)

2.56 Ac-KEELEKQKAELQQE-NH_2, RP-HPLC (RT = 8.9 min)

2.5.3.2 MALDI-TOF

MALDI-TOF: Performed on a Shimazdu Axima Performance MALDI-TOF in positive ionisation mode. All samples were calibrated to a ProteoMass™ ACTH Fragment 18-39 MALDI-MS Standard. Mol Wt 2,464.1989 Da.

2.46 Ac-KEELEKQKAELQQEVEKLASENASMKLE-NH_2 3300.7 $[M+H]^+$ expected, 3300.7 observed.

2.58 Ac-KAELQQEVEKLASENAS(Nle)KLE-NH_2 2398.2 expected $[M+H]^+$, 2398.6 observed.

2.57 Ac-KAELQQEVEKLASENASM(O)KLE-NH_2 2432.2 $[M+H]^+$ expected, 2432.9 observed.

2.52 Ac-KAELQQEVEKLASENASMKLE-NH_2 2416.2 expected $[M+H]^+$ 2416.2 observed.

2.53 Ac-KEELEKQKAELQQEVEKLASE-NH_2 2527.3 expected $[M+H]^+$ 2527.0 observed.

2.54 Ac-VEKLASENASMKLE-NH_2 1589.8 expected $[M+H]^+$, 1590.2 observed.

2.55 Ac-KAELQQEVEKLASE-NH_2 1642.8 expected $[M+H]^+$, 1642.2 observed.

2.56 Ac-KEELEKQKAELQQE-NH_2 1770.9 expected $[M+H]^+$, 1771.1 observed

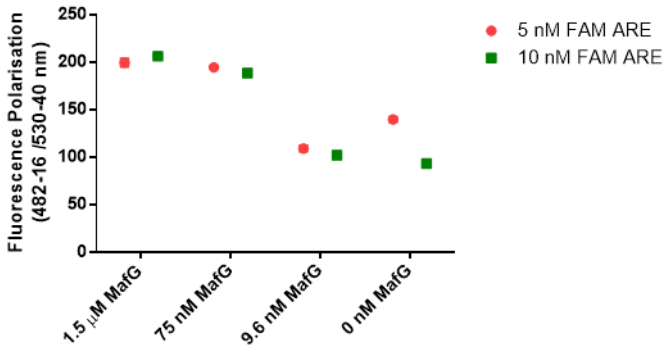
2.5.4 Circular dichroism

CD spectra were recorded on a Jasco J-810 spectropolarimeter using a 1 mm path length quartz cuvette (Starna Scientific). Peptides were suspended in 50:50 Trifluoroethanol: Potassium Phosphate Buffer, (10mM Potassium Phosphate Buffer at pH 7.4 in dH_2O) at 0.2 mg/mL. The scans were recorded at 20 °C between 185 and 260 nm. Data was then analysed using Dichroweb using an CDSSTR Method.

2.5.5 Fluorescence polarisation

Corning ™ 384-well solid black polystyrene microplates containing 40 µL/ well. Final concentration of 10 nM per well of fluorescein labelled ARE DNA with 50 nM of His-MafG and 62.5 nM of His-Halo-Nrf2 in PBS containing 10 mM DTT and 0.1% Tween-20 at pH 7.4. Peptides stored in DMSO were added to the plate ensuring no higher than 5% DMSO was used. Plate was incubated for 16 h at 4 °C, polarisation was measured using ex: 482-16 em: 530-40 Dichroic: LP 504 on Fluorescence Polarisation Mode on a ClarioStar Plate Reader.

To examine if the assay was functional, a titration of the proteins was performed in PBS at pH 7.4 containing 0.05% Tween-20, against 5 and 10 nM of the fluorescent DNA ranging the protein concentration from 1.5 µM in 1:5 dilutions in the presence and absence of Nrf2. 5 nM fluorophore produced an unexpected increase in polarisation in the absence of protein (**Figure 2-31**). 10 nM of fluorescein labelled DNA was subsequently used. It was subsequently decided to proceed experiments at 10 nM of fluorescent DNA to maintain low background interference with the assay.



Appendix 2-1 *Fluorescence polarisation signals of 5 and 10 nM fluorescein ARE DNA at different concentrations of MafG.*

2.5.6 Electrophoretic mobility shift assay (EMSA)

20% TBE gels (Novex™) was pre-run in 1x TBE buffer (Invitrogen™ UltraPure™) at 4 °C at 100 V for 30 minutes (PowerEase™ 120 W Power Supply). Fluorophore labelled ARE DNA (Sigma-Aldrich) was incubated with Nrf2 and MafG protein for 16 h with peptides at 4 °C and loaded with 1X BlueJuice Loading Dye (Invitrogen™). Gel run at 200 V for 135 minutes at 4 °C in darkness. Gels were visualised using the appropriate fluorophore filter with ImageQuant LAS 4000 (GE Healthcare).

2.5.7 Surface plasmon resonance

Biacore S200 SPR Instrument was used with Biacore Sensor Chip CM5 (Cytiva) for immobilisation of His_6HaloNrf2 by amide coupling. pH scouting was performed using 10 mM sodium acetate at pH 4, 4.5, 5 and 5.5, the estimated PI of Nrf2 is 4.67, consequently buffer at pH 4 was used. His_6HaloNrf2 was flown over the chip at 5 µL/min immobilising at 4640 response units (RU) in Sodium Acetate pH 4. HBS-EP+ pH 7.4 was used as the running buffer. A reference cell was blocked with ethanolamine.

Experiments to trial buffer conditions as described in **Chapter 2.3.6**, utilised buffers with contents provided in **Table 7**. Peptides **2.46** and **2.52-43** were injected at a flow rate of 30 µL/min for periods of 60 s followed by a dissociation period of 300 s at 250-3.125 µM. The chip surface was generated using 10 mM NaOH for 30 s. A DMSO correction (4.5-5.5%) was applied to the recorded sensogram and a blank subtracted from the signal. Results were analysed using Biacore S200 evaluation software using steady state kinetics to determine KD values. ML385, a known Nrf2 inhibitor, was flown over the chip at 30 µL/min for 60 seconds at 125-0.288 µM.

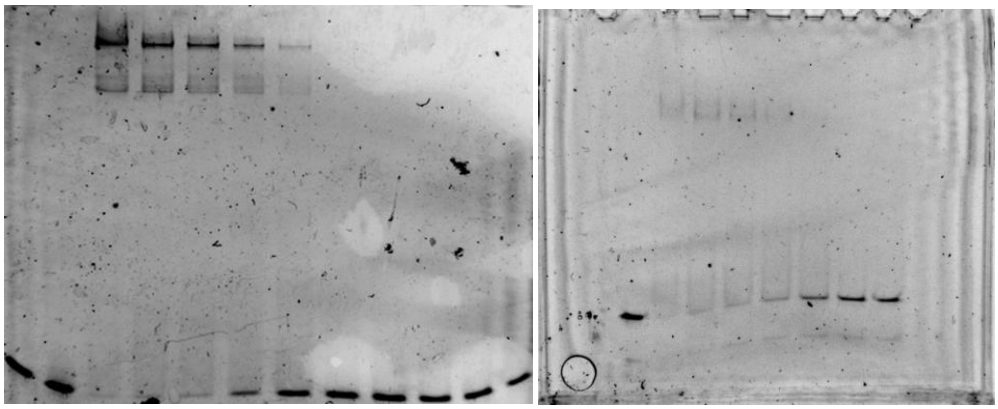
SPR Buffer compositions:

#	Buffer Name	Contents
A	HEPES-Buffered Saline (HBS-EP+)	10 mM HEPES pH 7.4, 150 mM NaCl, 3 mM EDTA, 0.01% P20
B	Tris-Buffered Saline (TBS-EP+)	50 mM TRIS-HCl pH 7.4, 150 mM NaCl, 3 mM EDTA, 0.01% P20
C	Phosphate Buffered Saline (PBS-EP+)	10.1 mM Na_2PO_4, 1.8 mM KH_2PO_4, 137 mM NaCl, 2.7 mM KCl, pH 7.4, 3 mM EDTA, 0.01% P20

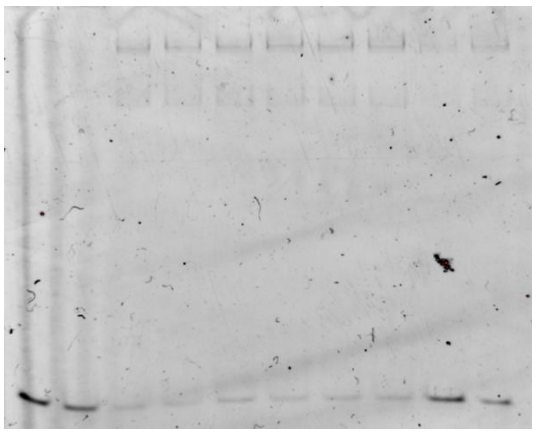
Table 2-7 *Buffer compositions tested*

2.6 Appendix

2.6.1.1 Complete gels from EMSA

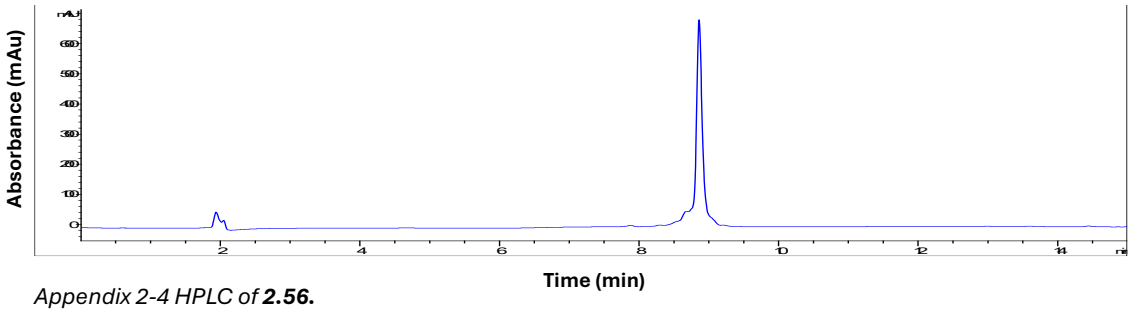


Appendix 2-2 Full gel image from **Figure 2-28.**

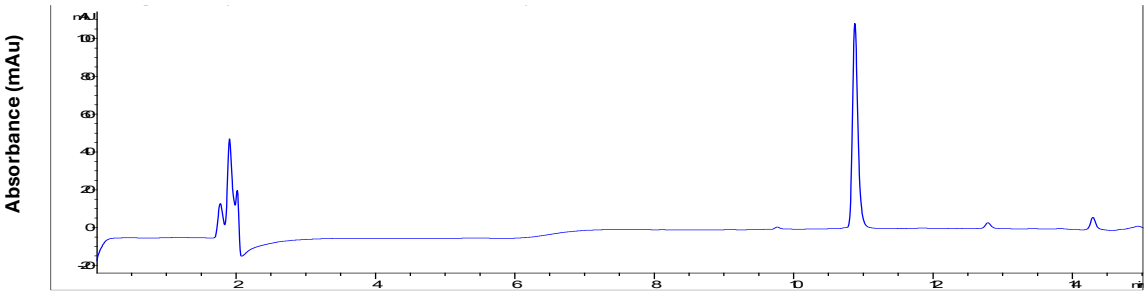


Appendix 2-3 *Full gel image from Figure 2-29.*

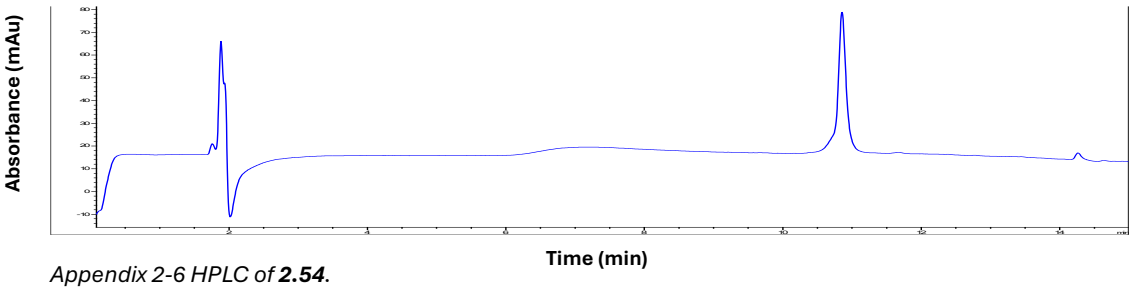
2.6.1.2 HPLC traces of peptides synthesised



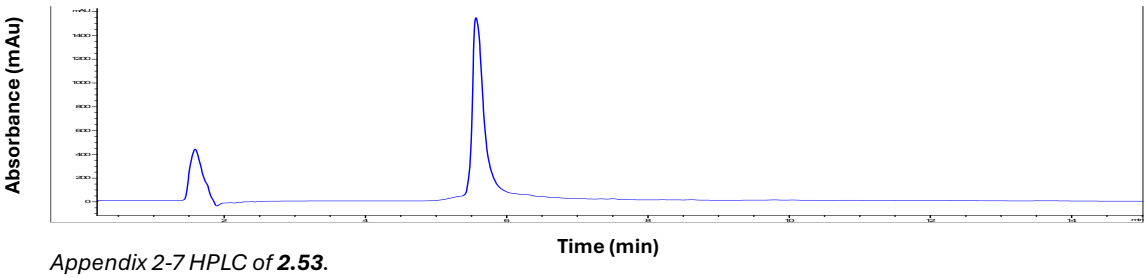
*Appendix 2-4 HPLC of **2.56.***



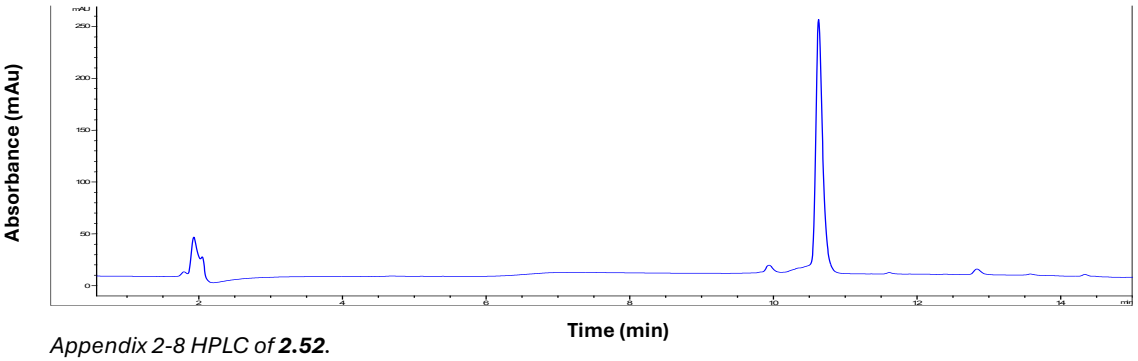
*Appendix 2-5 HPLC of **2.55**.*



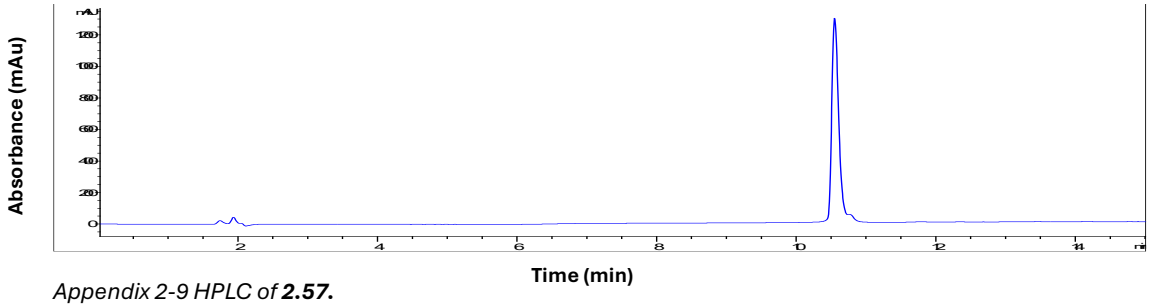
*Appendix 2-6 HPLC of **2.54**.*



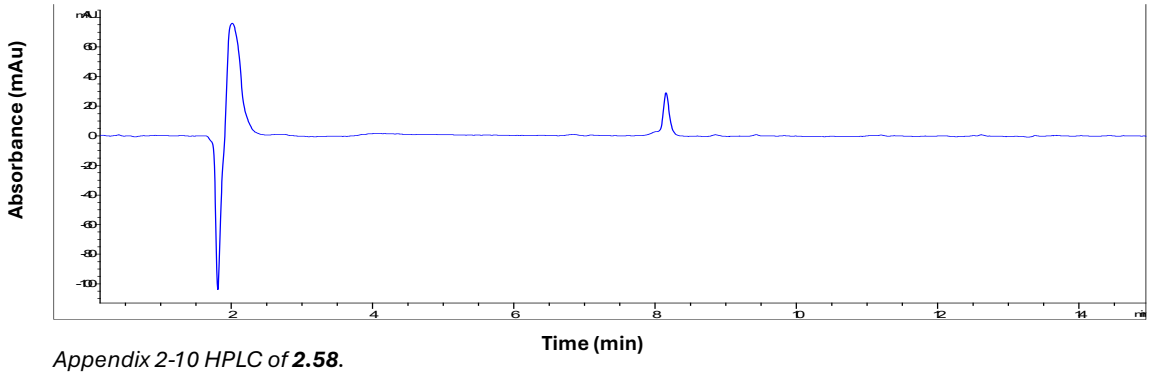
*Appendix 2-7 HPLC of **2.53**.*



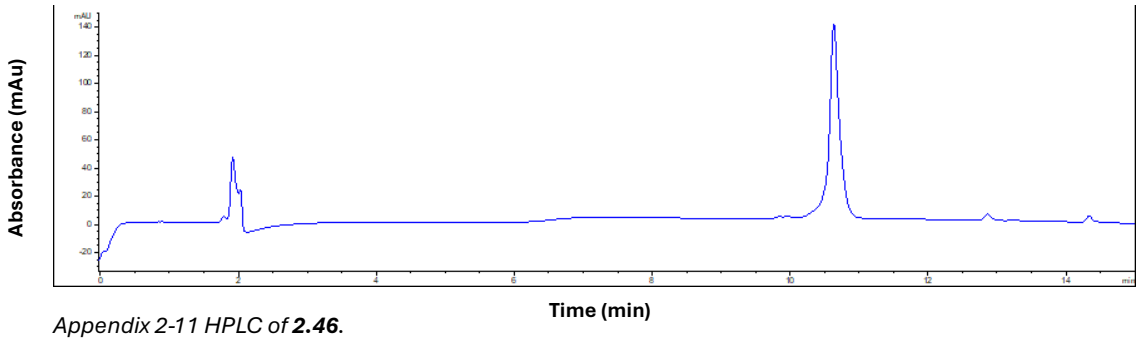
*Appendix 2-8 HPLC of **2.52**.*



*Appendix 2-9 HPLC of **2.57**.*



*Appendix 2-10 HPLC of **2.58**.*



*Appendix 2-11 HPLC of **2.46**.*

2.7 References

1 F. Friedberg, T. Winnick and D. M. Greenberg, *J Biol Chem*, 1947, **169**, 763.

2 J. Forbes and K. Krishnamurthy, in *Peptide Biochemistry*, StatPearls Publishing, Treasure Island (FL), 2025.

3 T. Kükenshöner, D. Wohlwend, C. Niemöller, P. Dondapati, J. Speck, A. V. Adeniran, A. Nieth, S. Gerhardt, O. Einsle, K. M. Müller and K. M. Arndt, *J Struct Biol*, 2014, **186**, 335–348.

4 J. M. Mason, M. A. Schmitz, K. M. Müller and K. M. Arndt, *Proc Natl Acad Sci USA*, 2006, **103**, 8989–8994.

5 U. B. Hagemann, J. M. Mason, K. M. Müller and K. M. Arndt, *J Mol Biol*, 2008, **381**, 73–88.

6 V. Potapov, J. B. Kaplan and A. E. Keating, *PLoS Comput Biol*, 2015, **11**, e1004046.

7 M. R. Hicks, S. S. Al-Mehairi and A. J. Sinclair, *J Virol*, 2003, **77**, 8173–8177.

8 R. O. Crooks, T. Rao and J. M. Mason, *J Biol Chem*, 2011, **286**, 29470–29479.

9 J. N. Pelletier, K. M. Arndt, A. Plückthun and S. W. Michnick, *Nat Biotechnol*, 1999, **17**, 683–690.

10 J. M. Mason, K. M. Müller and K. M. Arndt, *Biochem Soc Trans*, 2008, **36**, 1442–1447.

11 J. Speck, C. Räuber, T. Kükenshöner, C. Niemöller, K. J. Mueller, P. Schleberger, P. Dondapati, J. Hecky, K. M. Arndt and K. M. Müller, *Protein Eng Des Sel*, 2013, **26**, 225–242.
12 J. M. Mason, K. M. Müller and K. M. Arndt, *Biochemistry*, 2007, **46**, 4804–4814.
13 G. Grigoryan, A. W. Reinke and A. E. Keating, *Nature*, 2009, **458**, 859–864.
14 J. K.-L. Sinclair and A. Schepartz, *Org Lett*, 2014, **16**, 4916–4919.
15 D. C. Chan, C. T. Chutkowski and P. S. Kim, *Proc Natl Acad Sci U S A*, 1998, **95**, 15613–15617.
16 C. M. McDonald, C. Petosa and P. J. Farrell, *J Virol*, 2009, **83**, 3397–3401.
17 M. R. Hicks, S. Balesaria, C. Medina-Palazon, M. J. Pandya, D. N. Woolfson and A. J. Sinclair, *J Virol*, 2001, **75**, 5381–5384.
18 X. Ran, E. T. Burchfiel, B. Dong, N. J. Rettko, B. M. Dunyak, H. Shao, D. J. Thiele and J. E. Gestwicki, *Bioorg Med Chem*, 2018, **26**, 5299–5306.
19 J. K.-L. Sinclair, E. V. Denton and A. Schepartz, *J Am Chem Soc*, 2014, **136**, 11232–11235.
20 R. Modi, N. McKee, N. Zhang, A. Alwali, S. Nelson, A. Lohar, R. Ostafe, D. D. Zhang and E. I. Parkinson, *J Med Chem*, 2023, **66**, 6184–6192.
21 F. Katsuoka and M. Yamamoto, *Gene*, 2016, **586**, 197–205.
22 J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli and D. Hassabis, *Nature*, 2021, **596**, 583–589.
23 M. Baek, F. DiMaio, I. Anishchenko, J. Dauparas, S. Ovchinnikov, G. R. Lee, J. Wang, Q. Cong, L. N. Kinch, R. D. Schaeffer, C. Millán, H. Park, C. Adams, C. R. Glassman, A. DeGiovanni, J. H. Pereira, A. V. Rodrigues, A. A. van Dijk, A. C. Ebrecht, D. J. Opperman, T. Sagmeister, C. Buhlheller, T. Pavkov-Keller, M. K. Rathinaswamy, U. Dalwadi, C. K. Yip, J. E. Burke, K. C. Garcia, N. V. Grishin, P. D. Adams, R. J. Read and D. Baker, *Science*, 2021, **373**, 871–876.
24 M. Mirdita, K. Schütze, Y. Moriwaki, L. Heo, S. Ovchinnikov and M. Steinegger, *Nat Methods*, 2022, **19**, 679–682.
25 Z. Lin, H. Akin, R. Rao, B. Hie, Z. Zhu, W. Lu, N. Smetanin, R. Verkuil, O. Kabeli, Y. Shmueli, A. dos Santos Costa, M. Fazel-Zarandi, T. Sercu, S. Candido and A. Rives, *Science*, 2023, **379**, 1123–1130.
26 N. J. Fowler and M. P. Williamson, *Structure*, 2022, **30**, 925-933.e2.
27 *Nat Methods*, 2023, **20**, 163–163.
28 T. C. Terwilliger, D. Liebschner, T. I. Croll, C. J. Williams, A. J. McCoy, B. K. Poon, P. V. Afonine, R. D. Oeffner, J. S. Richardson, R. J. Read and P. D. Adams, *Nat Methods*, 2024, **21**, 110–116.
29 J. Abramson, J. Adler, J. Dunger, R. Evans, T. Green, A. Pritzel, O. Ronneberger, L. Willmore, A. J. Ballard, J. Bambrick, S. W. Bodenstein, D. A. Evans, C.-C. Hung, M. O'Neill, D. Reiman, K. Tunyasuvunakool, Z. Wu, A. Žemgulytė, E. Arvaniti, C. Beattie, O. Bertolli, A. Bridgland, A. Cherepanov, M. Congreve, A. I. Cowen-Rivers, A. Cowie, M. Figurnov, F. B. Fuchs, H. Gladman, R. Jain, Y. A. Khan, C. M. R. Low, K. Perlin, A. Potapenko, P. Savy, S. Singh, A. Stecula, A. Thillaisundaram, C. Tong, S. Yakneen, E. D. Zhong, M. Zielinski, A. Žídek, V. Bapst, P. Kohli, M. Jaderberg, D. Hassabis and J. M. Jumper, *Nature*, 2024, **630**, 493–500.
30 W. Zhu, A. Shenoy, P. Kundrotas and A. Elofsson, *Bioinformatics*, 2023, **39**, btad424.
31 R. L. Dunbrack, *bioRxiv*, 2025, 2025.02.10.637595.

32 A. S. Konagurthu, J. C. Whisstock, P. J. Stuckey and A. M. Lesk, *Proteins Struct, Funct Bioinf* 2006, **64**, 559–574.
33 S. Brüschweiler, J. E. Fuchs, G. Bader, D. B. McConnell, R. Konrat and M. Mayer, *ChemMedChem*, 2021, **16**, 3576–3587.
34 P. Bryant, G. Pozzati and A. Elofsson, *Nat Commun*, 2022, **13**, 1–11.
35 E. K. O'Shea, J. D. Klemm, P. S. Kim and T. Alber, *Science*, 1991, **254**, 539–544.
36 D. E. Casteel, E. V. Smith-Nguyen, B. Sankaran, S. H. Roh, R. B. Pilz and C. Kim, *J Biol Chem*, 2010, **285**, 32684–32688.
37 Y. Fujii, T. Shimizu, T. Toda, M. Yanagida and T. Hakoshima, *Nat Struct Mol Biol*, 2000, **7**, 889–893.
38 J. Zhao, J. R. Stagno, L. Varticovski, E. Nimako, V. Rishi, K. McKinnon, R. Akee, R. H. Shoemaker, X. Ji and C. Vinson, *Mol Pharmacol*, 2012, **82**, 814–823.
39 M. Miller, J. D. Shuman, T. Sebastian, Z. Dauter and P. F. Johnson, *J Biol Chem*, 2003, **278**, 15178–15184.
40 H. Kurokawa, H. Motohashi, S. Sueno, M. Kimura, H. Takagawa, Y. Kanno, M. Yamamoto and T. Tanaka, *Mol Cell Biol*, 2009, **29**, 6232–6244.
41 T. Sengoku, M. Shiina, K. Suzuki, K. Hamada, K. Sato, A. Uchiyama, S. Kobayashi, A. Oguni, H. Itaya, K. Kasahara, H. Moriwaki, C. Watanabe, T. Honma, C. Okada, S. Baba, T. Ohta, H. Motohashi, M. Yamamoto and K. Ogata, *Nucleic Acids Res.*, 2022, **50**, 12543–12557.
42 W. Li and A.-N. Kong, *Mol Carcinog*, 2009, **48**, 91–104.
43 M. Theodore, Y. Kawai, J. Yang, Y. Kleshchenko, S. P. Reddy, F. Villalta and I. J. Arinze, *J Biol Chem*, 2008, **283**, 8984–8994.
44 A. K. Jain, D. A. Bloom and A. K. Jaiswal, *J Biol Chem*, 2005, **280**, 29158–29168.
45 F. W. Lichtenthaler, *Eur J Org Chem*, 2002, **2002**, 4095–4122.
46 E. Fischer and E. Fourneau, *Chemische Berichte*, 1901, **34**, 2868–2877.
47 M. Bergmann, *Science*, 1934, **79**, 439–445.
48 K. Morihara, *Trends Biotechnol*, 1987, **5**, 164–170.
49 F. Guzmán, S. Barberis and A. Illanes, *Electron J Biotechnol*, 2007, **10**, 279–314.
50 B. Merrifield, *Methods Enzymol*, 1997, **289**, 3–13.
51 J. M. Collins, S. K. Singh, T. A. White, D. J. Cesta, C. L. Simpson, L. J. Tubb and C. L. Houser, *Nat Commun*, 2023, **14**, 8168.
52 J. A. Moss, *Curr Protoc Protein Sci*, 2005, **40**, 1–19.
53 W. C. Chan and S. L. Mellor, *J Chem Soc, Chem Commun*, 1995, **0**, 1475–1477.
54 H. Rink, *Tetrahedron Lett*, 1987, **28**, 3787–3790.
55 D. Tymecka and A. Misicka, in *Peptide Synthesis*, Humana, New York, NY, 2020, pp. 1–11.
56 P. Sieber and B. Riniker, *Tetrahedron Lett*, 1991, **32**, 739–742.
57 H. Gausepohl, M. Kraft and R. W. Frank, *Int J Pept Protein Res*, 1989, **34**, 287–294.
58 P. R. Hansen and A. Oddo, *Methods Mol Biol*, 2015, **1348**, 33–50.
59 F. Albericio and L. A. Carpino, *Methods Enzymol*, 1997, **289**, 104–126.
60 O. F. Luna, J. Gomez, C. Cárdenas, F. Albericio, S. H. Marshall and F. Guzmán, *Molecules*, 2016, **21**, 1542.
61 G. B. Fields, in *Peptide Synthesis Protocols*, eds M. W. Pennington and B. M. Dunn, Humana Press, Totowa, NJ, 1995, pp. 17–27.
62 V. Borghesani, *J Pept Sci*, 2025, **31**, e3649.
63 D. A. Pearson, M. Blanchette, M. L. Baker and C. A. Guindon, *Tetrahedron Lett*, 1989, **30**, 2739–2742.
64 A. Teixeira, W. E. Benckhuijsen, P. E. de Koning, A. R. P. M. Valentijn and J. W. Drijfhout, *Protein Pept Lett*, **9**, 379–385.
65 A. R. Kunys, W. Lian and D. Pei, *Curr Protoc Chem Biol*, 2012, **4**, 331–355.

66 Merck, NovaPEG Rink Amide resin Novabiochem® | Sigma-Aldrich, https://www.sigmaaldrich.com/GB/en/product/mm/855047, (accessed 11 March 2025).
67 T. Arnesen, *PLOS Biology*, 2011, **9**, e1001074.
68 R. J. Wallace, *Br J Nutr*, 1992, **68**, 365–372.
69 I. Dikiy and D. Eliezer, *J Biol Chem*, 2014, **289**, 3652–3665.
70 N. J. Greenfield, W. F. Stafford and S. E. Hitchcock-DeGregori, *Protein Sci*, 1994, **3**, 402–410.
71 I. Nicolau, A. Paun, C. C. Popescu, N. D. Hădade and M. Matache, *J Chem Educ*, 2023, **100**, 2430–2434.
72 E. Jung, Ed., *Peptides 1988: proceedings of the 20th European Peptide Symposium, September 4 - 9, 1988, University of Tübingen, Tübingen, Federal Republic of Germany*, 1989.
73 T. Haack and M. Mutter, *Tetrahedron Lett*, 1992, **33**, 1589–1592.
74 M. Mutter, A. Nefzi, T. Sato, X. Sun, F. Wahl and T. Wöhr, *Pept Res*, 1995, **8**, 145–153.
75 T. Wöhr and M. Mutter, *Tetrahedron Lett*, 1995, **36**, 3847–3848.
76 S. Szaniszló, K. Ferentzi, A. Perczel and V. Farkas, *Org Process Res Dev*, 2023, **27**, 1053–1060.
77 V. Goncalves, B. Gautier, F. Huguenot, P. Leproux, C. Garbay, M. Vidal and N. Inguimbert, *J Pept Sci*, 2009, **15**, 417–422.
78 American Peptide Symposium, Pierce Chemical Co., Rockford, Ill, 1983.
79 *13.1 Template-Assembled Synthetic Proteins (II)*, Thieme Verlag, 4th Edition., 2003.
80 M. Amblard, J.-A. Fehrentz, J. Martinez and G. Subra, *Mol Biotechnol*, 2006, **33**, 239–254.
81 T. Johnson, M. Quibell, D. Owen and R. C. Sheppard, *J Chem Soc Chem Commun*, 1993, 369–372.
82 H. M. Yu, S. T. Chen and K. T. Wang, *J Org Chem*, 1992, **57**, 4781–4784.
83 L. Malik, A. P. Tofteng, S. L. Pedersen, K. K. Sørensen and K. J. Jensen, *J Pept Sci*, 2010, **16**, 506–512.
84 I. Coin, M. Beyermann and M. Bienert, 2007, **2**, 3247–3256.
85 K. Sikora, M. Jaśkiewicz, D. Neubauer, D. Migoń and W. Kamysz, *Pharmaceuticals (Basel)*, 2020, **13**, 442.
86 S. M. Kelly, T. J. Jess and N. C. Price, *Biochim. Biophys. Acta*, 2005, **1751**, 119–139.
87 K. Bakshi, M. R. Liyanage, D. B. Volkin and C. R. Middaugh, *Methods Mol Biol*, 2014, **1088**, 247–253.
88 J. Kaminsk, J. Kubelka and P. Bou, *J Phys Chem A*, 2011, **115**, 1734–1742.
89 N. Sreerama and R. W. Woody, *Anal Biochem*, 2000, **287**, 252–260.
90 A. J. Miles, S. G. Ramalli and B. A. Wallace, *Protein Science*, 2022, **31**, 37–46.
91 L. A. Compton and W. C. Johnson, *Anal Biochem*, 1986, **155**, 155–167.
92 G. Holzwarth and P. Doty, *J Am Chem Soc*, 1965, **87**, 218–228.
93 A. G. Fort and D. C. Spray, *Biopolymers*, 2009, **92**, 173–182.
94 A. Bastida, J. Zúñiga, A. Requena, B. Miguel and J Cerezo, *J Chem Inf Model*, 2020, **60**, 6523–6531.
95 A. J. Doig, in *Progress in Molecular Biology and Translational Science*, ed. P. M. Conn, Academic Press, 2008, vol. 83, pp. 1–52.
96 A. M. Rossi and C. W. Taylor, *Nat Protoc*, 2011, **6**, 365–387.
97 O. D. Hendrickson, N. A. Taranova, A. V. Zherdev, B. B. Dzantiev and S. A. Eremin, *Sensors (Basel)*, 2020, **20**, 7132.

98 R. Hancock, H. C. Bertrand, T. Tsujita, S. Naz, A. El-Bakry, J. Laoruchupong, J. D. Hayes and G. Wells, *Free Radic Biol Med*, 2012, **52**, 444–451.
99 D. Inoyama, Y. Chen, X. Huang, L. J. Beamer, A.-N. T. Kong and L. Hu, *J Biomol Screen*, 2012, **17**, 435–447.
100 V. Simov, M. D. Altman, E. Bianchi, S. DelRizzo, E. N. DiNunzio, G. Feng, P. Goldenblatt, R. Ingenito, S. A. Johnson, M. S. Mansueto, T. Mayhood, J. D. Mortison, V. Serebrov, C. Sondey, V. Sriraman, T. J. Tucker, A. Walji, H. Wan, Y. Yue, A. Stoeck and E. F. DiMauro, *Eur J Med Chem*, 2021, **224**, 113686.
101 N. C. Karunatilleke, C. S. Fast, V. Ngo, A. Brickenden, M. L. Duennwald, L. Konermann and W.-Y. Choy, *Int J Mol Sci*, 2021, **22**, 7434.
102 A. Singh, S. Venkannagari, K. H. Oh, Y.-Q. Zhang, J. M. Rohde, L. Liu, S. Nimmagadda, K. Sudini, K. R. Brimacombe, S. Gajghate, J. Ma, A. Wang, X. Xu, S. A. Shahane, M. Xia, J. Woo, G. A. Mensah, Z. Wang, M. Ferrer, E. Gabrielson, Z. Li, F. Rastinejad, M. Shen, M. B. Boxer and S. Biswal, *ACS Chem Biol*, 2016, **11**, 3214–3225.
103 F. Katsuoka, A. Otsuki, M. Takahashi, S. Ito and M. Yamamoto, *Mol Cell Biol*, 2019, **39**, e00273-19.
104 J. D. Neufeld and W. W. Mohn, *Appl Environ Microbiol*, 2005, **71**, 4893–4896.

Chapter 3

Peptide-Directed Ligand Design to Identify Peptides-Small Molecule Hybrids to Inhibit Nrf2/MafG

3.1 Introduction

3.1.1 Targeting PPIs and coiled-coils with small molecules

Designing small molecules that can inhibit PPIs is challenging, as often binding sites are shallow and lack well-defined binding pockets. Consequently, small molecules are most commonly identified through high-throughput screening techniques, which can be expensive and resource intensive.[1] Here we will discuss methods for designing small molecules based on α-helical peptide scaffolds to afford inhibitors of α-helix mediating PPIs.

3.1.1.1 Small molecules designed from α-helices

Utilising structural information of an interaction can support design strategies of small molecule inhibitors.[2] An α-helix mediating a PPI on a singular face can be replaced with non-biological scaffolds to produce small molecules that can interact at a PPI interface. Within a helical peptide the *i, i+3/4* and *i*+7 positions can be modified to impose key residues onto an organic backbone.[3,4] Rational inhibitor design for the helical p53/*h*DM2 (**3.1**) and Bcl-x_L/Bad (**3.3**) interactions has been achieved with μM activity using benzamide (**3.2**) or imidazole-phenyl (**3.4**) structures to mimic the α-helical backbone of a known peptide inhibitor (**Figure 3-1**).[2,5]

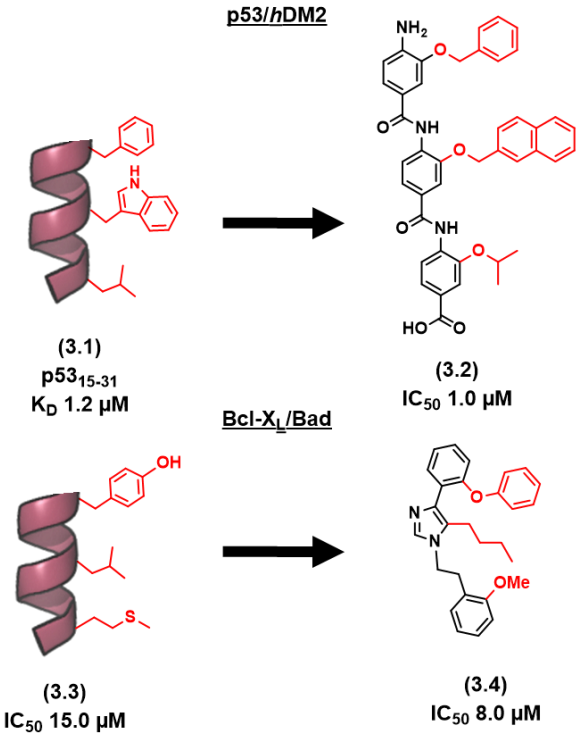


Figure 3-1 *α-helix mimicking backbones for designing small molecule inhibitors of the p53/hDM2 and Bcl-X_L interactions.*[2,5,6]

A more directed approach can be achieved by performing fragment based screening, helping to identify potential PPI hotspots to be exploited by further SAR.[7] When a hotspot residue of a PPI is known, anchor based small molecule design can be performed, searching for bioisosteres to design small molecule inhibitors, a technique that has been applied to α-helical PPIs. An example of this is the p53 α-helix binding to *h*DM2 (**Figure 3-2**), where a key tryptophan residue on p53 was used to perform substructure modifications leading to a hit spiro-oxindole (**3.5**) structure with a potent K_i of 0.88 nM.[8]

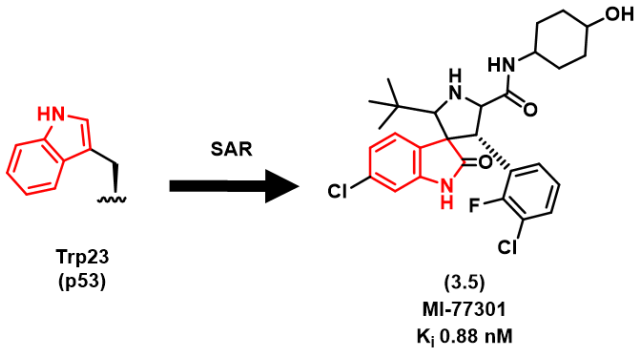


Figure 3-2 *Anchor based design of p53/hDM2 inhibitor MI-77301 (**3.5**) containing spiro-oxindole core (red) originating from hotspot residue Trp23 on p53.*

An alternative drug design strategy utilising the knowledge of known key residues within an α-helix mediated interaction is the REplacement with Partial Ligand Alternatives through Computational Enrichment (REPLACE) strategy.[9] First described in 2006, REPLACE was used to identify inhibitors of CDK2/cyclin A, applying structural information of a peptide based on the cyclin binding protein P21 (**3.7**). Sequence truncation (**3.8**) identified an arginine residue responsible for key charged interactions, non-peptidic structures were subsequently explored (**3.9**), generating peptide-small molecule hybrids with comparable inhibitory activity to the parent peptide sequence (Arg-Arg-Leu-Asn-(*p*-F-Phe)-NH_2) (**Figure 3-3**).[9]

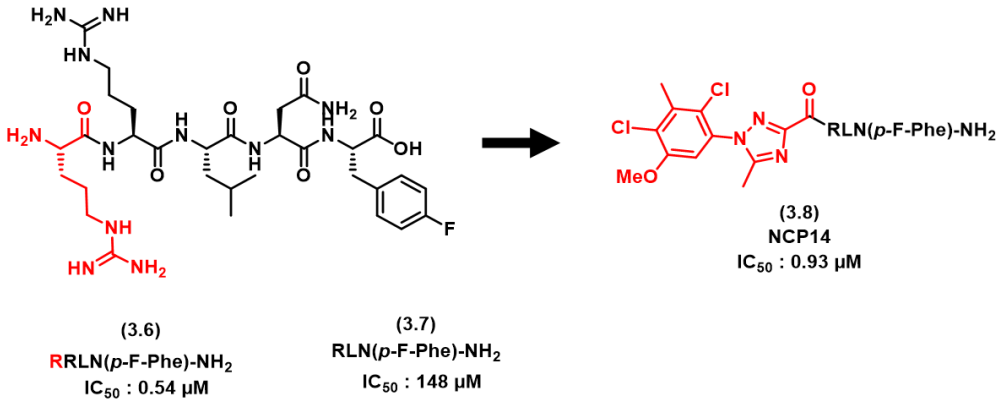


Figure 3-3 *Structure of peptide and peptide-small molecule hybrid identified through REPLACE technique substitution (red) of arginine for an aromatic scaffold for inhibition of the Cyclin2/CDK2 PPI.*

3.1.1.2 Coiled-coil small molecule inhibitors

Despite advances in small molecule design for PPI inhibition, success in targeting coiled-coil interactions has been limited. Existing small molecules target larger coiled-coil structures such as helical bundles, primarily the six-helical bundle of the viral entry proteins found in HIV-1 (**Figure 3-4**).[10–12] Helical bundle formation has been prevented by stabilising dimeric coiled-coil interactions with small molecules inhibitors. **3.9** stabilises the dimeric coiled-coil interactions between helical repeat 1 and 2 by a key π-π stack and hydrogen bonding interactions. **3.9** repels helical bundle formation through residue displacement to prevent hydrogen bond formation with additional helical repeat 1 units (**Figure 3-4, A**).[12]

8-anilinonaphthalene-1-sulfonic acid (ANS, **3.10**) has been found to bind a conserved hydrophobic pocket on the N-terminal helical bundle, displacing C-terminal peptide binding by displacing key tryptophan residues (**Figure 3-4, B**).[10,11]

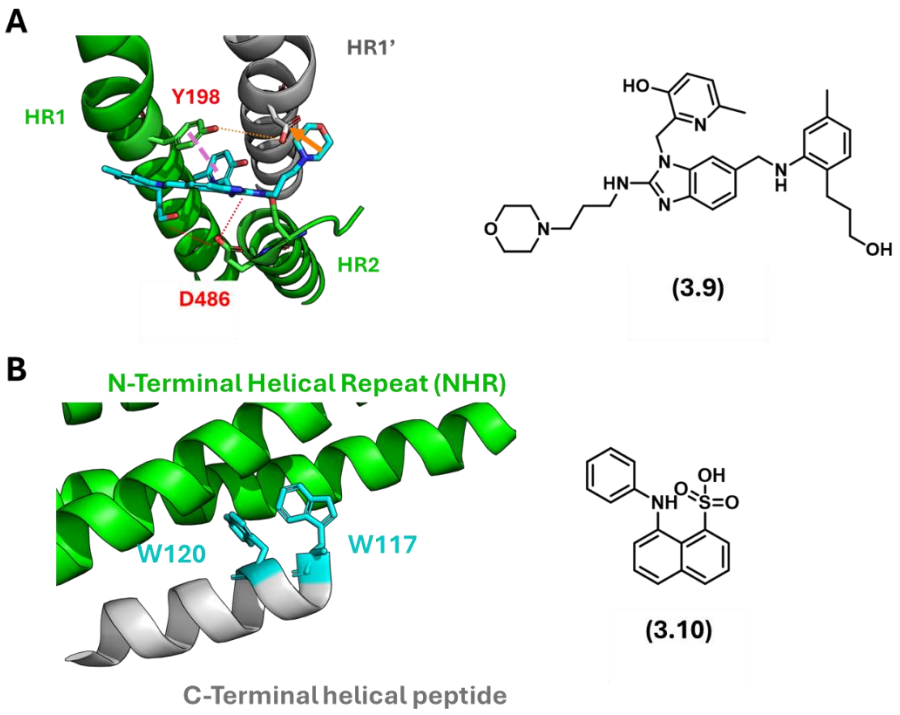


Figure 3-4 *Small molecules that disrupt helical bundles. (**A**) TMC353121 (**3.9**) (cyan) a benzimidazole containing structure that disrupts helical bundle formation between HR1 and HR2 of gp41, HIV-1, (Merge of PDB: 3KPE and 1G2C) interacting helices shown in green, HR1' displaced shown in grey, hydrogen bond interactions shown in red, π- π interaction in purple, displaced residue of HR1' indicated by an orange arrow. (**B**) 8-anilinonaphthalene-1-sulfonic acid (ANS, **3.10**) a fluorescent probe that binds to a hydrophobic pocket of the N-terminal helical bundle in gp41 (green), preventing C-terminal helix binding (grey) by displacing binding sites of tryptophan residues (cyan) (PDB: 6R2G).*

Inhibition of dimeric coiled-coils in transcription factors has also been achieved by small molecules against the Myc/Max interaction (**Figure 3-5**). **3.11** was found through HTS reaching low nM affinity. There is no experimentally determined structural evidence

confirming if small molecules directly target the coiled-coil or the loop regions of Myc to prevent protein dimerisation.[13,14] Induced fit docking simulations have suggested that the KJ-Pyr-9 scaffold binds at the coiled-coil region primarily to Myc in close proximity to the loop-helix domain (**Figure 3-5**).[15]

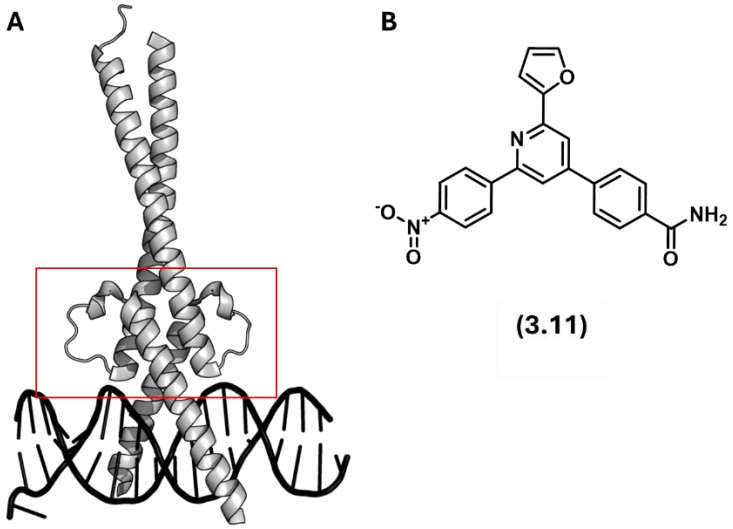


Figure 3-5 (A) *Myc/Max/DNA structure (PDB: 1NKP),* ***(B)*** *Myc inhibitor KJ-Pyr-9 (**3.11**) proposed binding site highlighted in red.*

Inhibitors against the MDB2/p66α dimeric coiled-coil have been identified using an *in silico* screening platform.[16] A compound library was used to perform molecular docking and dynamics simulations against the MBD2 α-helix (**Figure 3-6**). Of ten compounds validated experimentally, two were inhibitory, achieving low μM IC_{50} values against the PPI. This is an example of streamlining the efficiency of high throughput screening to identify inhibitors towards a coiled-coil interaction.

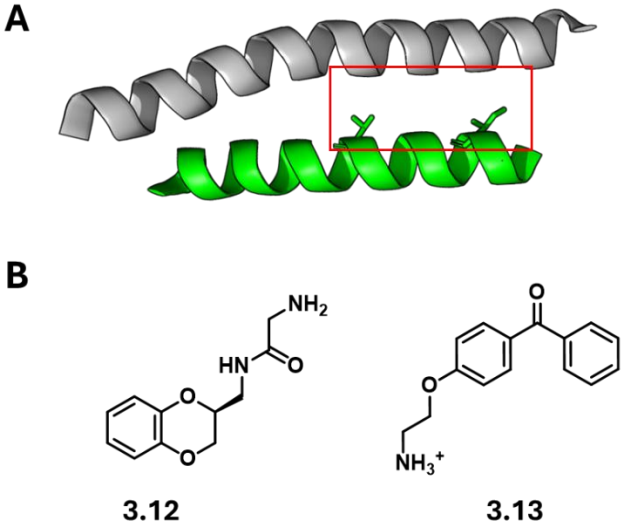


Figure 3-6 *(**A**) MDB2/p66α interaction (PDB: 2L2L), (**B**) Small molecule inhibitors (**3.12-3.13**) proposed binding site highlighted in red.*

3.1.2 Peptide-directed ligand design

An emerging technique for efficient small molecule fragment discovery is the use of peptide-directed ligand design (PDLD).[9–12] Described in **Figure 3-7**, PDLD has been utilised to perform *in silico* screens of alkyne and azide fragments against α-helical peptide scaffolds to afford potent inhibitors against Noxa/Mcl-1 and p53/*h*DM2 or p53/*h*DMx helix in groove interactions.[17,18]

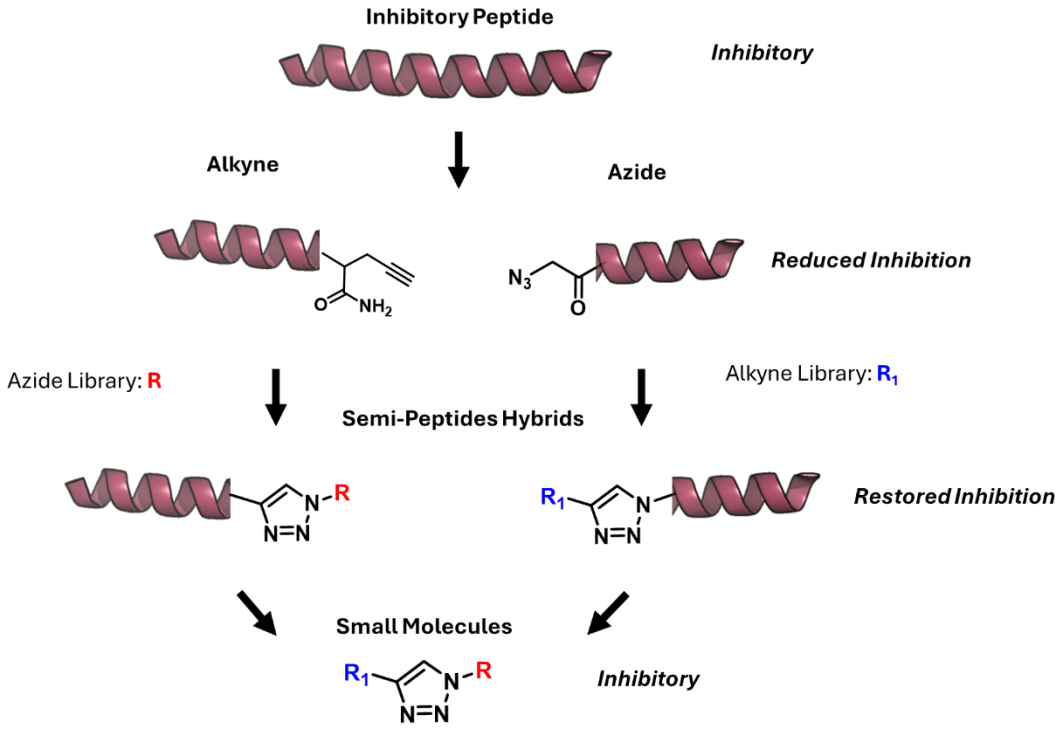


Figure 3-7 *Overview of peptide-directed ligand design for α-helix mediated PPI inhibitors.*

PDLD against Noxa/Mcl1 was performed using the α-helical NoxaB peptide (AAQLRRIGDKVNLRQKLLN, IC_{50} of 650 nM).[19] By dividing the peptide in half, two key binding residues were conserved in each peptide. Reactive handles were installed to perform copper-catalysed azide-alkyne cycloaddition (CuAAC) of small molecule fragments, creating a library of peptide-hybrids for investigating inhibitory activity against. Guided by *in silico* peptide-hybrid screening, 35 small molecules were assembled, 50% producing an IC_{50} of <100 μM as determined by fluorescence anisotropy.[17]

Similarly, the *h*DM2/*h*DMX interaction with p53 was explored utilising an 8-mer peptide (Ac-Phe-Met-Aib-Pmp-6-Cl-Trp-Glu-Ac3c-Leu-NH_2,[*] K_D 36 nM for *h*DM2, 7 nM for *h*DMX). Dividing the sequence in half (four amino acids in each reactive half), performing virtual CuAAC docking and subsequent triazole small molecule docking. This led to the synthesis of the top ten small molecules for the two proteins identifying hit small molecules. Again, 50% of the twenty small molecules prepared disrupted the *h*DMX/p53 interaction as determined by competition fluorescence polarisation assay against the p53 peptide.[18]

The lack of well-evidenced small molecule inhibitors of dimeric coiled-coil interactions highlights a need to utilise novel structural design methods to identify inhibitor scaffolds for this type of protein-protein interaction. The ability of PDLD to identify small molecules with highly efficient hit rates against a target protein-protein interaction is exciting, as it offers huge improvements against standard high throughput screening methods to identify PPI small molecule inhibitors. Consequently, it is an aspiring technique to identify inhibitors against PPIs where no small molecules modulators have been reported.

3.2 Chapter Aims

Using crystal structure information of the Nrf2/MafG/ARE ternary complex (PDB: 7X5F), this chapter reports the use of peptide-directed ligand design to develop peptide-small molecule hybrids through triazole click chemistry.

Utilising the leucine zipper sequence of MafG, $_{82}$KAELQQEVEKLASENASMKLE$_{104}$, found to have inhibitory properties against the Nrf2/MafG interaction in **Chapter 2,** small molecule peptide hybrids were investigated *in silico*. High scoring fragments were then synthesised as peptide-hybrids and experimentally evaluated by the fluorescence polarisation and electrophoretic mobility shift assays described in **Chapter 2.**

A small library of fragments, representative of a broad chemical space were also prepared for peptide-small molecule hybrid synthesis and investigated by fluorescence polarisation and electrophoretic mobility shift assays.

[*] Aib: α-aminoisobutyric acid. Ac_3c: 1-amino-cyclopropanecarboxylic acid. Pmp: phosphonomethylphenylalanine

3.3 Results and Discussion

3.3.1 *In silico* design of peptide-small molecule hybrids

3.3.1.1 Non-covalent docking of small molecule hybrids

Using the Nrf2/MafG crystal structure (PDB: 7X5F), Nrf2 and interacting MafG sequence $_{82}$KAELQQEVEKLASENASMKLE$_{104}$ were prepared in Schrodinger Maestro suite to support the docking of small molecule fragments to the Nrf2 protein.[20] As hot-spot residues of the MafG peptide sequence have not been identified, the sequence was divided into two halves $_{82}$KAELQQEVEK$_{93}$ and $_{94}$LASENASMKLE$_{104}$ to provide even distribution of *a* and *d* hydrophobic positions of the coiled-coil (**Figure 3-8**).

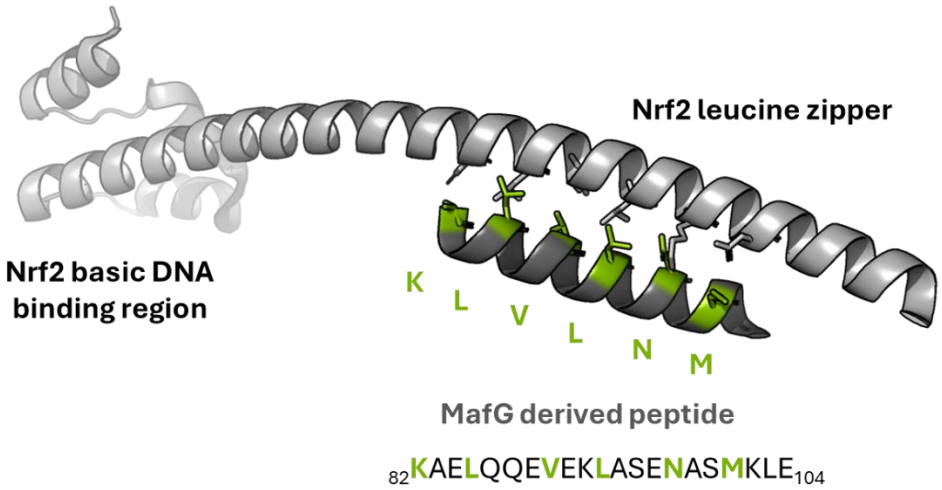


Figure 3-8 *Structure of Nrf2 Neh1 domain (light grey) with MafG truncated to peptide sequence $_{82}$KAELQQEVEKLASENASMKLE$_{104}$ (Dark Grey,)* ***a*** *and* ***d*** *positions are shown as sticks, highlighted in green, prepared in PyMOL from PDB 7X5F.*

The C-terminus of $_{82}$KAELQQEVEK$_{93}$ was modified with the addition of propargylglycine after Lys93, similarly $_{94}$LASENASMKLE$_{104}$ was modified at the N-terminus of Leu94 with azidoacetic acid (**Figure 3-9**) in preparation for covalent docking of small molecule fragments. [21]

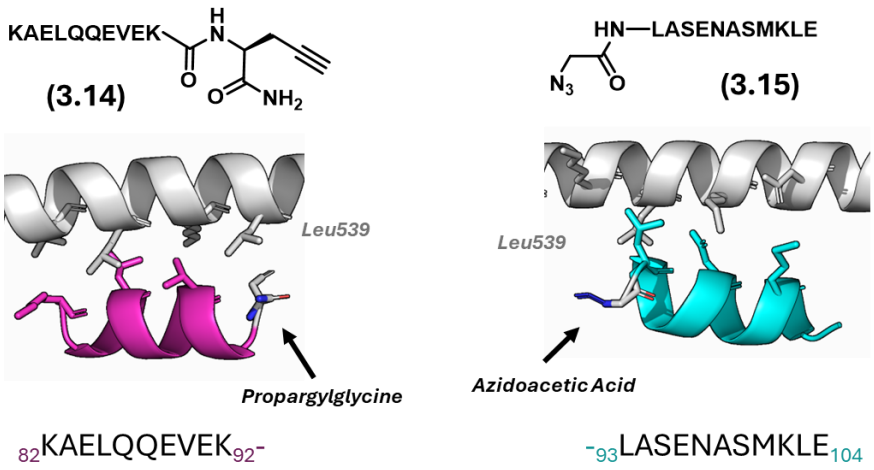


Figure 3-9 *Structures of KAELQQEVEK-Pra (**3.14**) and N_3Ac-LASENASMKLE (**3.15**)[†] Nrf2/MafG (PDB: 7X5F) Nrf2 (white), $_{82}$KAELQQEVEK$_{92}$ (magenta) and $_{93}$LASENASMKLE$_{104}$ (cyan).*

Azide and alkyne libraries were selected using the Enamine and ZINC library collections, with a preference for compounds that cover a large chemical space and structural diversity.[22,23] In total, 33,418 alkyne fragments and 3,347 azide fragments were prepared using the LigPrep function in Maestro to capture the relevant 3D structural conformations of each fragment.[24]

Non-covalent docking was used to screen the fragment libraries using grid based ligand docking with energetics (glide) scoring, a fragment screening tool reported to provide rapid ranking of ligand docking poses.[25] This allowed us to reduce the selection pool to the top 200 azide and alkyne fragments, where highly negative docking scores are representative of a strong binding interaction between the protein and the small molecule.

The 200 azides or alkynes were then covalently docked to the corresponding peptide, forming a triazole bond between the peptide and small molecule. The top ten azides and alkynes produced are summarised in **Table 3-1** and **3-2.** Example docking interactions of the top azide and alkyne fragment are highlighted in **Figure 3-10**. Docking of **3.26** to KAELQQEVEK suggests a π-stacking interaction between the triazole and Lys_{543} of Nrf2, which would occupy an "*a*" position of the leucine zipper. Additionally, a H-bond interaction was suggested to form with Glu_{542} occupying a "g" position of Nrf2 during coiled-coil formation. Docking proposes that **3.16** could form a salt-bridge with Lys_{536} which also occupies an "*a*" position within the Nrf2 leucine zipper (**Figure 3-10, B**).

[†] **Pra**: Propargylglycine **N_3Ac:** Azidoacetic acid

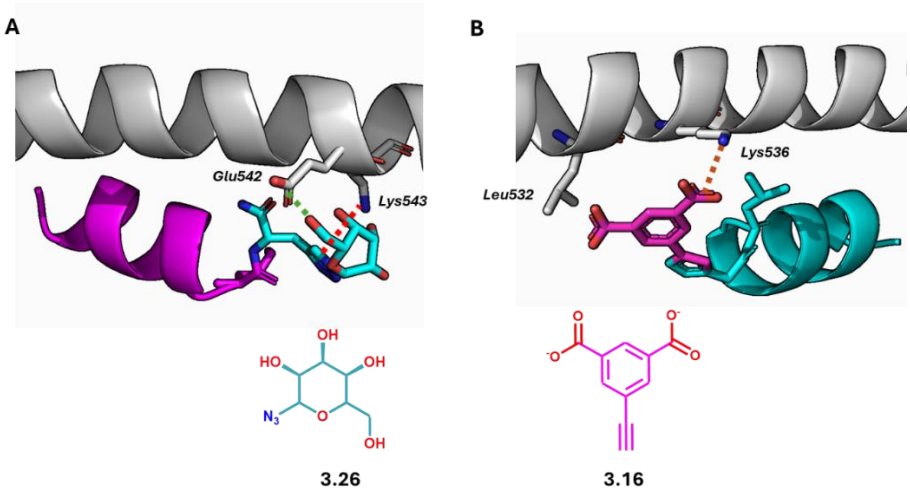


Figure 3-10 *Top binding poses of (**A**) **3.26** (cyan) MafG peptide KAELQQEVEK in magenta. (**B**) **3.16** (magenta) MafG peptide LASENASMKLE (cyan) Nrf2 in grey. H bonding (green), π-stacking (red), salt bridge (orange).*

Table 3-1 *Top 10 from small molecule covalent docking simulations across alkyne fragments.*

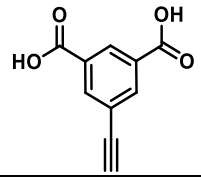
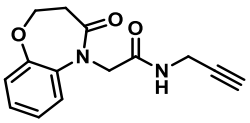
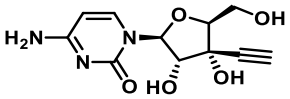
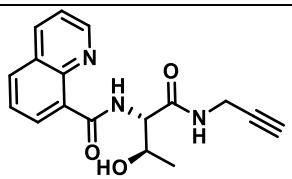
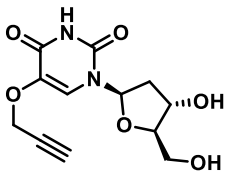
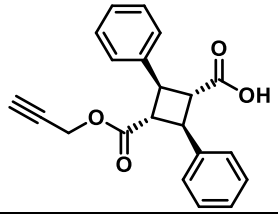
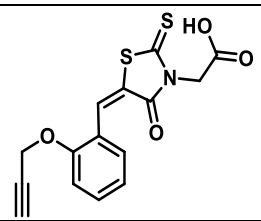
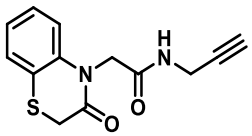
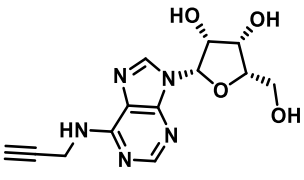
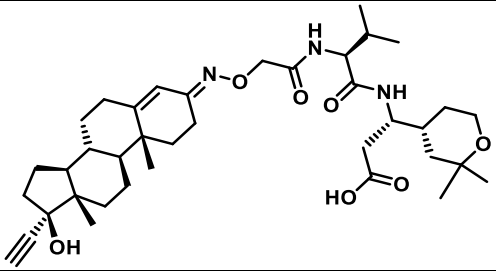
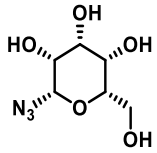
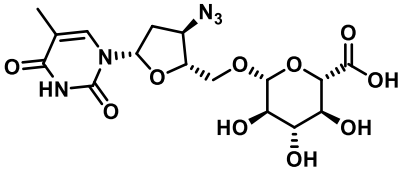
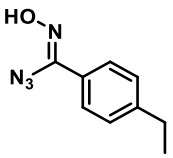
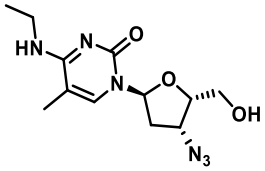
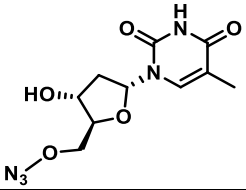
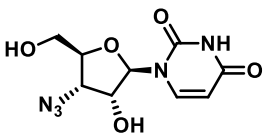
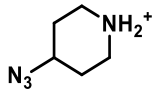
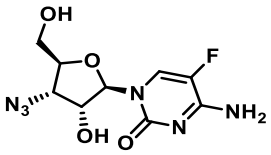
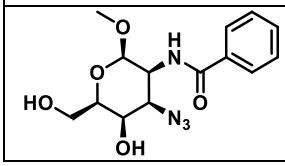
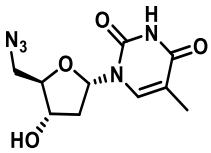
Ac-*KAELQQEVEK-R*			
Structure	**(Nr) cdock Score**	**Structure**	**(Nr) cdock Score**
	**(3.16)** -3.895		**(3.21)** -3.041
	**(3.17)** -3.809		**(3.22)** -2.975
	**(3.18)** -3.657		**(3.23)** -2.947
	**(3.19)** -3.522		**(3.24)** -2.913
	**(3.20)** -3.195		**(3.25)** -2.782

Table 3-2 *Top 10 azides from covalent docking simulations performed across 3347 azide fragments.*

R-LASENASMKLE-NH$_2$			
Structure	**(Nr) cdock Score**	**Structure**	**(Nr) cdock Score**
	**(3.26)** -4.546		**(3.31)** -2.898
	**(3.27)** -3.175		**(3.32)** -2.824
	**(3.28)** -3.104		**(3.33)** -2.622
	**(3.29)** -3.061		**(3.34)** -2.609
	**(3.30)** -2.912		**(3.35)** -2.595

3.3.1.2 Selected small molecule fragments from in silico screening

Following an assessment of commercial availability of the top docking fragments, four alkynes and three azides (**Figure 3-11**) were identified to be cost feasible (under £250 for 50 mg). Consequently, the library of peptide hybrids was expanded by utilising an in-house inventory of alkynes and azides previously used for identification of helix-in-groove mediated interactions **(Figure 3-12).**[17–19]

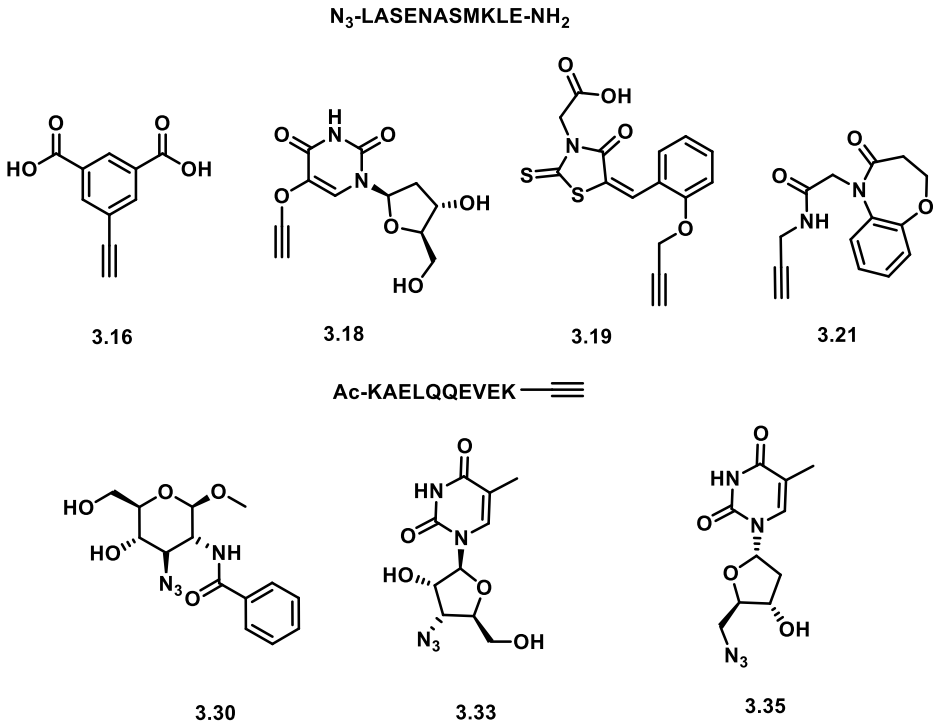


Figure 3-11 *Commercially viable docking fragments for synthesis.*

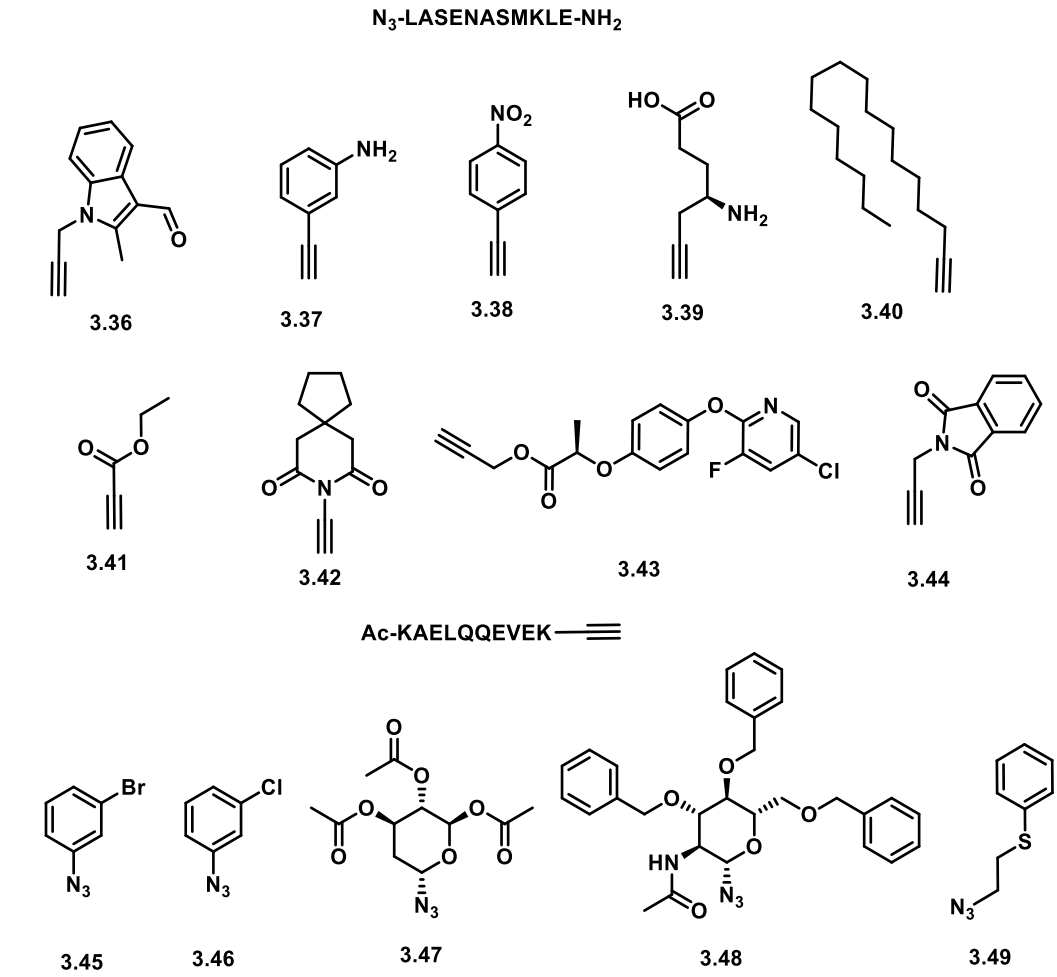


Figure 3-12 *In-house selected alkyne and azide fragments selected for peptide-small molecule hybrid synthesis.*

3.3.2 Peptide-small molecule hybrid synthesis

3.3.2.1 Synthesis of peptide (Ac-KAELQQEVEK-) and small molecule hybrids

Synthesis of (Ac-KAELQQEVEK**Pra**-NH_2[‡]**)** **(Figure 3-13, 3.14)** was achieved using microwave peptide synthesis on Tentagel S RAM resin (0.24 mmol/g) of the ten amino acid sequence incorporating Fmoc-propargylglycine (Fmoc-Pra-OH) at the C-terminus. To confirm synthesis a small resin cleavage was performed using TFA as previously described, the expected mass of 1336.7 Da $[M+H]^+$ was found, correlating to a peak observed by analytical HPLC at 9.9 min.

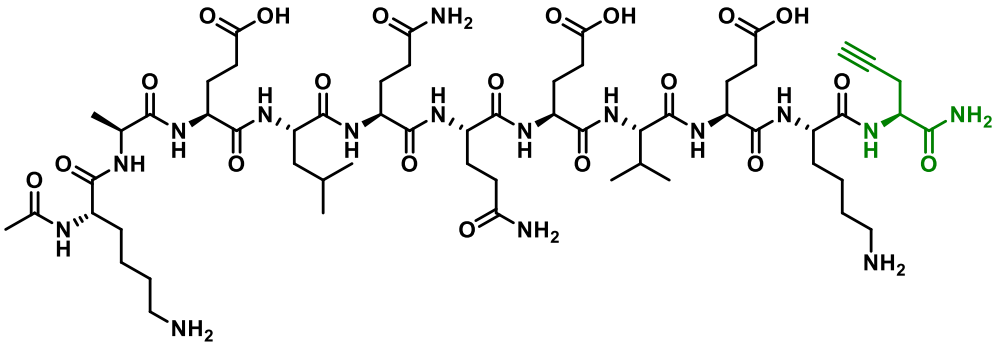


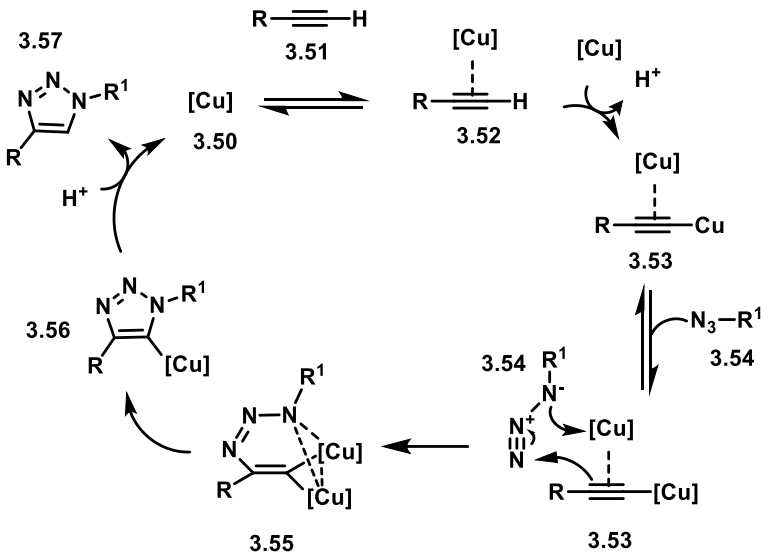
Figure 3-13 *Structure of **3.14** with incorporated Pra residue (green)*

Following the synthesis of **3.14**, azide fragments were reacted with the peptide on resin. Using the N-terminal alkyne handle using click chemistry to form a triazole linker between the peptide and small molecule fragment.

Click chemistry, first described by Sharpless *et al* in 2001. is a method of convenient covalent bond formation under mild conditions.[26] The best known example of click chemistry is the Cu(I) catalysed azide alkyne-cycloaddition (CuAAC) first reported in 2002 (**Scheme 3-1**).[27] CuAAC describes the use of a regioselective Cu(I) catalyst, generated by Cu(I) salts or a reduction of Cu(II) using sodium ascorbate to support the formation of a 1,4-disubstituted 1,2,3-triazole between a monosubstituted alkyne (**3.51**) and an organic azide (**3.54**). Although the mechanism is somewhat contentious, the literature commonly uses the mechanism presented by Fokin.[28–30] Cu(I) (**3.50)** coordinates with the alkyne to generate a copper acetylide (**3.52**) through a deprotonation, which further forms a weak π bond interaction with a second Cu(I) (**3.53).** This produces a catalytically favourable complex, in which reversible coordination with an organic azide (**3.54**) leads to formation of a copper(I) triazolide intermediate (**3.55**). A

[‡] Pra = Propylargylglycine

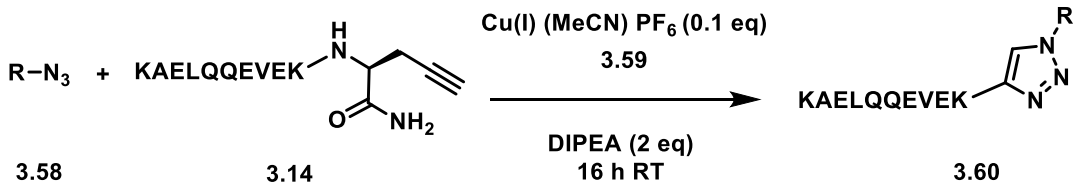
ring contraction on the triazolyl stabilised copper derivative (**3.56**) leads to protonolysis yielding a 1,4-triazole (**3.57**) and the release of Cu(I) (**3.50**).[28]



Scheme 3-1 *CuAAC reaction adapted from Fokin et al. 2013.*[28]

There is evidence of peptide degradation by active copper species during CuAAC reactions, adding a level of difficulty in achieving peptide conjugation by CuAAC.[31] The use of ascorbic acid to reduce Cu(I) can also generate reactive oxygen species that create off target reactions with amino acids during bioconjugation.[32] As such, the development of Cu(I) organic complexes such as $Cu(MeCN)_4PF_4$ (**Scheme 3-2, 3.59**) have been reported to remove the need for reducing agents and harsh basic conditions for CuAAC to proceed.[33]

The CuAAC reaction between on resin **3.14** and azide fragments (**Scheme 3-2, 3.58)** was performed at room temperature overnight with agitation, followed by peptide cleavage in TFA with DODT, H_2O and TIPS.



Scheme 3-2 *Click reaction of azide fragments with **3.14** to form peptide-small molecule hybrids.*

Peptides were synthesised, cleaved and analysed by HPLC to determine T_R values; the correct expected masses were found by MALDI-TOF, in total seven peptide-small molecule hybrids were prepared (**Figure 3-14, Table 3-3**). Problems were found with handling the product of benzamide (**3.30**) cycloaddition, which frequently precipitated

in MeCN, MeOH or H_2O. This perhaps suggesting the peptide-small molecule hybrid would not have been viable for use in biophysical assays prepared, as such it was not explored further.

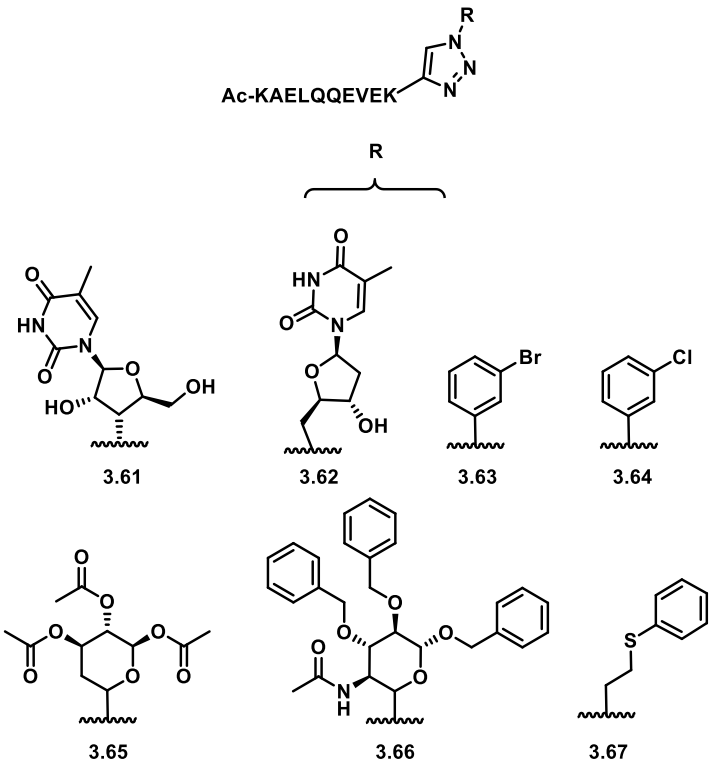
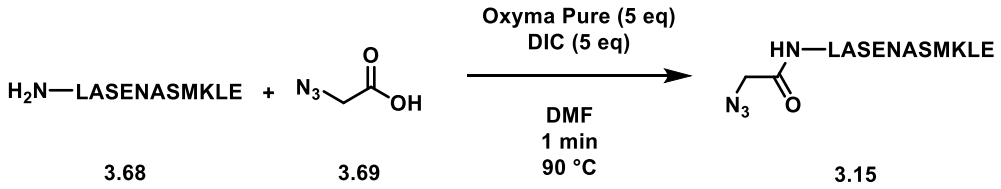


Figure 3-14 *Peptides-small molecule hybrids synthesised by CuAAC.*
Table 3-3 *Peptide small molecule hybrids synthesised by CuAAC with* ***3.14.***

(Ac-KAELQQEVEK-R)	
R-	**Yield (%)**
3.61	11%
3.62	17%
3.63	21%
3.64	24%
3.65	28%
3.66	23%
3.67	12%

3.3.2.2 Synthesis of peptide (N_3Ac-LASENASMKLE-NH_2) and small molecule hybrids

Synthesis of *N_3Ac-LASENASMKLE-NH_2* **(3.15**, **Figure 3-15)** was achieved using microwave peptide synthesis of the 11 amino acid sequence (**3.68**) followed by an acetylation reaction (**Scheme 3-3**), capping the N-terminus using 4 equivalents of azidoacetic acid (**3.69**) in a microwave assisted coupling at 90 °C for 1 min using Oxyma Pure (5 eq), DIC (5 eq) in DMF. The reaction successfully proceeded, resolving the expected $[M+H]^+$ of 1274.4 m/z by MALDI-TOF and singular peak by analytical HPLC at 11.1 min.



Scheme 3-3 *Azidoacetic acid (**3.69**) coupling to the N terminus of peptide (**3.68**) to form **3.15***

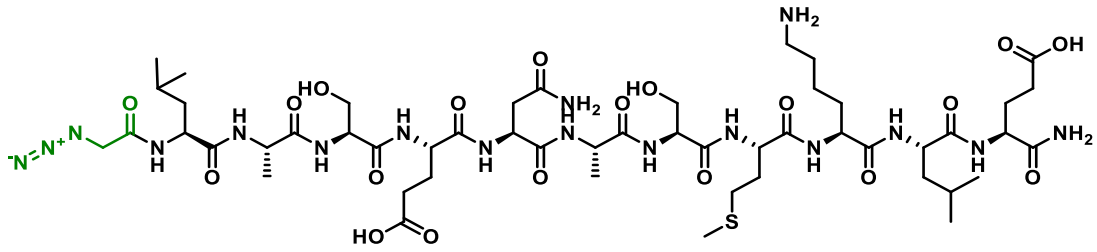
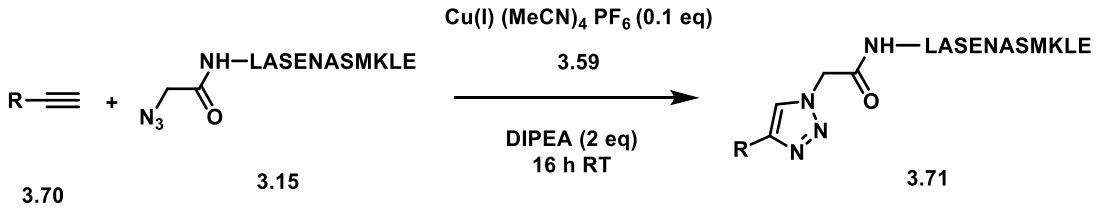


Figure 3-15 *Structure of **3.15** capped with azidoacetic acid.*

The synthesis of peptide-small molecule hybrids using alkyne fragments **(Scheme 3-4, 3-9.53)** was achieved using 0.1 equivalents of tetrakis(acetonitrile) copper(I) hexafluorophosphate (**3.75**) and DIPEA (2 eq) agitated overnight with **3.15** on resin, the resin was then washed extensively with DMF, DCM and MeOH following a cleavage in TFA with DODT, H_2O and TIPS for 3 hours. After cleavage, the peptide-small molecule hybrids (**Figure 3-16**) were subjected to semi-preparative purification to analyse the crude content for the desired product. Correct mass and corresponding T_R were found for thirteen hybrid molecules with sufficient yields to proceed with an investigation of peptide-small molecule hybrids by fluorescence polarisation assay (**Table 3-4**).



Scheme 3-4 *Click reaction of docking fragments to **3.15** to form peptide-small molecule hybrid (**3.74**).*

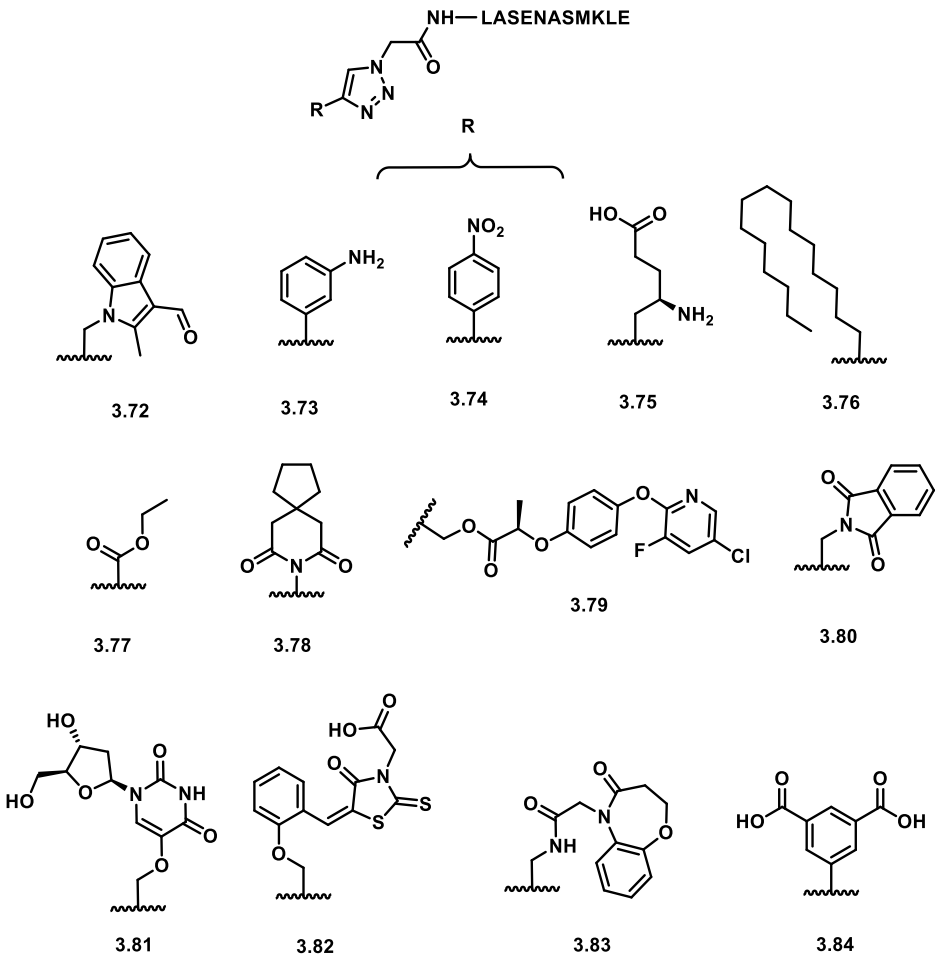
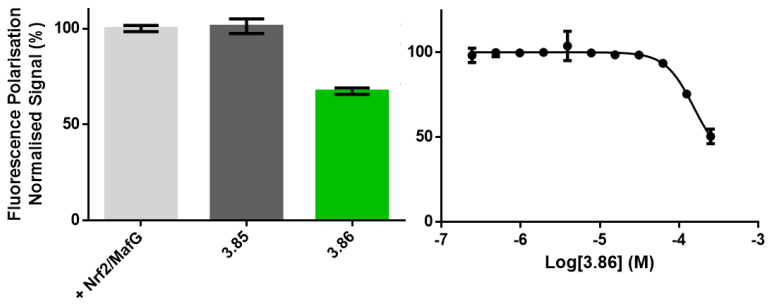


Figure 3-16 *Peptide-small molecule hybrids synthesised by CuAAC.*
Table 3-4 *Peptide-small molecule hybrids synthesised by CuAAC with* ***3.15.***

(R-LASENASMKLE-NH_2)	
R-	**Yield (%)**
3.72	13%
3.73	10%
3.74	25%
3.75	21%
3.76	5%
3.77	12%
3.78	22%
3.79	21%
3.80	13%
3.81	16%
3.82	13%
3.83	17%
3.84	14%

3.3.3 Effect of peptide-small molecule hybrids on Nrf2/MafG/DNA complex

The peptide-small molecule hybrids prepared were assessed using the fluorescence polarisation assay and EMSA established in **Chapter 2** to determine if inhibitory activity could be observed against the Nrf2/MafG/DNA interaction. Initially, Ac-KAELQQEVEK-NH_2 (**3.85**) and Ac-LASENASMKLE-NH_2 (**3.86**) were screened at 125 µM to investigate if inhibitory activity was lost by truncating the peptide sequence. Interestingly **3.86** continued to inhibit the Nrf2/MafG/DNA complex (**Figure 3-17**) producing an 18% reduction compared to Nrf2/MafG at 125 µM. This was unexpected, as the 14-mer peptide **2.41** investigated in **Chapter 2** generated no inhibitory activity, despite containing only three additional residues than **3.86.** Subsequently, an IC_{50} titration was performed of **3.86** at 16 h (**Figure 3-17**) which determined an IC_{50} of approximately 149.3 µM (95% CI 98.3 -226.7 µM). **3.86** produced 3-fold weaker inhibition, compared to **2.52** (Ac-KAELQQEVEKLASENASMKLE-NH_2) (IC_{50} of 36 µM).



3.85 Ac-KAELQQEVEK-NH_2 **3.86** Ac-LASENASMKLE-NH_2

Figure 3-17 *(left) FP screen of* ***3.85*** *and* ***3.86*** *at 125 µM (right)* ***3.86*** *titrated from 250 µM after 16 h, 4 °C incubation with 62.5 nM Nrf2, 50 nM MafG and 10 nM FAM ARE DNA.*

Twenty peptide-small molecules hybrids were screened at 125 µM. Two promising peptide hybrids derived from **3.15** were identified, **3.72** produced an 85% reduction in protein-DNA binding and **3.78** reduced binding by over 50% (**Figure 3-18**). Of the six *in silico* peptide-small molecule hybrids screened, inhibition of the Nrf2/MafG interaction was found in one hybrid, **3.82** produced a 10% reduction in DNA binding. The minimal success of *in silico* screened hybrids was perhaps limited by high cost of lead fragments limiting availability for synthesis. Peptides derived from **3.14** were also less successful in inhibiting ternary complex formation, and **3.65** produced a 45% reduction whilst **3.63** reduced the signal by 10%. Investigation of whether the hybrid molecules influence

baseline fluorescence found no significant change compared to the fluorophore control signals **(Appendix 3-1)**.

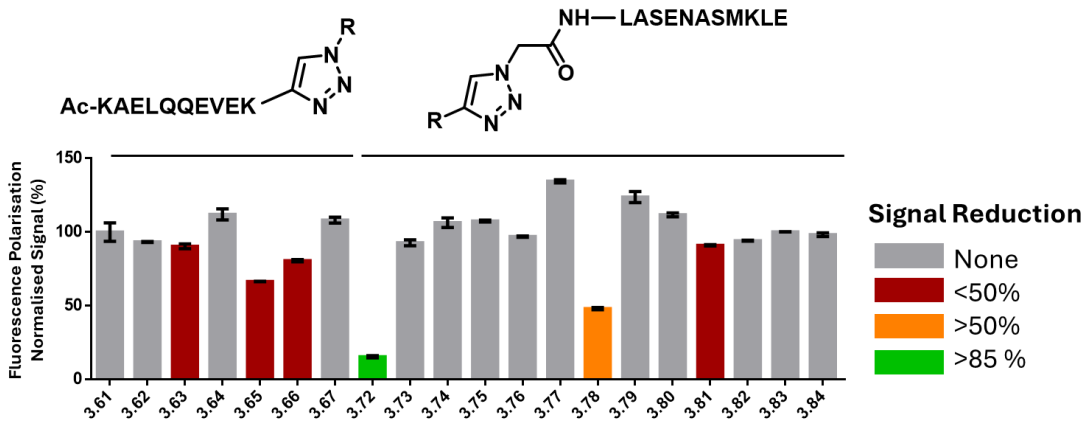


Figure 3-18 *Small-molecule peptide hybrids screened at 125 µM against Nrf2/MafG/ARE at 16 h 4 °C normalised against the average signal of Nrf2/MafG/ARE binding complex (100%)*

Hybrids demonstrating inhibition at 125 µM were then titrated from 250 µM, **3.72** and **3.78** demonstrated a concentration-dependent response allowing for calculation of an IC_{50} (**Figure 3-19). 3.72** and **3.78** demonstrated an IC_{50} of 62.14 µM (95% CI 45.24 -76.6 µM) and 127.1 µM (95% CI 88.18 -183.2 µM) respectively. Improving on the IC_{50} of **3.86**, suggesting the fragments have an effect. It is noted that **3.72** and **3.78** led to a complete reduction of protein-DNA produced fluorescence, implying inhibition of MafG homodimer/DNA complex as well. Despite an initial decrease observed by **3.63, 3.65, 3.66 and 3.82,** no concentration dependent inhibition was observed from 250 µM.

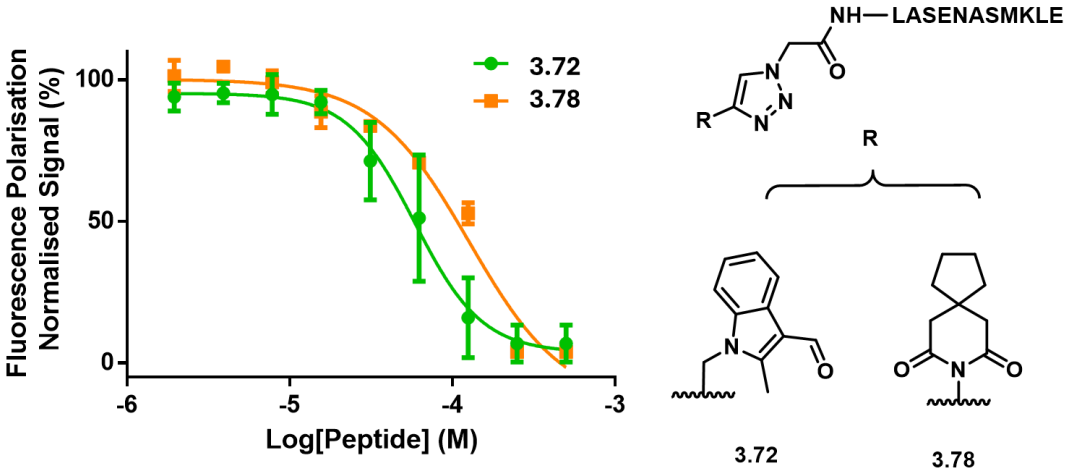


Figure 3-19 *Fluorescence polarisation assay of **3.72** and **3.78**. Measurements taken at 16 h, 4 °C against 62.5 nM Nrf2, 50 nM MafG and 10 nM FAM labelled ARE.*

An EMSA was carried out, to determine visually if the peptide-small molecule hybrids reduced the Nrf2/MafG complex band intensity. Gel images of **3.14** hybrids (**Figure 3-20**) found Nrf2/MafG band visualisation was weak. The DNA band however was absent in all

four peptides suggesting, in line with the fluorescence polarisation assay, that no inhibition of the ternary complex was found. The EMSA of peptide-small molecule hybrids based on **3.15** showed greater protein complex visibility (**Figure 3-20**) and found **3.72** and **3.78** did indeed reduce the protein/DNA complex causing a complete restoration of the free DNA band. Further implying that both Nrf2/MafG and MafG/MafG complexes are inhibited by these hybrid compounds. Interestingly, **3.81** and **3.82** also appeared to produce inhibition of the protein/DNA complex band, inconsistent with the results found in the FP assay.

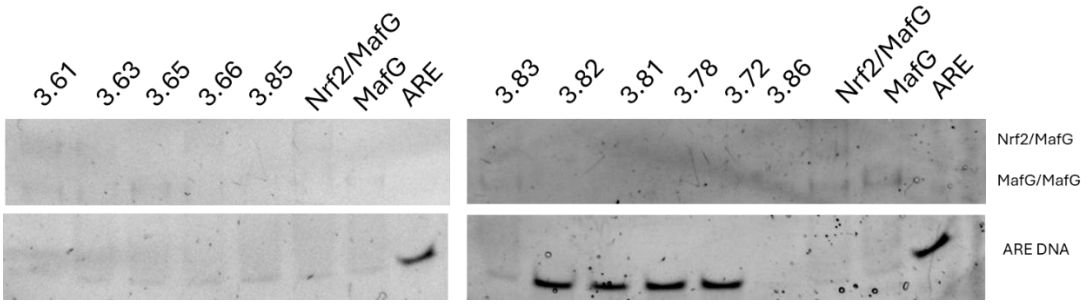


Figure 3-20 *EMSA of peptide* ***3.85*** *and peptide-small molecule hybrids(right). Peptide* ***3.86,*** *and peptide-small molecule hybrids (left). 125 μM peptide, 62.5 nM Nrf2, 50 nM MafG and 10 nM FAM labelled DNA.*

3.4 Conclusion

Peptide-directed ligand design has previously demonstrated success in the disruption of helix in groove interactions to identify small molecules capable of competing with α-helical peptides derived from p53 and Noxa proteins.[17–19] Currently, there are only a few small molecules that directly disrupt dimeric coiled-coil interactions, found by high-throughput screening of large compound libraries.[11,13,16] However, peptide-directed ligand design offers a promising strategy to explore small molecule inhibitor development with greater efficiency.

In **Chapter 2**, **2.52** (Ac-KAELQQEVEKLASENASMKLE-NH_2) was presented as an inhibitor of the Nrf2/MafG coiled-coil interaction. **2.52** disrupts Nrf2/DNA binding, demonstrated by fluorescence polarisation and EMSA biophysical assessments, most likely due to coiled-coil PPI inhibition. As such, the peptide can be used as a scaffold to perform peptide-directed ligand design to afford small molecule fragments capable of disrupting the coiled-coil interaction.

Identifying peptide small molecule hybrids was first investigated by an *in silico* covalent docking screen using alkyne and azide fragments designed to cover a wide range of chemical space.[22] Efforts to select potent inhibitory fragments by *in silico* screening were hampered by cost viability of top fragments. Future work would look to employ

further *in silico* screening with ligands of greater commercial availability or those that can be readily synthesised.

The preparation of **2.52** for peptide-directed ligand design revealed that **3.86 (**Ac-LASENASMKLE-NH_2) possessed inhibitory activity for the Nrf2/MafG interaction (IC_{50} of 149 μM), providing an interesting insight into the minimum peptide length for peptide based coiled-coil inhibition **(Figure 3-17)**. Future work could involve further exploration of peptide-directed ligand design against the shorter peptide, **3.86**.

Derived from peptide **3.86** (IC_{50} of 149 μM) two peptide-small molecule hybrids demonstrated improved activity towards the ternary complex of Nrf2/MafG/DNA**. 3.72** (IC_{50} of 62 μM) and **3.78** (IC_{50} of 125 μM) (**Figure 3-19**). These peptides demonstrated complete inhibition of all protein binding to DNA, which suggests the MafG homodimer binding to DNA was also abolished. Analysis of these samples by EMSA further confirms this as incubation with peptide hybrids **3.72** or **3.78** as well as **3.82**-**81** restored DNA migration suggesting these hybrids disrupt both MafG/MafG and Nrf2/MafG coiled-coils **(Figure 3-20)**. These fragments offer potential leads for small molecule inhibitors. Future work to identify hit fragments towards **3.14** will allow for synthesis of small molecules that may be capable of disrupting the Nrf2/MafG PPI. Discovery of these novel peptide-small molecule hybrids offers unique candidates to investigate against Nrf2 activity in cellular modes of chemoresistance in cancer.

3.5 Experimental

3.5.1 Molecular docking

The partial crystal structure of the Nrf2/MAFG/CsMBE complex was utilised for docking calculations PBD ID 7X5F.[20] MafG was truncated to: $_{80}$KAELQQEVEKLASENASMKLE$_{104}$ and the protein preparation wizard was applied in Schrodinger Suite 2023-2. MafG derived sequences $_{80}$KAELQQEVEK$_{89}$ and $_{90}$LASENASMKLE$_{104}$ were prepared and modified using Maestro build tools to contain azide and alkyne handles and the structure was minimised. Alkyne and azide small molecule fragments were taken from Enamine and ZINC Virtual Libraries and prepared using Schrodinger Suite 2023-2 LigPrep application.

Covalent docking was prepared by centering the binding site around the terminal alkyne/azide residue of the peptides using the covalent docking tool in Schrodinger Suite

2023-2 and azido-alkynyl condensation was instructed by using the following custom reaction file

3.5.1.1 Peptide: Azide, Small Molecule: Alkyne

RECEPTOR_SMARTS_PATTERN 3,N=[N+]=[N-]

LIGAND_SMARTS_PATTERN 2,[C,c]C#C

CUSTOM_CHEMISTRY ("<1>",("charge", 0, 1))

CUSTOM_CHEMISTRY ("<1>=[N+]",("charge", 0, 2))

CUSTOM_CHEMISTRY ("<1>|<2>",("bond",1,(1,2)))

CUSTOM_CHEMISTRY ("<2>#C",("bond",2,(1,2)))

CUSTOM_CHEMISTRY ("<1>=[N]=[N]",("bond",1,(2,3)))

CUSTOM_CHEMISTRY ("<1>[N]",("bond",2,(1,2)))

CUSTOM_CHEMISTRY ("<1>([C]=[C])=[N][N]",("bond",1,(3,5)))

3.5.1.2 Peptide: Alkyne, Small Molecule: Azide

RECEPTOR_SMARTS_PATTERN 2,C#CC

LIGAND_SMARTS_PATTERN 3,N=[N+]=[N-]

CUSTOM_CHEMISTRY ("<2>",("charge", 0, 1))

CUSTOM_CHEMISTRY ("<2>=[N+]",("charge", 0, 2))

CUSTOM_CHEMISTRY ("<1>|<2>",("bond",1,(1,2)))

CUSTOM_CHEMISTRY ("<1>#C",("bond",2,(1,2)))

CUSTOM_CHEMISTRY ("<1>[N]=[N]=N",("bond",1,(3,4)))

CUSTOM_CHEMISTRY ("<1>(=C)[N]=[N]N",("bond",1,(2,5)))

3.5.2 Peptide-small molecule hybrid synthesis

3.5.2.1 Microwave Assisted Solid Phase Peptide Synthesis

Peptides were synthesised using a Liberty Lite automated microwave peptide synthesiser, using Tentagel S RAM resin (0.26 mmol/g). Resin was manually swollen in DMF for 10 min before addition to the reaction vessel. Fmoc deprotection steps were carried out in 20% piperidine in DMF with microwave heating of 90 °C for 1 min. Wash steps were performed by draining the reaction vessel and washing with DMF (3 x 3 mL). Coupling reactions of all amino acids was achieved using Oxyma Pure (5 eq), DIC (5 eq) and Fmoc-protected amino acids (5 eq) at 90 °C for 5 min. The process of deprotection and coupling was carried out until the desired amino acid sequence was completed.

Upon the final deprotection all peptides were acetylated in acetic anhydride (4 eq) and DIPEA (4 eq) in DMF with agitation for 45 minutes. The final wash step rinsed the vessel

with DMF (3 x 5 mL), CH_2Cl_2 (3 x 5 mL) and MeOH (3 x 5 mL) to exhaustively remove remaining DMF.

A microcleavage of peptides from resin support was achieved using TFA (2 mL) with addition of reaction scavenger's H_2O, DODT and TIPS at 2.5% each in a reaction vessel and agitated for 3 h. The cleavage solution was then collected and concentrated *in vacuo*, followed by precipitation in diethyl ether. Peptides were purified using RP-HPLC and analysed by analytical HPLC and MALDI-TOF.

3.5.2.2 Azido-acetylation

The azido-peptide *N_3Ac-LASENASMKLE-NH_2* (**3.15**) was achieved following the coupling of the final leucine residue and subsequent deprotection with piperidine, resin was treated with azidoacetic acid (N_3Ac) (4 eq), Oxyma Pure (5 eq), DIC (5 eq) and heated for 1 min at 90 °C. To confirm synthesis, a small amount of resin was cleaved and analysed by HPLC and MALDI-TOF.

3.5.2.3 Alkynyl peptide synthesis

The alkynyl-peptide *Ac-KAELQQEVEK(Pra)-NH_2* (**3.14**) was prepared by microwave assisted solid phase synthesis, Fmoc-L-propargylglycine was coupled as the first residue to the resin using Oxyma Pure (5 eq), DIC (5 eq) Fmoc-amino acid (5 eq), microwaved at 90 °C for 5 min and deprotection or 20% in DMF with microwave heating of 90 °C for 1 min. To confirm synthesis, a small amount of resin was cleaved and analysed by HPLC and MALDI-TOF.

3.5.2.4 On resin CuAAC reaction

Peptide on resin (1 eq), alkyne or azide fragment (1.1 eq), $Cu(MeCN)_4PF_6$ (0.1 eq), DIPEA (2 eq) were shaken in DMF for 16 h. Resin was then washed with DMF, DCM, and MeOH. Peptides were cleaved and purified using RP-HPLC and analysed by analytical HPLC and MALDI-TOF.

3.5.3 Fluorescence polarisation assay

Corning ™ 384-well solid black polystyrene microplates containing 40 µL/ well PBS containing 10 mM DTT and 0.1% Tween-20 at pH 7.4. Final concentration in well of 10 nM FAM labelled ARE DNA, 50 nM of His-MafG and 62.5 nM of His-Halo-Nrf2 was used. Peptides stored in DMSO were added to the plate ensuring no higher than 5% DMSO was used. Plate was incubated for 16 h at 4 °C, polarisation was measured using ex: 482-16

em: 530-40 Dichroic: LP 504 on Fluorescence Polarisation Mode on a ClarioStar Plate Reader.

3.5.4 EMSA assay

20% TBE gels (Novex™) was pre-run in 1x TBE buffer (Invitrogen™ UltraPure™) at 4 °C at 100 V for 30 minutes (PowerEase™ 120 W Power Supply). 10 nM FAM labelled DNA (Sigma-Aldrich) was incubated with 62.5 nM Nrf2 and 50 nM MafG and for 16 h with peptides at 4 °C and loaded with 1X BlueJuice Loading Dye (Invitrogen™). Gel was run at 200 V for 135 minutes at 4 °C in darkness. Gels were visualised using the appropriate fluorophore filter with ImageQuant LAS 4000 (GE Healthcare).

3.5.5 Analytical techniques

3.5.5.1 RP-HPLC

An Agilent 1200 with an Agilent eclipse XDB-C18 column, 4.6 x 150mm, 5µM and a flow rate of 1 mL/min was used for performing Analytical RP-HPLC. Solvent System of Solvent A = Water + 0.05% TFA and Solvent B = MeCN + 0.05% TFA. Gradient 5% B - 95% B over 20 min. Detection at wavelength 214 nm reported:

3.14 *Ac-KAELQQEVEK(Pra)-NH_2*, t_R at 9.9 min

3.85 *Ac-KAELQQEVEK-NH_2* t_R at 9.4 min

3.61 t_R at 9.4 min

3.62 t_R at 6.5 min

3.63 t_R at 10.8 min

3.64 t_R at 10.7 min

3.65 t_R at 8.1 min

3.66 t_R at 10.2 min

3.67 t_R at 10.6 min

3.15 *N_3Ac-LASENASMKLE-NH_2*, t_R at 11.1 min

3.86 *Ac-LASENASMKLE-NH_2* t_R at 10.1 min

3.72 t_R at 9.7 min

3.73 t_R at 10.4 min

3.74 t_R at 10.8 min

3.75 t_R at 10.9 min

3.76 t_R at 10.8 min

3.77 t_R at 10.7 min

3.78 t_R at 10.1 min

3.79 t_R at 10.8 min

3.80 t_R at 10.6 min

3.81 t_R at 9.2 min

3.82 t_R at 9.7 min

3.83 t_R at 9.9 min

3.84 t_R at 9.3 min

3.5.5.2 MALDI-TOF

Performed on a Shimadzu Axima Performance MALDI-TOF in positive ionisation mode. All samples were calibrated to a ProteoMass™ Peptide standard $P_{14}R$ (Cat. # P2613, Sigma-Aldrich) MALDI-MS standard: 1,533.8 $[M+H]^+$

3.14 *Ac-KAELQQEVEK-Pra-*NH_2 expected mass 1336.70 Da $[M+H]^+$observed at 1336.8

3.85 *Ac-KAELQQEVEK-*NH_2 expected 1242.6 $[M+H]^+$ observed 1241.5

3.61 expected mass of 1617.8 Da $[M+H]^+$observed at 1617.3

3.62 expected mass of 1603.8 Da $[M+H]^+$observed 1604.8

3.63 expected mass of 1535.5 Da $[M+H]^+$ found at 1535.0

3.64 expected mass 1491.1 Da $[M+H]^+$observed 1491.0

3.65 expected mass of 1538.8 Da $[M+H]^+$observed at 1538.3

3.66 expected mass 1840.1 Da $[M+H]^+$observed 1840.6

3.67 expected mass 1502.8 Da $[M+H]^+$observed 1502.7

3.15 N_3*Ac-LASENASMKLE-*NH_2 expected 1274.4 Da $[M+H]^+$ observed at 1275.0

3.86 *Ac-LASENASMKLE-*NH_2 expected 1233.6 Da $[M+H]^+$ observed 1233.4

3.72 expected mass 1470.7 Da $[M+H]^+$observed at 1473.6

3.73 expected mass 1391.6 Da $[M+H]^+$observed at 1391.2

3.74 expected mass 1421.5 Da $[M+H]^+$observed 1421.3

3.75 expected mass 1387.5 Da $[M+H]^+$observed 1387.5

3.76 expected mass 1482.8 Da $[M+H]^+$observed 1482.5

3.77 expected mass 1371.7 Da $[M+H]^+$observed 1371.5

3.78 expected mass 1478.7 Da $[M+H]^+$observed at 1478.9

3.79 expected mass 1624.2 Da $[M+H]^+$observed at 1624.5

3.80 expected mass 1444.6 Da $[M+H]^+$observed 1444.5

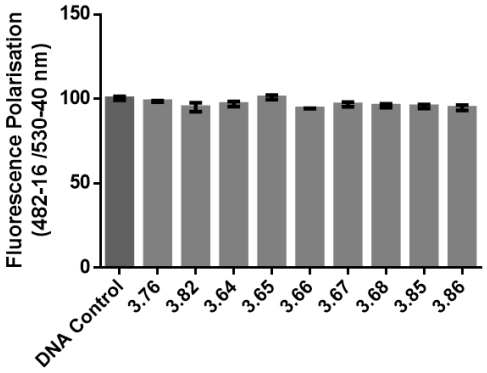
3.81 expected 1606.6 Da $[M+H]^+$observed 1606.0

3.82 expected 1531.7 Da $[M+H]^+$observed at 1531.7

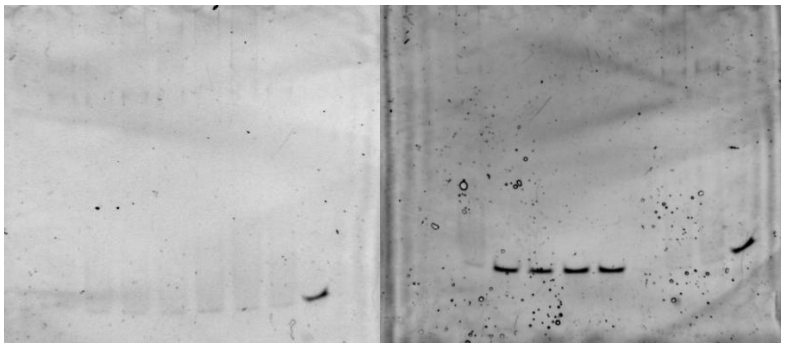
3.83 expected mass 1555.7 Da $[M+H]^+$observed 1557.9

3.84 expected mass 1464.5 Da $[M+H]^+$observed at 1464.7

3.6 Appendix

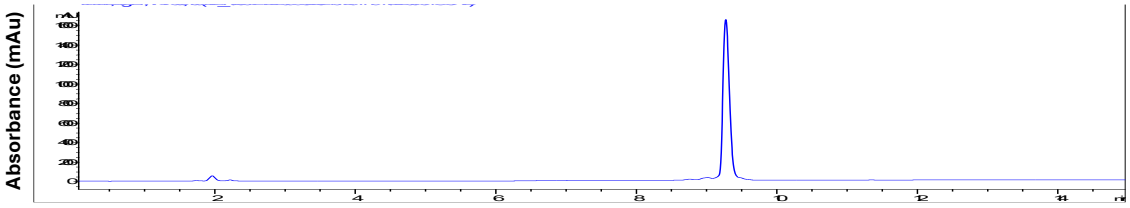


Appendix 3-1 *Fluorescence Polarisation of peptide-small molecule hybrids at 250 µM against 10 nM FAM ARE +/- Nrf2/MafG at 1 h RT and 16 h 4 °C.*

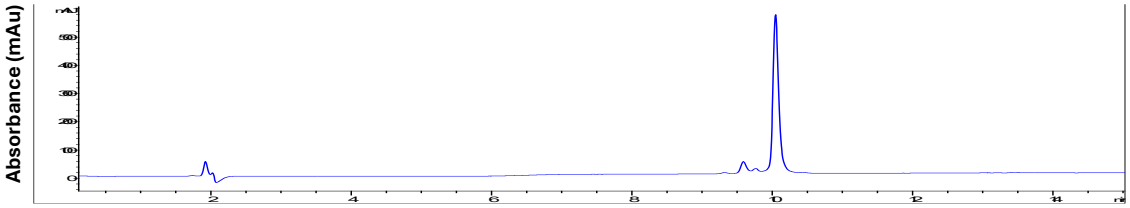


Appendix 3-2 *Full images of EMSAs on peptide-small molecule hybrids shown in* ***Figure 3-20***

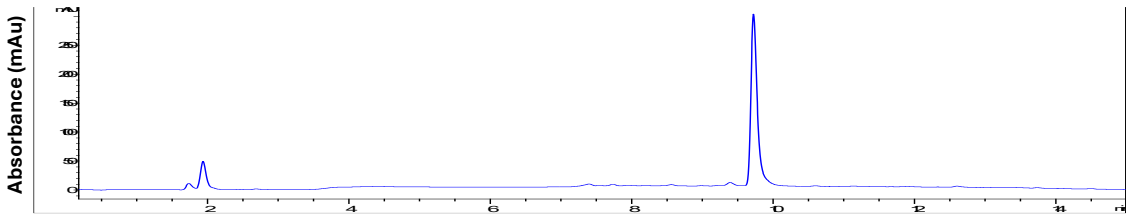
3.6.1.1 HPLC traces of compounds taken forward to cellular studies



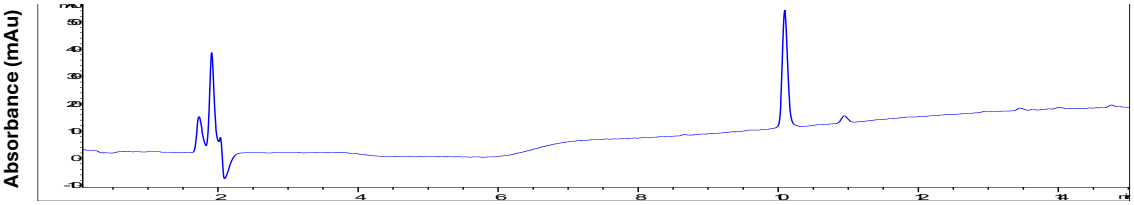
Appendix 3-3 *HPLC of* ***3.85***



Appendix 3-4 *HPLC of* ***3.86***



Appendix 3-5 *HPLC of* ***3.72***



Appendix 3-6** HPLC of **3.78

3.7 References

1 C. Zhuang and C. Sheng, in *Targeting Protein-Protein Interactions by Small Molecules*, Springer, Singapore, 2018, pp. 29–51.
2 C. Sheng, G. Dong, Z. Miao, W. Zhang and W. Wang, *Chem Soc Rev*, 2015, **44**, 8238–8259.
3 P. Tošovská and P. S. Arora, *Org Lett*, 2010, **12**, 1588–1591.
4 S. Marimganti, M. N. Cheemala and J.-M. Ahn, *Org Lett*, 2009, **11**, 4418–4421.
5 J. P. Plante, T. Burnley, B. Malkova, M. E. Webb, S. L. Warriner, T. A. Edwards and A. J. Wilson, *Chem Commun*, 2009, 5091–5093.
6 W. Antuch, S. Menon, Q.-Z. Chen, Y. Lu, S. Sakamuri, B. Beck, V. Schauer-Vukasinović, S. Agarwal, S. Hess and A. Dömling, *Bioorg Med Chem Lett*, 2006, **16**, 1740–1743.
7 P. J. Hajduk and J. Greer, *Nat Rev Drug Discov*, 2007, **6**, 211–219.
8 S. Wang, W. Sun, Y. Zhao, D. McEachern, I. Meaux, C. Barrière, J. Stuckey, J. Meagher, L. Bai, L. Liu, C. G. Hoffman-Luca, J. Lu, S. Shangary, S. Yu, D. Bernard, A. Aguilar, O. Dos-Santos, L. Besret, S. Guerif, P. Pannier, D. Gorge-Bernat and L. Debussche, *Cancer Res*, 2014, **74**, 5855–5865.
9 M. J. I. Andrews, G. Kontopidis, C. McInnes, A. Plater, L. Innes, A. Cowan, P. Jewsbury and P. M. Fischer, *Chembiochem*, 2006, **7**, 1909–1915.
10 P. Buchwald, *IUBMB life*, 2010, **62**, 724–731.
11 M. Cano-Muñoz, S. Jurado, B. Morel and F. Conejero-Lara, *Int J Biol Macromol*, 2021, **192**, 90–99.
12 D. Roymans, H. L. De Bondt, E. Arnoult, P. Geluykens, T. Gevers, M. Van Ginderen, N. Verheyen, H. Kim, R. Willebrords, J.-F. Bonfanti, W. Bruinzeel, M. D. Cummings, H. van Vlijmen and K. Andries, *Proc Natl Acad Sci U S A*, 2010, **107**, 308–313.
13 J. R. Hart, A. L. Garner, J. Yu, Y. Ito, M. Sun, L. Ueno, J.-K. Rhee, M. M. Baksh, E. Stefan, M. Hartl, K. Bister, P. K. Vogt and K. D. Janda, *Proc Natl Acad Sci U S A*, 2014, **111**, 12556–12561.
14 X. Yin, C. Giap, J. S. Lazo and E. V. Prochownik, *Oncogene*, 2003, **22**, 6151–6159.
15 R. J. Shirey, J. R. Hart, B. Sridharan, S. J. Novick, L. D. Turner, B. Zhou, A. L. Nielsen, L. M. Eubanks, L. Ueno, M. S. Hixon, L. L. Lairson, T. P. Spicer, L. D. Scampavia, P. R. Griffin, P. K. Vogt and K. D. Janda, *Bioorg Med Chem*, 2021, **42**, 116246.
16 M. Y. Kim, I. Na, J. S. Kim, S. H. Son, S. Choi, S. E. Lee, J.-H. Kim, K. Jang, G. Alterovitz, Y. Chen, A. van der Vaart, H.-S. Won, V. N. Uversky and C. G. Kim, *Sci Adv*, 2019, **5**, eaav9810.
17 L. A. Howell and A. M. Beekman, *RSC Chem Biol*, 2021, **2**, 215–219.
18 A. M. Beekman, M. M. D. Cominetti, S. J. Walpole, S. Prabhu, M. A. O'Connell, J. Angulo and M. Searcey, *Chem Sci*, 2019, **10**, 4502–4508.
19 A. M. Beekman, M. A. O'Connell and L. A. Howell, *Angew Chem Int Ed*, 2017, **56**, 10446–10450.
20 T. Sengoku, M. Shiina, K. Suzuki, K. Hamada, K. Sato, A. Uchiyama, S. Kobayashi, A. Oguni, H. Itaya, K. Kasahara, H. Moriwaki, C. Watanabe, T. Honma, C. Okada, S. Baba, T. Ohta, H. Motohashi, M. Yamamoto and K. Ogata, *Nucleic Acids Res*, 2022, **50**, 12543–12557.

21 C. S. Rapp, C. Kalyanaraman, A. Schiffmiller, E. L. Schoenbrun and M. P. Jacobson, *J Chem Inf Model*, 2011, **51**, 2082–2089.
22 J. J. Irwin, K. G. Tang, J. Young, C. Dandarchuluun, B. R. Wong, M. Khurelbaatar, Y. S. Moroz, J. Mayfield and R. A. Sayle, *J Chem Inf Model*, 2020, **60**, 6065–6073.
23 Azides for 'Click Chemistry' - Enamine, https://enamine.net/building-blocks/functional-classes/azides-for-click-chemistry, (accessed 21 July 2025).
24 J. R. Greenwood, D. Calkins, A. P. Sullivan and J. C. Shelley, *J Comput Aided Mol Des*, 2010, **24**, 591–604.
25 R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis and P. S. Shenkin, *J Med Chem*, 2004, **47**, 1739–1749.
26 H. C. Kolb, M. G. Finn and K. B. Sharpless, *Angew Chem Int Ed*, 2001, **40**, 2004–2021.
27 V. V. Rostovtsev, L. G. Green, V. V. Fokin and K. B. Sharpless, *Angew Chem Int Ed Engl*, 2002, **41**, 2596–2599.
28 B. T. Worrell, J. A. Malik and V. V. Fokin, *Science*, 2013, **340**, 457–460.
29 M. Meldal and C. W. Tornøe, *Chem Rev*, 2008, **108**, 2952–3015.
30 H. Ben El Ayouchia, L. Bahsis, H. Anane, L. R. Domingo and S.-E. Stiriba, *RSC Adv*, 2018, **8**, 7670–7678.
31 A. A. H. Ahmad Fuaad, F. Azmi, M. Skwarczynski and I. Toth, *Molecules*, 2013, **18**, 13148–13174.
32 S. Li, C. Schöneich, G. S. Wilson and R. T. Borchardt, *Pharm Res*, 1993, **10**, 1572–1579.
33 Peter W. Seavill, Katherine B. Holt, and Jonathan D. Wilden, *RSC Adv*, 2019, **9**, 29300–29304.

Chapter 4

Cellular Assessment of Peptides and Peptide-Small Molecule Hybrids Targeting Nrf2/MafG

4.1 Introduction

4.1.1 Nrf2 and chemoresistance

Nrf2-activated gene transcription leads to the production of drug metabolising enzymes, such as NAD(P)H dehydrogenase quinone 1 (NQO1) and heme oxygenase-1 (HO-1).[1,2] Nrf2 can also increase expression of drug efflux transporters, consequently, overexpression of Nrf2 in cancer cells has been linked to chemoresistance.[3] Nrf2 siRNA knockdown models have demonstrated increased cytotoxicity to cisplatin in ovarian cancer cell lines with cisplatin resistance.[4]

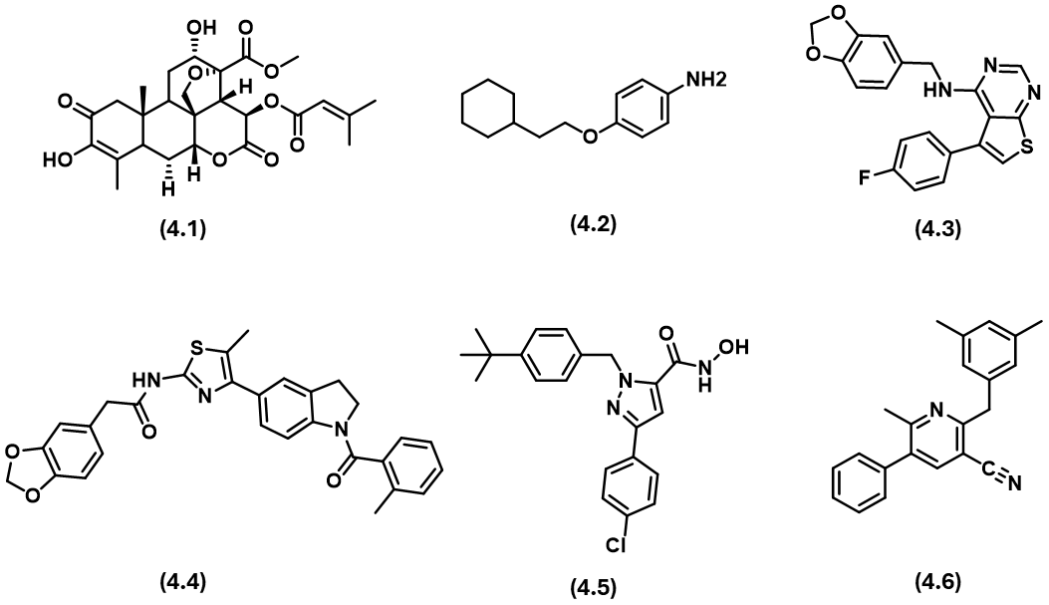
Nrf2 signalling has been linked to development of resistance to doxorubicin, an anthracycline topoisomerase II inhibitor capable of intercalation with DNA and responsible for the production of free radicals, causing cell apoptosis.[5,6] Inhibition of Nrf2 in models of triple negative breast cancer has been correlated with the down-regulation of p-glycoprotein, a multidrug-resistance protein (MRP), restoring doxorubicin sensitivity.[7] Similarly, in non-small cell lung cancer (NSCLC), resistance to doxorubicin has been associated with increased expression of Nrf2. Over-expression of Nrf2 is correlated with increased downstream MRP activity , lipoprotein receptor-related protein and the antioxidant glutathione pathway.[8]

Nrf2 has also been implicated in the generation of resistance against gemcitabine, a first line treatment for NSCLC and pancreatic cancer.[9,10] Gemcitabine is a pyrimidine nucleoside antimetabolite capable of inhibiting ribonucleotide reductase for nucleotide production. Metabolised gemcitabine can also incorporate into DNA to terminate DNA synthesis.[9,11] Treatment of pancreatic cell lines with gemcitabine has been shown to activate Nrf2, NQO1 and HO-1 protein levels. Co-treatment with brusatol (**4.1**), a natural product known to suppress Nrf2 activity enhanced antitumour effects in models of pancreatic cancer.[12]

4.1.2 Reported inhibitors of Nrf2

A number of Nrf2 inhibitors have been reported to inhibit downstream Nrf2 activity in cancer cell lines, producing synergistic combination therapies with cytotoxic agents (**Scheme 4-1**). Several of these are small molecule inhibitors, identified through a high throughput screen using an Nrf2/ARE luciferase reporter assay. This assay quantifies Nrf2 activity by introducing an ARE promoter before a luciferase gene, resulting in a quantifiable luminescence in cells with active Nrf2 transcription. The Nrf2/ARE-luciferase activity assay was first used to screen natural product compounds leading to

the identification of brusatol (**4.1**).[12] The natural product was found to enhance Nrf2 degradation and sensitise A549 lung adenocarcinoma cells to carboplatin, 5-fluorouracil, etoposide and paclitaxel. The combination of cisplatin and brusatol treatment reduced A549 tumour volume in nude mice.[12]



Scheme 4-1 *Small molecule inhibitors of Nrf2, Brusatol (**4.1**) IM3828 (**4.2**), AEM1 (**4.3**), ML385 (**4.4**) 4f (**4.5**) and MSU38225 (**4.6**).*[12–18]

IM3828 (**4.2**) was the first small molecule Nrf2 inhibitor to be found using a luciferase reporter assay, producing a concentration dependent inhibition of ARE-Luciferase activity to an IC_{50} of 2.9 μM in HEK293 cells treated with 40 μM of Nrf2-inducing agent tert-butylhydroquinone (tBHQ).[13] Expression of Nrf2 was downregulated, as was HO-1 expression, determined by western blot analysis from H1299 lung cancer cells treated with 10 μM of IM3828.[13] RT-PCR determined a concentration-dependent inhibition of Nrf2 mRNA expression and downstream gene activation of HO-1 and NQO1, suggesting inhibition of Nrf2 expression in cancer.[13] Subsequently, it was found that IM3828 could reduce radioresistance, enhancing radiation induced apoptosis in three lung cancer cell lines (H1299, A549, H460).[13] Interestingly, in comparison to more recent Nrf2 inhibitors described, IM3828 produced Nrf2 inhibition without significantly impacting cell viability.[13,14,17]

In 2015, ARE expression modulator 1 (AEM1) (**4.3**) was found by HTS using the same luciferase reporter assay, with additional controls to confirm inhibition of Nrf2-driven gene transcription against compounds inhibiting Nrf2 expression.[14] Treatment of AEM1 at 10 μM decreased heme oxygenase 1 (HO-1) mRNA expression and glutathione

content was decreased at 5 μM in A549 cells.[14] Cell viability in A549 cells was subsequently assessed in combination with doxorubicin, etoposide and 5-fluorouracil, and was found to produce synergistic toxicity in the NSCLC cell line. Profiling the relative expression of HO-1 in a panel of cell lines found Nrf2 inhibition was limited to cell lines that over-expressed Nrf2. AEM1 was found to have no effect on Nrf2 nuclear entry and Keap1 protein expression. To date, the mechanism of action to produce Nrf2 inhibition remains unknown.[14]

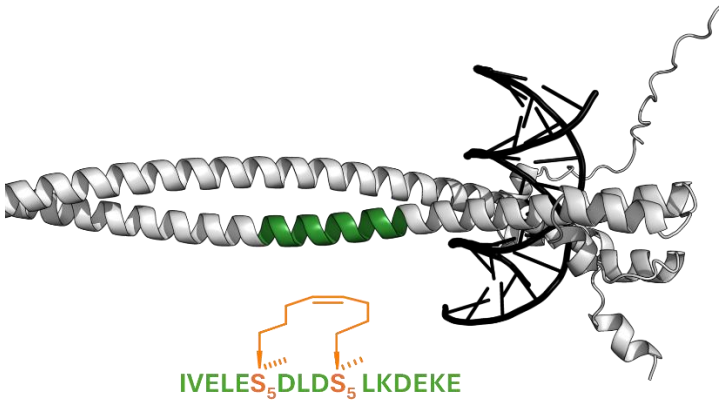
ML385 (**4.4**), reported in 2016, was the first Nrf2 inhibitor to provide evidence of direct inhibition by binding to the Neh1 domain of Nrf2, responsible for ARE binding and sMAF dimerisation.[17] Interestingly, SAR development highlighted the importance of the 1,3-benzodioxole group for inhibition of ARE binding, a functional group that AEM1 (**4.3**) also contains. Nrf2 inhibition was confirmed in A549 cells demonstrating concentration-dependent inhibition of Nrf2, NQO1 and HO-1 mRNA expression. NQO1 activity and GSH content were also found to be reduced at 5 and 10 μM of ML385. Combination treatment of ML385 with doxorubicin, carboplatin and paclitaxel was assessed in A549 and H460 lung adenocarcinoma cell lines and was found to sensitise the cell lines compared to single line treatment.[17]

In 2017, a pyrazolyl hydroxamic acid compound referred to as 4f (**4.5**), was found through HTS using the Nrf2/ARE luciferase assay and reduced mRNA expression of HO-1 and another Nrf2 target gene, glutamate–cysteine ligase catalytic subunit (GCLC) in HeLa cells.[18] 4f was subsequently investigated in acute myeloid leukaemia cell lines (THP-1, HL-60 and U937).[18] It was found that Nrf2 expression was reduced by 4f treatment as determined by western blot, cell viability of AML cells lines was reduced and flow cytometry found growth inhibition to be caused by cell apoptosis.[18]

The most recently reported small molecule inhibitor of Nrf2, MSU38225 (**4.6**) was shown to be selective for cell lines that overexpress Nrf2, inhibiting cell proliferation whilst having no effect on normal epithelial cell growth.[16] Nrf2 expression was reduced in cytosolic and nuclear contents of A549 cells and a panel of downstream genes were found to be significantly reduced at 5 μM of MSU38225, without impacting Nrf2 mRNA expression.[15] Keap1 and β-TrCP knock-out cell lines retained sensitivity to MSU38225 treatment at 5 μM, however further investigation into the compounds mechanism of action has not been explored.[15,16] Combination index calculations found the small

molecule to have synergism on co-treatment of carboplatin, doxorubicin and 5-fluorouracil.[15]

Despite the emergence of the small molecule inhibitors, confirmation of their target pathway, specificity and mechanism of action still require investigation. Rationally designed PPI inhibitors of Nrf2/MafG have also been identified through the Nrf2/ARE luciferase reporter assay. The first report of a peptide based Nrf2 inhibitor in 2023 achieved this through designing a MafG interacting peptide, N1S (**4.7**) based on a section of the Nrf2 leucine zipper $_{522}$IVELEQDLDHLKDEKE$_{537}$ (**Figure 4-1**).



(4.7)

Figure 4-1 *N1S stapled peptide (**4.7**) derived from Nrf2/MafG/ARE model, section of Nrf2 leucine zipper N1S derives from, highlighted in green.*

N1S was found to significantly reduce the expression of five Nrf2-driven genes (ABCC1, TXN, NQO1, HO-1 and GCLC).[19] Glutathione expression was also quantified and found to be significantly reduced compared to DMSO controls.[19] Cell growth was inhibited in the A549 cell line and combination treatment with cisplatin found N1S to have an additive effect on cell death from 100 μM.[19] Nrf2-derived peptides bind to the MafG protein, which could potentially block other sMaf interactions as well, creating off-target effects. Consequently, the effects of a peptide inhibitor designed to bind directly to Nrf2 offers significant promise but is yet to be explored.

4.2 Chapter Aims

This chapter describes the impact of Nrf2 inhibitory compounds identified in **Chapters 2** and **3** in cancer cells, specifically the non-small cell lung cancer derived A549 cell line. Firstly, the impact of peptides on cell viability were investigated using an enzymatic MTS assessment. Nrf2-related activity was investigated by quantifying downstream NQO1 activation by enzymatic activity. The impact of the compounds on chemotherapeutic

agents doxorubicin and gemcitabine was also investigated to determine if co-treatment could induce greater cytotoxicity in A549 cells.

4.3 Results and Discussion

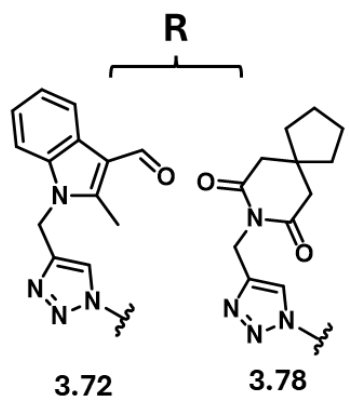
Peptides, outlined in **Table 4-1**, were selected to investigate Nrf2 inhibitory activity in the A549 cell line based on their promising activity in *in vitro* assays in **Chapter 2 and 3**. Peptide **2.53** proved inactive in biophysical assays and was used as a control compound. A549 lung adenocarcinoma cells were chosen as the cell line for investigating Nrf2 inhibition, as it is reported that Nrf2 is constitutively activated in A549 cells, due to a loss of function mutation on Keap1.[2]

Table 4-1 (A) *Peptides investigated in A549 cells for Nrf2 inhibitory activity. (**B**) Structures of small molecule fragments of **3.72** and **3.78** peptide hybrids.*

A

Peptide	*Sequence*
2.53	Ac-KEELEKQKAELQQEVEKLASE-NH_2
2.52	Ac-KAELQQEVEKLASENASMKLE-NH_2
2.58	Ac-KAELQQEVEKLASENAS(**Nle**)KLE-NH_2
3.86	Ac-LASENASMKLE-NH_2
3.72	**R**-LASENASMKLE-NH_2
3.78	**R**-LASENASMKLE-NH_2

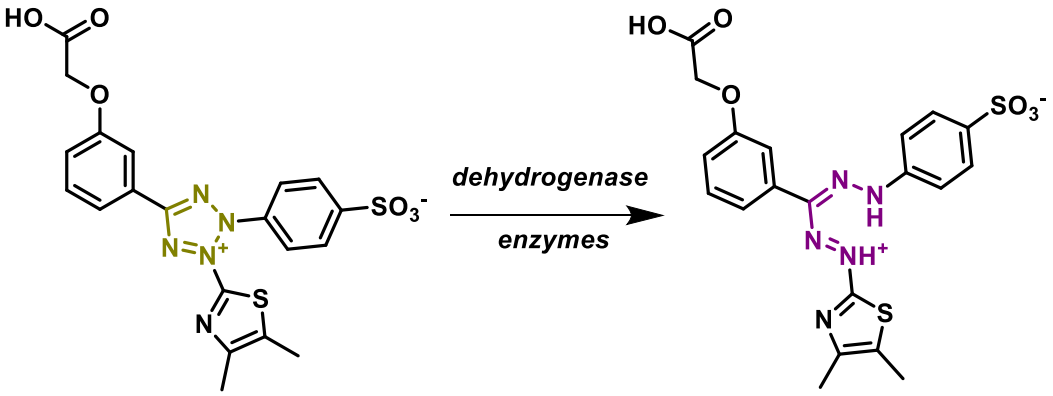
B



3.72 **3.78**

4.3.1 Cell viability assay

To understand the effects peptide inhibitors, have on proliferation, a cell viability assay was performed. Dimethylthiazol-carboxymethoxyphenyl-sulfophenyl tetrazolium (MTS) can be used for evaluating cellular viability by measuring the metabolic activity of cells through colorimetric analysis of MTS dye metabolism (**Figure 4-2**). Tetrazolium salts (**4.8**) undergo reduction by mitochondrial dehydrogenases upon cellular uptake, leading to the production of formazan crystals (**4.9**) which can be quantified by spectrophotometry at a wavelength of 490 nm. Reduced MTS dye metabolism is subsequently used as a measure of cell viability.



(4.8) (4.9)

Figure 4-2: *Metabolism of MTS (**4.8**) to formazan crystals producing an intense colour change due to the formazan functional group [-N=N-C(R)=N-NH-] (**4.9**).*

A549 cells were treated with either 100 µM or 10 µM of peptide over 72 h. Investigation by MTS assay revealed that none of the peptides significantly reduced cell viability relative to a DMSO vehicle control, suggesting the peptides and peptide-small molecule hybrids tested do not have potent anti-proliferative effects in A549 cells (**Figure 4-3**). Consequently, these peptides could be investigated for activity from 100 µM in an enzymatic activity assay related to Nrf2.

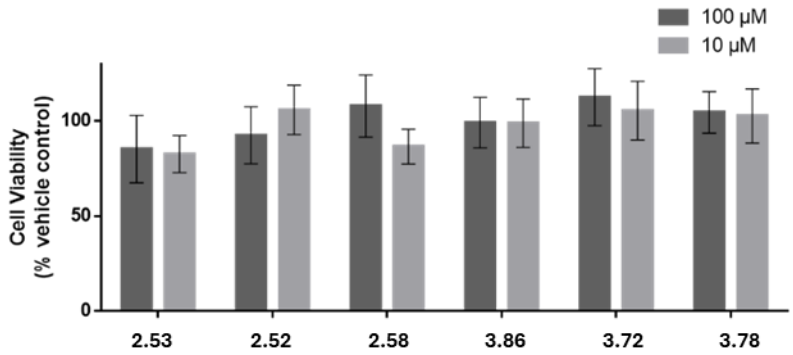


Figure 4-3 *Graph of MTS metabolism* following *100 or 10 µM of peptide treatment over 72 h. (±SEM) (N=3).*

4.3.2 NQO1 assay

The detoxification enzyme NQO1, is a target gene activated by Nrf2/ARE binding. NQO1 activity has been shown to correlate with the development of NSCLC.[20–22] NQO1 activity can be measured by oxidation of menadiol (**4.10**), a quinone metabolite product of menadione-specific reduction by NQO1 (**Figure 4-4**). Spontaneous oxidation of menadiol (**4.10**) to menadione (**4.11**) occurs rapidly, in a non-enzymatic manner promoting MTT (**4.12**) reduction to formazan upon co-incubation (**4.13**).[23,24] Assay specificity is provided by the NQO1-specific oxidation of NADPH to NADP+, catalysing the reduction of menadione to menadiol, which has been shown to be linearly

proportional to the generation of menadione by Prochaska et al. during the assay's development. [23]

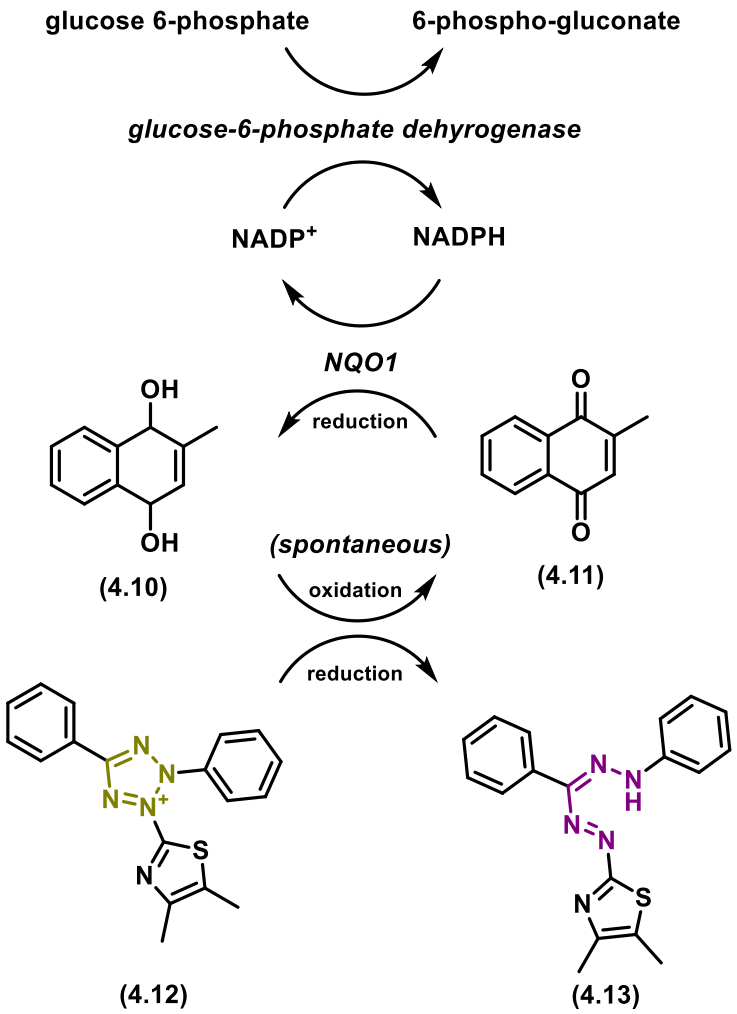


Figure 4-4: *Principle of NQO1 activity assay adapted from Prochaska et al. Addition of glucose-6-phosphate and glucose-6-phosphate dehydrogenase supports NADPH regeneration preventing substrate depletion. NQO1-catalysed reduction of menadione (**4.11**) to menadiol (**4.10**), which undergoes subsequent spontaneous re-oxidation back to menadione with coupled reduction of MTT (**4.12**) to formazan (**4.13**).*

A549 cells were treated with 25, 50 or 100 μM of peptide for 24 h, before measuring the cells NQO1 activity. Variable NQO1 activity was observed across peptides assessed (**Figure 4-5**). The best result was found from **3.72,** which decreased the NQO1 signal by 75% at 100 μM and continued to suppress activity up to 40% at 25 μM. **2.52** and **2.58** produced reduction of up to 50% in NQO1 activity compared to the vehicle control at 100 and 50 μM. **3.78**, **3.86** and control **2.53** produced minimal changes in response after 24 h. This result suggests peptides mimicking the MafG leucine zipper could be capable of causing Nrf2 inhibition leading to a functional decrease in associated cytoprotective activity.

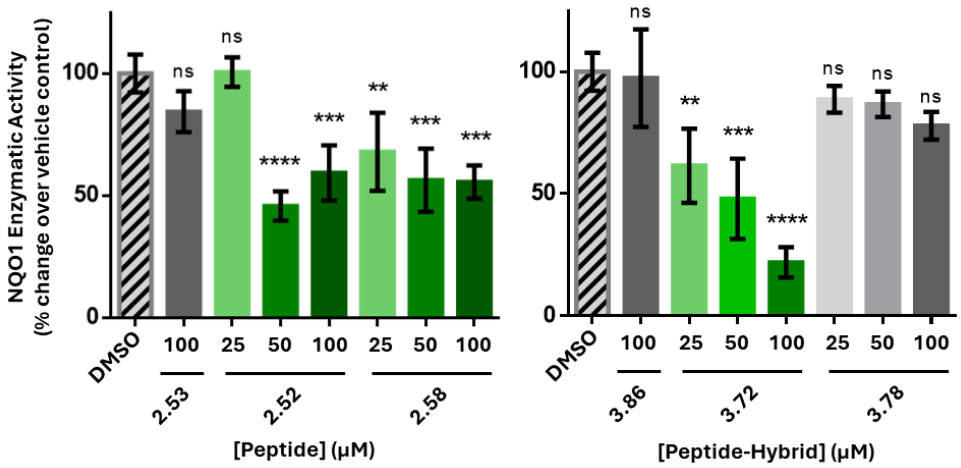


Figure 4-5: *NQO1 enzymatic activity in A549 cells, treated with MafG peptides. 5 x 10^3 A549 cells treated with 25, 50 or 100 μM of peptides for 24 h. Results presented as the mean and are representative of three individual experiments (±SEM) (N = 3). Statistical analysis was conducted via one way-ANOVA and post-hoc Dunnett's test (P**≤0.01,P***≤0.001, P****≤0.0001) comparison to vehicle control (DMSO).*

4.3.3 Synergism with chemotherapeutics agents

2.52 and **3.86** were subsequently taken forward for investigation of synergistic properties against A549 cells treated with doxorubicin from 100 μM over 72 hours. Cell viability following treatment with doxorubicin at eight different concentrations was compared at five different concentrations of **2.52** or **3.72**. No change in cell viability was observed in combination with either peptide (**Figure 4-6**). Treatment with **2.52** or **3.72** was not found to be significantly different from the vehicle control. Additionally, no significant difference was found between treatment with doxorubicin alone and co-treatment with 25-1.5 μM **2.52** or **3.72**.

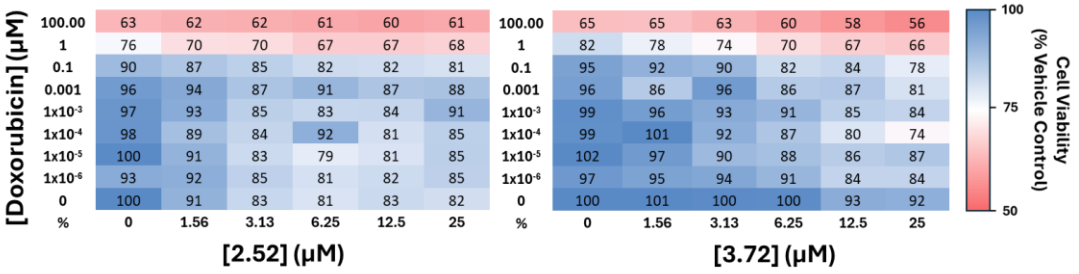


Figure 4-6: *Cell viability in A549 cells assessed by MTS, co-treated with doxorubicin and MafG peptides. Heat map plots of A549 cells treated with doxorubicin from 100 and 1 μM (1:10 dilution) combination with **2.52** or **3.72** from 25 μM (1:2 dilution) for 72 h at 1% DMSO final concentration. Represented as mean cell viability (% of vehicle control) (N=3).*

It is possible that the resistance of doxorubicin in A549 cells is not caused by Nrf2 overexpression. Consequently, gemcitabine was also investigated for synergistic activity with the **2.52** and **3.72.** It is known that A549 cells can become resistance to gemcitabine, this was observed consistently in our cultures.[25,26] Despite a literature reported IC_{50} of <10 μM in A549 cells, our cultures remained unresponsive to

gemcitabine at the top concentration of 200 μM.[27,28] To help validate this, the gemcitabine used was confirmed to be potently active in alternative cell lines (HL-60 and MCF-7) (**Appendix 4-1**).

Co-treatment of A549 cells with gemcitabine and **2.52** or **3.72** was investigated by MTS assay**.** At 100 μM gemcitabine, cell viability was found to be at 84-89% of the DMSO control. The most pronounced change was found with combination treatment of 25 μM of **3.72,** reducing viability to 47% of the vehicle control (**Figure 4-7, B**). Furthermore, at 25 μM of **3.72** and concentrations as low as 0.7 μM of gemcitabine reduced cell viability was observed from 89% to 63% of the vehicle control. This suggests the peptide-small molecule hybrid compound can chemo-sensitise A549 cells with resistance to gemcitabine.

2.52 produced no synergistic response in combination with gemcitabine (**Figure 4-7, A**) The lack of observed synergy may be influenced by the stability of the peptide over the 72-hour incubation period. It is possible that degradation over time diminished its effectiveness. Moreover, the NQO1 assay was conducted after 24 h, in line with the expected lifetime for the transient activity of Nrf2. Consequently, peptides that interact with Nrf2 may only temporarily block Nrf2 gene transcription. Upon peptide binding, Nrf2 may undergo translocation into the cytoplasm, making both the peptide and the inhibited protein susceptible to degradation. This could limit the sustained efficacy of the peptide treatment over extended periods.

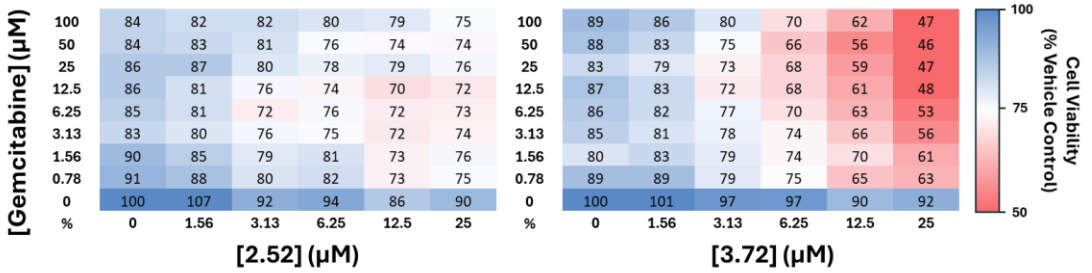


Figure 4-7 *Cell viability in A549 cells assessed by MTS, co-treated with gemcitabine and MafG peptides. Heat map plots of A549 cells treated with gemcitabine from 100 μM (1:2 dilution) combination with* ***2.52*** *or* ***3.72*** *from 25 μM (1:2 dilution) for 72 h at 1% DMSO final concentration. Represented as mean cell viability (% of vehicle control) (N=3).*

4.4 Conclusion

Inhibitors of Nrf2 have predominantly been identified by screening compound libraries against an ARE luciferase reporter assay, however this method does not distinguish if inhibitors work by disrupting expression or transcription of Nrf2. Through identifying inhibitors of the Nrf2/MafG interaction by fluorescence polarisation assay (**Chapter**

2.3.4), we have been able to investigate the impact of direct Nrf2 inhibition on cell viability, NQO1 activity and synergism with chemotherapeutic agents in A549 cells. Peptides derived from the MafG leucine zipper are anticipated to inhibit Nrf2/DNA binding without blocking the Nrf2 NES domain, responsible for nuclear export of the protein, further promoting inhibition of Nrf2.[29]

In A549 cells, known to over-express Nrf2, MafG peptide mimetics with inhibitory activity described in **Chapter 2 and 3** were investigated for Nrf2 related activity. Demonstrated in **Figure 4-3**, the peptides in isolation do not significantly impact cell viability from concentrations of 100 μM. Whilst many Nrf2 inhibitors report cell death in cancer cell lines, limited understanding of their modes of action raises the possibility that the effects of these inhibitors might enact off-target activity.[12–18]

The MafG derived peptides were taken forward for investigation in an NQO1 enzymatic assay. They demonstrated concentration-dependent inhibition of NQO1 activity in cells treated with **2.52**, **2.58** and **3.72** from 25 to 100 μM (**Figure 4-5**), suggesting these peptides could modulate the cytoprotective effects derived from Nrf2 overexpression. This was in line with results previously found in **Chapter 2** and **3,** where fluorescence polarisation assays found these peptides to inhibit ternary complex formation with IC_{50} values below 100 μM. Meaningful inhibition was not found from treatment with **3.86** or **3.78** correlating with the weaker activity found by these compounds by fluorescence polarisation assay in **Chapter 3.** Future work would investigate reasons for the instability of these compounds, including assessing the proteolytic stability against peptide digesting enzymes. **3.86** was the shortest peptide (10-mer) to demonstrate activity and may not retain stability in a cellular environment. On the other hand, peptide-hybrid **3.78** may be experiencing reduced cell permeability. As such further assays to assess cellular uptake of the peptides could provide valuable insight. The functional response observed by this enzymatic activity could be further solidified by monitoring mRNA expression of NQO1 to confirm inhibition is directly correlated with a decrease in expression.

Nrf2 is known to promote chemoresistance in NSCLC.[15,30] Inhibition of Nrf2 can sensitise chemotherapeutic agents through reduced expression of cytoprotective gene expression including DNA damage repair enzymes and drug efflux pumps.[3,15] To determine if MafG derived Nrf2 inhibitors acted as a chemotherapy sensitising agent, synergism was explored by co-treatment impact on cell viability by MTS assay. **2.52** and **3.72** were treated in combination with doxorubicin, a topoisomerase II inhibitor reported

to be impacted by chemo-resistance in NSCLC.[6] It was found that peptide treatment did not influence cell viability with doxorubicin **(Figure 4-6).** It is possible to consider that doxorubicin resistance in A549 cells may not be potentiated by Nrf2 expression, therefore investigation of further chemotherapeutic agents was warranted. An alternative investigation of combination treatment with gemcitabine, a DNA anti-metabolite, was conducted. **3.72** was found to reduce cell viability by up to 42% in combination with gemcitabine.

Overall, these peptides offer an interesting research probe for exploring the impact of Nrf2 expression on chemoresistance and can be explored further for their potential therapeutic benefit by disrupting the Nrf2/MafG interaction in cancer.

4.5 Experimental

4.5.1 Cell culture

The human A549 adenocarcinoma cell line is derived from a 58-yr old male (European Collection of Cell Cultures). Cells were cultured in RPMI-1640 supplemented with 2 mM L-Glutamine, 100 U/mL penicillin, 100 µg/mL streptomycin and 10% FBS.

A549 cells were passaged every 3-4 days until 70-80% confluent, determined qualitatively by white light microscopy. Cells were detached from the flask using pre-warmed (37 °C) Dulbecco's phosphate buffered saline (DPBS) (Fisher) to wash the cells and TrypLE Express (GIBCO) to trypsinise cells into suspension. TrypLE was subsequently diluted 1:4 using RPMI-1640 media and cells were diluted between 1:10 and 1:20 for maintenance at 37 °C, 5% CO_2 in a humidified incubator

4.5.2 MTS assay

Assessment of cell viability by MTS assay was performed using a CellTiter 96® Aqueous One Solution Cell Proliferation Assay Kit (Promega). Cells were seeded at a density of 5 x 10^3 in 100 µL in 96 well plates. Cells were treated in triplicate (1% DMSO v/v). 1% DMSO was used as a growth control and 100 µM doxorubicin was used as a death control. Plates were incubated for 72 h at 37 °C, 5% CO_2 in a humidified incubator. Media was replaced and each well treated with 10 µL of MTS reagent for 90 min. Absorbance was measured at 492 nm using a BMG Labtech POLARstar OPTIMA plate reader. Background absorbance was subtracting using a media control, viability was calculated as a percentage of the growth control (1% DMSO v/v).

4.5.3 Assessment of synergism by MTS

Cells were seeded as at a density of 5 x 10^3 in 100 µL in 96 well plates overnight. RPMI media was replaced to produce a final concentration matrix of 1.5-25 µM (1:2 dilution) of peptide and 0.76-100 µM (1:2 dilution) gemcitabine or 1x10^{-6}, 1 and 100 µM doxorubicin (1:10 dilution), final DMSO 1% (v/v). Plates treated with concentration matrix were each prepared in triplicate. Plates were incubated for 72 h at 37 °C, 5% CO_2 in a humidified incubator. RPMI media was replaced and each well treated with 10 µL of MTS reagent for 90 min at 37 °C, 5% CO_2. Absorbance was measured at 492 nm using a BMG Labtech POLARstar OPTIMA plate reader. Background absorbance was subtracting using a media control, viability was calculated as a percentage of the growth control (1% DMSO v/v).

4.5.4 NQO1 enzymatic assay

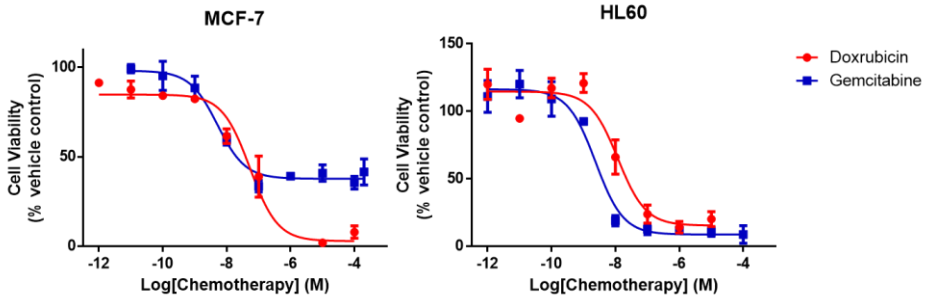
All reagents were purchased from Merck or Apollo Scientific.

Cells were seeded at a density of 5 x 10^3 in 100 µL in 96 well plates and incubated overnight at 37 °C, 5% CO_2 in a humidified incubator. Cells were then treated in triplicate (1% DMSO v/v) for 24 h at 37 °C, 5% CO_2 in a humidified incubator. Media was removed, and wells were treated with 50 µL of 0.8% digitonin (v/v) in 2 mM EDTA pH 7.8 for 10 min at 37 °C, followed by 10 min at room temperature.

Reaction buffer containing a final concentration of 25 mM Tris-HCl, 0.65 mg/mL bovine serum albumin (BSA), 0.65% Tween-20, 5 µM flavin adenine dinucleotide disodium salt hydrate (FAD), 1 mM glucose-6-phosphate (G-6-P), 30 µM nicotinamide adenine dinucleotide phosphate ($NADP^+$), 2 U/mL Yeast G-6-P dehydrogenase, 500 µM thiazolyl blue tetrazolium blue (MTT), 50 µM menadione. Reaction buffer was prepared fresh immediately before use and kept on ice.

To each well, 200 µL of reaction buffer was added and the plate was read immediately at 610 nm on a CLARIOstar Plus microplate reader (BMG Labtech). NQO1 enzyme standards were used as a positive control. Background absorbance was subtracted from a well containing reaction buffer without menadione, activity was calculated as a percentage of the vehicle control (1% DMSO v/v).

4.6 Appendix



Appendix 4-1 *IC_{50} plots of gemcitabine (blue) and doxorubicin (red) in (**A**) MCF-7 cell line (**B**) HL-60 cell line.*

4.7 References

1 Gil-Tae Gang, Yong-Hoon Kim, Jung-Ran Noh, Kyoung-Shim Kim, Ju-Young Jung, Minho Shong, Jung Hwan Hwang, and Chul-Ho Lee, *Toxicol Lett*, 2013, **221**, 165–175.

2 Mee-Hyang Kweon, Vaqar Mustafa Adhami, Jeong-Sang Lee, and Hasan Mukhtar, *J Biol Chem*, 2006, **281**, 33761–33772.

3 X.-J. Wang, Z. Sun, N. F. Villeneuve, S. Zhang, F. Zhao, Y. Li, W. Chen, X. Yi, W. Zheng, G. T. Wondrak, P. K. Wong and D. D. Zhang, *Carcinogenesis*, 2008, **29**, 1235–1243.

4 J.-M. Cho, S. Manandhar, H.-R. Lee, H.-M. Park and M.-K. Kwak, *Cancer Lett*, 2008, **260**, 96–108.

5 D. A. Gewirtz, *Biochem Pharmacol*, 1999, **57**, 727–741.

6 C. F. Thorn, C. Oshiro, S. Marsh, T. Hernandez-Boussard, H. McLeod, T. E. Klein and R. B. Altman, *Pharmacogenet Genomics*, 2011, **21**, 440–446.

7 D. Carlisi, A. De Blasio, R. Drago-Ferrante, R. Di Fiore, G. Buttitta, M. Morreale, C. Scerri, R. Vento and G. Tesoriere, *Cell Death Discov*, 2017, **3**, 17078.

8 P. Paramasivan, J. D. Kumar, R. Baskaran, C. F. Weng and V. V. Padma, *Cancer Drug Resist*, 2020, **3**, 647–665.

9 L. Toschi and F. Cappuzzo, *Onco Targets Ther*, 2009, **2**, 209–217.

10 Y. Xiang, Wen Ye, Chaohao Huang, Dinglai Yu, Hao Chen, Tuo Deng, Fan Zhang, Bin Lou, Jie Zhang, Keqing Shi, Bicheng Chen, and Mengtao Zhou, *Oxid Med Cell Longev*, 2018, **2018**, 2360427.

11 H. Hayashi, T. Kurata and K. Nakagawa, *Clin Med Insights Oncol*, 2011, **5**, 177–184.

12 D. Ren, N. F. Villeneuve, T. Jiang, T. Wu, A. Lau, H. A. Toppin and D. D. Zhang, *Proc Natl Acad Sci U S A*, 2011, **108**, 1433–1438.

13 S. Lee, M.-J. Lim, M.-H. Kim, C.-H. Yu, Y.-S. Yun, J. Ahn and J.-Y. Song, *Free Radic Biol Med*, 2012, **53**, 807–816.

14 M. J. Bollong, H. Yun, L. Sherwood, A. K. Woods, L. L. Lairson and P. G. Schultz, *ACS Chem Biol*, 2015, **10**, 2193–2198.

15 D. Zhang, Z. Hou, K. E. Aldrich, L. Lockwood, A. L. Odom and K. T. Liby, *Mol Cancer Ther*, 2021, **20**, 1692–1701.

16 Z. Hou, L. Lockwood, D. Zhang, C. J. Occhiuto, L. Mo, K. E. Aldrich, H. E. Stoub, K. A. Gallo, K. T. Liby and A. L. Odom, *RSC Med Chem*, 2023, **14**, 74–84.

17 A. Singh, S. Venkannagari, K. H. Oh, Y.-Q. Zhang, J. M. Rohde, L. Liu, S. Nimmagadda, K. Sudini, K. R. Brimacombe, S. Gajghate, J. Ma, A. Wang, X. Xu, S. A. Shahane, M. Xia, J. Woo, G. A. Mensah, Z. Wang, M. Ferrer, E. Gabrielson, Z. Li, F. Rastinejad, M. Shen, M. B. Boxer and S. Biswal, *ACS Chem Biol*, 2016, **11**, 3214–3225.

18 J. Zhang, L. Su, Q. Ye, S. Zhang, H. Kung, F. Jiang, G. Jiang, J. Miao and B. Zhao, *Oncotarget*, 2016, **8**, 7625–7636.

19 R. Modi, N. McKee, N. Zhang, A. Alwali, S. Nelson, A. Lohar, R. Ostafe, D. D. Zhang and E. I. Parkinson, *J Med Chem*, 2023, **66**, 6184–6192.
20 Z. Li, Y. Zhang, T. Jin, J. Men, Z. Lin, P. Qi, Y. Piao and G. Yan, *BMC Cancer*, 2015, **15**, 207.
21 A. E. M. A. Khan, V. Arutla and K. S. Srivenugopal, *Cells*, 2024, **13**, 1272.
22 Z. Yuan, X. Wang, B. Qin, R. Hu, R. Miao, Y. Zhou, L. Wang and T. Liu, *Drug Resist Updat*, 2024, **77**, 101160.
23 H. J. Prochaska and A. B. Santamaria, *Anal Biochem*, 1988, **169**, 328–336.
24 H. Sies, L. Packer, The "Prochaska" Microtiter Plate Bioassay for Inducers of NQO1, Quinone and Quinone Enzymes, Part B, **382**, 2004
25 C. E. Denlinger, B. K. Rundall, M. D. Keller and D. R. Jones, *Ann Thorac Surg*, 2004, **78**, 1207–1214; discussion 1207-1214.
26 R. Ikeda, L. C. Vermeulen, E. Lau, Z. Jiang, K. Sachidanandam, K. Yamada and J. M. Kolesar, *Int J Oncol*, 2011, **38**, 513–519.
27 W. Cao, Q. Yang, Z. Yuan, H. Li, W. Wang, X. Xiao, Z. Wang, L. Liang, P. Zhou, J. Liu, X. Hu and B. Zhang, *Biochem Biophys Res Commun*, 2020, **524**, 549–554.
28 J. Li, Y.-Y. Pan and Y. Zhang, *Oncol Lett*, 2013, **5**, 440–446.
29 A. K. Jain, D. A. Bloom and A. K. Jaiswal, *J Biol Chem*, 2005, **280**, 29158–29168.
30 W. D. Kim, Y. W. Kim, I. J. Cho, C. H. Lee and S. G. Kim, *J Cell Sci*, 2012, **125**, 1284–1295.

Chapter 5

Recombinant Protein Expression of Nrf2 and MafG

5.1 Introduction

Protein expression of the Nrf2 and MafG proteins was necessary to perform biophysical measurements of inhibitory compounds against the coiled-coil interaction discussed in **Chapters 2 and 3**. The most effective method for obtaining the proteins was to develop in-house recombinant protein expression, leading to the development of our own optimised protocol for producing each protein.

5.1.1 Methods for recombinant protein expression

Recombinant protein expression uses genetic engineering to produce a desired protein in a host organism that may not naturally express it. There are three main methods, summarised in **Table 5-1**, for recombinant protein expression. The most described method uses *Escherichia coli* (*E. coli),* a Gram negative bacterium, as the host organism.[1,2] For protein expression of the transcription factors described in this thesis, bacterial expression is preferable due to the literature precedent and unsuitability of eukaryotic or mammalian methods to produce high protein yields at a low cost.

Table 5-1 *Description of host models for recombinant expression of protein.*

Host Species	***Example Species***	***Benefits***	***Limitations***
Prokaryotic	*E. coli*	• Low cost • Scalable [3] • Fast growth time • mg/L yields	• No post translational modifications. • Solubility issues • Inclusion bodies.
Eukaryotic	*Saccharomyces cerevisiae*	• Simple post translational modifications. (glycosylation).[4] • mg to g/L yields. [5]	• Slow growth time • 30 kDa maximum protein size
Mammalian	Chinese Hamster Ovarian Cells	• Complex post translational modifications. • Industrial scale, high g/L yields.[6]	• Expensive • Transformation and cloning process time consuming

5.1.2 Recombinant protein expression in *E. coli*

Recombinant protein expression in bacterial cells was first achieved in the 1970's for the production of human hormones, leading the way in *E. coli* expression systems for protein production.[7,8] *E. coli* can be engineered to express a desired protein through gene transformation of a plasmid. Plasmids benefit from an ability to independently replicate, without incorporating into a host cells chromosomal DNA. The most used plasmid backbone for protein expression is the pET vector (**Figure 5-1)**, which contains several key features for gene expression and subsequent protein production.[9] The presence of

multiple cloning sites allows insertion of a gene of interest into the plasmid backbone (**5.4**). The inserted gene may be fused to a sequence encoding an affinity tag for purification **(5.3)**; tags are often enzymatically cleavable and can be positioned at either terminus of the gene for the protein of interest (**5.4**). Plasmids require a promoter to begin gene transcription; in pET this is a T7 promoter (**5.1)**. T7 promoters are regulated by the LacI operon (**5.7)** to suppress basal expression of the protein, allowing for strong user control over the expression activity. pET plasmids also carry an origin of replication (**5.6)** to allow the plasmid to multiply within *E. coli*. An antibiotic resistance gene (**5.5)** must also be present to provide an internal selection pressure of the host bacteria that have successfully incorporated the plasmid.

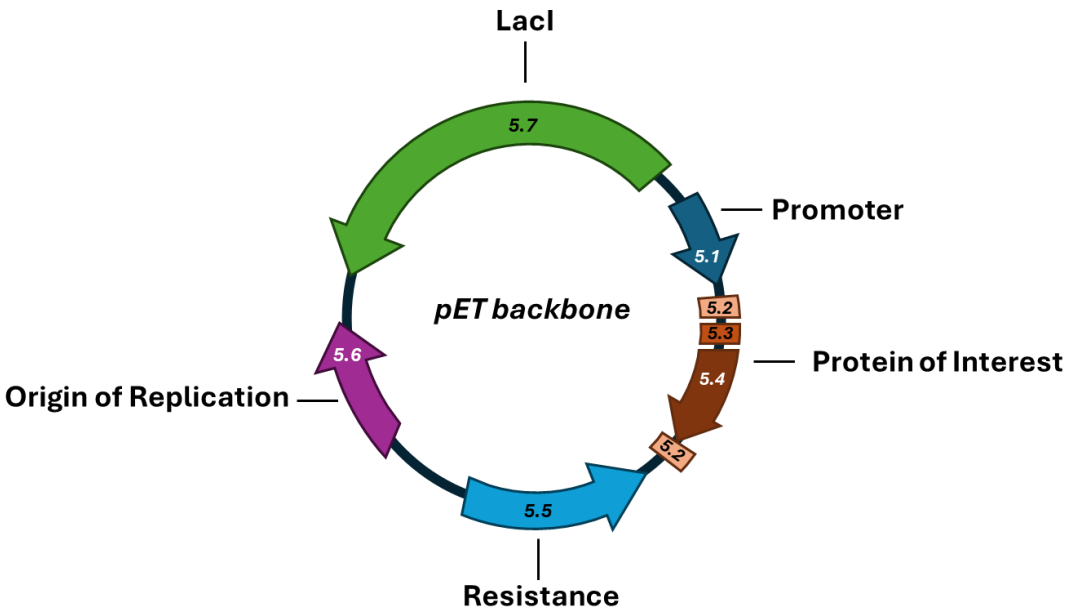


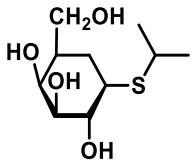
Figure 5-1 *Vector map of key features in a pET plasmid vector: Promoter (**5.1)** Restriction sites (**5.2**) Affinity Tag **(5.3)** Protein of interest **(5.4)** Gene for antibiotic resistance **(5.5)** Origin of replication **(5.6)** LacI gene **(5.7).***

Recombinant pET plasmid can be transformed into a competent *E. coli* cell line to perform expression. Competency is produced by chemical treatment, such as $CaCl_2$ or $MgCl_2$ to weaken the cell wall, improving cell permeability for plasmid uptake. After incubating the plasmid with competent *E. coli*, transformation is then further encouraged, by a physical method. Either electroporation, which uses electrical pulses, or heat shock at 42 °C to produce transient pores, allowing a plasmid to enter the cell.

After transformation, cells are incubated in nutrient rich super optimal medium (SOC) to recover from the cellular stresses caused by transformation. The bacteria are then grown on a solid agar support, in the presence of antibiotics to select for the competent *E. coli* cells containing the transformed plasmid. Individual colonies can then be

extracted for growth in a liquid culture which can be used for inducing protein expression. Bacterial growth is measured by optical density at a wavelength of 600 nm, an OD_{600} value of 0.6 corresponds to logarithmic phase of cell growth where cells are at optimum conditions for metabolic activities, such as protein expression.[10,11]

The T7 promoter utilised by pET vectors induces protein expression through repression of LacI.[12] LacI binds to the lac operon, preventing T7 RNA polymerase transcription. LacI also binds to the LacUV5 gene, which also supresses T7 RNA polymerase gene transcription, offering tight control of plasmid transcription. Inhibition of LacI is achieved using isopropyl β-D-1-thiogalactopyranoside (IPTG) (**Figure 5-2, 5.8**) an analogue of allolactose which induces removal of lac operon repressing proteins.



5.8

Figure 5-2 *IPTG chemical structure.*

Upon IPTG induction, cultures are incubated with agitation over an optimised time and temperature to yield efficient quantities of protein. Proteins are recovered from expressing cell cultures by lysing the bacterial cell wall. Common lysis methods use sonication and lysozyme treatment to recover soluble protein whilst limiting degradation. Lysates can subsequently be taken forward for protein purification to yield recombinantly expressed protein.

5.1.3 Protein purification

Expressing proteins within a host organism requires purification to separate them from the proteins expressed natively by the host organism. Consequently, chromatographic techniques are employed to achieve a high purity of the desired protein. Purification can include affinity tag, size exclusion or ion exchange chromatography relying on the unique properties of a protein to achieve separation.

5.1.3.1 His_6 tag purification

Ni^{2+} in complex with nitriloacetic acid (NTA) functionalised agarose (**Figure 5-3**) can be used for affinity purification of histidine tagged protein. Ni^{2+} contains six coordination sites supporting four interactions with carboxylate groups and a tertiary amine on NTA and two sites for coordination of neighbouring histidine residues. This coordination allows for impurities to be eluted separately to the desired protein. Protein can then be

collected by increasing the imidazole concentration within the buffer, as it is structurally similar to histidine (**Figure 5-3**), competitively binding to the Ni^{2+} complex and eluting the affinity tagged protein.

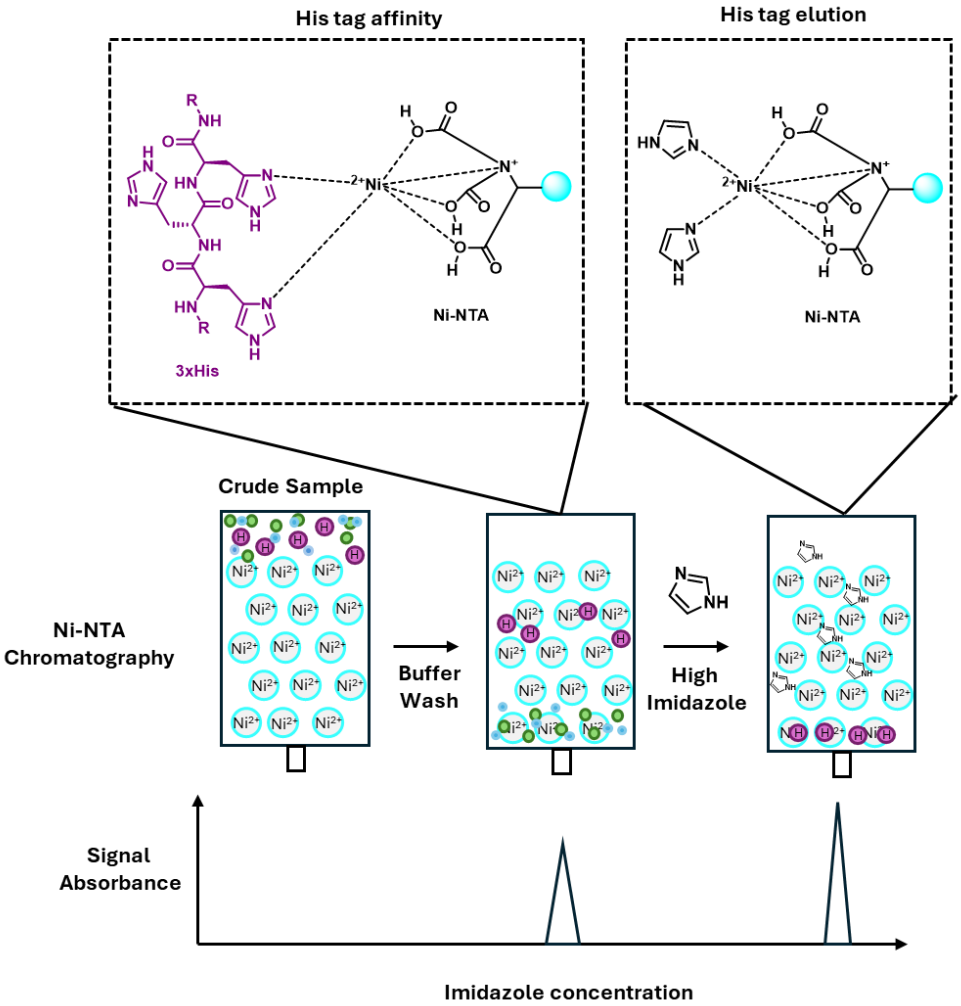


Figure 5-3 *Ni-NTA affinity chromatography schematic.*

5.1.3.2 Size exclusion chromatography

Where affinity tag purification fails to isolate a target protein from impurities of different sizes, size exclusion chromatography (SEC) can be used to achieve separation. In SEC columns, the gel matrix contains pores that large biomolecules are unable to enter, allowing faster travel times through the column. Smaller molecules are slowed by diffusion through the pores of the solid support, creating separation by size **(Figure 5-4)**.

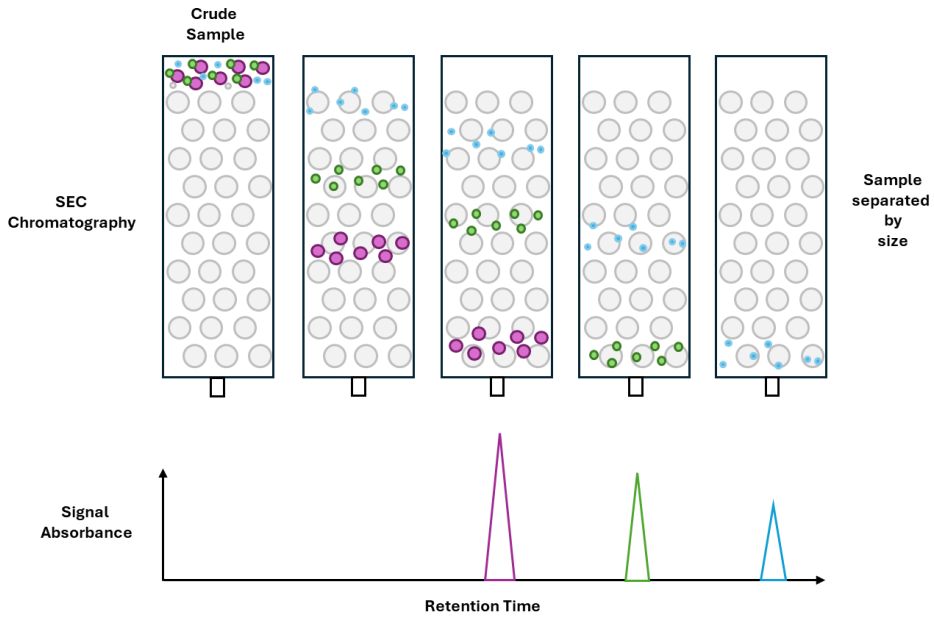


Figure 5-4 *Process of size exclusion chromatography for eluting size-varied impurities.*

5.1.3.3 Ion exchange chromatography

Protein purification can be achieved by separating molecules based on charge using ion exchange chromatography (IEC). Cationic or anionic exchange columns can be used to separate contaminants from the protein of interest. For example, a positively charged resin will interact more readily with negatively charged proteins whilst positively charged protein will pass through the column **(Figure 5-5)**. An interacting protein can be eluted from the column by increasing the ionic concentration with salt, effectively competing off the protein to create separation of proteins by charge.

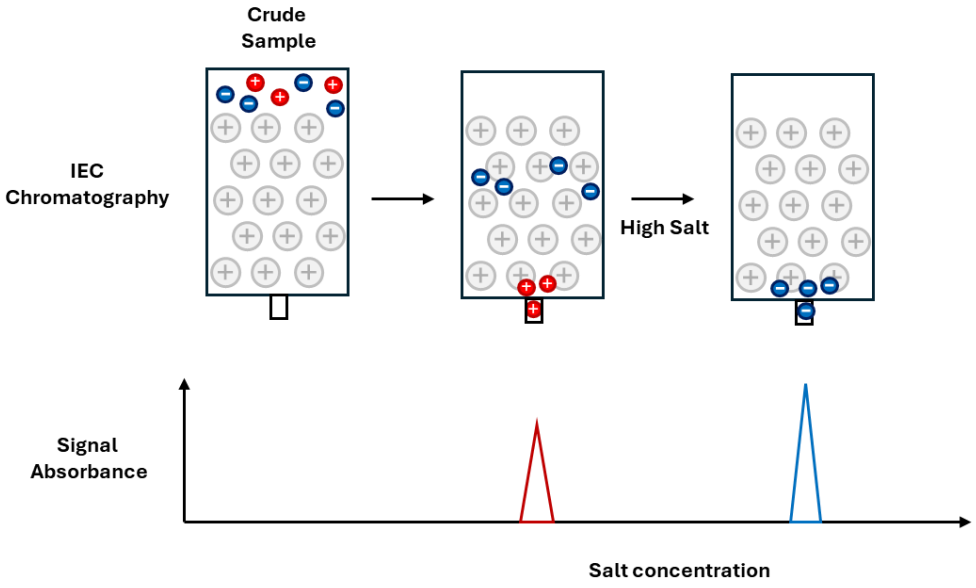


Figure 5-5 *Schematic of IEX purification of a charged protein across a positively charged resin.*

5.2 Chapter Aims

Establishing an in-house system of protein expression is the most cost-viable method for running an investigation of protein-protein interaction (PPI) disruption by biophysical assays. This chapter aims to describe the expression and purification of Nrf2 and MafG protein for the assays discussed in **Chapters 2** and **3**. The expression of Nrf2 and MafG was performed using *E. coli* and plasmids previously described within the literature.[18,19] Optimisation of transformation, expression and purification was explored. Additionally, this chapter will discuss the use of Nrf2 in surface plasmon resonance experiments for observing binding interactions with MafG derived peptides.

5.3 Results and Discussion

5.3.1 His-Halo-Nrf2

5.3.1.1 Expression of Nrf2 in literature

Nrf2 is expressed through transcription of the *NFE2L2* gene.[13] Stable recombinant expression of Nrf2 is challenging, as it is predicted to be partially disordered due to the high net charge across the full length protein.[14] This lack of secondary structure has been confirmed by circular dichroism, which found Nrf2 to be 27% alpha helical.[15] Consequently, the structural organisation of Nrf2 has largely been unknown, with only crystal structures of the DNA binding domain and parts of Neh2 (containing ETGE and DLG motif) (PDB: 2DYH).[1] As discussed in **Chapter 1.5.1** a crystal structure was most recently published in December 2022 on the Neh1 domain of Nrf2 in complex with MafG and ARE DNA (PDB: 7X5F).[16]

Despite challenges in crystallising the disordered protein, expression of the full length Nrf2 protein has been achieved for multiple studies of Nrf2 activity.[15–19] These studies have highlighted a range of challenges in the protein purification process including solubility issues, yield, and purification from aggregating protein molecules.[15,17,18] Herein, we describe our own methods of optimising the yield and purification of recombinantly expressed Nrf2.

5.3.1.2 His-Halo-Nrf2

Expression of full length Nrf2 containing a His_6Halotag was trialled with Nrf2 cloned into a pET28A plasmid designed by Professor Yimon Aye.[18] Halotag was included as whilst cleavable, it provided the option to utilise fluorescent Halotag ligands for additional biophysical assay exploration. Obtained from Addgene, plasmid #62455 was provided in a strain of DH5α cells. This cell line is used for plasmid propagation, to yield a high

quantity of stable plasmid that can be easily extracted. Plasmid quality is maintained in DH5α *E. coli* through the presence of a RecA1 mutation which reduces homologous recombination of the circular DNA and an *end*A mutation to prevent plasmid degradation by endonuclease enzymes.

5.3.1.3 Plasmid purification and sequencing of pET28A His-Halo-Nrf2

A transformed culture of *E.coli* DH5α containing the pET28A plasmid was inoculated on solid agar containing 50 μg/mL of kanamycin to maintain selection pressure against the loss of the plasmid. Overnight incubation at 37 °C yielded individual colonies that were isolated and grown in 10 mL luria broth (LB) to produce a liquid culture. Cells were prepared for plasmid extraction using a QIAprep spin Miniprep kit to yield 72.3-97.3 ng/μL of DNA with a purity absorbance ratio of 1.86 which was then sent for sequencing.

Initially, sequencing of the plasmid was trialled in the presence of a primer complementary to the T7 promoter of the pET28A vector (**Figure 5-1**). Sequencing analysis found the sample to be poorly resolved with ambiguous overlapping peaks at each nucleotide base. Analysis by basic local alignment search tool (BLAST) did not find successful matches against the Nrf2 gene within the Protein Data Bank (PDB). Sequencing may fail to match with the protein of interest if the T7 promoter is positioned too far from desired gene.[20] As such, sequencing of the sample was repeated using a primer for the Nrf2 gene (5'- AAC CAC CCT GAA AGC ACA GC-3') providing a 100% match for the human NFE2L2 gene within the extracted plasmid sample. Successful extraction of the pET28A plasmid containing the Nrf2 gene permitted subsequent transformation of the plasmid into *E. coli* for protein expression.

5.3.1.4 Protein expression of His-Halo-Nrf2

For successful expression of the Nrf2 gene, cell lines with machinery for the T7 promoter transcription were used. The most commonly used *E. coli* cell line is competent BL21 (DE3), an engineered strain, lacking proteases to prevent protein degradation.[21,22] BL21 (DE3) cells contain the gene for T7 RNA polymerase, indicated by the presence of a λDE3 lysogen encoding its expression. pLysS containing strains express a T7 lysozyme to suppresses basal expression of the T7 promoter, which is particularly useful if a protein proves toxic to the host. [22]

Rosetta2 cells are a variation on BL21, with an additional pRARE plasmid to encode for rare tRNA codons, designed to enhance expression efficiency of mammalian or eukaryotic proteins.[23] Rosetta2 also utilises a lacY mutation that prevents expression of

lactose permease, which may otherwise induce basal expression of the transformed plasmid.[24]

Optimisation of conditions to stably transform bacterial cells were trialled, varying *E. coli* cell lines, heat shock conditions and SOC incubation time (**Table 5-2**). The addition of β-mercaptoethanol (β-ME) was also trialled, as it has been reported to improve the transformation efficiency of plasmids in *E. coli* by inactivating surface nuclease that might degrade migrating DNA.[25] Variation on heat shock times have been reported, and increasing heat shock exposure to 90 seconds proved successful in producing colonies upon agar. Inoculation failed to produce growth in liquid culture, suggesting the plasmid was not stably transformed. Increasing incubation time in SOC can allow better cellular recovery and establishes the bacteria's ability for antibiotic resistance.[26] Ultimately, leading to the successful growth of liquid culture for transformed BL21 (DE3) pLysS cells, conditions trialled summarised in **Table 5-2.**

Table 5-2 *Conditions trialled for plasmid transformation of E.coli host cells with pET28A Nrf2 plasmid.*

E. coli	*Heat Shock*	*SOC Incubation*	*Agar Culture*	*Liquid Culture*
BL21 (DE3) pLysS	30 s	1 h	No Growth	-
BL21 (DE3) pLysS	45 s	1 h	No Growth	-
BL21 (DE3) pLysS	90 s	1 h	**Growth**	No Growth
BL21 (DE3) pLysS	45 s + 48 mM βME	1 h	**Growth**	No Growth
BL21	45 s	1 h	No Growth	-
BL21	90 s	1h	No Growth	-
BL21	90 s + 48 mM βME	1h	No Growth	-
Rosetta 2	45 s + 48 mM βME	1 h	No Growth	-
Rosetta 2	90 s + 48 mM βME	1 h	No Growth	-
Rosetta 2	45 s + 48 mM βME	2 h	**Growth**	No Growth
BL21 (DE3) pLysS	45 s + 48 mM βME	2 h	**Growth**	**Growth**

Successful liquid cultures from the transformed BL21 (DE3) were prepared for large scale protein expression at 500 mL. Cultures were grown in LB medium with kanamycin 15 µg/mL for 4 hours to reach a log growth phase of OD_{600} of 0.6 at 37 °C at 180 rpm. Expression was induced using 1 mM IPTG for 4 hours at 25 °C and the resulting culture was pelleted and prepared as a cell lysate for protein purification, a process described in **5.5.2**.

5.3.1.5 Purification of His-Halo-Nrf2

As the Nrf2 protein was expressed with an N-terminal Hisx6 tag, affinity chromatography was used to selectively interact with the histidine residues of the tag. The Nrf2 protein was purified with Ni-NTA functionalised resin (**Figure 5-6**)

producing a low purity preparation, with co-eluting impurities from 10 – 100 kDa. Proteins with high histidine, cysteine or tryptophan content can also undergo coordination with Ni^{2+}, limiting the ability of Ni-NTA resin to yield high purity isolations of a desired protein.[27,28]

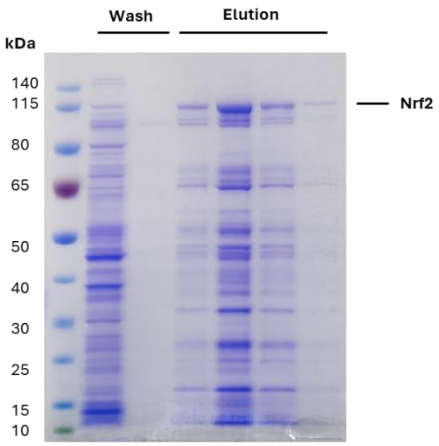


Figure 5-6 *SDS-PAGE analysis of 4-12% Bis-Tris gel after Ni-NTA purification. Wash buffer composed of 50 mM Tris, 20 mM Imidazole, 300 mM NaCl at pH 8. Elution buffer composed of 50 mM Tris, 250 mM Imidazole, 300 mM NaCl at pH 8. Gel ran in 1x MOPS at 200 V over 35 minutes.*

The low ratio of Nrf2 to impurities observed by gel electrophoresis suggested it would first be preferable to explore methods to increase the desired protein expression relative to the impurities. This was explored by altering the time and temperature of expression conditions. Inducing protein expression at a lower temperature can improve protein solubility through reduced aggregation of proteins, which has previously been reported as a challenge in Nrf2 protein expression.[29,30] Addition of glycerol can provide an improved source of carbon for bacterial cell growth, which has been shown to correlate with higher protein yields.[31]

Protein was prepared as a lysate and purified by Ni-NTA chromatography, elution fractions containing the desired Nrf2 protein were collected and concentrated into 50 mM Tris at pH 7.4 and measured by Nanodrop Spectrophotometer-1000 measuring 280 nm absorbance to quantify protein concentration (**Table 5-3**). The addition of glycerol did not change yield and was not included moving forwards. Expression was significantly higher over 16 h at 18 °C, so was chosen for subsequent protein expression of the Nrf2 protein **(Figure 5-7)**.

Time	*Temperature*	*Additive*	*Concentration*
4 h	37 °C	-	0.12 mg/mL
4 h	37 °C	+ 5% Glycerol	0.06 mg/mL
16 h	18 °C	-	0.58 mg/mL
16 h	18 °C	+ 5% Glycerol	0.45 mg/mL

Table 5-3 *Summary of experimental conditions of HisHaloNrf2 in BL21 cells*

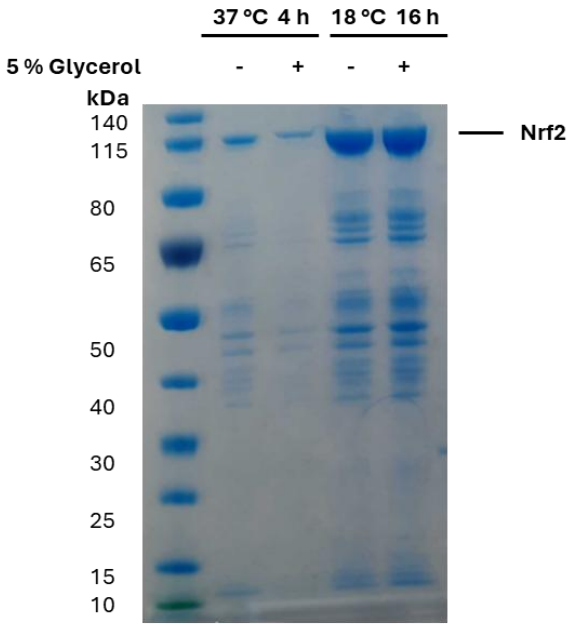


Figure 5-7 *SDS PAGE analysis of concentrated elution buffer fractions by 4-12% bis-tris gel in 1x MOPS from experimental conditions referred to in* ***Table 5-3.***

High concentrations of IPTG can adversely influence bacterial cell growth, reducing the expression of a desired recombinant protein.[32] Expression conditions were trialled using a reduced concentration of IPTG, however this appeared to induce stronger co-expression of impurities compared to 1 mM IPTG (**Figure 5-8**). Whilst Ni-NTA purification was performed on a gradient to improve separation, impurities were still evident at both IPTG concentrations conditions trialled.

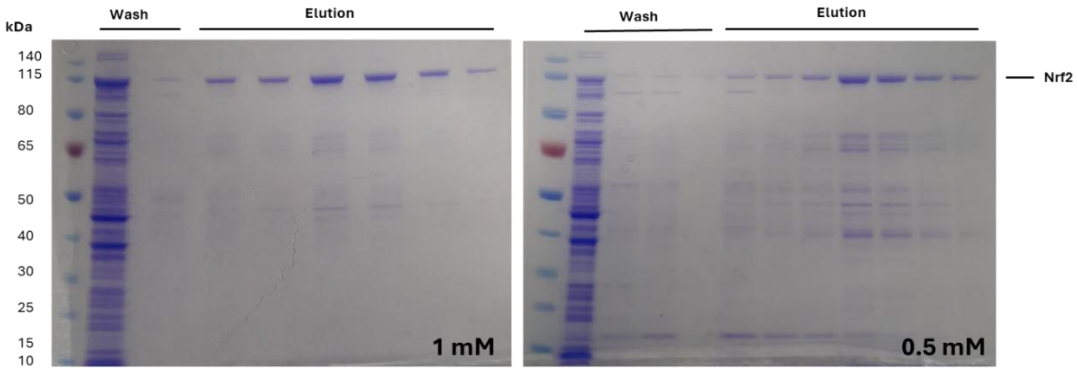


Figure 5-8 *SDS-PAGE 4-12% Bis-tris analysis of Nrf2 protein expression using 1 mM (left) and 0.5 mM IPTG (right). Wash buffer composed of 50 mM Tris, 20 mM Imidazole, 300 mM NaCl at pH 8. An elution gradient was created diluting Elution buffer (50 mM Tris, 250 mM Imidazole, 300 mM NaCl at pH 8) in wash buffer from 15 -100%.*

The removal of co-eluting impurities across a range of sizes was investigated by size exclusion chromatography. Firstly, a culture of BL21 (DE3) pLysS transformed with pET28A Nrf2 expressing plasmid was induced at 1 mM IPTG overnight at 18 °C and prepared for Ni-NTA purification. Crude lysate was loaded onto a Ni-NTA column and washed with lysate buffer (50 mM tris, 10 mM imidazole 300 mM NaCl, pH 8) until the absorbance signal had returned to baseline (**Figure 5-9, A**). An elution gradient was subsequently performed from 0 – 100% of elution buffer (50 mM tris, 250 mM imidazole 300 mM NaCl, pH 8) and fractions were visualised by gel electrophoresis shown in **Figure 5-9, B.**

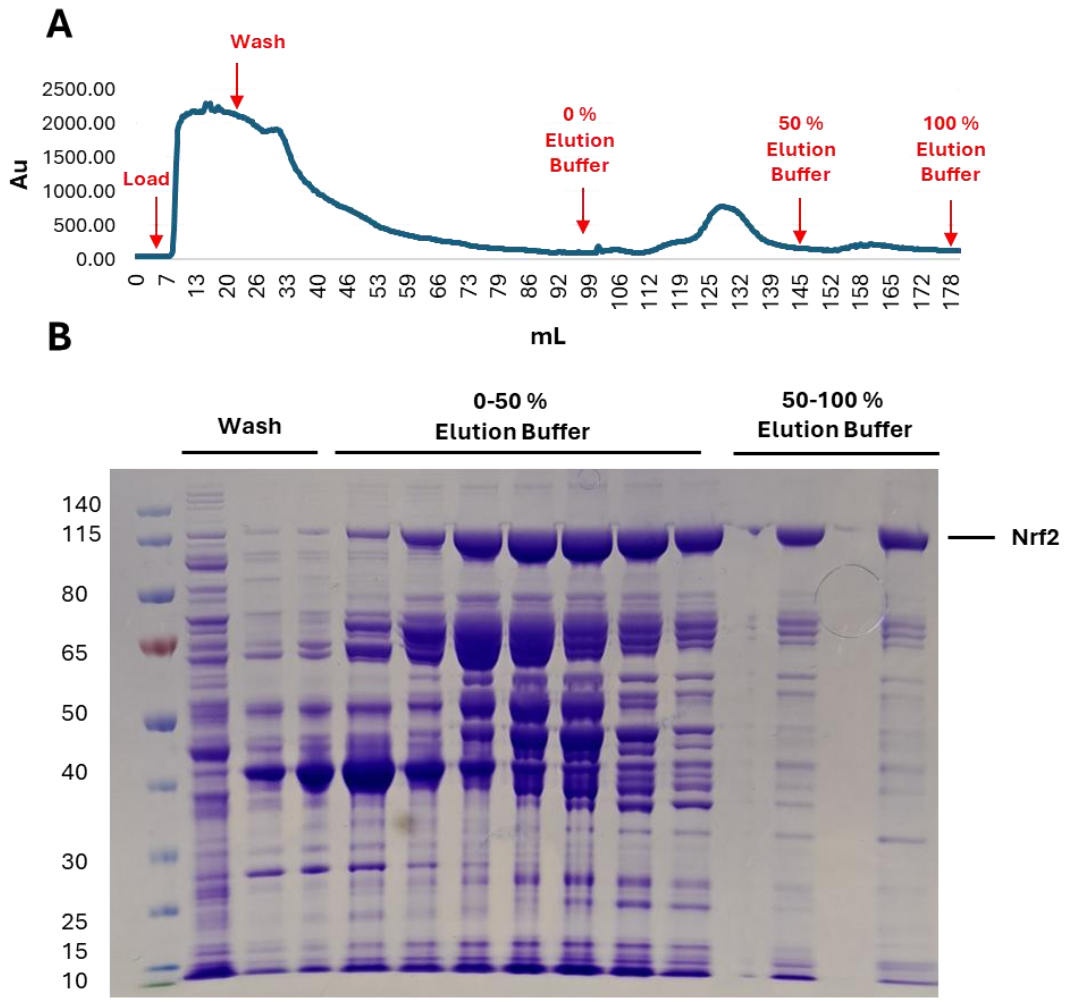


Figure 5-9 (A) *AKTA purification of Nrf2 protein by HisTrap™ FF on AKTA Prime Pure using 17/70 mL of lysate collected from 4 L of culture.* **(B)** *4-12% Bis-tris SDS-PAGE analysis of AKTA PURE Ni-NTA purification of Nrf2 protein.*

Fractions from 50-100% elution buffer from the Ni-NTA column produced the highest Nrf2 to impurity ratio and were subjected to SEC to separate the remaining proteins by eluting on molecule size across a porous gel matrix.

A superdex 200 SEC column was used for its optimal separation range, capable of separations of globular proteins from 600 kDa to 10 kDa in size. Fractions from 50-100% elution buffer, indicated from **Figure 5-9,** were injected onto the column, protein elution monitored by mAu absorbance at 280 nm. Analysis of the fractions by SDS-PAGE gel found impurities were still co-eluting with the Nrf2 protein after SEC purification, shown in **Figure 5-10**. Failure to separate molecules by size through SEC can be an indication that the impurities are aggregating with the desired protein, causing co-elution of the crude mixture.[33] As such further investigation was required to identify a purification method capable of separating the impurities.

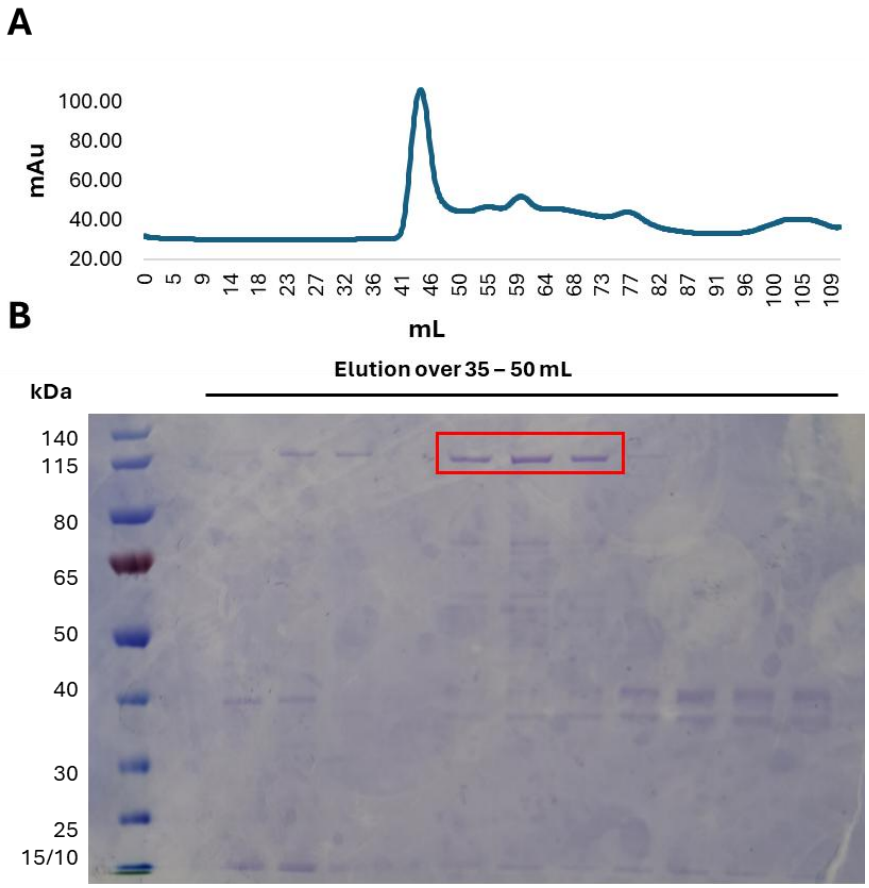


Figure 5-10 (A) *AKTA purification using a HiLoad™ 16/60 Superdex ™ 200 Prep Grade SEC at 0.1 mL/min for 18 h of Nrf2 after Ni-NTA purification.* ***(B)*** *4-12% Bis-tris SDS-PAGE of fractions collected from size exclusion chromatography.*

As Nrf2 has a pI of approximately 4.7, ionic exchange purification was investigated. Anionic exchange columns were prioritised for the capture and separation of protein bands, as the net charge of Nrf2 is negative in a phosphate buffer of pH 7.4.[34]

Protein lysate from Ni-NTA purification was prepared by buffer exchange into Tris buffer at pH 7.4 to remove the salt and imidazole, concentrating the sample to 2 mg/mL. The

sample was loaded onto an ion exchange chromatography (IEC) stack passing through a cationic, weak and strong anionic exchange column, order summarised in **Table 5-4**. After loading the columns as a set, the columns were separated and individually eluted by increasing the salt concentration of the column from 100 mM to 1 M of NaCl monitoring the 280 nm absorbance for protein elution. Ion exchange purification did not remove the co-eluting impurities with the Nrf2 protein (**Figure 5-11**). This perhaps suggests that the impurities were creating non-specific interactions with the protein causing their co-elution by SEC and perhaps are hydrophobic in nature, rationalising why separation could not be achieved by IEC.[35,36]

Table 5-4 *IEC columns employed for purification of HisHaloNrf2 fractions after Ni-NTA purification.*

Order	*Column*	*Exchange Type*	*Result*
1	Sulfopropyl High Performance (SP HP)	Strong Cationic Exchanger	Impurities Eluted
2	Diethylaminoethyl fast flow (DEAE FF)	Weak Anionic Exchanger	Nrf2 eluted with impurities
3	Quaternary Ammonium Flow (Q F)	Strong Anionic Exchanger	Nrf2 eluted with impurities

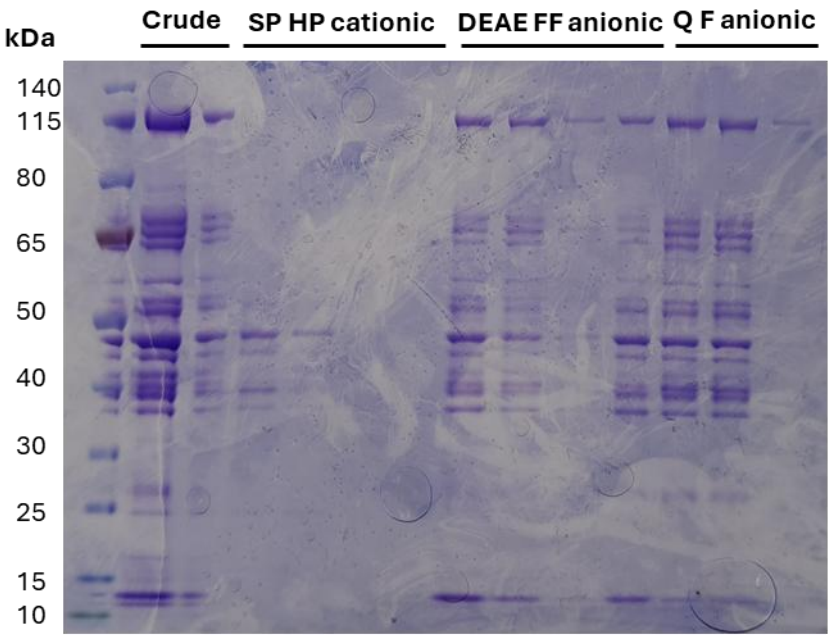
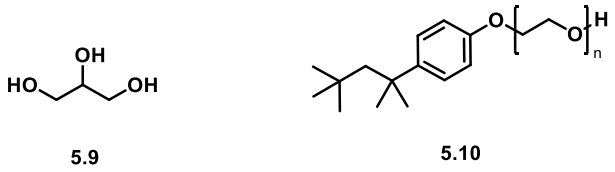


Figure 5-11 *4-12% SDS-PAGE analysis of IEX fractions collected off the SP HP cationic exchanger, DEAE FF weak anionic exchanger and Q F exchanger.*

When facing challenges with removing aggregates in protein purification alterations to the purification buffer system can be made through additives such as detergents. The use of non-ionic surfactant in protein purification can support protein solubility and limit aggregation through reduction of non-specific binding of contaminating proteins

through ionic or hydrophobic interactions.[37–39] Triton X-100 (**Figure 5-12, 5.9**) is an amphiphilic organic compound derived from a polyoxyethylene with a hydrophilic alkylphenyl head, it is a commonly used surfactant in protein purification. Similarly, the addition of glycerol (**Figure 5-12, 5.10**) to wash steps of protein purification can support protein stabilisation through the ability of alcohols to mediate amphiphilic and electrostatic interactions between protein and the glycerol to prevent protein aggregation at high concentrations.



5.9 5.10

Figure 5-12 *Chemical structure of glycerol (**5.9**) and triton x-100 (**5.10**).*

A second purification of the Nrf2 protein by Ni-NTA was performed with 0.1% Triton X-100 and 5% glycerol in the wash buffer, performing an extended wash at 1 mL/min for 150 mL to exhaustively strip co-eluting impurities from the Nrf2 protein. This optimisation successfully eluted clean recombinant Nrf2 protein requiring only Ni-NTA purification to achieve efficient yields of 1.95 mg from 4 L of culture of Nrf2 after buffer exchange into PBS pH 7.4 (**Figure 5-13**).

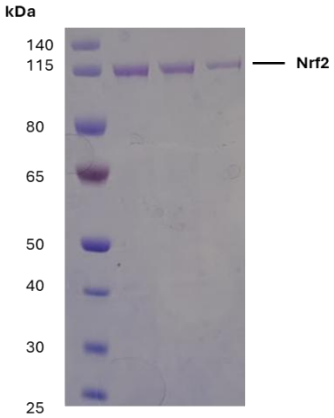


Figure 5-13 *4-12% Bis-tris SDS-PAGE gel of purified Nrf2**.** Achieved by Ni-NTA affinity purification (twice), washed with 20 mM Imidazole 50 mM NaH_2PO_4, 500 mM NaCl, 0.01% Triton x-100, 5% Glycerol pH 7.5 at pH 8. Eluted over a gradient of 40-250 mM Imidazole, Nrf2 containing fractions concentrated and buffer exchanges into PBS pH 7.4.*

Validation of protein expression was achieved using peptide mass fingerprinting to identify sequence coverage across both the attached affinity tag and the Nrf2 protein (**Appendix 5-1**), providing confidence in the expression of the desired protein. Protein was stored at –80 °C until required for biophysical assays used in **Chapters 2** and **3.**

5.3.2 His-MafG

5.3.2.1 Expression of MafG in literature

MafG is a considerably smaller protein in comparison to Nrf2 (162 residues) and has proven easier to express for functional activity-related studies. As it lacks extended domains for transcriptional recruitment the protein observes a more stable secondary structure. The full structure has been experimentally resolved in complex with a DNA consensus sequence (PDB: 3A5T).[40–42] In the expression and purification of MafG, literature suggests the purification, and yields, of MafG are highly efficient, however there is reference to MafG expression proving toxic in *E. coli*.[16,19,40]

5.3.2.2 pET15B His-MafG expression

We designed an expression plasmid in-house using a pET15B plasmid, as reported for expression of MafG in the MafG/DNA complex and Nrf2/MafG/DNA complex crystal structures.[40] Recombinant pET15B plasmid containing full length $MafG_{1\text{-}165}$ was designed and ordered from Genscript inserting the *MAFG* gene at excision sites for BamH I and NdeI (**Figure 5-14**).

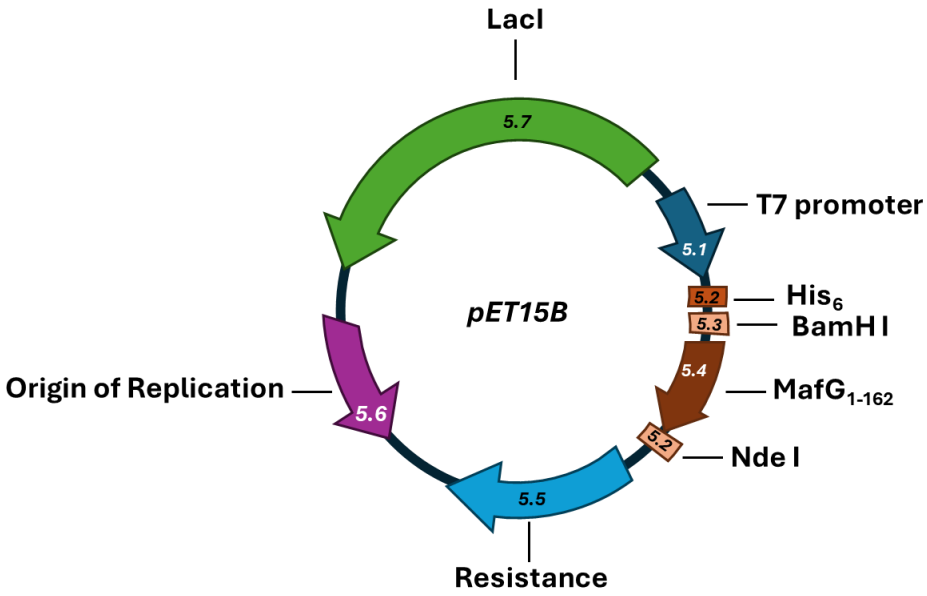


Figure 5-14 *pET15B plasmid vector map.*

Lyophilised plasmid was received from genscript, nuclease-free water was added to solubilise the circular DNA. Following the recommended protocol by genscript, 2 µL of the plasmid at 100 ng/µL was used for transformation into BL21 (DE3) and Rosetta 2 (DE3) pLysS cells following the successful transformation procedure previously described in **Table 5-5**. Colonies were found in both BL21 and Rosetta 2 cell lines (**Table 5-5**).

Table 5-5 *Cell lines and agar conditions used for transformation of His-MafG PET15B plasmid.*

Cell line	***Antibiotics Used***	***Result***
BL21	50 µg/mL ampicillin	Growth
Rosetta 2	50 µg/mL ampicillin + 34 µg/mL chloramphenicol	Growth

10 mL cultures were grown in LB overnight at 37 °C 180 rpm and used for large-scale culture at 500 mL LB with 50 µg/mL ampicillin at 37 °C for 3-4 h until the cultures reached log phase growth at OD_{600} 0.6. Gene expression was induced with 1 mM IPTG for 4 hours at 37 °C 180 rpm before the culture was prepared for lysate purification by Ni-NTA affinity column. As the molecular weight of the recombinant MafG protein (19 kDa) was relatively small, SDS-PAGE analysis was performed using 12% polyacrylamide Bis-Tris in MES buffer to improve visualisation. This revealed multiple bands around 20 kDa (**Figure 5-15)**.

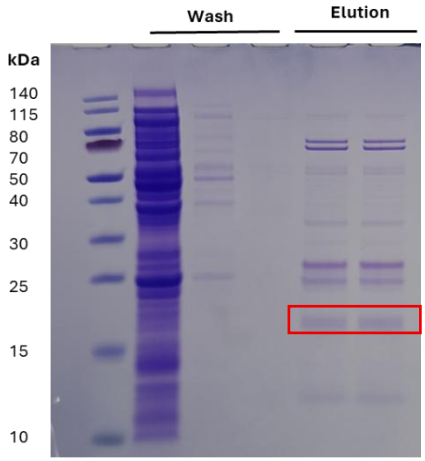


Figure 5-15 *12% Bis-Tris SDS-PAGE analysis of Ni-NTA resin purification of HisMafG expression in Rosetta 2 (DE3) cells.*

Protein bands can sometime migrate unexpectedly in gel electrophoresis.[43,44] To efficiently confirm expression of the His-tagged MafG, a western blot was run on crude lysates. To find bands corresponding to the tagged protein expression, an anti-His_6 tag antibody was used with an anti-rabbit HRP conjugated secondary antibody (**Figure 5-16**). Expression conditions were repeated for both BL21 and Rosetta 2 cell lines transformed with the pET15B testing conditions in LB, 2xYT and TB growth medium in 500 mL cultures to determine an optimum growth medium. Western blot confirmed His tag protein expression at 25 kDa in the Rosetta 2 cell line with the pET15B plasmid with strongest expression found in TB growth medium (**Figure 5-16**).

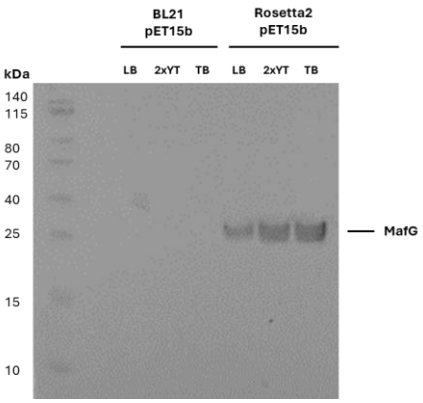


Figure 5-16 *His-MafG protein expression identified via western blot using whole cell lysate from BL21 and Rosetta 2 transformed cells expressed with 1 mM IPTG incubated for 4 h at 37 °C (1) Protein Ladder (3) BL21 PET15B MafG grown in LB (4) BL21 PET15B MafG in TB (5) BL21 PET15B MafG in 2xYT (6) Rosetta 2 PET15B MafG in LB (7) Rosetta 2 PET15B MafG in 2xYT (8) Rosetta 2 PET15B MafG in TB.*

Mass spectrometry was performed on the protein band isolated at 25 kDa by gel electrophoresis, it is unclear why the band runs at this height as this is inconsistent with the band migration shown later in **Figure 5-17**. Analysis of the band excised matched a mass of 19.7 kDa by MALDI-TOF and matching sequence coverage by peptide mass fingerprinting (PMF) (**Appendix 5-2**), providing confidence in protein identification.[45]

5.3.2.3 Protein purification of pET15B expressed His-MafG

In a reduced supply of glucose within a growth medium, bacteria will begin to metabolise lactose sugars instead, a process controlled by the lac operon. Ensuring good supply of glucose in the growth medium can prevent basal expression from reduced repression by the dual-functioning lac operon. As MafG expression has previously been reported to be toxic to bacterial cells, 0.5% glucose was added to the growth medium of *E. coli* cultures in large scale purification to suppress basal suppression.[16]

Expression was performed in 4 L of bacterial culture, prepared as a cell lysate for Ni-NTA column chromatography on an elution gradient of 50-250 mM imidazole. Fractions were collected for SDS-PAGE (**Figure 5-17, left**) producing a clean band of the desired MafG protein on elution at 250 mM imidazole in elution fraction 5. Elution fractions containing MafG were buffer exchanged into lysis buffer for a second Ni-NTA column to achieve further elution of clean bands of His-MafG (**Figure 5-17, right**), which was then subject to a final buffer exchange into PBS pH 7.4 and stored in aliquots at –80 °C at a final concentration of 20 – 25 µM.

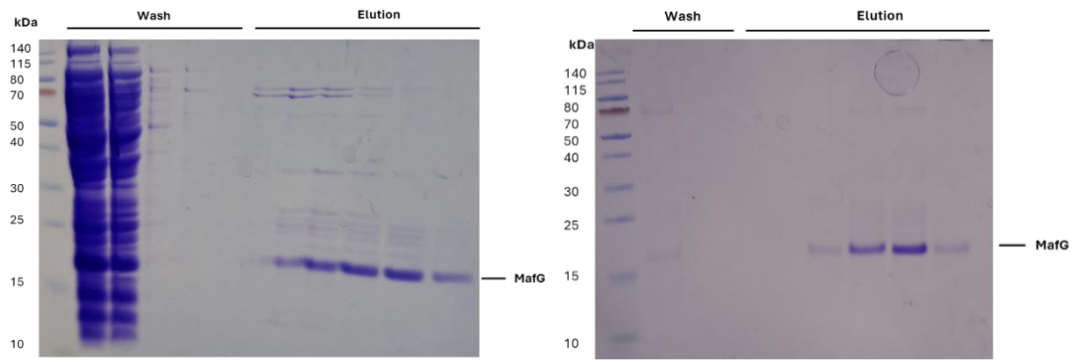


Figure 5-17 *12% Bis-Tris SDS-PAGE analysis of Ni-NTA resin purification of HisMafG expression in Rosetta 2 E. coli. (left) First Ni-NTA column. (right) Second Ni-NTA column.*

5.4 Conclusion

The expression of the BZIP transcription factors Nrf2 and MafG have been reported in the literature for use in biophysical assays such as fluorescence polarisation, electrophoretic mobility shift assays and protein crystallography. Nrf2 expression has previously demonstrated challenges with solubility and difficulty in purification from aggregating impurities resulting in low yield.[15,17,18]

Herein we have described optimised expression conditions of HisHaloNrf2 protein expression. This was successfully achieved by transformation into variants of the BL21 *E. coli* cell line. Optimisation of IPTG-induced protein expression conditions found Nrf2 was best induced at 18 °C overnight. After exhaustive efforts to purify Nrf2 through Ni-NTA (**Figure 5-8**), SEC (**Figure 5-10**) and IEC purification techniques (**Figure 5-11**), it was thought that impurities were forming complexes with the protein, reducing the efficiency of these techniques. Adding non-ionic detergent and glycerol to the wash steps in affinity tag purification generated successful protein purification from a singular purification method, yielding Nrf2 at concentrations of 6.5 µM in PBS at pH 7.4 (**Figure 5-13**).

MafG protein expression has been achieved for measurement of protein-DNA interactions in protein crystallography studies, disclosing efficient expression to achieve high yields of protein.[16,19,40] Recombinant expression of MafG was achieved using a pET15B plasmid containing the full length *MAFG* gene leading to successful transformation of Rosetta 2 cell line. Expression conditions were optimised to 30 °C over 3-4 h with an efficient purification of purified protein at yields of His_6MafG of 1.5 mg per 4 L culture stored in PBS at pH 7.4 (**Figure 5-17**).

Future work would include further recombinant protein expression of Nrf2, without the expression of the Hisx6 and Halotag and instead employ different affinity tags for

expanding the availability of the protein for biophysical assay development for FRET, HTRF or ligand affinity-based fluorescence polarisation assays.

5.5 Experimental

5.5.1 Reagents

Reagents were purchased from Fisher Scientific, Merck or Cytiva.

5.5.2 Protein expression

5.5.2.1 His_6Halo-Nrf2

pET28a-His6-Halo-TEV-Nrf2 was kindly provided by Dr Yimon Aye.[18] Plasmid was replicated in NEB® 5-alpha cells by 10 mL overnight cultures of LB and 15 µg/mL kanamycin at 37 °C 180 RPM. Cells were then pelleted at 6000 *g* for 18 min at 4 °C. Pellet was then prepared for extraction using a QIAprep® Spin Miniprep Kit following the manufacturers protocol. Plasmid was then confirmed by Mix2Seq sequencing from Eurofin, analysis performed in chromas, BLAST sequence comparison found 100% match from human *NFE2L2* gene.

10 ng of plasmid was incubated with chemically competent BL21 (DE3) *E. coli* further treated with 48 mM β-ME for 30 minutes at 4 °C. The *E. coli* was then heatshocked at 42 °C for 45 seconds followed by a 2-hour incubation in SOC media. Overnight cultures were prepared in LB containing 15 µg/mL kanamycin, further diluted 1:50 in LB media with 15 µg/mL Kanamycin. After 4-5 hrs incubation at 37 °C 180 RPM, expression was induced at OD_{600} 0.6-0.8 with 1 mM IPTG at 18 °C for 16 hrs, 180 RPM. Cell culture was pelleted by centrifugation 14000 *g* for 18 min at 4 °C. The pellet was then resuspended in buffer (50 mM NaH_2PO_4, 500 mM NaCl, 10 mM Imidazole, 0.01% Triton x-100, 5% Glycerol pH 7.5) and sonicated for 8.2 min at 60 V on a 3 x pulse cycle. Lysed solution was then centrifuged at 5000 *g* for 30 min at 4 °C. Lysate supernatant was then loaded onto Ni-NTA resin (Thermo Scientific™ HisPur™ Ni-NTA) at 4 °C, Ni-NTA resin was washed with excess buffer (20 mM Imidazole 50 mM NaH_2PO_4, 500 mM NaCl, 0.01% Triton x-100, 5% Glycerol pH 7.5 at pH 8). His_6Halo-Nrf2 was then eluted over an imidazole gradient of 40 mM – 250 mM in 50 mM NaH_2PO_4 pH 8 over 10 CV. Purification by Ni-NTA resin was then repeated. His_6Halo-Nrf2 was concentrated and resuspended in PBS pH 7.4, with a final concentration of 0.66 mg/mL and stored at -80 °C. Purification was confirmed by SDS-PAGE on a 4-12% Bis-Tris gel (Invitrogen) in 1x MOPS buffer. Sequence was confirmed by mass spectrometry. His_6Halo-Nrf2, 104.7 kDa:

(His$_6$_Halo_Nrf2)[§]

HHHHHH_SSGLVPRGSHMAEIGTGFPFDPHYVEVLGERRMHYVDVGPRDGTPVLFLHGNPT
SSYVWRNIIPHVAPTHRCIAPDLIGMGKSDKPDLGYFFFDDHVRFMDAFIEALGLEEVVLVIHD
WGSALGFHWAKRNPERVKGIAFMEFIRPIPTWDEWPEFARETFQAFRTTDVGRKLIIDQNVFIE
GTLPMGVVRPLTEVEMDHYREPFLNPVDREPLWRFPNELPIAGEPANIVALVEEYMDWLHQS
PVPKLLFWGTPGVLIPPAEAARLAKSLPNCKAVDIGPGLNLLQEDNPDLIGSEIARWLSTLEISG
SGENLYFQGSG_MMDLELPPPGLPSQQDMDLIDILWRQDIDLGVSREVFDFSQRRKEYELEK
QKKLEKERQEQLQKEQEKAFFAQLQLDEETGEFLPIQPAQHIQSETSGSANYSQVAHIPKSDAL
YFDDCMQLLAQTFPFVDDNEVSSATFQSLVPDIPGHIESPVFIATNQAQSPETSVAQVAPVDLD
GMQQDIEQVWEELLSIPELQCLNIENDKLVETTMVPSPEAKLTEVDNYHFYSSIPSMEKEVGN
CSPHFLNAFEDSFSSILSTEDPNQLTVNSLNSDATVNTDFGDEFYSAFIAEPSISNSMPSPATLS
HSLSELLNGPIDVSDLSLCKAFNQNHPESTAEFNDSDSGISLNTSPSVASPEHSVESSSYGDTL
LGLSDSEVEELDSAPGSVKQNGPKTPVHSSGDMVQPLSPSQGQSTHVHDAQCENTPEKELP
VSPGHRKTPFTKDKHSSRLEAHLTRDELRAKALHIPFPVEKIINLPVVDFNEMMSKEQFNEAQL
ALIRDIRRRGKNKVAAQNCRKRKLENIVELEQDLDHLKDEKEKLLKEKGENDKSLHLLKKQLST
LYLEVFSMLRDEDGKPYSPSEYSLQQTRDGNVFLVPKSKKPDVKKN

5.5.2.2 His$_6$-MafG

The human MAFG sequence (1-162) was inserted into a pET-15b cloning vector containing a C-terminal His$_6$ tag attached by a thrombin cleavable linker. Plasmid was produced by GenScript. 200 ng of the plasmid was transformed into Rosetta 2 (DE3) pLysS Competent Cells (Novagen), incubated for 30 minutes, 42 °C heat shock for 120 seconds, 2 hour incubation in SOC media. Single colonies were picked and overnight cultures grown in LB with 50 µg/mL Ampicillin, further diluted 1:50 in Terrific Broth (TB) media supplemented with 0.05% glucose with 50 µg/mL Ampicillin incubated at 37 °C 180 RPM for 4.5 hours. Protein expression was then induced at OD_{600} 0.6 with 1 mM IPTG at 37 °C for 4 hours. Cells were pelleted by centrifugation at 14000 *g* for 18 min at 4 °C. The pellet was then resuspended in buffer (50 mM Tris, 10 mM Imidazole, 100 mM NaCl pH 8) and sonicated for 8.2 min at 60 V on a 3 x pulse cycle. Lysed solution was then centrifuged at 5000 *g* for 30 min at 4 °C. Lysate supernatant was loaded onto Ni-NTA resin (Thermo Scientific™ HisPur™ Ni-NTA) at 4 °C, Ni-NTA resin was washed with excess buffer (20 mM Imidazole 50 mM NaH_2PO_4, 500 mM NaCl, pH 7.5 at pH 8) and was then eluted over an imidazole gradient of 40 mM – 250 mM in 50 mM NaH_2PO_4 pH 8 over 10 CV. Elution fractions containing MafG protein were combined and re-purified by Ni-NTA

[§] "_" is included in sequence to highlight separation between His, Halotag and the protein.

purification. Protein expression and purity was confirmed by SDS-PAGE on a 12% Bis-Tris gel in MES buffer.

Expression of His tagged protein was confirmed by western blot with His-Tag Antibody #2365 and Anti-rabbit IgG, HRP-linked Antibody #7074 from Cell Signalling Technology, PDVF membrane visualised using Invitrogen™ Novex™ ECL Chemiluminescent Substrate Reagent Kit. MafG was purified using Ni-NTA resin in buffer (50 mM NaH_2PO_4, 500 mM NaCl, 10 mM Imidazole 0.01% Triton x-100, 5% Glycerol pH 7.5) His_6MafG protein was eluted over an imidazole gradient 10mM – 250 mM in 10 CV. Eluted His_6-MafG underwent buffer exchanged into 1x PBS, 10 mM DTT yielding approximately 0.39 mg/mL from 4 L of culture.

(His_6_MAFG)

HHHHHH_MTTPNKGNKALKVKREPGENGTSLTDEELVTMSVRELNQHLRGLSKEEIVQLKQR
RRTLKNRGYAASCRVKRVTQKEELEKQKAELQQEVEKLASENASMKLELDALRSKYEALQTFAR
TVARSPVAPARGPLAAGLGPLVPGKVAATSVITIVKSKTDARSGGATCC

5.5.3 SDS-PAGE

Protein samples were prepared by incubating 10 µL of protein with 10 µL tris-glycine SDS buffer and 2 µL of 0.5 M DTT at 95 °C for 15 minutes. Cooled samples were then loaded onto Invitrogen pre-cast NuPAGE 1.0 mm Mini-Protein gels containing either 4-12% Bis-Tris (for HisHalo-Nrf2) or 12% Bis-Tris (for HisMafG) accompanied by a lane loading 3 µL of Thermo Scientific PageRuler prestained protein ladder for size determination. Gels then underwent electrophoresis at a constant voltage of 200 V for 30-45 minutes in a running buffer of either MOPS or MES at 50 mM with 50 mM Tris, 0.1% SDS, 0.01-0.09% DMF and 1 mM EDTA at pH 7.7.

Gels were subsequently washed in deionised water and incubated with a Coomassie stain (1 mM Coomassie with 20% (v/v) methanol 10% (v/v) acetic acid) for 10 minutes followed by washes in deionised water and incubation in a destaining solution (30% (v/v) methanol and 10% (v/v) glacial acetic acid) until band visualisation was clear.

5.5.4 Western blot

Pre-run SDS-PAGE gels were used to transfer protein content onto a Polyvinylidene Difluoride (PVDF) membrane using XCell SureLock Mini Cell Electrophoresis System using XCell II™ Blot Module for 1 h 30 min in 1 X Transfer buffer (20 mM Tris base, 150 mM Glycine, 3 mM SDS, 20% MeOH). The membrane was then washed with TBST (137 mM

NaCl, 2.7 mM KCl, 19 mM Tris Base, 0.5% Tween-20, pH 7.4) then blocked in milk solution (5% milk powder in TBST) overnight at 4 °C. Membrane was then incubated with desired rabbit primary antibody, following manufacturer instructions on dilution, incubation time and temperature. Membrane then washed thoroughly in TBST before incubation with anti-rabbit HRP conjugated antibody for 45 min at room temperature. After another thorough washing with 5% milk solution, followed by TBST 1x, band visualisation carried out using Novex™ ECL Chemiluminescence Kit on ImageQuant LAS 4000 chemiluminescence camera (GE Healthcare).

Antibodies used:

Anti-Nrf2 Antibody #AF0639 (Affinity Biosciences)
Anti-His Antibody #2365 (Cell Signalling Technology)
Anti-Rabbit IgG HRP-linked Antibody #7074 (Cell Signalling Technology)

5.5.5 Peptide mass fingerprinting

Protein collected from SDS-PAGE gel in either 4-12% or 12% Bis-Tris in MOPS or MES and gel band for sequencing was isolated and washed in buffer (50% of 50 mM Triethylammonium bicarbonate buffer (TEAB) with 50% Acetonitrile), samples were incubated in 10 mM DTT for 30 min at 55 °C. Sample was then incubated in the dark with agitation in 30 mM Iodoacetamide (IAA) in 50 mM TEAB for 30 minutes then washed in 50 mM TEAB in 50% acetonitrile. Sample was then prepared in 100% acetonitrile and dried thoroughly in preparation for trypsin digestion and MALDI-TOF performed at the John Innes Centre proteomics facility using a Bruker Autoflex™ Speed Maldi-TOF/TOF, protein identification was performed by database search using Mascot Server 3.0 and Mascot Distiller 2.8.

5.6 Appendix

5.6.1 Peptide mass fingerprinting sequence coverage

Protein sequence coverage: 49%

Matched peptides shown in ***bold red***.

```
   1 HHHHHHSSGL VPRGSHMAEI GTGFPFDPHY VEVLGERRMH YVDVGPRDGT
  51 PVLFLHGNPT SSYVWRNIIP HVAPTHRCIA PDLIGMGKSD KPDLGYFFFD
 101 DHVRFMDAFI EALGLEEVVL VIHDWGSALG FHWAKRNPER VKGIAFMEFI
 151 RPIPTWDEWP EFARETFQAF RTTDVGRKLI IDQNVFIEGT LPMGVVRPLT
 201 EVEMDHYREP FLNPVDREPL WRFPNELPIA GEPANIVALV EEYMDWLHQS
 251 PVPKLLFWGT PGVLIPPAEA ARLAKSLPNC KAVDIGPGLN LLQEDNPDLI
 301 GSEIARWLST LEISGSGENL YFQGSGMMDL ELPPPGLPSQ QDMDLIDILW
 351 RQDIDLGVSR EVFDFSQRRK EYELEKQKKL EKERQEQLQK EQEKAFFAQL
 401 QLDEETGEFL PIQPAQHIQS ETSGSANYSQ VAHIPKSDAL YFDDCMQLLA
 451 QTFPFVDDNE VSSATFQSLV PDIPGHIESP VFIATNQAQS PETSVAQVAP
 501 VDLDGMQQDI EQVWEELLSI PELQCLNIEN DKLVETTMVP SPEAKLTEVD
 551 NYHFYSSIPS MEKEVGNCSP HFLNAFEDSF SSILSTEDPN QLTVNSLNSD
 601 ATVNTDFGDE FYSAFIAEPS ISNSMPSPAT LSHSLSELLN GPIDVSDLSL
 651 CKAFNQNHPE STAEFNDSDS GISLNTSPSV ASPEHSVESS SYGDTLLGLS
 701 DSEVEELDSA PGSVKQNGPK TPVHSSGDMV QPLSPSQGQS THVHDAQCEN
 751 TPEKELPVSP GHRKTPFTKD KHSSRLEAHL TRDELRAKAL HIPFPVEKII
 801 NLPVVDFNEM MSKEQFNEAQ LALIRDIRRR GKNKVAAQNC RKRKLENIVE
 851 LEQDLDHLKD EKEKLLKEKG ENDKSLHLLK KQLSTLYLEV FSMLRDEDGK
 901 PYSPSEYSLQ QTRDGNVFLV PKSKKPDVKK N
```

Appendix 5-1 *Protein sequencing of His_6Halo-Nrf2 protein by trypsin digest and mass spectrometry compared against a database of commonly found E. coli contaminants confirmed correct protein expression*

Protein sequence coverage: 36%

Matched peptides shown in ***bold red***.

```
   1 HHHHHHMTTP NKGNKALKVK REPGENGTSL TDEELVTMSV RELNQHLRGL
  51 SKEEIVQLKQ RRRTLKNRGY AASCRVKRVT QKEELEKQKA ELQQEVEKLA
 101 SENASMKLEL DALRSKYEAL QTFARTVARS PVAPARGPLA AGLGPLVPGK
 151 VAATSVITIV KSKTDARSGG ATCC
```

Appendix 5-2 *Peptide mass fingerprint results from PET15B His-MafG confirming sequence coverage (36%) from expression induced in Rosetta 2 (DE3) cells in TB medium expression induced at OD_{600} 0.6 for 3-4 hours with 1 mM IPTG at 37˚C.*

5.7 References

1 A. Schütz, F. Bernhard, N. Berrow, J. F. Buyel, F. Ferreira-da-Silva, J. Haustraete, J. van den Heuvel, J.-E. Hoffmann, A. de Marco, Y. Peleg, S. Suppmann, T. Unger, M. Vanhoucke, S. Witt and K. Remans, *STAR Protoc.*, 2023, **4**, 102572.

2 O. Khow and S. Suntrarachun, *Asian Pac J Trop Biomed*, 2012, **2**, 159–162.

3 L. G. Morão, L. R. Manzine, L. O. D. Clementino, C. Wrenger and A. S. Nascimento, *PLOS ONE*, 2022, **17**, e0271403.

4 M. Tian, X. Li, L. Yu, J. Qian, X. Bai, J. Yang, R. Deng, C. Lu, H. Zhao and Y. Liu, *Cell Commun Signal CCS*, 2025, **23**, 214.

5 R. Weis, F. Hartner and A. Glieder, in *Encyclopedic Reference of Genomics and Proteomics in Molecular Medicine*, Springer, Berlin, Heidelberg, 2005, pp. 1620–1625.

6 F. M. Wurm and D. Hacker, *Nat Biotechnol*, 2011, **29**, 718–720.

7 K. Itakura, T. Hirose, R. Crea, A. D. Riggs, H. L. Heyneker, F. Bolivar and H. W. Boyer, *Science*, 1977, **198**, 1056–1063.
8 D. V. Goeddel, D. G. Kleid, F. Bolivar, H. L. Heyneker, D. G. Yansura, R. Crea, T. Hirose, A. Kraszewski, K. Itakura and A. D. Riggs, *Proc Natl Acad Sci U. S. A.*, 1979, **76**, 106–110.
9 A. Singh, *Nat. Methods*, 2020, **17**, 655–655.
10 K. Stevenson, A. F. McVey, I. B. N. Clark, P. S. Swain and T. Pilizota, *Sci Rep*, 2016, **6**, 38828.
11 G. H. Rajacharya, A. Sharma and S. S. Yazdani, *Sci. Rep.*, 2024, **14**, 12271.
12 J. W. Dubendorff and F. W. Studier, *J Mol Biol*, 1991, **219**, 45–59.
13 P. Moi, K. Chan, I. Asunis, A. Cao and Y. W. Kan, *Proc Natl Acad Sci USA*, 1994, **91**, 9926–9930.
14 P. Tompa, *Trends Biochem Sci*, 2002, **27**, 527–533.
15 N. C. Karunatilleke, C. S. Fast, V. Ngo, A. Brickenden, M. L. Duennwald, L. Konermann and W.-Y. Choy, *Int J Mol Sci*, 2021, **22**, 7434.
16 T. Sengoku, M. Shiina, K. Suzuki, K. Hamada, K. Sato, A. Uchiyama, S. Kobayashi, A. Oguni, H. Itaya, K. Kasahara, H. Moriwaki, C. Watanabe, T. Honma, C. Okada, S. Baba, T. Ohta, H. Motohashi, M. Yamamoto and K. Ogata, *Nucleic Acids Res*, 2022, **50**, 12543–12557.
17 R. S. Chumanov and R. R. Burgess, *Protein Expr. Purif.*, 2010, **74**, 280–288.
18 S. Parvez, Y. Fu, J. Li, M. J. C. Long, H. Y. Lin, D. K. Lee, G. S. Hu and Y. Aye, *J Am Chem Soc*, 2015, **137**, 10–13.
19 R. Modi, N. McKee, N. Zhang, A. Alwali, S. Nelson, A. Lohar, R. Ostafe, D. D. Zhang and E. I. Parkinson, *J Med Chem*, 2023, **66**, 6184–6192.
20 D. G. Calvopina-Chavez, M. A. Gardner and J. S. Griffitts, *G3 Bethesda*, 2022, **12**, jkac070.
21 G. L. Rosano and E. A. Ceccarelli, *Front. Microbiol.*, 2014, **5**, 172.
22 F. W. Studier and B. A. Moffatt, *J Mol Biol*, 1986, **189**, 113–130.
23 S. Agrawal, M. H. Padmaswari, A. L. Stokes, D. Maxenberger, M. Reese, A. Khalil and C. E. Nelson, *Biomedicines*, 2024, **12**, 1226.
24 Joo Shun Tan, Ramanan, R.N, Tau Chuan Ling., Shuhaimi Mustafa, and Arbakariya B. Ariff, *Ann Microbiol*, 2011, **62**, 1427–1435.
25 J. Sambrook, *Molecular cloning : a laboratory manual*, Cold Spring Harbor, New York, 4th edn., 2001.
26 S. Chen, M. Larsson, R. C. Robinson and S. L. Chen, *Sci Rep*, 2017, **7**, 4788.
27 C. Arrowsmith, P. Nordlund, J. Weigelt, B. M. Hallberg, J. Bray, O. Gileadi, S. Knapp, U. Oppermann, C. Arrowsmith, R. Hui, J. Ming, S. dhe-Paganon, H. Park, A. Savchenko, A. Yee, A. Edwards, R. Vincentelli, C. Cambillau, R. Kim, S.-H. Kim, Z. Rao, Y. Shi, T. C. Terwilliger, C.-Y. Kim, L.-W. Hung, G. S. Waldo, Y. Peleg, S. Albeck, T. Unger, O. Dym, J. Prilusky, J. L. Sussman, R. C. Stevens, S. A. Lesley, I. A. Wilson, A. Joachimiak, F. Collart, I. Dementieva, M. I. Donnelly, W. H. Eschenfeldt, Y. Kim, L. Stols, R. Wu, M. Zhou, S. K. Burley, J. S. Emtage, J. M. Sauder, D. Thompson, K. Bain, J. Luz, T. Gheyi, F. Zhang, S. Atwell, S. C. Almo, J. B. Bonanno, A. Fiser, S. Swaminathan, F. W. Studier, M. R. Chance, A. Sali, T. B. Acton, R. Xiao, L. Zhao, L. C. Ma, J. F. Hunt, L. Tong, K. Cunningham, M. Inouye, S. Anderson, H. Janjua, R. Shastry, C. K. Ho, D. Wang, H. Wang, M. Jiang, G. T. Montelione, D. I. Stuart, R. J. Owens, S. Daenke, A. Schütz, U. Heinemann, S. Yokoyama, K. Büssow and K. C. Gunsalus, *Nat Methods*, 2008, **5**, 135–146.
28 A. Spriestersbach, J. Kubicek, F. Schäfer, H. Block and B. Maertens, *Methods Enzymol*, 2015, **559**, 1–15.

29 L. Li, H. Li, Q. Tian, B. Ge, X. Xu, Y. Chi, H. Zhao, Y. Liu, N. Jia, T. Zhou, Y. Zhu and Y. Zhou, *Microb Cell Factories*, 2022, **21**, 244.
30 T.-J. Choi and T. T. Geletu, *J. Genet. Eng. Biotechnol.*, 2018, **16**, 347–355.
31 C.-J. Chiang, M.-C. Hu and Y.-P. Chao, *J. Agric. Food Chem.*, 2020, **68**, 8883–8889.
32 L. Gomes, G. Monteiro and F. Mergulhão, *Int. J. Mol. Sci.*, 2020, **21**, 576.
33 P. Hong, S. Koza and E. S. P. Bouvier, *J Liq Chromatogr Relat Technol*, 2012, **35**, 2923–2950.
34 A. Kopacz, A. I. Rojo, C. Patibandla, D. Lastra-Martínez, A. Piechota-Polanczyk, D. Kloska, A. Jozkowicz, C. Sutherland, A. Cuadrado and A. Grochot-Przeczek, *Free Radic Biol Med*, 2022, **192**, 37–49.
35 Richard R. Burgess, *Protein Expr Purif*, 2018, **150**, 81–85.
36 S. Fekete, A. Beck, J.-L. Veuthey and D. Guillarme, *J Pharm Biomed Anal*, 2015, **113**, 43–55.
37 D. Yu, G. Huang, F. Xu, M. Wang, S. Liu and F. Huang, *Photosynth Res*, 2014, **120**, 311–321.
38 R. Brahma and H. Raghuraman, *Curr Protoc*, 2022, **2**, e452.
39 A. Stetsenko and A. Guskov, *Crystals*, 2017, **7**, 197.
40 H. Kurokawa, H. Motohashi, S. Sueno, M. Kimura, H. Takagawa, Y. Kanno, M. Yamamoto and T. Tanaka, *Mol Cell Biol*, 2009, **29**, 6232.
41 M. Kyo, T. Yamamoto, H. Motohashi, T. Kamiya, T. Kuroita, T. Tanaka, J. D. Engel, B. Kawakami and M. Yamamoto, *Genes Cells*, 2004, **9**, 153–164.
42 H. Motohashi, F. Katsuoka, C. Miyoshi, Y. Uchimura, H. Saitoh, C. Francastel, J. D. Engel and M. Yamamoto, *Mol Cell Biol*, 2006, **26**, 4652–4663.
43 N. J. Carruthers, G. C. Parker, T. Gratsch, J. A. Caruso and P. M. Stemmer, *J Biomol Tech JBT*, 2015, **26**, 103–112.
44 Y. Shi, R. A. Mowery, J. Ashley, M. Hentz, A. J. Ramirez, B. Bilgicer, H. Slunt-Brown, D. R. Borchelt and B. F. Shaw, *Protein Sci Publ Protein Soc.*, 2012, **21**, 1197–1209.
45 S. Damodaran, T. D. Wood, P. Nagarajan and R. A. Rabin, *Genomics Proteomics Bioinformatics*, 2007, **5**, 152–157.
46 R. P. Sparks, J. L. Jenkins and R. Fratti, *Methods Mol Biol Clifton NJ*, 2019, **1860**, 199–210.
47 B. D. Bowes, H. Koku, K. J. Czymmek and A. M. Lenhoff, *J Chromatogr A*, 2009, **1216**, 7774–7784.
48 B. Johnsson, S. Löfås and G. Lindquist, *Anal Biochem*, 1991, **198**, 268–277.
49 M. I. Henderson, Y. Eygeris, A. Jozic, M. Herrera and G. Sahay, *Mol Pharm*, 2022, **19**, 4275–4285.
50 P. M. Wax, *J Med Toxicol*, 2013, **9**, 303–307.
51 A. Singh, S. Venkannagari, K. H. Oh, Y.-Q. Zhang, J. M. Rohde, L. Liu, S. Nimmagadda, K. Sudini, K. R. Brimacombe, S. Gajghate, J. Ma, A. Wang, X. Xu, S. A. Shahane, M. Xia, J. Woo, G. A. Mensah, Z. Wang, M. Ferrer, E. Gabrielson, Z. Li, F. Rastinejad, M. Shen, M. B. Boxer and S. Biswal, *ACS Chem Biol*, 2016, **11**, 3214–3225.
52 J.-F. Masson, T. M. Battaglia, J. Cramer, S. Beaudoin, M. Sierks and K. S. Booksh, *Anal Bioanal Chem*, 2006, **386**, 1951–1959.
53 J. Y. Lichtenberg, Y. Ling and S. Kim, *Sensors*, 2019, **19**, 2488.
54 C. Situ, A. R. G. Wylie, A. Douglas and C. T. Elliott, *Talanta*, 2008, **76**, 832–836.
55 A. G. Ngounou Wetie, I. Sokolowska, A. G. Woods, U. Roy, J. A. Loo and C. C. Darie, *Proteomics*, 2013, **13**, 10.1002/pmic.201200328.

Chapter 6
Conclusion

6.1 General Discussion

Protein-protein interactions have long been considered undruggable due to their large and shallow surface area at their binding interface.[1,2] Whilst PPIs across a range of shapes and sizes can now be targeted with peptides and small molecules, the majority of PPI inhibitors reported over the last fifty years disrupt interactions involving α-helix recognition motifs. Interestingly within this secondary structure categorisation, coiled-coil interactions are under-represented, with only a handful of inhibitors reported (**Chapter 1.2.3**). This presents a unique opportunity to explore novel peptide and small molecule inhibitors against disease-relevant dimeric coiled-coil PPIs.

The transcription factor Nrf2 was a prime candidate for this research, forming a dimeric coiled-coil with its obligatory binding partner MafG, necessary for DNA binding to initiate gene transcription. Nrf2 has been extensively studied for its conflicting cytoprotective role, reducing inflammation in cells but it is also capable of promoting hallmarks of cancer.[3,4] Concerningly, Nrf2 has been implicated in cancer metastasis, growth and chemo-resistance (**Chapter 1.5.3**).[5,6]

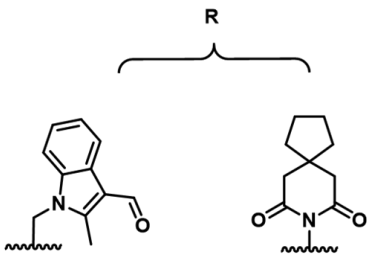
Existing Nrf2 inhibitors consist mostly of natural products or small molecules.[7–11] These inhibitors often display off-target effects or do not directly target Nrf2 or the PPI with MafG. At the outset of this research in 2021, no inhibitors had been reported against the Nrf2/MafG PPI, compounded by the absence of a crystal structure to confirm the residues involved within this interaction. The advent of machine learning programs such as AlphaFold has helped to overcome this structural dilemma, demonstrating an ability to produce highly accurate models of protein-protein and protein-DNA interactions.[12] Moreover, the crystal structure of the Nrf2/MafG heterodimer reported at the end of 2022 provided reassuring similarity to the AlphaFold2 homology model initially used to guide the peptide design within this work (**Chapter 2.3.1**).

The leucine zippers of Nrf2 and MafG form a parallel coiled-coil with regularly repeating hydrophobic residues at *a* and *d* positions within the heptadic motif $(abcdefg)_n$. Upon the discovery of the Nrf2 protein, it was predicted that residues 505-552 created the leucine zipper.[13] Analysis of crystal structure and AlphaFold models suggests residues 522-568 more accurately describes the zipper domain of Nrf2.[12,14] Nrf2 residues 553-562 have also been reported to constitute a nuclear exportation sequence, promoting cytosolic migration of the protein.[15–17] Heterodimerisation with MafG masks this sequence, maintaining nuclear retention of Nrf2.[15] Utilising this information, residue

alignment with MafG demonstrates sequence $_{73}$KEELEKQKAELQQEVELASENASMKLE$_{104}$ interacts directly with Nrf2$_{522\text{-}552}$, without shielding nuclear exportation (**Chapter 2.3.2**).

Herein we have described the design and synthesis of novel peptides mimicking the leucine zipper of MafG$_{73\text{-}104}$ to afford inhibition of the Nrf2/MafG PPI. Peptides were synthesised conserving the heptadic motif of the MafG leucine zipper, required for coiled-coil formation with Nrf2. This led to the solid-phase peptide synthesis of six peptide sequences of 28, 21 or 14 residues in length (**Figure 6-1, 2.46, 2.52-2.56**). Synthesis of peptides (**2.46, 2.52** and **2.54**) underwent optimisation due to sequence aggregation (**Chapter 2.3.2.4**). The addition of lithium chloride, a chaotropic salt, improved synthesis of **2.54, 2.52** and **2.46.**

2.46: Ac-KEELEKQKAELQQEVEKLASENASMKLE-NH_2
2.52: Ac-KAELQQEVEKLASENASMKLE-NH_2
2.58: Ac-KAELQQEVEKLASENAS(Nle)KLE-NH_2
2.53: Ac-KEELEKQKAELQQEVEKLASE-NH_2
2.54: Ac-VEKLASENASMKLE-NH_2
2.55: Ac-KAELQQEVEKLASE-NH_2
2.56: Ac-KEELEKQKAELQQE-NH_2
3.85: Ac-KAELQQEVEK-NH_2
3.86: Ac-LASENASMKLE-NH_2
3.72: **R**-LASENASMKLE-NH_2
3.78: **R**-LASENASMKLE-NH_2



3.72 **3.78**

Figure 6-1 *MafG peptides and peptide-small molecule hybrids of interest.*

Secondary structure analysis of the six peptides confirmed their capacity to adopt the α-helical structure, essential for coiled-coil formation (**Chapter 2.3.3).** 21-mer sequences, **2.53** and **2.52** exhibited the highest degree of helicity (71% and 81%), suggesting a strong propensity for forming a stable coiled-coil interaction with Nrf2.

Characterising the binding affinity of Nrf2-binding inhibitors towards its leucine zipper is inherently difficult, as the protein is predominantly disordered and is believed to only fold transiently in the presence of a binding partner.[18] This limitation became evident in our attempts to measure binding interactions by SPR (**Chapter 2.3.6**). **2.46** and **2.52-2.56** displayed weak interactions with the immobilised Nrf2 with K_d values predicted outside of the concentration ranges tested. Moreover, the protein itself began to demonstrate signs of instability, with baseline erosion on the SPR chip hindering further

experimental condition trials. Despite these challenges, to our knowledge, this represents the first reported attempt to measure Nrf2-ligand binding kinetics by SPR.

Previous research into Nrf2/MafG PPI inhibitors that bind to Nrf2 is an underexplored area of study. The only small molecule inhibitor with direct evidence of Nrf2 binding is ML385, identified by HTS using an Nrf2/DNA binding luciferase assay.[19] ML385 has been shown to bind to the bZIP domain of Nrf2 by a biotin pull down assay.[19] Additionally, fluorescence polarisation was used to confirm disruption of the ternary complex of Nrf2/MafG with DNA.[19] However, this does not confirm the binding site of ML385, which could cause disruption by engaging the DNA binding domain of Nrf2.

As transcription factor binding interactions are often transient and dynamic, they can be challenging to quantify. A fluorescence polarisation (FP) assay to monitor ternary complex formation between Nrf2, MafG, and ARE-promoter DNA has helped to overcome this. Originally developed to characterise the small-molecule inhibitor ML385, this assay has since been employed to validate the inhibitory activity of a stapled peptide (N1S) reported in 2023.[19,20] N1S is the first documented example of a peptide capable of disrupting the PPI between Nrf2 and MafG.

Guided by the existing literature, a three-part fluorescence polarisation assay was established, using Nrf2 and MafG proteins expressed in **Chapter 5** and a fluorescently labelled DNA sequence containing an ARE promoter. Interestingly, only two peptides, 28-mer **2.46 (**Ac-KEELEKQKAELQQEVEKLASENASMKLE-NH_2) and 21-mer **2.52**, **(**Ac-KAELQQEVEKLASENASMKLE-NH_2**)** demonstrated an ability to disrupt ternary complex formation, further confirmed by an electrophoretic mobility shift assay (**Chapter 2.3.4** and **2.3.5**).

Both **2.52** and **2.46** peptides conserve three out of four heptad repeats, which suggested a minimum length may be necessary for stable coiled-coil formation and subsequent disruption of the protein/DNA interaction. In contrast, shorter peptides were thought to lack the structural stability required to engage in productive interactions, as circular dichroism analysis suggested they less than 50% helical. Notably, **2.53** does not contain the 'NASMKLE' motif, potentially indicating that this sequence is critical for binding. Alternatively, it is possible that some peptides self-assemble into homomeric coiled-coils, sequestering themselves and preventing interaction with full-length Nrf2.

It was found that **2.52** was capable of undergoing methionine oxidation, which noticeably decreased the peptides activity within the FP assay established. Synthesis of the sequence with norleucine substitution (**2.58**) was found to retain binding inhibitory activity in the FP assay. Disruption of the MafG homodimer by **2.58** was also evident, as at 250 μM DNA binding of either protein was completely abolished.

It was interesting to discover later in the project, that 11-mer peptide **3.86,** (Ac-LASENASMKLE-NH_2) was still capable of disrupting the ternary complex formation, at a much-reduced affinity (**Chapter 3.3.3**). Especially given 14-mer **2.54** (Ac-VEKLASENASMKLE-NH_2) demonstrated no inhibitory activity within our FP or EMSA assay. This finding suggests conservation of a full heptad repeat may not be strictly necessary to disrupt the Nrf2/MafG coiled-coil. Further investigation into the secondary structure characterisation of the shorter peptides, **3.85** and **3.86** may offer insight if the helical content appears more stable in comparison to **2.54**.

Building on the discovery of our novel peptide inhibitor, we aimed to initiate a peptide-directed ligand design strategy using **2.52** as a scaffold. There are limited reports of small molecule inhibitors capable of disrupting dimeric coiled-coils. We investigated the use of peptide-directed ligand design with *in silico* fragment screening to investigate small molecule interactions with the Nrf2 leucine zipper (**Chapter 3**). Fragments derived from *in silico* screening were supplemented with additional fragments sourced from existing laboratory collections, to achieve synthesis of twenty peptide small molecule hybrids. Interestingly no hits were found from the *in silico* screen. Future efforts to utilise *in silico* methodology may benefit from curating more targeted screening libraries to improve efficiency and hit rates.

Exploration of available fragments across a chemical space identified two promising candidates. **3.72**, an aromatic indole featuring an aldehyde sidechain, and **3.78**, an azaspirodecane structure featuring two ketone functional groups (**Figure 6-1**). Both peptides demonstrated greater activity compared to the parent peptide (**3.72**), but also an ability to disrupt both hetero and homodimeric complex binding. This would be the first example of a peptide-small molecule hybrid capable of disrupting the Nrf2/MafG interaction. Future work could investigate development of small molecule scaffolds with specificity to the Nrf2 leucine zipper over MafG.

Over the last two decades Nrf2 has become synonymous with chemoresistance.[3,5,10,21] Few direct inhibitors of the Nrf2/MafG PPI have been reported and our understanding of

the impact of Nrf2 inhibition in a cellular context is limited.[19,20] In particular, there are mixed reports on the impact of Nrf2 inhibition on cell viability and the potency of all reported small molecules, which are all active at micromolar concentrations.[7–11] Consequently, investigation of the impact of Nrf2/MafG disrupting peptides and peptide-small molecule hybrids adds to our understanding of the transcription factor. In **Chapter 4,** our biological investigation found MafG peptides did not significantly decrease cell viability compared to the growth control at 100 μM in A549 NSCLC. Comparatively, at 100 μM, small molecule inhibitor ML385 causes 100% growth inhibition and hydrocarbon stapled peptide inhibitor of MafG reduces cell viability by 37%.[19,20]

Inhibition of Nrf2 hopes to achieve a decrease in the cytoprotective genes transcribed, including the enzyme NQO1. This enzyme supports drug detoxification processes, linked with Nrf2 overexpression, and the development of chemoresistance in NSCLC.[22,23] Our investigation into the impact of our Nrf2 inhibitors on NQO1 activity found a noticeable, dose-dependent (100-25 μM) decrease in activity from **2.52, 2.58** and **3.72** in A549 cells, suggesting these peptides were capable of a functional response, relevant to the inhibition of Nrf2 transcription. This response from **2.52** and **3.72** correlated with the results found by our competition fluorescence polarisation assay, achieving inhibition with IC_{50} of <100 μM. As a result, **2.52** and **3.72** were investigated further for synergism with doxorubicin, a DNA intercalator and a topoisomerase II inhibitor, linked to Nrf2-driven resistance.[24] Disappointingly, the inhibitors did not improve the IC_{50} of doxorubicin treated A549 cells. Whilst it was anticipated that doxorubicin resistance in A549 cells would be driven by Nrf2 expression, this resistance may produce an over-expression of drug efflux pump associated P-glycoprotein (P-gp), which might also act on the MafG peptides.[25] Further investigation of the peptides on other Nrf2 transcribed genes by mRNA expression, such as P-gp, may provide further mechanistic understanding of the transcription factors influence.

Alternatively, doxorubicin resistance may occur by a different mechanism than Nrf2 overexpression. Consequently, co-treatment of **2.52** and **3.72** was also explored with gemcitabine, an anti-metabolite that incorporates into DNA, blocking enzymes from completing DNA synthesis.[26] Excitingly, **3.72** generated a response in the A549 cells towards gemcitabine at concentrations as low as 0.7 μM. Whilst **2.52** did not produce an improved response with gemcitabine, it is possible that the peptide is more susceptible to degradation compared to the hybrid inhibitor.

6.2 Conclusion

The study of the Nrf2 transcription factors can provide valuable insight into the mechanisms of chemoresistance in cancer. Despite numerous reports of small molecule and natural products capable of inhibiting transcription factor activity, none have progressed through clinical trials.[27] Furthermore, the inhibitors described do not fully characterise the mechanism of Nrf2 inhibition, raising questions on the selectivity of these compounds and the risk of off-target effects.[7–11] Consequently, there is a great need to develop research probes towards direct Nrf2 inhibition. Within this thesis, design and synthesis of rational peptide inhibitors of the Nrf2 PPI with MafG has been explored, and we have highlighted an important section, $MafG_{73\text{-}104}$. Use of a fluorescence polarisation assay and novel use of an electrophoretic mobility shift assay confirms the ability of **2.52** to disrupt Nrf2 activity. To the best of our knowledge, this is the first example of peptides designed to bind to Nrf2 to disrupt the Nrf2/MafG PPI.

This research aimed to investigate if peptide-directed ligand design could offer a route to designing small molecule inhibitors against a coiled-coil interaction. Our peptide-small molecule hybrids offer a promising step towards achieving this. Utilising $MafG_{73\text{-}104}$ as a scaffold, a peptide to small molecule drug discovery campaign was performed, leading to the identification of two novel peptide-small molecule hybrids (**3.72** and **3.78**). **3.72** demonstrated improved efficacy compared **2.52,** decreasing NQO1 enzymatic activity in A549 cells and an ability to sensitise the NSCLC cells to gemcitabine, a chemotherapeutic agent.

6.3 References

1 Neklesa TK, Winkler JD, Crews CM., *Pharmacol Ther*, **174**, 138–144.
2 S. Pathmanathan, I. Grozavu, A. Lyakisheva and I. Stagljar, *Curr Opin Chem Biol*, 2022, **66**, 102079.
3 A. Lau, N. F. Villeneuve, Z. Sun, P. K. Wong and D. D. Zhang, *Pharmacol Res*, 2008, **58**, 262–270.
4 S. Saha, B. Buttari, E. Panieri, E. Profumo and L. Saso, *Molecules*, 2020, **25**, 5474.
5 X.-J. Wang, Z. Sun, N. F. Villeneuve, S. Zhang, F. Zhao, Y. Li, W. Chen, X. Yi, W. Zheng, G. T. Wondrak, P. K. Wong and D. D. Zhang, *Carcinogenesis*, 2008, **29**, 1235–1243.
6 R. Grossman and Z. Ram, *World Neurosurg*, 2013, **80**, 284–286.
7 D. Ren, N. F. Villeneuve, T. Jiang, T. Wu, A. Lau, H. A. Toppin and D. D. Zhang, *Proc Natl Acad Sci U S A*, 2011, **108**, 1433–1438.
8 S. Lee, M.-J. Lim, M.-H. Kim, C.-H. Yu, Y.-S. Yun, J. Ahn and J.-Y. Song, *Free Radic Biol Med*, 2012, **53**, 807–816.
9 M. J. Bollong, H. Yun, L. Sherwood, A. K. Woods, L. L. Lairson and P. G. Schultz, *ACS Chem Biol*, 2015, **10**, 2193–2198.

10 D. Zhang, Z. Hou, K. E. Aldrich, L. Lockwood, A. L. Odom and K. T. Liby, *Mol Cancer Ther*, 2021, **20**, 1692–1701.
11 Z. Hou, L. Lockwood, D. Zhang, C. J. Occhiuto, L. Mo, K. E. Aldrich, H. E. Stoub, K. A. Gallo, K. T. Liby and A. L. Odom, *RSC Med Chem*, 2023, **14**, 74–84.
12 P. Bryant, G. Pozzati and A. Elofsson, *Nat Commun*, 2022, **13**, 1–11.
13 P. Moi, K. Chan, I. Asunis, A. Cao and Y. W. Kan, *Proc Natl Acad Sci U S A*, 1994, **91**, 9926–9930.
14 T. Sengoku, M. Shiina, K. Suzuki, K. Hamada, K. Sato, A. Uchiyama, S. Kobayashi, A. Oguni, H. Itaya, K. Kasahara, H. Moriwaki, C. Watanabe, T. Honma, C. Okada, S. Baba, T. Ohta, H. Motohashi, M. Yamamoto and K. Ogata, *Nucleic Acids Res.*, 2022, **50**, 12543–12557.
15 M. Theodore, Y. Kawai, J. Yang, Y. Kleshchenko, S. P. Reddy, F. Villalta and I. J. Arinze, *J Biol Chem*, 2008, **283**, 8984–8994.
16 A. K. Jain, D. A. Bloom and A. K. Jaiswal, *J Biol Chem*, 2005, **280**, 29158–29168.
17 W. Li and A.-N. Kong, *Mol Carcinog*, 2009, **48**, 91–104.
18 N. C. Karunatilleke, C. S. Fast, V. Ngo, A. Brickenden, M. L. Duennwald, L. Konermann and W.-Y. Choy, *Int J Mol Sci*, 2021, **22**, 7434.
19 A. Singh, S. Venkannagari, K. H. Oh, Y.-Q. Zhang, J. M. Rohde, L. Liu, S. Nimmagadda, K. Sudini, K. R. Brimacombe, S. Gajghate, J. Ma, A. Wang, X. Xu, S. A. Shahane, M. Xia, J. Woo, G. A. Mensah, Z. Wang, M. Ferrer, E. Gabrielson, Z. Li, F. Rastinejad, M. Shen, M. B. Boxer and S. Biswal, *ACS Chem Biol*, 2016, **11**, 3214–3225.
20 R. Modi, N. McKee, N. Zhang, A. Alwali, S. Nelson, A. Lohar, R. Ostafe, D. D. Zhang and E. I. Parkinson, *J Med Chem*, 2023, **66**, 6184–6192.
21 A. Demirezen and O. Erbaş, *D J Med Sci*, 2024, **10**, 123–128.
22 Z. Li, Y. Zhang, T. Jin, J. Men, Z. Lin, P. Qi, Y. Piao and G. Yan, *BMC Cancer*, 2015, **15**, 207.
23 A. E. M. A. Khan, V. Arutla and K. S. Srivenugopal, *Cells*, 2024, **13**, 1272.
24 P. Paramasivan, J. D. Kumar, R. Baskaran, C. F. Weng and V. V. Padma, *Cancer Drug Resist*, 2020, **3**, 647–665.
25 L. Xu, H. Li, Y. Wang, F. Dong, H. Wang and S. Zhang, *Oncol Lett*, 2014, **7**, 387–392.
26 L. Toschi and F. Cappuzzo, *Onco Targets Ther*, 2009, **2**, 209–217.
27 J. Zhang, H.-X. Xu, J.-Q. Zhu, Y.-X. Dou, Y.-F. Xian and Z.-X. Lin, *Int J Biol Sci*, 2023, **19**, 3029–3041.