The design, synthesis and evaluation of peptides and small molecules targeting the dyskerin-dyskerin protein-protein interaction

Suzanne Pauline van Wier

School of Pharmacy

University of East Anglia

October 2023

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

© This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with the author and that use of any information derived therefrom must be in accordance with current UK Copyright Law. In addition, any quotation or extract must include full attribution.

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.

Suzanne van Wier

Abstract

One of the hallmarks of cancers is their ability to replicate limitlessly making them immortal. In 80-90% of cancer cells this is due to the reactivation of telomerase, a protein complex which elongates telomeres at the end of chromosomes, protecting the chromosomes from degradation and preventing cell senescence. The Cryo-EM structure of telomerase published in 2021 provided an opportunity to identify new ways to target telomerase. In patients with Dyskeratosis Congenita, a disease characterised by shortened telomeres, the structure showed that genetic mutations are transcribed to the dyskerin-dyskerin protein-protein interaction (PPI) in telomerase.

Herein, we investigate the inhibition of the dyskerin-dyskerin PPI with peptides and small molecules. Based on the telomerase Cryo-EM structure, linear, hydrocarbon and lactam stapled peptides were designed and synthesised. Using recombinantly-expressed MBP-dyskerin, fluorescence polarisation and surface plasmon resonance binding assays were developed to investigate the binding affinity of the peptides towards dyskerin. The peptides were further evaluated in a Homogeneous Time Resolved Fluorescence assay to assess their ability to disrupt the dimer formation with six peptides found to disrupt the dyskerin-dyskerin PPI at 200 μM.

When compared to small molecule drugs, peptide therapeutics can have drawbacks such as their limited stability and cell permeability. Therefore, peptide-directed binding was used to go from peptide to small molecule inhibitor by computationally identifying fragments able to replace parts of the peptide. Nineteen small molecules were synthesised with six compounds able to disrupt the dyskerin-dyskerin dimer at 100 μM.

Both the peptides and small molecules were further evaluated for their cytotoxicity in an MTS assay. The peptides were also explored for their cellular uptake and proteolytic stability, two properties which could limit their cellular effects.

Access Condition and Agreement

Each deposit in UEA Digital Repository is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the Data Collections is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form. You must obtain permission from the copyright holder, usually the author, for any other use. Exceptions only apply where a deposit may be explicitly provided under a stated licence, such as a Creative Commons licence or Open Government licence.

Electronic or print copies may not be offered, whether for sale or otherwise to anyone, unless explicitly stated under a Creative Commons or Open Government license. Unauthorised reproduction, editing or reformatting for resale purposes is explicitly prohibited (except where approved by the copyright holder themselves) and UEA reserves the right to take immediate ‘take down’ action on behalf of the copyright and/or rights holder if this Access condition of the UEA Digital Repository is breached. Any material in this database has been supplied on the understanding that it is copyright material and that no quotation from the material may be published without proper acknowledgement.

Contents

Declaration ii
Abstract iii
Contents iv
Abbreviations viii
Acknowledgements x

Chapter 1 Introduction 1
1.1 The discovery of peptide inhibitors of PPIs derived from the sequence of the protein partner 2
1.1.1 Introduction 2
1.1.2 Selection of an initial sequence 3
1.1.3 Analysing the sequence 10
1.1.4 Increasing peptide affinity 13
1.1.5 Increasing cell permeability 18
1.1.6 Conclusion and future outlook 19
1.2 The role of dyskerin and telomerase in cancer 21
1.2.1 The structure and function of telomerase 21
1.2.2 The inhibition of telomerase 22
1.2.3 The function of dyskerin and its link to cancer 24
1.2.4 The dyskerin-dyskerin PPI 26
1.3 Thesis outline 27
1.4 References 28

Chapter 2 The design, synthesis and *in vitro* biological evaluation of peptides targeting the dyskerin-dyskerin protein-protein interaction 40
2.1 Introduction 41
2.1.1 The principles of solid-phase peptide synthesis 41
2.1.2 SPPS of difficult sequences 46
2.1.3 *In vitro* biophysical assays for protein-protein interactions 48
2.2 Chapter aims 55
2.3 Results and discussion 56

2.3.1 Initial peptide design and synthesis 56
2.3.2 The development of a fluorescence polarisation binding assay 62
2.3.3 Investigating the α-helicity of peptide 2.47 63
2.3.4 The design, synthesis and evaluation of hydrocarbon stapled peptides 64
2.3.5 The development of a surface plasmon resonance assay for MBP-dyskerin 73
2.3.6 The development of an HTRF assay to investigate the inhibition of the dyskerin-dyskerin PPI 80
2.3.7 The design, synthesis, and evaluation of lactam stapled peptides 84
2.4 Conclusion and Future work 90
2.5 References 91

Chapter 3 Design, synthesis, and *in vitro* biological evaluation of small molecules targeting the dyskerin-dyskerin PPI 98
3.1 Introduction 99
3.1.1 Targeting PPIs with small molecules 99
3.1.2 Peptide-directed binding 100
3.1.3 Triazole click chemistry 101
3.2 Chapter aims 104
3.3 Results and discussion 104
3.3.1 *In silico* design of semi-peptides 104
3.3.2 Covalent docking of azide fragments 106
3.3.3 Covalent docking of alkyne fragments 108
3.3.4 Triazole induced-fit docking 113
3.3.5 Synthesis of alkyne fragments 115
3.3.6 Synthesis of azide fragment 118
3.3.7 Triazole synthesis 119
3.3.8 Effect of triazoles on the dyskerin-dyskerin PPI 120
3.3.9 Determination of binding affinity through SPR 125
3.4 Conclusion and future work 126
3.5 References 128

Chapter 4 The cellular evaluation of peptides and small molecules targeting the dyskerin-dyskerin PPI 133
4.1 Introduction 134
4.1.1 Biophysical vs cellular activity 134
4.1.2 Cell permeability assays 135

4.1.3 Drug stability 136
4.2 Chapter aims 137
4.3 Results and discussion 137
4.3.1 The effect of the linear and stapled peptides on cell viability 137
4.3.2 The effect of peptide stapling on the cellular uptake 139
4.3.3 The effect of peptide stapling on the proteolytic stability 140
4.3.4 Small molecule cell viability 142
4.4 Conclusion and future work 143
4.5 References 145

Chapter 5 Recombinant expression of dyskerin 148
5.1 Introduction 149
5.1.1 Host organisms for protein expression 149
5.1.2 Protein expression in *E. coli.* 149
5.1.3 Protein purification methods 151
5.2 Chapter aims 153
5.3 Results and discussion 153
5.3.1 His-dyskerin protein expression 153
5.3.2 MBP-dyskerin expression 158
5.4 Conclusion and future work 162
5.5 References 164

Chapter 6 Experimental 167
6.1 General procedures 168
6.1.1 Reagents and solvents 168
6.1.2 Instrumental techniques 168
6.2. The design, synthesis and *in vitro* biological evaluation of peptides targeting the dyskerin-dyskerin protein-protein interaction 169
6.2.1 General synthesis of linear peptides 169
6.2.2 General synthesis of hydrocarbon stapled peptides 172
6.2.3 General synthesis of lactam stapled peptides 174
6.2.4 Synthesis of N-α-Fmoc-L-aspartic acid β-2-phenylisopropyl ester (Fmoc-L-Asp(O-2PhiPr)-OH) (2.84) 176
6.2.5 Fluorescence polarisation 178
6.2.6 HTRF 179
6.2.7 Surface plasmon resonance 180

6.2.8 Circular dichroism 180
6.3 Design, synthesis, and in vitro biological evaluation of small molecules targeting the dyskerin-dyskerin PPI 180
6.3.1 Computational modelling 180
6.3.2 Alkyne synthesis 182
6.3.3 Azide synthesis 186
6.3.4 Triazole synthesis 186
6.3.5 Surface plasmon resonance 196
6.4 The cellular evaluation of peptides and small molecules targeting the dyskerin-dyskerin PPI 196
6.4.1 General cell culture 196
6.4.2 MTS 197
6.4.3 Flow cytometry 197
6.4.4 Proteolytic stability 197
6.5 Recombinant expression of dyskerin 198
6.5.1 Recombinant expression of MBP-dyskerin 198
6.5.2 SDS-PAGE 198
6.5.3 Plasmid extraction 199
6.6 References 200

Abbreviations

CuAAC	Copper(I)-catalyzed azide alkyne cycloaddition
DBF	Dibenzofulvene
Boc	tert-butyloxycarbonyl
DIC	*N,N*'-diisopropylcarbodiimide
DIPEA	*N*,*N*-Diisopropylethylamine
DKC	Dyskeratosis congenita
DODT	2,2'-(Ethylenedioxy)diethanethiol
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
FITC	Fluorescein isothiocyanate
FP	Fluorescence polarisation
FRET	Förster resonance energy transfer
HATU	Hexafluorophosphate azabenzotriazole tetramethyl uronium
HBTU	Hexafluorophosphate benzotriazole tetramethyl Uronium
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HH	Hoyeraal-Hreidarsson syndrome
HOBt	1-hydroxybenzotriazole hydrate
hTR	Telomerase RNA
HTRF	Homogeneous Time Resolved Fluorescence
LC-MS	Liquid chromatography-mass spectrometry
MALDI	Matrix-assisted laser desorption/ionisation
MOPS	3-(*N*-morpholino)propanesulfonic acid
MTS	3-[4,5,dimethylthiazol-2-yl]-5-[3-carboxymethoxy-phenyl]-2-[4-sulfophenyl]-2H-tetrazolium
MWCO	Molecular weight cut off
Oxyma	Ethyl 2-cyano-2-(hydroxyamino)acetate

Pbf	2,2,4,6,7-Pentamethyldihydrobenzofuran-5-sulfonyl
PEG	Polyethylene glycol
PPI	Protein-protein interaction
PyBOP	Benzotriazol-1-yloxytripyrrolidinophosphonium hexafluorophosphate
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SPAAC	Strain promoted azide alkyne cycloaddition
SPPS	Solid-phase peptide synthesis
SPR	Surface plasmon resonance
RNP	Ribonucleoprotein
RPMI	Roswell Park Memorial Institute 1640 Medium
***t*Bu**	tert-butyl
TERT	Telomerase reverse transcriptase
TFA	Trifluoroacetic acid
TFE	Trifluoroethanol
TIPS	Triisopropylsilane
TRAP	Telomerase Repeated Amplification Protocol
TRIS	Tris(hydroxymethyl)aminomethane

Acknowledgements

I would like to thank my supervisors Dr. Andrew Beekman and Prof. Mark Searcey for all their support throughout my PhD. Thank you for this amazing opportunity, for guiding me throughout my research and for always being there to chat to, I couldn't imagine having done my PhD anywhere else.

I would also like to thank the Beekman and Searcey group, past and present, you have all made this the most fun experience and I am so happy to have spent this time with you all. A special thanks goes to Dr. Zoë Goddard and Dr. Marco Cominetti for all your help, support and advice day-to-day. Thank you to Prof. Maria O'Connell and Emily Hobson for all your help with the cell culture performed throughout my PhD. Thank you to everyone else at UEA who has helped me throughout the past three years.

Thank you to all of my friends outside of the lab for their support and for a welcome distraction from my PhD when needed. Lastly, I'd like to thank my family for all their support throughout my life and encouraging me to pursue my dreams.

Chapter 1

Introduction

1.1 The discovery of peptide inhibitors of PPIs derived from the sequence of the protein partner

1.1.1 Introduction

Protein-protein interactions (PPIs) modulate many biological processes, and dysregulation of these interactions leads to disease states, such as cancer. Targeting PPIs is challenging in comparison to classical drug targets, such as enzymes, receptors and ion channels as PPIs often take place across large (1500-3000 $Å^2$) and flat surfaces compared to the smaller surfaces (300-1000 $Å^2$) with well-defined binding pockets of classical targets.[1] Because of this, PPIs have previously been classed as 'undruggable', which has meant that many PPIs have still been largely unexplored in drug discovery and it is thus important to find new methods and improve existing methods to target PPIs.

PPIs can be identified *in vitro* using methods such as affinity chromatography using immobilised protein, and co-immunoprecipitation with both techniques able to detect protein partners.[2,3] *In vivo* yeast two-hybrid screens can identify PPIs in which the protein complex needs to be formed for the transcription of reporter genes to occur which can allow growth on specific media or for a colour change to occur.[4,5] To understand the importance of the discovered PPIs assays utilising silencing RNA or knockout technology can remove one of the proteins and the cellular effect can then be observed.[6,7] With the progress in Cryo-EM as well as in artificial intelligence techniques, such as Alphafold, which allow for the prediction of multimeric protein structures, there are now many opportunities for finding structural information about PPIs.[8,9] This information can therefore be utilised in the drug discovery process.

PPIs are commonly modulated using antibodies, peptides, and small molecules. Humanised antibodies are highly selective and effective cancer treatments, exemplified in targeting the PD-1/PD-L1 interaction in the immune system, for which multiple antibody treatments have been approved.[10] However, negatives to antibody therapies are their high production cost, low tissue penetration and adverse immune responses.[11] Their use has also been limited to extracellular targets, although research to develop intrabodies, intracellularly active antibodies, is ongoing.[11,12]

Peptides provide a good alternative to antibodies as they are easier to produce and lower in cost but still allow for the specificity in targeting interactions unlike small molecules. Phage display and mRNA display techniques have identified peptide inhibitors for several PPIs, but are generally non-specific for the protein interaction site, or indeed the target protein.[13–16]

Techniques such as split-intein circular ligation of proteins and peptides (SICLOPPS) incorporate a selection bias that identifies inhibitors of the target interaction, but require several layers of sifting to achieve selectivity. SICLOPPS is often linked to a bacterial reverse two-hybrid system in which, like in yeast two-hybrid systems, the assembly of the PPI is required for the transcription of reporter genes.[17,18] Identifying peptides from the secondary structures responsible for the targeted PPI promises efficacy and selectivity, and a conceptually simple starting point for modulator development. In this review we will describe approaches for the selection of an initial peptide sequence, and the processes for optimising the peptide for affinity and cell permeability.

1.1.2 Selection of an initial sequence

The choice of sequence starting point can be made by analysing the binding interface if structural information is available. However, without this structural information or without the knowledge of which part of the interaction is most important, methods including protein mutagenesis, sequence conservation analysis and peptide arrays can be used.

1.1.2.1 Selection of peptides through structural information

When the structure of a protein is known a direct approach to designing a peptide inhibitor is through analysing the structure and identifying the interaction face of the protein partner. The interacting section can be taken from this as a starting point before further optimisations take place.

If the interaction is mediated through an α-helix the whole helix can be used as a starting point for the interaction. For example, Kirsten rat sarcoma viral oncogene homolog (KRAS) is a small GTPase which is involved cell survival and proliferation with mutations found in cancers leading to decreased GTPase activity which in turn leads to increased RAS pathway signalling.[19] KRAS is negatively regulated by son of sevenless 1 (SOS1) and therefore the inhibition of the KRAS/SOS1 PPI can increase GTPase activity.[20] The SOS1 derived section $_{929}$FFGIYLTNILKTEEGN$_{944}$ (**1.1**) forming an α-helix (**Figure 1.1A**) was extracted from the crystal structure and used as the starting point.[21,22] The peptide was not reported until stapled analogues were synthesised as the linear peptide was inactive. Hydrocarbon stapling (RRFFGI**S$_5$**LTN**S$_5$**LKTEEGN (**1.2**)) increased the affinity up to a K_D of 60 nM recorded in a fluorescence polarisation (FP) assay with two arginine (R) residues added to increase cellular uptake.[21] Depending on the interaction site a region around the α-helix also may be included if it looks to be part of the interacting sequence in the structure, exemplified in targeting the nuclear transcription factor Y (NF-Y) trimer formation. NF-Y activates the genes associated

with cell cycle regulation and DNA repair with overexpression of NF-Y found in cancer.[23] Residues $_{267}$VNAKQYHRILKRRQARAKLEAEGKIPKER$_{295}$ (**1.3**) were taken from NF-YA encompassing an α-helical section which continues into an interacting disordered section (**Figure 1.1B**). This section was shown to bind to NF-YB/C dimer (K_D = 0.7 μM) in an FP assay. This sequence was used as a starting point allowing exploration of truncated sequences to determine if the disordered section is crucial for the interaction (see **Section 1.1.3.2**).[24] This approach was also taken in targeting the RbAp48/MTA1 interaction which consist of both an α-helical and disordered section.[25,26]

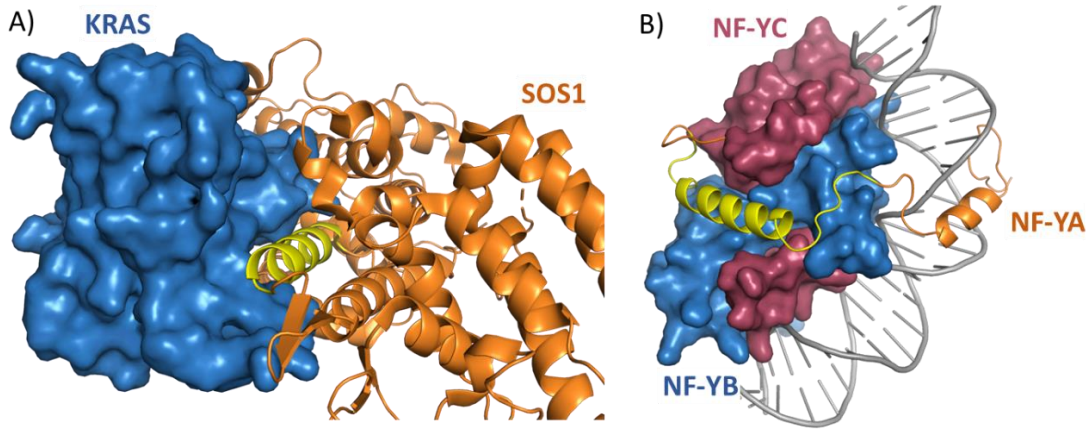


Figure 1.1: *A) Structure of SOS1 (orange)/KRAS (blue) (pdb: 1NVU) with the sequence $_{929}$FFGIYLTNILKTEEGN$_{944}$ (**1.1**) in yellow and B) Structure of NFY trimer (pdb: 4AWL) with NF-YA (orange) with peptide starting sequence $_{267}$VNAKQYHRILKRRQARAKLEAEGKIPKER$_{295}$ (**1.3**) (yellow), NF-YB (blue) and NF-YC (burgundy).*

The approach of deriving a peptide starting point from the sequence has now been taken for targeting many different α-helix mediated PPIs, such as Axin/β-catenin[27], BCL9/β-catenin[28], $Ca_v/Ca_v\beta$[29], Cullin3/KCTD11[30] and SARS-CoV-2 spike/ ACE2[31].

The protein structure may also reveal more than one potential interaction sites between the protein and its partner. Nucleus accumbens 1 (Nac1) is a repressor protein mediating the interactions between transcription factors with an essential role in carcinoma tumour growth.[32] In targeting Nac1 homodimerization two separate sections were explored as initial sequences: $_{12}$FGNSILECLNEQR$_{24}$ (**1.4**) and $_{44}$HRAVLAASSSYFRDLFN$_{60}$ (**1.5**) with the first peptide sequence showing weak binding in an FP assay (K_D = 360 μM) and the second peptide showing no binding. (**Figure 1.2**). Regardless of this both sections were further explored as constrained peptides to examine their ability to disrupt the PPI.[33]

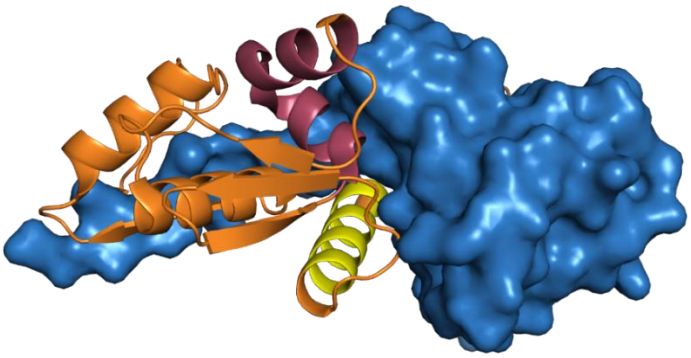


***Figure 1.2**: Structure of the Nac1 dimer (pdb: 3GA1) with section $_{12}$FGNSILECLNEQR$_{24}$ (**1.4**) in yellow and $_{44}$HRAVLAASSSYFRDLFN$_{60}$ (**1.5**) in burgundy.*

If the target protein has multiple protein partners, these could all be explored as potential starting points. In targeting Ras-related in brain (Rab) proteins, which misregulation is implicated in neurodegenerative diseases and cancer, crystal structures with multiple different protein partners were explored to identify initial peptide sequences.[34,35] Nine different peptides from Rab protein interfaces were derived and tested against a set of seven Rab proteins in an FP assay. Four peptides (derived from R6IP, LidA, REP1 and Rabin8) with low micromolar binding affinities were then explored as stapled peptides to further increase their binding affinity (**Figure 1.3**).[35]

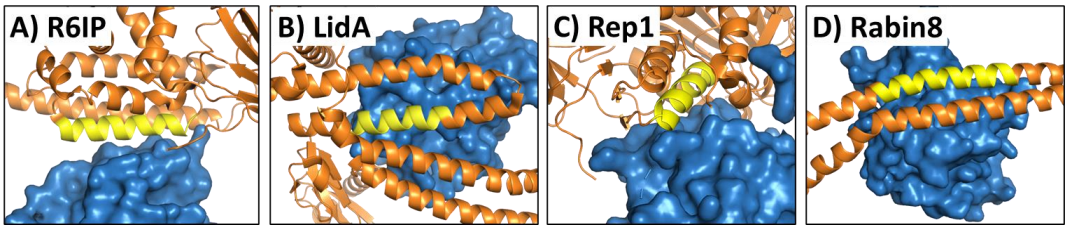


***Figure 1.3:** Structures of a Rab protein in complex with A) R6IP (pdb: 3CWZ), B) LidA (pdb: 3TNF), C) Rep1 (pdb: 1VG0) and D) Rabin8 (pdb: 4LHX) highlighting the peptide starting sequence in yellow.*

Using structural information is not limited to α-helical motifs with the same approach deployed for β-hairpin sections. This has been applied to targeting the epidermal growth factor receptor (EGFR) which is overexpressed in cancer.[36] EGFR forms an asymmetric dimer through the interaction of the C-lobe of one kinase domain with the N-lobe of another kinase domain at the PPI. Disruption of the dimerization could inactivate the kinase, which was explored through mimicking the EGFR dimerization arm residues $_{269}$YNPTTYQM$_{278}$ (**1.6**) which form a β-hairpin (**Figure 1.4**). This sequence was not able to inhibit the dimerization of EGFR as a linear sequence and therefore a head-to-tail disulfide cyclised peptide was designed, CYNPTTYQMC (**1.7**), which decreased dimer formation by 20% at 1 µM.[37]

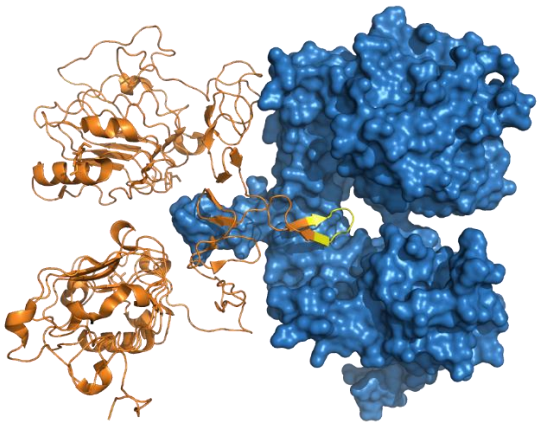


Figure 1.4: *Structure of EGFR dimer (pdb: 1IVO) with β-hairpin starting peptide sequence* $_{269}$*YNPTTYQM*$_{278}$ *highlighted in yellow.*

1.1.2.2 Using site-directed protein mutagenesis

Protein mutagenesis can be used to analyse which residues within the protein are most important for the interaction. The expression of alanine mutants that result in a loss of binding affinity suggest a hot spot residue. Protein mutagenesis is a useful tool to determine the best starting point when the PPI is comprised of multiple binding interfaces. This approach can be combined with computational techniques to select which residues to mutate minimising the number of proteins generated.

This approach was applied in targeting *Leishmania infantum*, a parasitic disease causing visceral leishmaniasis. Trypanothione reductase (TryR) maintains the intracellular redox state and is essential for parasitic activity.[38] The homodimer interface of TryR was explored computationally to identify potential hot spots by examining the burial of solvent-accessible surface area, the Van der Waals and electrostatic contributions of each residue to the interaction. Three mutant proteins (W81A, E436A and Q439A, **Figure 1.5**) were expressed and analysed using native PAGE showing that only the E436A mutation decreased dimer formation.[39] This residue is part of an α-helix ($_{435}$PEIIQSVGICMKM$_{447}$ (**1.8**)) chosen as the starting point for peptide design. The cysteine was replaced with serine to avoid disulfide bond formation leading to peptide **1.9** (Ac-PEIIQSVGISMKM-NH_2), which was able to disrupt the dimer formation in an ELISA assay (IC_{50} = 40.8 μM).[39]

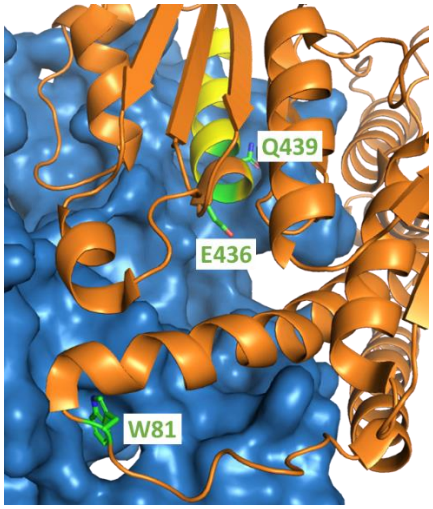


Figure 1.5: *Structure of the Leishmania infantum TryR dimer (pdb: 2JK6) highlighting residues W81, E436 and Q439 in green with the starting peptide section* $_{435}$*PEIIQSVGICMKM*$_{447}$ *(**1.8**) highlighted in yellow.*

Casein kinase 2 (CK2) is a protein kinase with functions in cell proliferation and apoptosis suppression. Overexpression of CK2 has been found in tumours, making it an attractive cancer target.[40] CK2 consists of two catalytic (α) and two regulatory (β) subunits and as an inhibitory strategy the CK2α /CK2β PPI was targeted. Three residues of CK2β (M166, Y188 and F190, **Figure 1.6**) were identified as pointing directly towards CK2α with Y188 and F190 found together on one of the CK2β subunits and M166 found on the other CK2β subunit. A single F190A mutant showed significantly reduced binding and a double (F190A + Y188A) and triple (F190A + Y188A + M166A) abolished binding as observed by surface plasmon resonance (SPR).[41] Y188 and F190 are part of the same β-hairpin loop ($_{186}$RLYGFKIH$_{193}$ (**1.10**)) which was taken as the starting point for the design of cyclic peptides. The sequence was not tested as a linear peptide, but a disulfide cyclised 13-mer peptide GCRLYGFKIHGCG (**1.11**) was shown to have moderate binding affinity by ITC (K_D = 1.75 µM).[41,42]

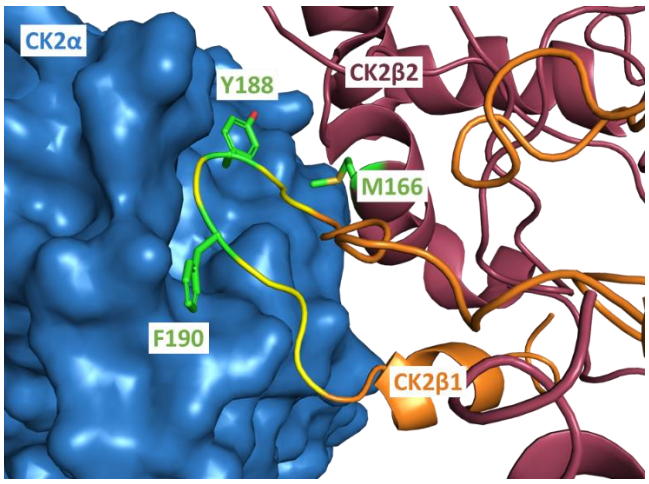


Figure 1.6: *Structure of CK2α (blue) in complex with CK2β1 (orange) and CK2β2 (purple) (pdb: 1JWH) highlighting residues M166, Y188 and F190 in green with the section* $_{186}$*RLYGFKIH*$_{193}$ *(**1.10**) used as the starting point in yellow.*

Site-directed mutagenesis can also confirm the importance of mutations found in disease, guiding the design of peptides. Phosphatidylinositol 3-kinase α (PI3Kα) is involved in cell proliferation forming interactions with adapter proteins such as insulin receptor substrate 1 (IRS1) in the presence of growth factor stimulation, stimulating the catalytic activity of PI3Kα.[43] The two most common mutations found in cancer (E545K and H1047R) were explored experimentally identifying IRS1 as a binder of E545K mutant protein but not of H1074R in pull-down experiments.[44] It was also shown that the interaction of E545K mutant PI3Kα with IRS1 is required for the growth of colon cancer cells *in vivo.* A 30-residue section consisting of two α-helices around E545 ($_{528}$EQLKAISTRDPLSEITEQEKDFLWSHRHYC$_{558}$ (**1.12**)) was chosen for peptide design, which inhibited the PPI formation in cell lysate as investigated by immunoprecipitation. Each helix was explored separately giving the N-terminal 21-mer peptide $_{528}$EQLKAISTRDPLSEITEQEKD$_{549}$ (**1.13**) and the C-terminal 18-mer peptide $_{541}$SEITEQEKDFLWSHRHYC$_{558}$ (**1.14**) with only the C-terminal peptide able to disrupt the PPI as shown by immunoprecipitation.[44]

Deletion scanning mutagenesis can be used to investigate the binding regions of a PPI through deleting sections of the protein. Eukaryotic translation initiation factor 4E (eIF4E) plays a role in the initiation of translation with overexpression inducing tumorigenesis.[45] eIF4E is negatively regulated by eIF4E-binding protein 1 (4E-BP1) through competition with eIF4G binding which is required for the formation of the active complex needed for translation.[46] Peptides mimicking 4E-BP1 were explored to inhibit translation through stopping the eIF4E-eIF4G PPI formation. Scanning deletion mutagenesis was used to remove sections of the 4E-BP1 protein residues 1-118. It was possible to remove 51 residues from the C-terminus and 50 residues from the N-terminus whilst conserving binding to eIF4E, giving the section $_{51}$RIIYDRKFLMECRNSPV$_{67}$ (**1.15**) as the starting point for peptide design which bound eIF4E with a K_D of 50 nM by ITC.[47]

1.1.2.3 Peptides spanning protein

If the interacting site of the PPI is not known peptides probing the entire protein can be generated. The synthesis and purification of many peptides can be time-consuming using solid phase peptide synthesis and so peptide arrays are applied. Many peptides are synthesised on a cellulose membrane which can be incubated with the target protein, after which the target protein is detected with an antibody and the results are visualised.[48] This approach allows a quick screening of overlapping peptide sequence to identify which section of the protein is able to bind the target protein.[49]

This technique was applied to find new antiviral peptides for human immunodeficiency virus (HIV) to target the PPI between human apolipoprotein-B mRNA-editing catalytic polypeptide-like 3G (A3G) and HIV-1 Virion infectivity factor (Vif) protein which leads to the degradation of A3G stopping the anti-viral activity of A3G.[50] A peptide array of both A3G and Vif was designed leading to the identification of A3G derived peptide $_{211}$WVRGRHETYLCYEVE$_{225}$ (**1.16**) able to inhibit HIV-1 propagation.[51] The same approach was also taken in targeting the BH3-interacting domain death agonist (BID) and mitochondrial carrier homologue 2 (MTCH2) PPI, an interaction involved in apoptosis. Two peptides ($_{59}$WTDGNRSSHSRLGRIE$_{73}$ (**1.17**) and $_{111}$WLQLRNTSRSEEDRNR$_{125}$ (**1.18**)) were found to induce near complete cell death at 50 μM in osteosarcoma cells as penetratin, a cell-penetrating peptide, conjugates.[52]

Although potentially time-consuming a similar approach as a peptide array can be performed with solid phase peptide synthesis, as demonstrated when targeting HIV-1 integrase (IN). HIV-1 IN is required for viral replication and therefore of interest as an HIV drug target. HIV-1 IN forms multimers and as well as this interacts with other proteins such as lens epithelium-derived growth factor (LEDGF) as well as DNA.[53] 16 peptides spanning the protein between residues 10-267 were created to probe the N-terminus, catalytic core and C-terminus for HIV-1 integrase inhibition (**Figure 1.7**).[54]

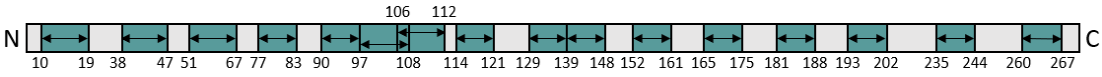


Figure 1.7: *Sections of HIV-1 IN selected for peptide synthesis*

The designed peptides were synthesised using solid-phase peptide synthesis and then tested for HIV integrase catalytic activity. This resulted in the identification of two peptide sequences: $_{97}$TAYFLLKLAGRW$_{108}$ (**1.19**) and $_{129}$ACWWAGIKQEF$_{139}$ (**1.20**) which showed inhibition of integration (also known as strand transfer) with an IC_{50} of 2.7 and 56 μM respectively with further work needed to determine their mode of action.[54]

1.1.2.4 Conserved regions across proteins or species

Many proteins are conserved across species or across groups of proteins with similar functions, which can indicate the importance of these regions, and suggest an initial sequence for binding and inhibiting target PPIs.

Targeting protein for Xklp2 (TPX2) is a spindle assembly factor which is required during cell mitosis. TPX2 localises Aurora-A, an essential mitotic kinase to spindle microtubules which has been highlighted as an oncogene.[55] Residue conservation across human, *Xenopus* and

zebrafish TPX2 is seen across residues 1-43 (**1.21**) and immunoprecipitation and pull-down experiments showed the interaction of **1.21** with Aurora-A whilst also showing that the shorter 15-43 section (**1.22**) was not able to do so.[56]

Protein kinase A (PKA) regulates different biological processes and consists of a tetramer of two catalytic subunits (PKA-C) and two regulatory subunits (PKA-R) with the two PKA-R units forming a dimer.[57] The activity of PKA is regulated by Akinase-anchoring proteins (AKAPs) with the deregulation of AKAPs associated with cancer as well as cardiovascular and immune diseases.[58] Among AKAPs there is a conserved α-helical Akinase binding (AKB) motif which binds at the interface of the PKA-R dimer.[59] AKB motifs from three different AKAPs (RIAD (EQYANQLADQIIKEATE (**1.23**)), AKAP220 (SGLANFLVSEALSNALK (**1.24**)) and smAKAP (LDYAQRLSEEIVRAVQQWA (**1.25**))) were explored. The peptides were modified by introducing lysine residues on the solvent-exposed face to increase solubility generating the following three peptides: **KK**YA**K**QLADQIIKEATE (**1.26**) (K_D = 605 nM, derived from RIAD), **KK**LA**K**FLVSEAL**K**NALK (**1.27**) (K_D = 22.4 nM, derived from AKAP220), and **KK**YAQRLS**KK**IVRAVQQWA (**1.28**) (K_D = >1000 nM, derived from smAKAP).[60] The binding affinity was then further improved by hydrocarbon stapling to 2.0 nM, 6.2 nM and 2.1 nM for the RIAD, AKAP220 and smAKAP derived peptides respectively.[60]

Nuclear factor erythroid 2–related factor 2 (Nrf2) is a transcription factor leading to the transcription of cytoprotective genes. Kelch-like ECH-associated protein 1 (Keap1) is a negative regulator of Nrf2 tagging Nrf2 for degradation by ubiquitination.[61] To target the Nrf2/Keap1 PPI the human sequence of Nrf2 was compared to that in mouse, chicken, zebrafish and drosophila with two highly conserved regions found: residues 17-32 (**1.29**) containing an LXXQDXDLG motif and residues 77-82 (**1.30**) containing an DXETGE motif. Trough ITC measurements of mutant proteins missing either section, it was shown that the binding affinity of **1.30** was higher than of **1.29**.[62] Therefore a 16-mer peptide ($_{69}$AFFAQLQLDEETGEFL$_{84}$ (**1.31**)) around the human $_{77}$DEETGE$_{82}$ motif was used as an initial sequence for inhibitor development, binding Keap1 with a K_D of 20 nM.[62,63]

1.1.3 Analysing the sequence

Once an initial sequence has been identified optimising the binding affinity and ligand efficiency can be achieved with truncation studies. An alanine scan provides information about the importance of each residue in the sequence allowing for sequence truncation, identifying residues to be optimised using mutations and potential stapling position.

1.1.3.1 Alanine scan

An alanine scan can be used to highlight the hot spot residues within a peptide sequence. In an alanine scan residues within the sequence are sequentially replaced with alanine and the binding affinity is analysed with a loss in binding affinity equating to a hot spot.

Inducible nitric oxide (NO) synthase (iNOS) produces NO which is required for the intracellular killing of pathogens.[64] iNOS is negatively regulated by SPRY domain–containing SOCS (suppressor of cytokine signaling) box protein 2 (SPSB2) and the linear peptide Ac-KEEKDINNNVKKT-NH_2 (**1.32**) derived from iNOS was shown to bind SPSB2 with a K_D of 13.3 nM (ITC). Through an alanine scan the most important residues in the sequence for binding were identified in the section $_{23}$DINNN$_{27}$, which was used in the design of cyclic peptides.[65,66] Disruptor of telomeric silencing 1-like (DOT1L) catalyses the methylation of histone H3 at lysine 79 and has been implicated in leukaemia.[67] In up to 10% of acute leukaemia cases the mixed lineage leukaemia (MLL) protein is fused to a partner protein such as AF9. The PPI of AF9 with DOT1L recruits DOT1L to AF9 targeted genes increasing their methylation therefore increasing their expression leading to leukaemia. An alanine scan was performed of peptide $_{865}$LPISIPLSTV$_{874}$ (**1.33**) (IC_{50} = 0.49 μM by competitive SPR) derived from DOT1L showing that three C-terminal residues were not essential for binding. Therefore the shortened 7-mer peptide $_{879}$LPVSIPL$_{886}$ (**1.34**) was generated with a small loss in binding affinity (IC_{50} = 3.9 μM).[68]

Alternatively, the unimportant residues can be mutated to improve the binding. The PPI between (Suppressor of Mothers against Decapentaplegic) Smad and Yes-associated protein (YAP) is involved in the activation of a signalling pathway leading to heterotopic ossification (HO) characterised by bone formation outside of the skeleton, making it a target for the treatment of HO.[69] Computationally an optimised peptide of the 205-214 section of Smad was identified $_{205}$DGWPPPYPRV$_{214}$ (**1.35**) (K_D = 2.5 μM). An alanine scan of **1.35** revealed three unimportant residues (G206, P212 and V214) which were computationally mutated to all other natural amino acids. Six peptides were explored experimentally with the combination of two mutations ($_{205}$D**Q**WPPPYPR**H**$_{214}$ (**1.36**)) leading to a nearly 10-fold improvement in binding (K_D = 0.34 μM).[70]

Similarly, an alanine scan can highlight macrocyclization positions for us in, for example, peptide stapling. Calpain is a calcium-regulated cysteine protease exploited in disease. Calpastatin is its endogenous inhibitor and was used to develop peptide inhibitors using section $_{201}$IPPKYRELLA$_{210}$ (**1.37**). Virtual alanine scanning mutagenesis identified P202 and

Y205 as important residues and, using the information of the alanine scan, were avoided as stapling positions in the design of m-xylyl stapled peptides. The stapled peptide showed inhibition of calpain (K_i = 10.2 µM in a protease activity assay) with the linear peptide having no effect.[71]

An alanine scan may also lead to the discovery of a peptide with enhanced binding affinity. An alanine scan of Ac-DEETGEF-OH (**1.38**) (IC_{50} = 5.39 µM), a shortened version of the Nrf2 derived peptide **1.31** (see **Section 1.1.2.4**), found that an E77A mutation increased binding (IC_{50} = 0.730 µM) nearly 10-fold. This approached the binding of the longer 9-mer peptide Ac-LDEETGEFL-OH (**1.39**) (IC_{50} = 0.389 µM).[72] YAP residues $_{86}$MRLRKLPDSFFKPPE$_{100}$ (**1.40**) form a twisted-coil with R87 and F96 forming a cation-π interaction. The section $_{81}$PQTVP$_{85}$ (**1.41**) was previously reported to be essential to binding and therefore the linear peptide $_{81}$PQTVPMRLRKLPDSFFKPPE$_{100}$ (**1.42**) was chosen as the starting point.[73] **1.42** gave an IC_{50} of 49 µM by competitive SPR and an alanine scan of **1.42** revealed that a D93A mutation, giving $_{81}$PQTVPMRLRKLP**A**SFFKPPE$_{100}$ (**1.43**) increased binding (IC_{50} of 25 µM).[74]

1.1.3.2 Sequence truncation

Shortening the peptide sequence to contain only the essential residues for the interaction is important in drug discovery to improve not only the ligand efficiency but also improve the synthesis. As described in **Section 1.1.3.1** an alanine scan be used to do so, but it can also be probed experimentally by sequentially removing residues from the sequence.

Truncation of the sequence derived from the structure of NF-YA (residues 267-295 (**1.3**)) (see **Section 1.1.2.1**) was explored (**Table 1.1**).

Table 1.1: *The NF-YA derived sequence 267-295 and the six truncated sequences based on this and their binding affinity (K_D) as measured by FP*

Nr.	K_D (µM)	Sequence
1.3	0.7	$_{267}$VNAKQYHRILKRRQARAKLEAEGKIPKER$_{295}$
1.45	1.3	$_{267}$VNAKQYHRILKRRQARAKLEAE$_{288}$
1.46	2.0	$_{267}$VNAKQYHRILKRRQARAKL$_{285}$
1.47	9.5	$_{267}$VNAKQYHRILKRRQAR$_{282}$
1.48	2.1	$_{268}$NAKQYHRILKRRQARAKL$_{285}$
1.44	2.9	$_{270}$KQYHRILKRRQARAKL$_{285}$
1.49	45	$_{271}$QYHRILKRRQARAKL$_{285}$

Six shortened peptides were generated, first removing residues from the C-terminus until no longer tolerated followed by shortening of the N-terminus until no longer tolerated (**Table**

1.1). In total, the sequence was shortened by 13 residues to 270-285 (**1.44**) with only a small loss in binding affinity in an FP assay (K_D = 0.7 μM to K_D = 2.9 μM).[24]

Repressor/activator protein 1 (RAP1) is part of the shelterin complex which plays a role in the regulation of telomeres.[75] The PPI of RAP1 with telomeric repeat-binding factor 2 (TRF2) represses the localisation of poly [ADP-ribose] polymerase 1 (PARP1) to telomers resulting in catastrophic telomere loss.[76] The original 41-residue TRF2 interacting sequence, consisting of two α-helices, bound RAP1 with a K_D of 16.5 nM. This was shortened to a 16-mer peptide ($_{281}$TTIGMMTLKAAFKTLS$_{296}$ (**1.50**)) containing only one of the α-helices as a starting point for stapling. A triazole stapled peptide, $_{281}$TTIGMMTLK**Z**AFK**X**LS$_{296}$, (**1.51**) was 10-fold more potent than the linear 16-mer peptide (K_i = 0.14 compared to K_i = 2.0 μM in a competitive FP assay). The binding affinity of the peptide was further improved through the introduction of mutations as discussed in **Section 1.1.4.1**.[77]

Apoptosis-inducing factor (AIF) regulates cellular survival through functions in the mitochondria and AIF mediates neuronal cell death under lethal cellular stress with inhibition of AIF having a neuroprotective effect.[78] Cytosolic cyclophilin A (CypA) is a protein partner of AIF which upon binding initiates the translocation of the proteins into the nucleus.[79] The linear AIF mimicking 25-mer peptide ($_{370}$QSVGVSSGKLLIKLKDGRKVETDHI$_{394}$ (**1.52**)) showed binding to CypA (K_D = 12 μM by SPR) and blocked nuclear translocation.[80] **1.52** was shortened to 9-mer section $_{381}$LIKLKDGRKVE$_{389}$ (**1.53**) as NMR studies had shown that was the most important section of the interaction.[81] An increased binding affinity was seen for **1.53** (K_D = 2.4 μM by label-free Corning Epic) compared to **1.52** (K_D = 12 μM by SPR).[82]

Although the removal of a large section may not be possible for most peptides, the removal of any residues is still beneficial to improve the drug-like properties of the peptide. The 16-mer peptide ($_{69}$AFFAQLQLDEETGEFL$_{84}$ (**1.31**)) derived from Nrf2 (see **Section 1.1.2.4**) was truncated to a 14-mer, 12-mer, 10-mer, 9-mer, 8-mer and 7-mer. From the 16-mer to a 10-mer ($_{75}$QLDEETGEFL$_{84}$ (**1.54**)) the binding affinity was maintained (K_D = 27.3 nM for **1.54** and K_D = 23.9 nM for **1.31**). The removal of another residue of the N-terminus to give the 9-mer peptide ($_{76}$LDEETGEFL$_{84}$ (**1.55**)) led to a big decrease in binding affinity (K_D = 352 nM) with any further truncations abolishing binding.[83,84]

1.1.4 Increasing peptide affinity

Having identified the essential residues the peptide can be optimised further for binding affinity, cell permeability and proteolytic stability. To increase the binding affinity, it may be possible to introduce mutations into the sequence or use peptide stapling or

macrocyclization. Peptide stapling and/or macrocylisation can also provide an opportunity to increase the peptides cell permeability and proteolytic stability.

1.1.4.1 Introducing mutations into the sequence

Improving the binding affinity of the designed peptide can be achieved through the introduction of a mutation into the sequence with the possibility to explore unnatural amino acids. It may be possible to increase the size of a hydrophobic residue by introducing substituents or through introducing larger aromatic rings.

As discussed in **Section 1.1.3.2** a triazole stapled peptide derived from TRF2 (**1.51**) had been established as a binder of RAP1 (K_i = 0.14 µM). The crystal structure of the RAP1/TRF2 complex showed a large hydrophobic space around I283 (**Figure 1.8A**). Therefore different natural and unnatural hydrophobic mutations were attempted at I283 (2,2-dimethylpropylalanine, cyclohexylalanine, phenylalanine, tryptophan, naphthalene and 2, 3 and 4-chlorophenylananine) with the 2-chlorophenylalanine (**1.56**) found to increase binding the most (K_i = 7 nM).[77] Similarly mutations were explored around F374 in a 35-mer linear peptide (347-381 (**1.57**)) peptide derived from BCL9 targeting β-catenin. The binding pocket of F374 is shallow but wide which could accommodate a larger group (**Figure 1.8B**). Smaller and larger hydrophobic groups were explored with the introduction of a 2-naphthylalanine (**1.58**) leading to a 3-fold increase in binding affinity.[85]

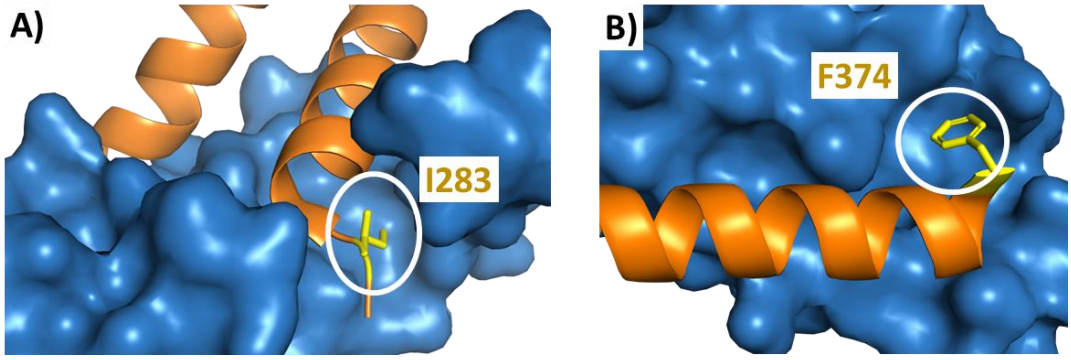


Figure 1.8: *Structure of A) RAP1 (blue) with TRF2 (orange) (pdb: 3K6G) highlighting the binding pocket of I283 with a white circle and B) β-catenin (blue) with BCL9 (orange) (pdb: 2GL7) highlighting the binding pocket of F374 with a white circle*

As mentioned in **Section 1.1.2.2** disulfide cyclised 13-mer peptide GCRLYGFKIHGCG (**1.11**) was shown to have moderate binding affinity by ITC (K_D = 1.75 µM). Docking experiments were performed based on the crystal structure. Meta-substitutions were introduced in F190 to explore hydrogen or halogen bond formation with a well-defined water molecule in the structure. A chloro and iodo-substituent version were synthesised with the introduction of the meta-iodine (**1.59**) resulting in an improved binding affinity (K_D = 0.239 µM by ITC).[86]

Multiple mutations can also be introduced into a sequence although the combination of individually beneficial mutations may not be additive. A FYCO1 derived linear 18-mer peptide spanning residues $_{1275}$GQGANTDYRPPDDAVFDIITDEELCQIQESGSSL$_{1298}$ (**1.60**) was shown to bind LC3B (K_D = 0.29 µM) in an FP assay.[87] The binding pocket of F1280 was shown to be able to accommodate a bigger hydrophobic group and a 2-naphthylalanine substitution (**1.61**) increased the binding affinity (K_D = 0.14 µM). A substitution of L1288 with a *tert*-butylalanine (**1.62**) led to a small improvement of binding affinity (K_D = 0.25 µM). The N-terminal section of the peptide is located near various negatively charged residues of LC3B and an introduction of an Arg to the N-terminus (**1.63**) showed a slight improvement of the binding affinity (K_D = 0.20 µM). A combination of these three mutations (**1.64**) led to a peptide with a K_D of 0.12 µM which was similar to **1.61**.[88]

1.1.4.2 Stapling of α-helical peptides

Peptide stapling can increase the binding affinity of a peptide due to the preorganisation of a peptide secondary structure.[89] In peptide stapling a bond is formed between two residues on the same side of the α-helix (**Figure 1.9**).

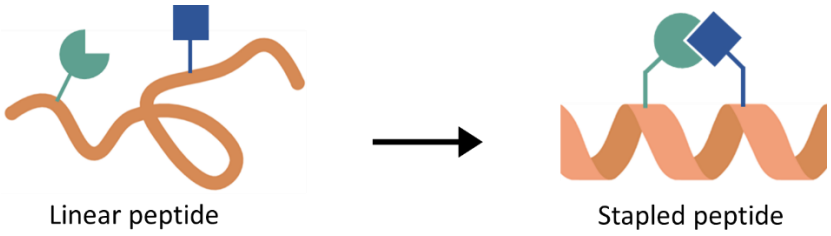


Figure 1.9: *Schematic overview of peptide stapling*

This is achieved with natural amino acids such as cysteines to form thioether stapled peptides, or lysine with aspartic or glutamic acid to create a lactam stapled peptide. Unnatural amino acids are also used in peptide stapling with both hydrocarbon stapling and triazole stapling described. Fairlie and co-workers compared the α-helicity of pentapeptides using six different stapling techniques finding that lactam-stapling induced the highest level of α-helicity followed by hydrocarbon stapling.[90]

Hydrocarbon stapling is commonly due to its ability to increase cell permeability.[91] Hydrocarbon stapling has been extensively reviewed with many examples of its successful application in increasing binding affinity, proteolytic stability and cell permeability.[27,92–94] Two cases will be highlighted here as examples of the use of a single hydrocarbon staple leading to the peptides with *in vivo* activity.

The PPI between β-catenin and its cofactor B-cell lymphoma 9 (BCL9) is part of the Wnt signalling pathway which plays a key role in cell proliferation with activation of this pathway found in different cancers.[95,96] From the crystal structure, a 35-mer linear peptide (residues 347-381 (**1.65**)) from BCL9 was selected initially and was shown to bind β-catenin with a K_D of 616 nM in an FP assay. Truncation to a 24-mer (residues 351-375) (**1.67**) resulted in only a small loss in binding affinity (Ki = 1.90 μM compared to 0.96 μM) as measured in a competitive FP assay.[97] Therefore this shortened sequence was used in the design of hydrocarbon stapled peptides. The best peptide, $_{351}$LSQEQLEHRERSL**S$_5$**TLR**S$_5$**IQRMLF$_{374}$ (**1.68)**, showed dissociation of the BCL9/ β-catenin complex with an IC_{50} of 135 nM. *In vivo* mouse xenograft models of Wnt-driven cancer showed suppression of tumour growth, angiogenesis and invasion with treatment of **1.68**.[98]

Similarly, peptides were derived from Bcl-2–interacting mediator of cell death (BIM) to target antiapoptotic B-cell lymphoma 2 (BCL2) family proteins.[99,100] The BIM BH3 helix ($_{146}$IWIAQELRRIGDEFNAYYARR$_{166}$ (**1.69**)) was modified with *i*, *i*+4 hydrocarbon stapling at positions 154 and 158, giving peptide **1.70** ($_{146}$IWIAQELR**S$_5$**IGD**S$_5$**FNAYYARR$_{166}$). **1.70** showed nanomolar binding against a range of BCL-2 family proteins including BCL-X_L, MCL-1 and BFL-1 and suppressed tumour growth *in vivo*.[101]

Hydrocarbon stapling may however not always prove the best option for structure restraint. In targeting Li-TryR dimer using peptides based on section $_{435}$PEIIQSVGISMKM$_{447}$ (**1.9**) (see **Section 1.1.2.2**) it was found that although hydrocarbon stapling gave the most α-helical peptides, only lactam stapled peptides were able to disrupt the dimer formation. Molecular dynamics studies suggested Q439 was not able to adopt the required conformation in the hydrocarbon stapled peptides.[102–104] Other stapling methods have also been chosen to increase the solubility of the peptides, such as was done in targeting the RAP1/TRF2 PPI in which triazole stapling was applied to a 16-mer peptide (281-296) giving peptide **1.51** (see **Section 1.1.3.2**). The triazole staple increased the binding affinity (K_i = 0.14 compared to K_i = 2.0 μM) whilst improving its solubility.[77]

Multiple staples can also be introduced into the sequence, which can be beneficial for longer sequences (**Figure 1.10**).

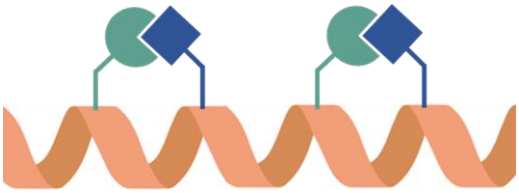


Figure 1.10: *Schematic overview of a double stapled peptide*

Double hydrocarbon *i*, *i*+4 stapling was applied to a JAZ9 derived 21-mer peptide (218-238) targeting MYC with both higher α-helicity (45% compared to 25%) and better binding affinity (K_D = 0.10 μM compared to 0.88 μM and 2.4 μM) seen for the double over the single stapled analogues.[105] The same approach taken in targeting the Rab protein, Rab6a, through deriving a peptide from R6IP[35], and for targeting DAP12 homodimerization[106], SNARE/synaptotagmin-1[107] and the p53 PPI with MDM2/MDMX.[108]

1.1.4.3 Macrocyclisation of peptides

For α-helical structures stapling can be used to ensure the helical structure of the peptide. For other secondary structures, such as β-hairpin peptides, this can be achieved with macrocylisation. The EGFR dimerization arm residues $_{269}$YNPTTYQM$_{278}$ (**1.6**) form a β-hairpin as was discussed in **Section 1.1.2.1** with a cyclic version of this peptide, CYNPTTYQMC (**1.7**), decreasing dimer formation by 20% at 1 μM (**Figure 1.11**).[37]

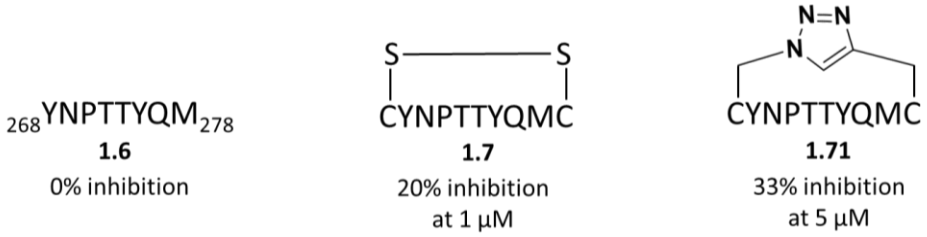


Figure 1.11: *The macrocylisation of EGFR derived peptide $_{269}$YNPTTYQM$_{278}$ (**1.6**) using disulfide (**1.7**) and triazole (**1.71**) linkers*

A triazole stapled version (**1.71**) inhibited EGFR dimerization by 33% at 5 μM, similar to the results from previous disulfide cyclised peptides. However, the proteolytic stability was better for the triazolyl-cyclised peptide over the disulfide version which is advantageous for further development.[109]

Vestigial-like protein (VGLL) is a coactivator of transcriptional enhanced associate domains (TEADs) and has been found to be a transcriptional repressor inhibiting tumorigenesis caused by YAP as part of the Hippo pathway with VGLL4 competing with YAP for binding to TEADs.[110] VGLL4 is a small protein consisting of an α-helix linked to a double-stranded β-sheet through a loop linker. To mimic the double-stranded β-sheet head-to-tail disulfide cyclised peptides

were synthesised spanning different lengths over residues 19-32 and conjugated to TAT. The 19-32 cyclic peptide (**1.72**) showed increased activity in a cell viability assay of 18.1 μM compared to 87.4 μM for its linear counterpart and the slightly shorter 20-31 (**1.73**) showed a similar activity of 25.0 μM with its linear counterpart showing >100 μM activity.[111]

The initial CK2β peptide $_{186}$RLYGFKIH$_{193}$ (**1.10**) forms a β-hairpin loop with Y188 at the turn (see **Section 1.1.2.2**). The peptide was extended and Cys residues were added which facilitated cyclisation through a disulfide bridge giving the 13-mer peptide GCRLYGFKIHGCG (**1.11**). This peptide inhibited the CK2α/CK2β complex with an IC_{50} of 3 μM, a ten-fold increase on the linear version (IC_{50} = 30 μM).[41] The replacement of the disulfide bridge with a two-component triazole bridge further increased binding (K_D = 460 nM) compared to **1.11** (K_D = 1000 nM) as measured by ITC.[112] An I192W mutation was introduced giving peptide **1.74** increasing binding further (K_D = 150 nM) thought to be due to π-π stacking between the Trp residue and the phenyl ring of the linker.[112]

The linear AIF mimicking 25-mer peptide (370-394) (**1.52**) (see **Section 1.1.3.2**) showed binding to CypA and blocked nuclear translocation.[80] This section forms a β-hairpin and so disulfide and triazole cyclisation were applied to create mono and bicyclic peptides. A peptide with a combination of disulfide and a triazole bridge, $_{370}$QSCGVSSGZLLCKLKDGRKVXTDHI$_{394}$ (**1.75**) (**Figure 1.12**) showed improved binding (K_D = 0.85 μM) compared to the **1.52** (K_D = 5.0 μM).[113]

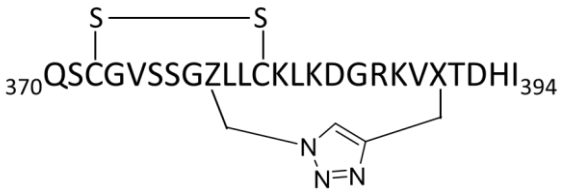


Figure 1.12: *Structure of bicylic peptide $_{370}$QSCGVSSGZLLCKLKDGRKVXTDHI$_{394}$ (**1.75**) derived from AIF*

1.1.5 Increasing cell permeability

One of the main problems which can be faced in the development of peptide therapeutics is their lack of cell permeability due to their hydrophilic character. Different methods to increase the cell permeability of peptides has been developed, such as hydrocarbon stapling discussed in **Section 1.1.4.2**. However, it may not be possible to insert a staple into any peptide sequence especially in the development of non α-helical peptides. Instead cell-penetrating peptides (CPPs) have been developed which can be conjugated to the main peptide, with this top having been reviewed extensively.[114–117]

This was applied to the a disulfide cyclised YAP derived peptide (see **Section 1.1.4.3**). Although the peptide showed high binding affinity (IC_{50} = 15 nM in competitive SPR), it was not able to penetrate cells. Therefore, the peptide was conjugated to TAT, a CPP derived from HIV, with cell permeability now seen.[73,74]

Many different CPPs have now been developed, but a similarity between them all is their positively charged nature. Therefore, alternatively the addition of two or three arginine residues to the end of the peptide has also been shown to increase cell permeability by changing the overall charge without the need to add a large sequence of amino acids. As mentioned in **Section 1.1.2.1** the hydrocarbon stapled peptide **1.2** derived from SOS1 required the addition of two Arg residues, changing the overall charge from -1 to +1, for cellular uptake.[22] The CK2 peptide inhibitor (**1.74**) discussed in **Section 1.1.4.2** showed a good binding affinity (K_D = 150 nM), but the cell permeability was limited. Through the introduction of three (D)-arginine residues to the sequence (**1.76**) the peptide was now cell permeable.[112]

Instead of improving the cell permeability of the peptides through the addition of a CPP or arginine residues, mutations can be introduced to change the overall charge. A double hydrocarbon-stapled R6IP derived peptide targeting Rab6a ($_{900}$DDE(S5)EWF(S5)YHL(S5)FFN(S5)V_{916} (**1.77**)) demonstrated a K_D of 7.8 μM, but cellular uptake was limited. Introduction of D901N and E904Q (**1.78**) removed the negative charges and increased cellular uptake (cellular uptake in range of TAT) with only a small loss in binding affinity (K_D = 12.7 μM).[118] This method was also applied to a p53 peptide ($_{17}$ETFSDLWKLLPE$_{28}$ (**1.79**)). Through a double staple of hydrocarbon stapling at positions 17 and 21 and lactam stapling at positions 24 and 28 as well as mutations L22K and P27R, giving peptide **1.80** (Ac-**S_5**TFS**S_5**KW**D**LLR**K**-NH_2), the overall charge was changed from -2 to +2. This peptide reduced cell viability whereas its counterpart without mutations (giving an overall 0 net charge) did not affect cell viability.[108]

1.1.6 Conclusion and future outlook

Peptides for the inhibition of protein-protein interactions offers several advantages over antibodies and small molecules. Peptides are inexpensive, structurally designed, and readily modified, while offering exquisite selectivity for their target. Peptides cover large protein interaction surfaces and have demonstrated high levels of success for the inhibition of biomacromolecule interactions. The design of peptides which can selectively bind their target is however not trivial. Display techniques such as phage display or mRNA display offer excellent starting points but limited selectivity. We have highlighted the conceptually elegant

and simple method of deriving a peptide to inhibit a PPI from one of the interacting protein partners. Using the information available about the interaction interface an initial sequence for investigation can be chosen. High amounts of structural information allow for the extraction of secondary motifs from the interface. With limited structural information available the interaction between the proteins can be analysed using protein mutagenesis to understand the importance of regions or residues for the interactions. When only sequence information is available peptide arrays allow for the synthesis of a large library of peptides which can span large sections of protein an identify binding motifs. Similarly, conserved regions across different proteins or species indicate the importance of that section of the sequence and offer an excellent initial sequence. After a binding sequence has been identified the sequence can be analysed using an alanine scan to provide information about the key residues for the interaction and the sequence may be shortened to improve the drug-like properties of the peptide. To further increase the binding affinity of the selected peptides mutations may be introduced or conformational constraints (peptide stapling or macrocyclization) may be applied. Further modifications can be made if necessary to improve the cellular uptake by increasing the positive charge of the peptide or by conjugation to a cell-penetrating peptide. These methods have resulted in many examples of peptide inhibitors for a wide array of interactions.

Recent developments in the identification of protein-protein interactions, and with structural information increasingly available through advances in protein structure prediction and cryo-EM, this method of identifying peptide inhibitors for biomacromolecular interactions has become more accessible and can allow for the exploration of previously unexplored interactions including protein-protein, protein-DNA and protein-RNA interactions.

1.2 The role of dyskerin and telomerase in cancer

1.2.1 The structure and function of telomerase

Telomerase creates and elongates telomeres by adding the repetitive sequence 5'-TTAGGG to the end of the chromosomes, which protects the chromosomes from degradation and fusion.[119] It is active in gametes and during early development, but not active in most somatic cells. During each cell division a number of base pairs are lost due to incomplete replication of the lagging strand and due to oxidative stress.[120] Therefore, without telomerase, during each cell division the telomeres will progressively shorten until eventually the DNA damage response (DDR) mechanism is activated. This prevents the DNA from being damaged and being incorrectly replicated into new cells and instead cell apoptosis and/or senescence is induced.[121] In up to 90% of cancer cells reactivation of telomerase is seen, which leads to the cancer cells being able to replicate limitlessly as the telomere length never reaches the point at which the DDR mechanism is induced.[122]

Due to the telomerase reactivation seen in most cancer cells and its lack of activity in most healthy cells, telomerase is a good selective cancer target. However, drug development targeting telomerase has been difficult due to a lack of structural information available. Telomerase is a large protein complex, containing many unexplored PPIs that could be used as cancer targets. Although parts of the structure were known and some homology modelling could be used to mimic the structure of human telomerase, the first complete structure was not published until 2018.[123]

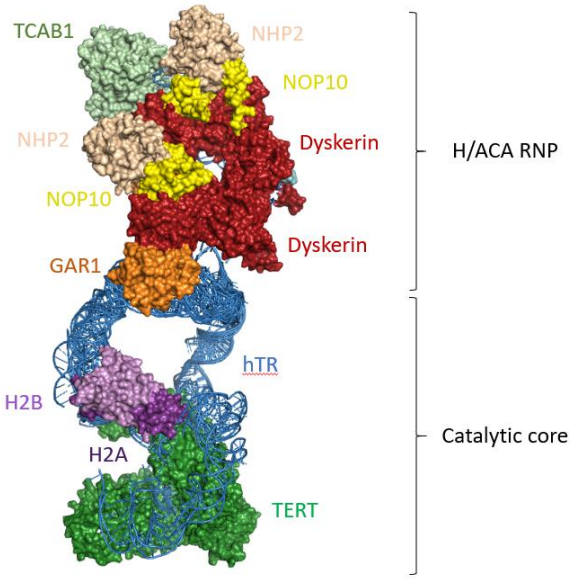


Figure 1.13: *Structure of human telomerase, highlighting the different components in different colours*

Before this cryo-EM structure some homology modelling still had to be used due to its resolution, but a higher resolution structure was published in 2021, giving a detailed view of DNA- and RNA-binding interfaces as well as of the protein-protein interactions in telomerase (**Figure 1.13**).[124]

The telomerase complex consists of three parts, telomerase RNA (hTR), telomerase reverse transcriptase (TERT) and a H/ACA ribonucleoprotein (RNP). hTR and TERT together form the catalytic core which is directly responsible for telomere elongation. The H/ACA RNP consists of two heterotetramers, each consisting of dyskerin, nucleolar protein 10 (NOP10), nucleolar human protein 2 (NHP2) and glycine- and arginine-rich nucleolar protein 1 (GAR1), and one Cajal body protein, Telomerase Cajal body protein 1 (TCAB1). Although the H/ACA RNP does not directly interact with the telomeres, it is needed for telomerase assembly and stabilisation.[123]

1.2.2 The inhibition of telomerase

The inhibition of telomerase has been difficult with only one candidate, Imetelstat, having advanced into clinical trials.[125] It has been shown that telomerase inhibition can be dependent on telomere length, with cell replication continuing until the critical telomere length has been reached. This could mean that the response of the cancer cells to treatment would take longer than for other anti-cancer agents.[126]

The inhibition of telomerase has focused on either directly targeting the catalytic core with a focus on the interaction between hTR and TERT or on targeting the interaction between telomerase and the telomeres.

One of the strategies in targeting the catalytic core of telomerase has been in targeting hTR with antisense oligonucleotides, such as was done for Imetelstat.[127] However, a drawback in oligonucleotide therapies is their low cell permeability. Therefore in the development of Imetelstat (**1.81**) the oligonucleotide consisting of 13 bases (5'-TAGGGTTAGACAA-3') was conjugated to a lipid (**Figure 1.14**). **1.81** was shown by Telomerase Repeated Amplification Protocol (TRAP) assay to lead to anti-telomerase activity in cancer cells with IC_{50} of ~150 nM.[127] Following on from this it was also shown that Imetelstat inhibited telomerase *in vivo* in a mouse model with a ~56% inhibition observed after 24 h.[127]

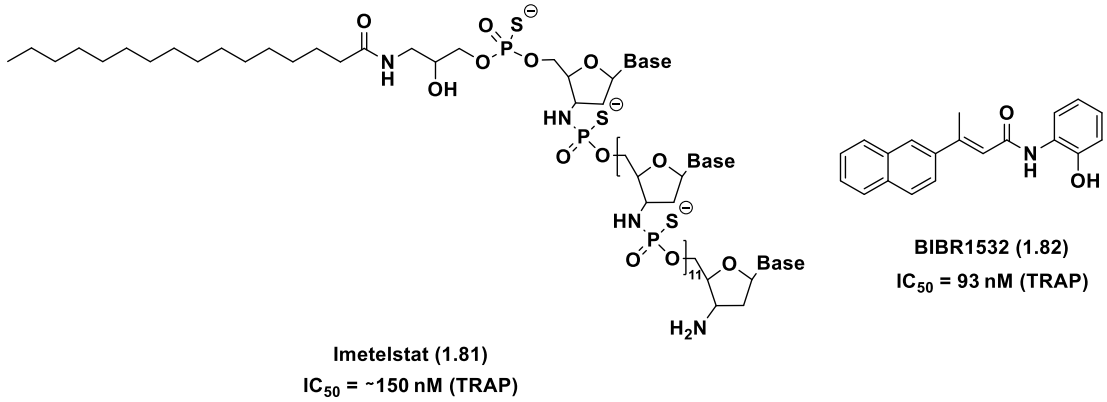


Figure 1.14: *Structures of Imetelstat (**1.81**) and BIBR1532 (**1.82**)*

BIBR1532 (**1.82**) is a small molecule (**Figure 1.14**) that also directly targets telomerase. **1.82** binds to the active site of TERT and was able to inhibit telomerase with an IC_{50} of 93 nM. It also showed selectivity for telomerase over other DNA and RNA polymerases.[128]

Alternatively for the inhibition of telomerase activity the G-quadruplexes found in telomeric DNA have been stabilised. G-quadruplexes are secondary DNA structures, which are found in guanine-rich sections of DNA, such as in telomeres (repeated TTAGGG sequence). The stabilisation of G-quadruplexes prevents telomerase from binding to them and this is therefore an attractive therapeutic strategy.[129,130] A number of G-quadruplex binders have been found with anti-telomerase activity. Telomestatin was first isolated from *Streptomyces analatus* in 2001 and is one of the most potent natural product inhibitors of telomerase.[131] Telomestatin was first explored as a racemic mixture, but a synthetic route has since been developed allowing for the synthesis of enantiopure compounds. The *S*-enantiomer (**1.83**) was found to be more active with an IC_{50} of 5 nM in a TRAP assay compared to 20 nM for the *R*-enantiomer (**1.84**) (**Figure 1.15**).[132]

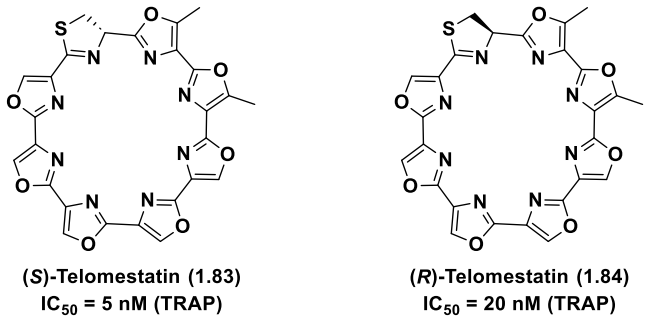


Figure 1.15: *Strucutures of (S)-telomestatin (**1.83**) and (R)-telomestatin (**1.84**) and their IC_{50} values as measured by TRAP assay*

Other G-quadruplex stabilisers with inhibitory activity against telomerase include Berberine (**1.85**) (IC_{50} = 70 μM)[133], Cryptolepine (**1.86**) (IC_{50} = 9.4 μM)[134] and Meridine (**1.87**)[135] (IC_{50} = 11 μM) (**Figure 1.16**), but none are as potent as Telomestatin.

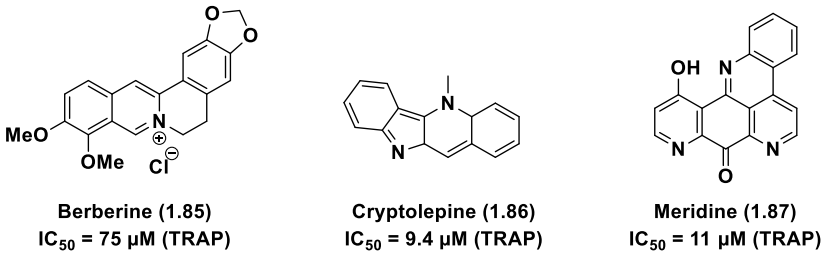


Figure 1.16: *Structures of Berberine (**1.85**), Cryptolepine (**1.86**) and Meridine (**1.87**) with their IC_{50} values as measured by TRAP assay*

Telomerase can also be targeted using nucleoside-based inhibitors, such as azidothymidine (**1.88**) through blocking the incorporation of nucleotides during telomere elongation.[136]

There are also other known telomerase inhibitors with yet to be identified mechanisms of action. MST-199 and MST-312 have been developed based on the natural product epigallocatechin gallate (EGCG, **1.89**). **1.89** showed anti-telomerase activity, but also exhibited limited stability.[137] Therefore a series of analogues was synthesised, leading to the identification of MST-199 (**1.90**) and MST-312 (**1.91**) (**Figure 1.17**). The replacement of the ester bond with an amide bond increased the stability and both showed enhanced activity in the TRAP assay compared to **1.89** (IC_{50} = 0.38 μM for **1.90** and 0.67 μM for **1.91**, compared to 1 μM for **1.89**).[138] However, their exact mechanism remains unclear.

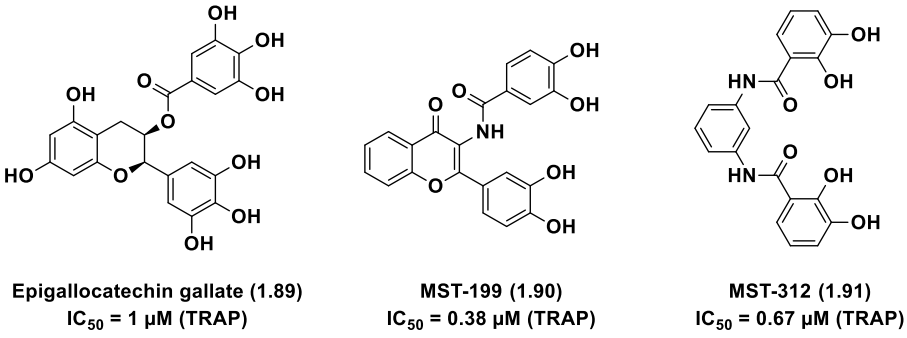


Figure 1.17: *Structures of Epigallocatechin gallate (EGCG), MST-199 and MST-312 and their IC_{50} values as measured by TRAP assay*

1.2.3 The function of dyskerin and its link to cancer

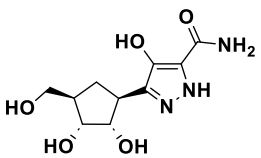
Dyskerin (also known as NAP57) is part of the telomerase H/ACA RNP interacting with the other proteins within it, but also with itself and hTR. It is not only part of the telomerase complex, but it also has another function as a component of H/ACA ribonucleoproteins.[139,140]

In this function it has a pseudouridine synthase activity on ribosomal RNAs (rRNAs), messenger RNAs (mRNAs) and small nuclear RNAs (snRNAs), which is important for ribosome biogenesis.[140,141] Around a 100-200 different H/ACA RNPs exist in mammalian cells with in each of these the two heterotetramers found in combination with a specific mRNA template.[140]

Mutations in dyskerin can cause Dyskeratosis Congenita (DKC), a rare X-linked disorder which affects highly proliferative tissues, resulting in nail dystrophy, bone marrow failure and premature aging.[142] In comparison to age-matched control cells, patients with DKC showed shorter telomeres and lower levels of telomerase activity.[143] These mutations can also lead to Hoyeraal-Hreidarsson (HH) syndrome, a severe variant of DKC, which is characterised by growth retardation, developmental delay and a short life span.[144]

The upregulation of dyskerin expression has also been associated with a number of cancers, including breast cancer, prostate cancer and lung cancer.[145–147] Although dyskerin upregulation has been associated with cancer its use as a drug target has been minimally explored. Only one study was published in 2014, which showed the catalytic activity of dyskerin could be inhibited with small molecules.[148]

A high resolution Cryo-EM structure of telomerase was only published 2021 and so before this the only option was to use a homology model of the protein. A virtual high throughput screen of such a homology model identified four potential small molecule inhibitors of the dyskerin catalytic activity targeting ribosome biogenesis.[148] One of the compounds, pyrazofurin (**1.92**) (**Figure 1.18**), was able to inhibit pseudouridylation with 77% inhibition observed at 100 µM. **1.92** also showed a significant cytotoxic effect against MCF7 cancer cells, with inhibition on cell growth seen from the 24 h time point with 50 µM compound. This shoes the promise of dyskerin as an anti-cancer drug target.[148]



Pyrazofurin (1.92)

Figure 1.18: *Structure of pyrazofurin (**1.92**)*

A different virtual screen was done focussing on the residue K314, a residue found mutated in severe cases of DKC, as the centre of the docking. Three compounds were found to inhibit telomerase activity by targeting the hTR-dyskerin interaction.[149]

1.2.4 The dyskerin-dyskerin PPI

The structure of telomerase revealed two different sections which mediate the PPI between the dyskerin proteins (**Figure 1.19**).[124]

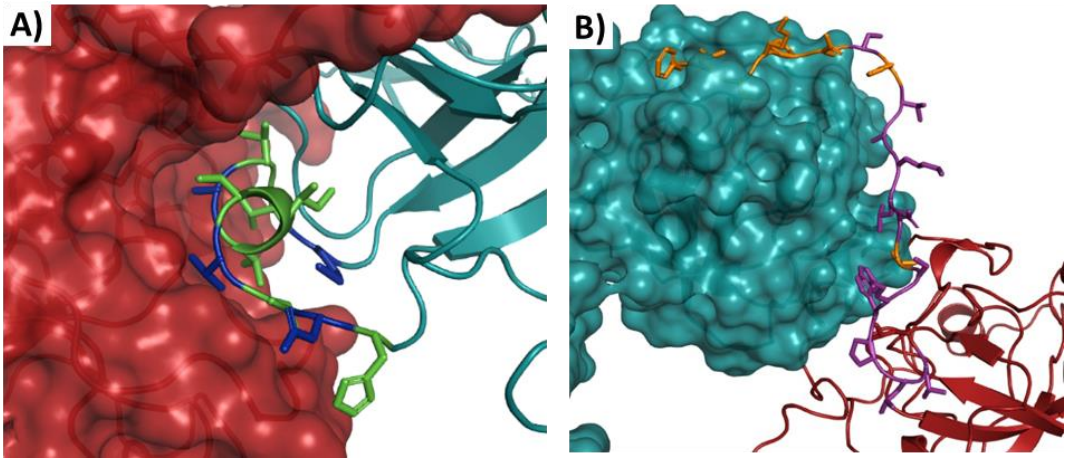


Figure 1.19: *The dyskerin-dyskerin PPI (PDB: 7BGB) with A) The interacting sequence ($_{350}$MTTAVISTCDH$_{360}$) highlighted in green with mutations found in patients with DKC and HH in dark blue. B) The interacting sequence ($_{36}$FLIKPESKVAKLDTSQWPLL$_{55}$) highlighted in purple with the mutations in orange.*

The first is formed through the section $_{350}$MTTAVISTCDH$_{360}$ as highlighted in green (**Figure 1.19A**), which forms an α-helical shape. Mutations M350T, A353V, T357A and D359N have commonly been found in dyskeratosis congenita (DKC) and Hoyeraal-Hreidarsson Syndrome (HH), conditions that suffer from insufficient telomerase. The A353V mutation is found in 40% of DKC cases, indicating this may be an important PPI within telomerase.[150] The second section of the PPI was identified as residues $_{36}$FLIKPESKVAKLDTSQWPLL$_{55}$ of the other dyskerin protein (**Figure 1.19B**). This is a longer disordered section of the protein in which DC and HH mutations are seen at F36V, L37del, I38T, K39E, P40R, E41K, K43E and T49S again indicating this may be an important interaction within telomerase.[124,150]

1.3 Thesis outline

The research in this thesis is focused on the design, synthesis and biological evaluation of peptides and small molecules targeting the dyskerin-dyskerin protein-protein interaction. In **chapter 2** the structural information available of the dyskerin-dyskerin PPI through the cryo-EM structure of telomerase is used to design peptides. The optimisation of the SPPS is discussed as well as the development of biophysical assays. The effect of both hydrocarbon and lactam stapling on the peptide is discussed.

Chapter 3 then discusses the use of an *in silico* process known as peptide-directed binding to use the structural information from the dyskerin-dyskerin PPI to design small molecules inhibitors. The synthesis of the top compounds from the *in silico* process is discussed as well as their evaluation in the assays developed in chapter 2.

The compounds synthesised in both chapters 2 and 3 are then evaluated for cell viability and cell permeability in **chapter 4** to assess their effect on cancer cells.

For the biophysical assay throughout this thesis the dyskerin protein was required and the expression process of dyskerin is described in **chapter 5**.

Finally, **chapter 6** outlines the experimental procedures, instruments and materials used throughout this thesis.

1.4 References

1 M. C. Smith and J. E. Gestwicki, Features of Protein-Protein Interactions that Translate into Potent Inhibitors: Topology, Surface Area and Affinity, *Expert. Rev. Mol. Med.*, 2012, **14**, 1–24.

2 S. Beeckmans, Chromatographic Methods to Study Protein–Protein Interactions, *Methods*, 1999, **19**, 278–305.

3 J. S. Lin and E. M. Lai, in *Bacterial Protein Secretion Systems: Methods and Protocols*, eds. L. Journet and E. Cascales, Springer, New York, NY, 2017, pp. 211–219.

4 S. Fields and O. Song, A novel genetic system to detect protein–protein interactions, *Nature*, 1989, **340**, 245–246.

5 A. Brückner, C. Polge, N. Lentze, D. Auerbach and U. Schlattner, Yeast Two-Hybrid, a Powerful Tool for Systems Biology, *Int. J. Mol. Sci.*, 2009, **10**, 2763–2788.

6 C. D. Curtis and A. M. Nardulli, in *The Nuclear Receptor Superfamily: Methods and Protocols*, ed. I. J. McEwan, Humana Press, Totowa, NJ, 2009, pp. 187–204.

7 M. Bouhaddou, M. Eckhardt, Z. Z. Chi Naing, M. Kim, T. Ideker and N. J. Krogan, Mapping the protein–protein and genetic interactions of cancer to guide precision medicine, *Curr. Opin. Genet. Dev.*, 2019, **54**, 110–117.

8 X. Benjin and L. Ling, Developments, applications, and prospects of cryo-electron microscopy, *Protein Sci.*, 2020, **29**, 872–882.

9 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, Highly accurate protein structure prediction with AlphaFold, *Nature*, 2021, **596**, 583–589.

10 J. Gong, A. Chehrazi-Raffle, S. Reddi and R. Salgia, Development of PD-1 and PD-L1 inhibitors as a form of cancer immunotherapy: A comprehensive review of registration trials and future considerations, *J. Immunother. Cancer*, 2018, **6**, 1–18.

11 P. Chames, M. Van Regenmortel, E. Weiss and D. Baty, Therapeutic antibodies: Successes, limitations and hopes for the future, *Br. J. Pharmacol.*, 2009, **157**, 220–233.

12 T. Böldicke, Therapeutic Potential of Intrabodies for Cancer Immunotherapy: Current Status and Future Directions, *Antibodies*, 2022, **11**, 49.

13 G. P. Smith, Filamentous Fusion Phage: Novel Expression Vectors That Display Cloned Antigens on the Virion Surface, *Science*, 1985, **228**, 1315–1317.

14 J. Pande, M. M. Szewczyk and A. K. Grover, Phage display: Concept, innovations, applications and future, *Biotechnol. Adv.*, 2010, **28**, 849–858.

15 In vitro virus: Bonding of mRNA bearing puromycin at the 3'-terminal end to the C-terminal end of its encoded protein on the ribosome in vitro, *FEBS Lett.*, 1997, **414**, 405–408.

16 R. W. Roberts and J. W. Szostak, RNA-peptide fusions for the in vitro selection of peptides and proteins, *Proc. Natl. Acad. Sci.*, 1997, **94**, 12297–12302.

17 F. Castillo and A. Tavassoli, in *Cyclic Peptide Design*, ed. G. Goetz, Springer, New York, NY, 2019, pp. 317–328.

18 A. Tavassoli and S. J. Benkovic, Split-intein mediated circular ligation used in the synthesis of cyclic peptide libraries in E. coli, *Nat. Protoc.*, 2007, **2**, 1126–1133.

19 K. Scheffzek, M. R. Ahmadian, W. Kabsch, L. Wiesmuller, A. Lautwein, F. Schmitz and A. Wittinghofer, The Ras-RasGAP Complex: Structural Basis for GTPase Activation and Its Loss in Oncogenic Ras Mutants | Science, *Science*, **227**, 333–338.

20 J. M. Rojas, J. L. Oliva and E. Santos, Mammalian Son of Sevenless Guanine Nucleotide Exchange Factors: Old Concepts and New Perspectives, *Genes Cancer*, 2011, **2**, 298–305.

21 A. Patgiri, K. K. Yadav, P. S. Arora and D. Bar-Sagi, An orthosteric inhibitor of the Ras-Sos interaction, *Nat. Chem. Biol.*, 2011, **7**, 585–587.

22 E. S. Leshchiner, A. Parkhitko, G. H. Bird, J. Luccarelli, J. A. Bellairs, S. Escudero, K. Opoku-Nsiah, M. Godes, N. Perrimon and L. D. Walensky, Direct inhibition of oncogenic KRAS by hydrocarbon-stapled SOS1 helices, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, 1761–1766.

23 D. Dolfini and R. Mantovani, Targeting the Y/CCAAT box in cancer: YB-1 (YBX1) or NF-Y?, *Cell Death Differ.*, 2013, **20**, 676–685.

24 S. Jeganathan, M. Wendt, S. Kiehstaller, D. Brancaccio, A. Kuepper, N. Pospiech, A. Carotenuto, E. Novellino, S. Hennig and T. N. Grossmann, Constrained Peptides with Fine-Tuned Flexibility Inhibit NF-Y Transcription Factor Assembly, *Angew Chem Int Ed*, 2019, **58**, 17351–17358.

25 E. E. Gallagher, J. M. Song, A. Menon, L. D. Mishra, A. F. Chmiel and A. L. Garner, Consideration of Binding Kinetics in the Design of Stapled Peptide Mimics of the Disordered Proteins Eukaryotic Translation Initiation Factor 4E-Binding Protein 1 and Eukaryotic Translation Initiation Factor 4G, *J. Med. Chem.*, 2019, **62**, 4967–4978.

26 P. Hart, P. Hommen, A. Noisier, A. Krzyzanowski, D. Schüler, A. T. Porfetye, M. Akbarzadeh, I. R. Vetter, H. Adihou and H. Waldmann, Structure Based Design of Bicyclic Peptide Inhibitors of RbAp48, *Angew. Chem. - Int. Ed.*, 2021, **60**, 1813–1820.

27 T. N. Grossmann, J. T.-H. Yeh, B. R. Bowman, Q. Chu, R. E. Moellering and G. L. Verdine, Inhibition of oncogenic Wnt signaling through direct targeting of β-catenin, *Proc. Natl. Acad. Sci.*, 2012, **109**, 17942.

28 K. Takada, D. Zhu, G. H. Bird, K. Sukhdeo, J.-J. Zhao, M. Mani, M. Lemieux, D. E. Carrasco, J. Ryan, D. Horst, M. Fulciniti, N. C. Munshi, W. Xu, A. L. Kung, R. A. Shivdasani, L. D. Walensky and D. R. Carrasco, Targeted Disruption of the BCL9/β-Catenin Complex Inhibits Oncogenic Wnt Signaling, *Sci. Transl. Med.*, 2012, **4**, 148ra117 LP-148ra117.

29 F. Findeisen, M. Campiglio, H. Jo, F. Abderemane-Ali, C. H. Rumpf, L. Pope, N. D. Rossen, B. E. Flucher, W. F. DeGrado and D. L. Minor, Stapled Voltage-Gated Calcium Channel (CaV) α-Interaction Domain (AID) Peptides Act As Selective Protein-Protein Interaction Inhibitors of CaV Function, *ACS Chem. Neurosci.*, 2017, **8**, 1313–1326.

30 I. De Paola, L. Pirone, M. Palmieri, N. Balasco, L. Esposito, L. Russo, D. Mazzà, L. D. Marcotullio, S. Di Gaetano, G. Malgieri, L. Vitagliano, E. Pedone and L. Zaccaro, Cullin3 - BTB interface: A novel target for stapled peptides, *PLoS ONE*, 2015, **10**, 1–21.

31 D. C. Morgan, C. Morris, A. Mahindra, C. M. Blair, G. Tejeda, I. Herbert, M. L. Turnbull, G. Lieber, B. J. Willett, N. Logan, B. Smith, A. B. Tobin, D. Bhella, G. Baillie and A. G. Jamieson, Stapled ACE2 peptidomimetics designed to target the SARS-CoV-2 spike protein do not prevent virus internalization, *Pept. Sci.*, 2021, **113**, e24217.

32 K. Nakayama, N. Nakayama, B. Davidson, J. J.-C. Sheu, N. Jinawath, A. Santillan, R. Salani, R. E. Bristow, P. J. Morin, R. J. Kurman, T.-L. Wang and I.-M. Shih, A BTB/POZ protein, NAC-1, is related to tumor recurrence and is essential for tumor growth and survival, *Proc. Natl. Acad. Sci.*, 2006, **103**, 18739–18744.

33 T. Wu, P. He, W. Wu, Y. Chen and F. Lv, Targeting oncogenic transcriptional corepressor Nac1 POZ domain with conformationally constrained peptides by cyclization and stapling, *Bioorganic Chem.*, 2018, **80**, 1–10.

34 S. Mitra, K. W. Cheng and G. B. Mills, Rab GTPases implicated in inherited and acquired disorders, *Semin. Cell Dev. Biol.*, 2011, **22**, 57–68.

35 J. Spiegel, P. M. Cromm, A. Itzen, R. S. Goody, T. N. Grossmann and H. Waldmann, Direct targeting of rab-GTPase-effector interactions, *Angew. Chem. - Int. Ed.*, 2014, **53**, 2498–2503.

36 N. Normanno, A. De Luca, C. Bianco, L. Strizzi, M. Mancino, M. R. Maiello, A. Carotenuto, G. De Feo, F. Caponigro and D. S. Salomon, Epidermal growth factor receptor (EGFR) signaling in cancer, *Gene*, 2006, **366**, 2–16.

37 T. Mizuguchi, H. Uchimura, T. Kakizawa, T. Kimura, S. Yokoyama, Y. Kiso and K. Saito, Inhibitory effect of a dimerization-arm-mimetic peptide on EGF receptor activation, *Bioorg. Med. Chem. Lett.*, 2009, **19**, 3279–3282.

38 C. Dumas, M. Ouellette, J. Tovar, M. L. Cunningham, A. H. Fairlamb, S. Tamar, M. Olivier and B. Papadopoulou, Disruption of the trypanothione reductase gene of Leishmania decreases its ability to survive oxidative stress in macrophages., *EMBO J.*, 1997, **16**, 2590–2598.

39 M. A. Toro, P. A. Sánchez-Murcia, D. Moreno, M. Ruiz-Santaquiteria, J. F. Alzate, A. Negri, M. J. Camarasa, F. Gago, S. Velázquez and A. Jiménez-Ruiz, Probing the Dimerization Interface of Leishmania infantum Trypanothione Reductase with Site-Directed Mutagenesis and Short Peptides, *ChemBioChem*, 2013, **14**, 1212–1217.

40 J. H. Trembley, G. Wang, G. Unger, J. Slaton and K. Ahmed, Protein Kinase CK2 in Health and Disease, *Cell. Mol. Life Sci.*, 2009, **66**, 1858–1867.

41 B. Laudet, C. Barette, V. Dulery, O. Renaudet, P. Dumy, A. Metz, R. Prudent, A. Deshiere, O. Dideberg, O. Filhol and C. Cochet, Structure-based design of small peptide inhibitors of protein kinase CK2 subunit interaction, *Biochem. J.*, 2007, **408**, 363–373.

42 J. Raaf, B. Guerra, I. Neundorf, B. Bopp, O.-G. Issinger, J. Jose, M. Pietsch and K. Niefind, First Structure of Protein Kinase CK2 Catalytic Subunit with an Effective CK2β-Competitive Ligand, *ACS Chem. Biol.*, 2013, **8**, 901–907.

43 M. F. White, The IRS-1 signaling system, *Curr. Opin. Genet. Dev.*, 1994, **4**, 47–54.

44 Y. Hao, C. Wang, B. Cao, B. M. Hirsch, J. Song, S. D. Markowitz, R. M. Ewing, D. Sedwick, L. Liu, W. Zheng and Z. Wang, Gain of Interaction with IRS1 by p110α-Helical Domain Mutants Is Crucial for Their Oncogenic Functions, *Cancer Cell*, 2013, **23**, 583–593.

45 J. R. Graff, B. W. Konicek, J. H. Carter and E. G. Marcusson, Targeting the Eukaryotic Translation Initiation Factor 4E for Cancer Therapy, *Cancer Res.*, 2008, **68**, 631–634.

46 D. Rousseau, A. C. Gingras, A. Pause and N. Sonenberg, The eIF4E-binding proteins 1 and 2 are negative regulators of cell growth, *Oncogene*, 1996, **13**, 2415–2420.

47 J. Marcotrigiano, A.-C. Gingras, N. Sonenberg and S. K. Burley, Cap-Dependent Translation Initiation in Eukaryotes Is Regulated by a Molecular Mimic of eIF4G, *Mol. Cell*, 1999, **3**, 707–716.

48 L. C. Szymczak, H.-Y. Kuo and M. Mrksich, Peptide Arrays: Development and Application, *Anal. Chem.*, 2018, **90**, 266–282.

49 H. Amartely, A. Iosub-Amir and A. Friedler, Identifying Protein-protein Interaction Sites Using Peptide Arrays, *JoVE J. Vis. Exp.*, 2014, e52097.

50 A. Mehle, B. Strack, P. Ancuta, C. Zhang, M. McPike and D. Gabuzda, Vif Overcomes the Innate Antiviral Activity of APOBEC3G by Promoting Its Degradation in the Ubiquitin-Proteasome Pathway, *J. Biol. Chem.*, 2004, **279**, 7792–7798.

51 T. H. Reingewertz, E. Britan-Rosich, S. Rotem-Bamberger, M. Viard, A. Jacobs, A. Miller, J. Y. Lee, J. Hwang, R. Blumenthal, M. Kotler and A. Friedler, Mapping the Vif–A3G interaction using peptide arrays: A basis for anti-HIV lead peptides, *Bioorg. Med. Chem.*, 2013, **21**, 3523–3532.

52 C. Katz, Y. Zaltsman-Amir, Y. Mostizky, N. Kollet, A. Gross and A. Friedler, Molecular Basis of the Interaction between Proapoptotic Truncated BID (tBID) Protein and Mitochondrial Carrier Homologue 2 (MTCH2) Protein, *J. Biol. Chem.*, 2012, **287**, 15016–15023.

53 R. Craigie, The molecular biology of HIV integrase, *Future Virol.*, 2012, **7**, 679–686.

54 H. Y. Li, Z. Zawahir, L. D. Song, Y. Q. Long and N. Neamati, Sequence-based design and discovery of peptide inhibitors of HIV-1 integrase: Insight into the binding mode of the enzyme, *J. Med. Chem.*, 2006, **49**, 4477–4486.

55 G. Neumayer, C. Belzil, O. J. Gruss and M. D. Nguyen, TPX2: of spindle assembly, DNA damage response, and cancer, *Cell. Mol. Life Sci.*, 2014, **71**, 3027–3047.

56 R. Bayliss, T. Sardon, I. Vernos and E. Conti, Structural Basis of Aurora-A Activation by TPX2 at the Mitotic Spindle, *Mol. Cell*, 2003, **12**, 851–862.

57 D. A. Johnson, P. Akamine, E. Radzio-Andzelm, Madhusudan and S. S. Taylor, Dynamics of cAMP-Dependent Protein Kinase, *Chem. Rev.*, 2001, **101**, 2243–2270.

58 P. Skroblin, S. Grossmann, G. Schäfer, W. Rosenthal and E. Klussmann, in *International Review of Cell and Molecular Biology*, ed. K. Jeon, Academic Press, 2010, vol. 283, pp. 235–330.

59 L. L. Burns-Hamuro, Y. Hamuro, J. S. Kim, P. Sigala, R. Fayos, D. D. Stranz, P. A. Jennings, S. S. Taylor and V. L. Woods Jr., Distinct interaction modes of an AKAP bound to two regulatory subunit isoforms of protein kinase A revealed by amide hydrogen/deuterium exchange, *Protein Sci.*, 2005, **14**, 2982–2992.

60 Y. Wang, T. G. Ho, D. Bertinetti, M. Neddermann, E. Franz, G. C. H. Mo, L. P. Schendowich, A. Sukhu, R. C. Spelts, J. Zhang, F. W. Herberg and E. J. Kennedy, Isoform-selective disruption of AKAP-Localized PKA using hydrocarbon stapled peptides, *ACS Chem. Biol.*, 2014, **9**, 1386.

61 K. Taguchi and M. Yamamoto, The KEAP1–NRF2 System in Cancer, *Front. Oncol.*

62 K. I. Tong, Y. Katoh, H. Kusunoki, K. Itoh, T. Tanaka and M. Yamamoto, Keap1 Recruits Neh2 through Binding to ETGE and DLG Motifs: Characterization of the Two-Site Molecular Recognition Model, *Mol. Cell. Biol.*, 2006, **26**, 2887–2900.

63 S.-C. Lo, X. Li, M. T. Henzl, L. J. Beamer and M. Hannink, Structure of the Keap1:Nrf2 interface provides mechanistic insight into Nrf2 signaling, *EMBO J.*, 2006, **25**, 3605–3617.

64 D. Chakravortty and M. Hensel, Inducible nitric oxide synthase and control of intracellular bacterial pathogens, *Microbes Infect.*, 2003, **5**, 621–627.

65 B. K. Yap, E. W. W. Leung, H. Yagi, C. A. Galea, S. Chhabra, D. K. Chalmers, S. E. Nicholson, P. E. Thompson and R. S. Norton, A potent cyclic peptide targeting SPSB2 protein as a potential anti-infective agent, *J. Med. Chem.*, 2014, **57**, 7006–7015.

66 Z. Kuang, R. S. Lewis, J. M. Curtis, Y. Zhan, B. M. Saunders, J. J. Babon, T. B. Kolesnik, A. Low, S. L. Masters, T. A. Willson, L. Kedzierski, S. Yao, E. Handman, R. S. Norton and S. E. Nicholson, The SPRY domain-containing SOCS box protein SPSB2 targets iNOS for proteasomal degradation, *J. Cell Biol.*, 2010, **190**, 129–141.

67 A. T. Nguyen and Y. Zhang, The diverse functions of Dot1 and H3K79 methylation, *Genes Dev.*, 2011, **25**, 1345–1358.

68 C. Shen, S. Y. Jo, C. Liao, J. L. Hess and Z. Nikolovska-Coleska, Targeting recruitment of disruptor of telomeric silencing 1-like (DOT1L): Characterizing the interactions between dot1l and mixed lineage leukemia (MLL) fusion proteins, *J. Biol. Chem.*, 2013, **288**, 30585–30596.

69 Y. Xu, M. Huang, W. He, C. He, K. Chen, J. Hou, M. Huang, Y. Jiao, R. Liu, N. Zou, L. Liu and C. Li, Heterotopic Ossification: Clinical Features, Basic Researches, and Mechanical Stimulations, *Front. Cell Dev. Biol.*

70 D. Chen, S. Liu, W. Zhang and L. Sun, Rational design of YAP WW1 domain-binding peptides to target TGFβ/BMP/Smad-YAP interaction in heterotopic ossification, *J. Pept. Sci.*, 2015, **21**, 826–832.

71 H. Jo, N. Meinhardt, Y. Wu, S. Kulkarni, X. Hu, K. E. Low, P. L. Davies, W. F. Degrado and D. C. Greenbaum, Development of α-helical calpain probes by mimicking a natural protein-protein interaction, *J. Am. Chem. Soc.*, 2012, **134**, 17704–17713.

72 R. Hancock, H. C. Bertrand, T. Tsujita, S. Naz, A. El-Bakry, J. Laoruchupong, J. D. Hayes and G. Wells, Peptide inhibitors of the Keap1-Nrf2 protein-protein interaction, *Free Radic. Biol. Med.*, 2012, **52**, 444–451.

73 L. Chen, S. W. Chan, X. Q. Zhang, M. Walsh, C. J. Lim, W. Hong and H. Song, Structural basis of YAP recognition by TEAD4 in the Hippo pathway, *Genes Dev.*, 2010, **24**, 290–300.

74 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, Structure-based design and synthesis of potent cyclic peptides inhibiting the YAP-TEAD protein-protein interaction, *ACS Med. Chem. Lett.*, 2014, **5**, 993–998.

75 Y. Cai, V. Kandula, R. Kosuru, X. Ye, M. G. Irwin and Z. Xia, Decoding telomere protein Rap1: Its telomeric and nontelomeric functions and potential implications in diabetic cardiomyopathy, *Cell Cycle*, 2017, **16**, 1765–1773.

76 R. Rai, Y. Chen, M. Lei and S. Chang, TRF2-RAP1 is required to protect telomeres from engaging in homologous recombination-mediated deletions and fusions, *Nat. Commun.*, 2016, **7**, 10881.

77 X. Ran, L. Liu, C. Y. Yang, J. Lu, Y. Chen, M. Lei and S. Wang, Design of High-Affinity Stapled Peptides to Target the Repressor Activator Protein 1 (RAP1)/Telomeric Repeat-Binding Factor 2 (TRF2) Protein-Protein Interaction in the Shelterin Complex, *J. Med. Chem.*, 2016, **59**, 328–334.

78 E. C. C. Cheung, L. Melanson-Drapeau, S. P. Cregan, J. L. Vanderluit, K. L. Ferguson, W. C. McIntosh, D. S. Park, S. A. L. Bennett and R. S. Slack, Apoptosis-Inducing Factor Is a Key Factor in Neuronal Cell Death Propagated by BAX-Dependent and BAX-Independent Mechanisms, *J. Neurosci.*, 2005, **25**, 1324–1334.

79 C. Candé, N. Vahsen, I. Kouranti, E. Schmitt, E. Daugas, C. Spahr, J. Luban, R. T. Kroemer, F. Giordanetto, C. Garrido, J. M. Penninger and G. Kroemer, AIF and cyclophilin A cooperate in apoptosis-associated chromatinolysis, *Oncogene*, 2004, **23**, 1514–1521.

80 N. Doti, C. Reuther, P. L. Scognamiglio, A. M. Dolga, N. Plesnila, M. Ruvo and C. Culmsee, Inhibition of the AIF/CypA complex protects against intrinsic death pathways induced by oxidative stress, *Cell Death Dis.*, 2014, **5**, 1–13.

81 B. Farina, M. Sturlese, F. Mascanzoni, A. Caporale, A. Monti, G. Di Sorbo, R. Fattorusso, M. Ruvo and N. Doti, Binding mode of AIF(370–394) peptide to CypA: insights from NMR, label-free and molecular docking studies, *Biochem. J.*, 2018, **475**, 2377–2393.

82 L. Russo, F. Mascanzoni, B. Farina, A. M. Dolga, A. Monti, A. Caporale, C. Culmsee, R. Fattorusso, M. Ruvo and N. Doti, Design, Optimization, and Structural Characterization of

an Apoptosis-Inducing Factor Peptide Targeting Human Cyclophilin A to Inhibit Apoptosis Inducing Factor-Mediated Cell Death, *J. Med. Chem.*, 2021, **64**, 11445–11459.

83 Y. Chen, D. Inoyama, A. N. T. Kong, L. J. Beamer and L. Hu, Kinetic analyses of Keap1-Nrf2 interaction and determination of the minimal Nrf2 peptide sequence required for Keap1 binding using surface plasmon resonance, *Chem. Biol. Drug Des.*, 2011, **78**, 1014–1021.

84 K. I. Tong, B. Padmanabhan, A. Kobayashi, C. Shang, Y. Hirotsu, S. Yokoyama and M. Yamamoto, Different Electrostatic Potentials Define ETGE and DLG Motifs as Hinge and Latch in Oxidative Stress Response, *Mol. Cell. Biol.*, 2007, **27**, 7511–7521.

85 S. A. Kawamoto, A. D. Thompson, A. Coleska, Z. Nikolovska-Coleska, H. Yi and S. Wang, Analysis of the interaction of BCL9 with β-catenin and development of fluorescence polarization and surface plasmon resonance binding assays for this interaction, *Biochemistry*, 2009, **48**, 9534–9541.

86 J. Hochscherf, D. Lindenblatt, M. Steinkrüger, E. Yoo, Ö. Ulucan, S. Herzig, O. G. Issinger, V. Helms, C. Götz, I. Neundorf, K. Niefind and M. Pietsch, Development of a high-throughput screening-compatible assay to identify inhibitors of the CK2α/CK2β interaction, *Anal. Biochem.*, 2015, **468**, 4–14.

87 H. L. Olsvik, T. Lamark, K. Takagi, K. B. Larsen, G. Evjen, A. Øvervatn, T. Mizushima and X. T. Johansen, FYCO1 contains a C-terminally extended, LC3A/B-preferring LC3-interacting region (LIR) motif required for efficient maturation of autophagosomes during basal autophagy, *J. Biol. Chem.*, 2015, **290**, 29361–29374.

88 R. A. Cerulli, L. Shehaj, H. Brown, J. Pace, Y. Mei and J. A. Kritzer, Stapled Peptide Inhibitors of Autophagy Adapter LC3B, *ChemBioChem*, 2020, **21**, 2777–2785.

89 J. A. Miles, D. J. Yeo, P. Rowell, S. Rodriguez-marin, C. M. Pask, S. L. Warriner, A. T. Edwards and A. J. Wilson, Hydrocarbon constrained peptides – understanding preorganisation and binding affini, *Chem Sci*, 2016, **7**, 3694–3702.

90 A. D. De Araujo, H. N. Hoang, W. M. Kok, F. Diness, P. Gupta, T. A. Hill, R. W. Driver, D. A. Price, S. Liras and D. P. Fairlie, Comparative α-helicity of cyclic pentapeptides in water, *Angew. Chem. - Int. Ed.*, 2014, **53**, 6965–6969.

91 Q. Chu, R. E. Moellering, G. J. Hilinski, Y.-W. Kim, T. N. Grossmann, J. T.-H. Yeh and G. L. Verdine, Towards understanding cell penetration by stapled peptides, *MedChemComm*, 2015, **6**, 111–119.

92 C. E. Schafmeister, J. Po and G. L. Verdine, An All-Hydrocarbon Cross-Linking System for Enhancing the Helicity and Metabolic Stability of Peptides, *J Am Chem Soc*, 2000, **122**, 5891–5892.

93 L. D. Walensky, A. L. Kung, I. Escher, T. J. Malia, S. Barbuto, R. D. Wright, G. Wagner, G. L. Verdine and S. J. Korsmeyer, Activation of apoptosis in vivo by a hydrocarbon-stapled BH3 helix, *Science*, 2004, **305**, 1466–1470.

94 L. D. Walensky, K. Pitter, J. Morash, K. J. Oh, S. Barbuto, J. Fisher, E. Smith, G. L. Verdine and S. J. Korsmeyer, A Stapled BID BH3 Helix Directly Binds and Activates BAX, *Mol. Cell*, 2006, **24**, 199–210.

95 P. Polakis, Wnt signaling and cancer, *Genes Dev.*, 2000, **14**, 1837–1851.

96 M. Mani, D. E. Carrasco, Y. Zhang, K. Takada, M. E. Gatt, J. Dutta-Simmons, H. Ikeda, F. Diaz-Griffero, V. Pena-Cruz, M. Bertagnolli, L. L. Myeroff, S. D. Markowitz, K. C. Anderson and D. R. Carrasco, BCL9 Promotes Tumor Progression by Conferring Enhanced Proliferative, Metastatic, and Angiogenic Properties to Cancer Cells, *Cancer Res.*, 2009, **69**, 7577–7586.

97 S. A. Kawamoto, A. D. Thompson, A. Coleska, Z. Nikolovska-Coleska, H. Yi and S. Wang, Analysis of the interaction of BCL9 with β-catenin and development of fluorescence polarization and surface plasmon resonance binding assays for this interaction, *Biochemistry*, 2009, **48**, 9534–9541.

98 K. Takada, D. Zhu, G. H. Bird, K. Sukhdeo, J. Zhao, M. Lemieux, D. E. Carrasco, J. Ryan, D. Horst, N. C. Munshi, W. Xu, A. L. Kung, R. A. Shivdasani, L. D. Walensky and D. R. Carrasco, Targeted Disruption of the BCL9/β-catenin Complex Inhibits Oncogenic Wnt Signaling, *Sci. Transl. Med.*, 2012, **4**, 117–136.

99 L. Chen, S. N. Willis, A. Wei, B. J. Smith, J. I. Fletcher, M. G. Hinds, P. M. Colman, C. L. Day, J. M. Adams and D. C. S. Huang, Differential Targeting of Prosurvival Bcl-2 Proteins by Their BH3-Only Ligands Allows Complementary Apoptotic Function, *Mol. Cell*, 2005, **17**, 393–403.

100 M. Certo, V. D. G. Moore, M. Nishino, G. Wei, S. Korsmeyer, S. A. Armstrong and A. Letai, Mitochondria primed by death signals determine cellular addiction to antiapoptotic BCL-2 family members, *Cancer Cell*, 2006, **9**, 351–365.

101 J. L. Labelle, A. L. Kung, L. D. Walensky, S. G. Katz, G. H. Bird, E. Gavathiotis, M. L. Stewart, C. Lawrence, J. K. Fisher, M. Godes and K. Pitter, A stapled BIM peptide overcomes apoptotic resistance in hematologic cancers, *J. Clin. Invest.*, 2012, **122**, 2018–2031.

102 P. A. Sánchez-Murcia, M. Ruiz-Santaquiteria, M. A. Toro, H. De Lucio, M. Á. Jiménez, F. Gago, A. Jiménez-Ruiz, M. J. Camarasa and S. Velázquez, Comparison of hydrocarbon-and lactam-bridged cyclic peptides as dimerization inhibitors of Leishmania infantum trypanothione reductase, *RSC Adv.*, 2015, **5**, 55784–55794.

103 M. Ruiz-Santaquiteria, P. A. Sánchez-Murcia, M. A. Toro, H. de Lucio, K. J. Gutiérrez, S. de Castro, F. A. C. Carneiro, F. Gago, A. Jiménez-Ruiz, M. J. Camarasa and S. Velázquez, First example of peptides targeting the dimer interface of Leishmania infantum trypanothione reductase with potent in vitro antileishmanial activity#, *Eur. J. Med. Chem.*, 2017, **135**, 49–59.

104 M. Ruiz-Santaquiteria, S. de Castro, M. A. Toro, H. de Lucio, K. J. Gutiérrez, P. A. Sánchez-Murcia, M. Á. Jiménez, F. Gago, A. Jiménez-Ruiz, M. J. Camarasa and S. Velázquez, Trypanothione reductase inhibition and anti-leishmanial activity of all-hydrocarbon stapled α-helical peptides with improved proteolytic stability, *Eur. J. Med. Chem.*, 2018, **149**, 238–247.

105 K. Suzuki, Y. Takaoka and M. Ueda, Rational design of a stapled JAZ9 peptide inhibiting protein-protein interaction of a plant transcription factor, *RSC Chem. Biol.*, 2021, **2**, 499–502.

106 W. Tang, C. Wang, Y. Zhou, J. Luo, T. Ye and B. Yang, Hydrocarbon-stapling stabilization of the reduced homodimerization interaction of hepatic cancer DAP12 transmembrane domain in water phase, *Chem. Pap.*, 2020, **74**, 2153–2161.

107 Y. Lai, G. Fois, J. R. Flores, M. J. Tuvim, Q. Zhou, K. Yang, J. Leitz, J. Peters, Y. Zhang, R. A. Pfuetzner, L. Esquivies, P. Jones, M. Frick, B. F. Dickey and A. T. Brunger, Inhibition of calcium-triggered secretion by hydrocarbon-stapled peptides, *Nature*, 2022, **603**, 949–956.

108 H. Li, X. Chen, M. Wu, P. Song and X. Zhao, Bicyclic stapled peptides based on p53 as dual inhibitors for the interactions of p53 with MDM2 and MDMX, *Chin. Chem. Lett.*, 2022, **33**, 1254–1258.

109 L. E. Hanold, K. Oruganty, N. T. Ton, A. M. Beedle, N. Kannan and E. J. Kennedy, Inhibiting EGFR dimerization using triazolyl-bridged dimerization arm mimics, *PLoS ONE*, 2015, **10**, 1–18.

110 S. Jiao, H. Wang, Z. Shi, A. Dong, W. Zhang, X. Song, F. He, Y. Wang, Z. Zhang, W. Wang, X. Wang, T. Guo, P. Li, Y. Zhao, H. Ji, L. Zhang and Z. Zhou, A Peptide Mimicking VGLL4 Function Acts as a YAP Antagonist Therapy against Gastric Cancer, *Cancer Cell*, 2014, **25**, 166–180.

111 W. Zheng, J. Lan, L. Feng, Z. Chen, S. Feng, Y. Gao, F. Ren and Y. Chen, Structure-Based Optimization of Conformationally Constrained Peptides to Target Esophageal Cancer TEAD Transcription Factor, *Int. J. Pept. Res. Ther.*, 2021, **27**, 923–930.

112 J. Iegre, P. Brear, D. J. Baker, Y. S. Tan, E. L. Atkinson, H. F. Sore, D. H. O'Donovan, C. S. Verma, M. Hyvönen and D. R. Spring, Efficient development of stable and highly functionalised peptides targeting the CK2α/CK2β protein-protein interaction, *Chem. Sci.*, 2019, **10**, 5056–5063.

113 A. Monti, M. Sturlese, A. Caporale, J. D. A. Roger, F. Mascanzoni, M. Ruvo and N. Doti, Design, synthesis, structural analysis and biochemical studies of stapled AIF(370-394) analogues as ligand of CypA, *Biochim. Biophys. Acta - Gen. Subj.*, 2020, **1864**, 129717.

114 M. Lindgren, M. Hällbrink, A. Prochiantz, Ü. Langel, M. Lindgren, M. Hällbrink, A. Prochiantz and Ü. Langel, Cell-penetrating peptides, *Trends Pharmacol. Sci.*, 2000, **21**, 99–103.

115 F. Heitz, M. C. Morris and G. Divita, Twenty years of cell-penetrating peptides: from molecular mechanisms to therapeutics, *Br. J. Pharmacol.*, 2009, **157**, 195–206.

116 D. M. Copolovici, K. Langel, E. Eriste and Ü. Langel, Cell-Penetrating Peptides: Design, Synthesis, and Applications, *ACS Nano*, 2014, **8**, 1972–1994.

117 G. Guidotti, L. Brambilla and D. Rossi, Cell-Penetrating Peptides: From Basic Research to Clinics, *Trends Pharmacol. Sci.*, 2017, **38**, 406–424.

118 P. M. Cromm, J. Spiegel, P. Küchler, L. Dietrich, J. Kriegesmann, M. Wendt, R. S. Goody, H. Waldmann and T. N. Grossmann, Protease-Resistant and Cell-Permeable Double-Stapled Peptides Targeting the Rab8a GTPase, *ACS Chem. Biol.*, 2016, **11**, 2375–2382.

119 Y.-S. Cong, W. E. Wright and J. W. Shay, Human Telomerase and Its Regulation, *Microbiol. Mol. Biol. Rev.*, 2002, **66**, 407–425.

120 E. Gilson and V. Géli, How telomeres are replicated, *Nat. Rev. Mol. Cell Biol.*, 2007, **8**, 825–838.

121 S. P. Jackson and J. Bartek, The DNA-damage response in human biology and disease, *Nature*, 2009, **461**, 1071–1078.

122 N. W. Kim, M. A. Piatyszek, K. R. Prowse, C. B. Harley, M. D. West, P. L. Ho, G. M. Coviello, W. E. Wright, S. L. Weinrich and J. W. Shay, Specific association of human telomerase activity with immortal cells and cancer, *Science*, 1994, **266**, 2011 LP – 2015.

123 T. H. D. Nguyen, J. Tam, R. A. Wu, B. J. Greber, D. Toso, E. Nogales and K. Collins, Cryo-EM structure of substrate-bound human telomerase holoenzyme, *Nature*, 2018, **557**, 190–198.

124 G. E. Ghanim, A. J. Fountain, A. M. M. van Roon, R. Rangan, R. Das, K. Collins and T. H. D. Nguyen, Structure of human telomerase holoenzyme with bound telomeric DNA, *Nature*, 2021, **593**, 449–453.

125 S. Bajaj, M. S. Kumar, G. J. Peters and Y. C. Mayur, Targeting telomerase for its advent in cancer therapeutics, *Med. Res. Rev.*, 2020, **40**, 1871–1919.

126 C. B. Harley, Telomerase and cancer therapeutics, *Nat. Rev. Cancer*, 2008, **8**, 167–179.

127 B.-S. Herbert, G. C. Gellert, A. Hochreiter, K. Pongracz, W. E. Wright, D. Zielinska, A. C. Chin, C. B. Harley, J. W. Shay and S. M. Gryaznov, Lipid modification of GRN163, an N3' → P5' thio-phosphoramidate oligonucleotide, enhances the potency of telomerase inhibition, *Oncogene*, 2005, **24**, 5262–5268.

128 K. Damm, U. Hemmann, P. Garin-Chesa, N. Hauel, I. Kauffmann, H. Priepke, C. Niestroj, C. Daiber, B. Enenkel, B. Guilliard, I. Lauritsch, E. Müller, E. Pascolo, G. Sauter, M. Pantic, U. M. Martens, C. Wenz, J. Lingner, N. Kraut, W. J. Rettig and A. Schnapp, A highly selective telomerase inhibitor limiting human cancer cell proliferation, *EMBO J.*, 2001, **20**, 6958–6968.

129 R. J. Harrison, A. P. Reszka, S. M. Haider, B. Romagnoli, J. Morrell, M. A. Read, S. M. Gowan, C. M. Incles, L. R. Kelland and S. Neidle, Evaluation of by disubstituted acridone derivatives as telomerase inhibitors: The importance of G-quadruplex binding, *Bioorg. Med. Chem. Lett.*, 2004, **14**, 5845–5849.

130 M. Read, R. J. Harrison, B. Romagnoli, F. A. Tanious, S. H. Gowan, A. P. Reszka, W. D. Wilson, L. R. Kelland and S. Neidle, Structure-based design of selective and potent G quadruplex-mediated telomerase inhibitors, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 4844–4849.

131 K. Shin-ya, K. Wierzba, K. Matsuo, T. Ohtani, Y. Yamada, K. Furihata, Y. Hayakawa and H. Seto, Telomestatin, a Novel Telomerase Inhibitor from Streptomyces anulatus, *J. Am. Chem. Soc.*, 2001, **123**, 1262–1263.

132 T. Doi, K. Shibata, M. Yoshida, M. Takagi, M. Tera, K. Nagasawa, K. Shin-ya and T. Takahashi, (S)-Stereoisomer of telomestatin as a potent G-quadruplex binder and telomerase inhibitor, *Org. Biomol. Chem.*, 2010, **9**, 387–393.

133 D. Yun, S. Y. Yoon, S. J. Park and Y. J. Park, The Anticancer Effect of Natural Plant Alkaloid Isoquinolines, *Int. J. Mol. Sci.*, 2021, **22**, 1653.

134 L. Guittat, P. Alberti, F. Rosu, S. Van Miert, E. Thetiot, L. Pieters, V. Gabelica, E. De Pauw, A. Ottaviani, J.-F. Riou and J.-L. Mergny, Interactions of cryptolepine and neocryptolepine with unusual DNA structures, *Biochimie*, 2003, **85**, 535–547.

135 L. Guittat, A. De Cian, F. Rosu, V. Gabelica, E. De Pauw, E. Delfourne and J.-L. Mergny, Ascididemin and meridine stabilise G-quadruplexes and inhibit telomerase in vitro, *Biochim. Biophys. Acta BBA - Gen. Subj.*, 2005, **1724**, 375–384.

136 D. E. Gomez, R. G. Armando and D. F. Alonso, AZT as a telomerase inhibitor, *Front. Oncol.*, 2012, **2**, 1–5.

137 I. Naasani, H. Seimiya and T. Tsuruo, Telomerase Inhibition, Telomere Shortening, and Senescence of Cancer Cells by Tea Catechins, *Biochem. Biophys. Res. Commun.*, 1998, **249**, 391–396.

138 H. Seimiya, T. Oh-hara, T. Suzuki, I. Naasani, T. Shimazaki, K. Tsuchiya and T. Tsuruo, Telomere shortening and growth inhibition of human cancer cells by novel synthetic telomerase inhibitors MST-312, MST-295, and MST-1991, *Mol. Cancer Ther.*, 2002, **1**, 657–665.

139 U. T. Meier and G. Blobel, NAP57, a mammalian nucleolar protein with a putative homolog in yeast and bacteria., *J. Cell Biol.*, 1994, **127**, 1505–1514.

140 U. T. Meier, The many facets of H/ACA ribonucleoproteins, *Chromosoma*, 2005, **114**, 1–14.

141 A. Angrisani, M. Turano and M. Furia, Human dyskerin : beyond telomeres, *Biol Chem*, 2014, **395**, 593–610.

142 S. A. Savage and B. P. Alter, Dyskeratosis Congenita, *Hematol. Oncol. Clin. North Am.*, 2009, **23**, 215–231.

143 J. R. Mitchell, E. Wood and K. Collins, A telomerase component is defective in the human disease dyskeratosis congenita, *Nature*, 1999, **402**, 551–555.

144 G. Glousker, F. Touzot, P. Revy, Y. Tzfati and S. A. Savage, Unraveling the pathogenesis of Hoyeraal–Hreidarsson syndrome, a complex telomere biology disorder, *Br. J. Haematol.*, 2015, **170**, 457–471.

145 K. A. Elsharawy, O. J. Mohammed, M. A. Aleskandarany, A. Hyder, H. L. El-gammal, M. I. Abou-dobara, A. R. Green, L. W. Dalton and E. A. Rakha, The nucleolar-related protein Dyskerin pseudouridine synthase 1 (DKC1) predicts poor prognosis in breast cancer, *Br. J. Cancer*, 2020, **1**, 1–10.

146 P. Sieron, C. Hader, J. Hatina, R. Engers, A. Wlazlinski, M. Müller and W. A. Schulz, DKC1 overexpression associated with prostate cancer progression, *Br. J. Cancer*, 2009, **101**, 1410–1416.

147 I. Fernandez-Garcia, T. Marcos, A. Muñoz-Barrutia, D. Serrano, R. Pio, L. M. Montuenga and C. Ortiz-de-Solorzano, Multiscale in situ analysis of the role of dyskerin in lung cancer cells, *Integr. Biol.*, 2013, **5**, 402–413.

148 J. M. Barbosa, C. Onofrillo, A. Del Rio, L. Rocchi and L. Montanaro, Inhibition of Human Dyskerin as a New Approach to Target Ribosome Biogenesis, *PLoS ONE*, 2014, **9**, 1–10.

149 R. G. Armando, D. Luis, G. Mengual, E. I. Juritz, P. L. Menna and D. E. Gomez, Homology Model and Docking-Based Virtual Screening for Ligands of Human Dyskerin as New Inhibitors of Telomerase for Cancer Treatment, *Int. J. Mol. Sci.*, 2018, **19**, 3216–3232.

150 G. Sarek, P. Marzec, P. Margalef and S. J. Boulton, Molecular basis of telomere dysfunction in human genetic diseases, *Nat. Struct. Mol. Biol.*, 2015, **22**, 867–874.

Chapter 2

The design, synthesis and *in vitro* biological evaluation of peptides targeting the dyskerin-dyskerin protein-protein interaction

2.1 Introduction

2.1.1 The principles of solid-phase peptide synthesis

Peptides are good starting points for the design of inhibitors for PPIs due to their selectivity and low cost. As discussed in **Chapter 1**, they can easily be derived from the protein partner. After an initial sequence has been selected, the peptide can be further optimised for affinity and cell permeability.

Solid-phase peptide synthesis (SPPS) was developed by Merrifield in 1963, greatly simplifying the synthesis of peptides.[1] The technique relies on the use of a solid support, also known as a resin, on which the peptide is built through consecutive coupling and deprotection steps (**Figure 2.1**). In between each step any reagents are washed away with only the growing peptide chain remaining on the resin. This allows for the use of excess reagents in each step of the synthesis driving the reaction to completion.

The most common type of SPPS is Fmoc SPPS which uses Fmoc-protected amino acids that can be deprotected using a base (i.e. piperidine or piperazine).[2] The final peptide is then cleaved from the resin using trifluoroacetic acid (TFA). SPPS can also be used with Boc-protected amino acids, but Boc-protecting groups are removed using TFA and therefore the resin on which the peptide is built cannot be TFA cleavable. Instead hydrofluoric acid (HF) is used for the cleavage of these peptides, which is a highly toxic reagent and therefore Fmoc SPPS is often preferred over Boc SPPS.[3]

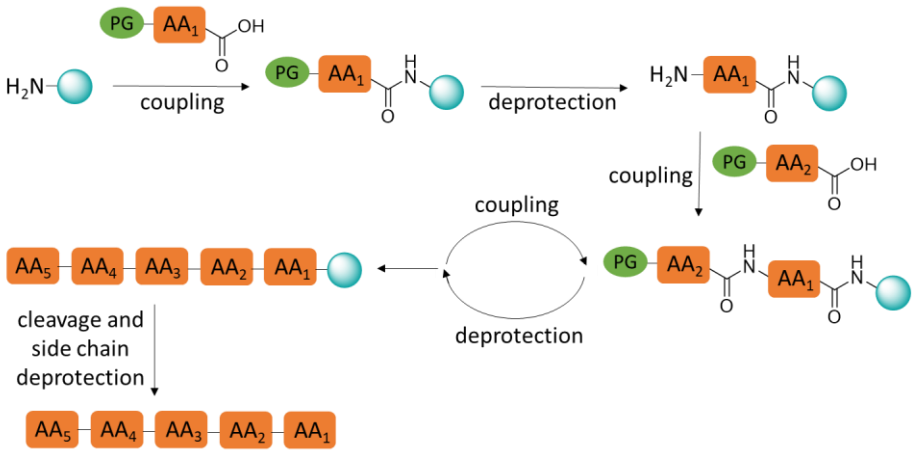
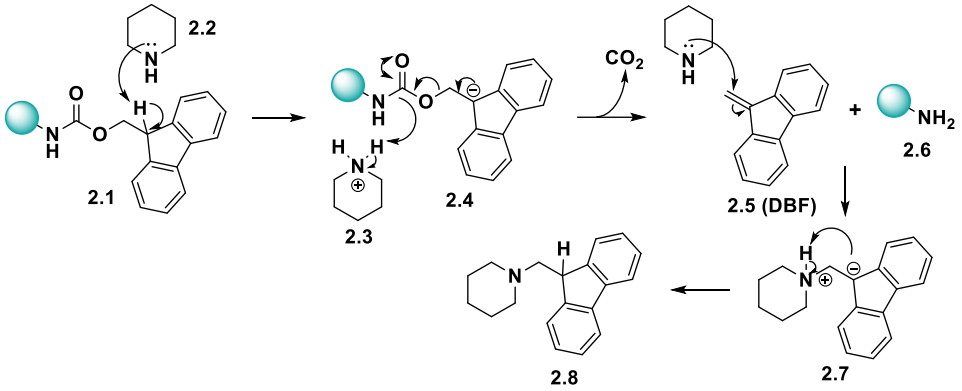


***Figure 2.1**: Schematic overview of SPPS, PG = protecting group and AA = amino acid*

2.1.1.1 Fmoc deprotection

In Fmoc SPPS piperidine is commonly used in the deprotection of the Fmoc group.[4] The lone pair on the nitrogen of the piperidine (**2.2**) removes the acidic proton at the 9-postion of the fluorene ring system of the Fmoc-group (**2.1**), which leads to a cascade of events in which a free amine is formed on the resin (**2.6**), carbon dioxide is released, and dibenzofulvene (DBF, **2.5**) is formed. DBF contains an highly reactive electrophilic double bond due to the stabilisation of the formed aromatic cyclopentadienide anion (**2.7**).[5,6] DBF reacts with piperidine to generate a piperidine adduct (**2.8**) which can then be washed away (**Scheme 2.1**).[7,8]



Scheme 2.1: *SPPS Fmoc deprotection mechanism using piperidine*

2.1.1.2 Amino acid coupling

In SPPS, coupling of the amino acids is done using a coupling agent in conjunction with a base.

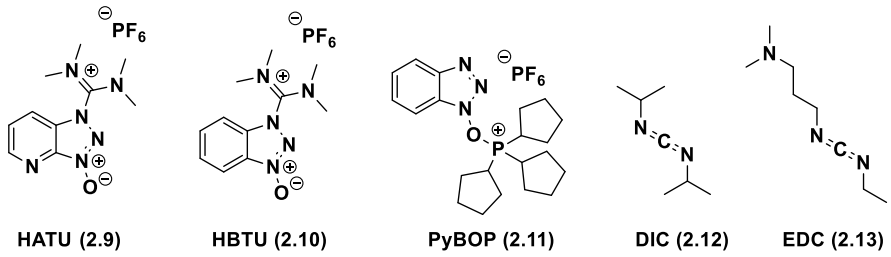
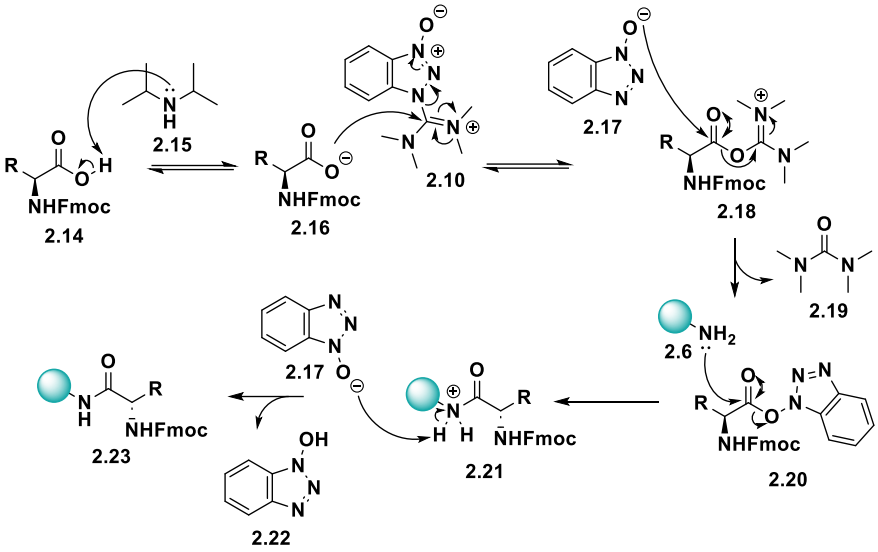


Figure 2.2: *Structures of coupling agents HATU (**2.9**), HBTU (**2.10**), PyBOP (**2.11**), DIC (**2.12**) and EDC (**2.13**)*

Common peptide coupling agents include the uronium based coupling agents hexafluorophosphate azabenzotriazole tetramethyl uronium (HATU) and hexafluorophosphate benzotriazole tetramethyl uronium (HBTU), the phosphonium based coupling agent benzotriazol-1-yloxytripyrrolidinophosphonium hexafluorophosphate (PyBOP) and the carbodiimide based coupling agents *N,N'*-diisopropylcarbodiimide (DIC) and 1-ethyl-3-(3-

dimethylaminopropyl)carbodiimide (EDC) (**Figure 2.2**).

All of these reagents work through a similar mechanism in which they react with the carboxylic acid of the next amino acid to form an activated ester, which reacts with the amine of the growing peptide chain. For example, in automated SPPS a combination of hydroxybenzotriazole (HOBt), hexafluorophosphate benzotriazole tetramethyl uronium (HBTU) and *N*,*N*-diisopropylethylamine (DIPEA) is commonly used (**Scheme 2.2**). DIPEA (**2.15**) deprotonates the carboxylic acid of the amino acid (**2.14**), which reacts with HBTU (**2.10**) to form an *O*-acylisourea (**2.18**) and HOBt (**2.17**). The HOBt then reacts with the *O*-acylisourea to form a HOBt activated ester (**2.20**), which the amine on the resin (**2.6**) reacts with to form the amide bond (**2.21**), reforming HOBt in the process. A proton transfer leads to the final product (**2.23**) (**Scheme 2.2**).



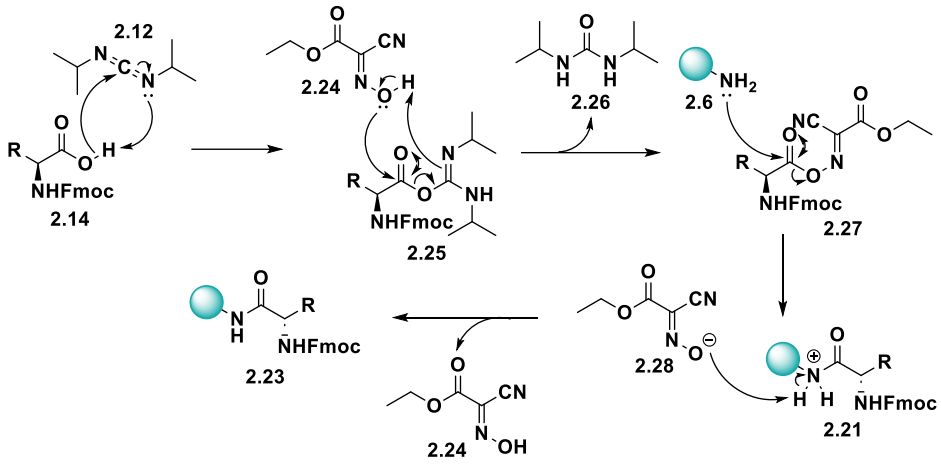
Scheme 2.2: *Fmoc SPPS coupling mechanism using HBTU (**2.10**) and DIPEA (**2.15**)*

In the coupling steps, additional HOBt is added to the reaction to prevent the epimerisation of the stereocentre in amino acids that can occur due to the acidic nature of the α-proton, which can be removed under basic conditions such as by HBTU. The additional HOBt forms an equilibrium with HBTU limiting this deprotonation from occurring.[9]

2.1.1.3 Microwave-assisted SPPS

Originally, SPPS was developed as a room temperature technique, but, since its inception, microwave-assisted SPPS has been developed to increase the speed of synthesis as well as minimise waste.[10] In microwave-assisted SPPS the coupling times can be lowered from 45 min

to 2 min and deprotection times lowered from 15 min to 1 min, with improvement in peptide purity shown by heating to 90 °C.[11] For these coupling conditions *N,N*'-diisopropylcarbodiimide (DIC) and ethyl 2-cyano-2-(hydroxyamino)acetate (oxyma) are used with oxyma preferred over HOBt due to the latter's explosive properties and DIC used because of its stability at high temperatures.[10,12,13] Not all amino acids can be coupled at 90 °C, with a reduction to 50 °C needed for histidine to avoid racemisation of its steric centre.[14] The mechanism for the coupling using DIC and oxyma is illustrated in **Scheme 2.3**. The amino acid is first activated with the DIC (**2.12**) after which it reacts with the oxyma (**2.24**) to form an activated ester (**2.27**). The lone pair of the amine on the resin (**2.6**) then reacts with this activated ester to form an amide bond (**2.21**) reforming oxyma in the process. A proton transfer leads to the final product (**2.23**).



***Scheme 2.3**: SPPS amino acid coupling mechanism using DIC (**2.12**) and oxyma (**2.24**)*

2.1.1.4 Resin material and linker selection

Resin beads can be made from a range of materials such as polystyrene, polyethylene glycol (PEG) and a mix of the two. Resins are functionalised with linkers for their use in SPPS. with different linkers available for the synthesis of peptides with C-terminal acids or amides as well as for the cleavage of peptides with their protecting groups in place. Wang (**2.29**) and 2-chlorotrityl (**2.30**) resins can be used for the synthesis of peptides with C-terminal acids, whereas Rink amide MBHA (**2.32**) and Sieber amide (**2.33**) resins are used for the synthesis of peptides with C-terminal amides (**Figure 2.3**).[15] If the peptides are needed fully protected off the resin, 2-chlorotrityl resin (**2.30**) or Sieber amide resin (**2.33**) can be used as the peptide can be with a low

percentage of TFA, leaving other orthogonal amino acid protecting groups intact. This allows for the further modification of the peptide in solution.[15]

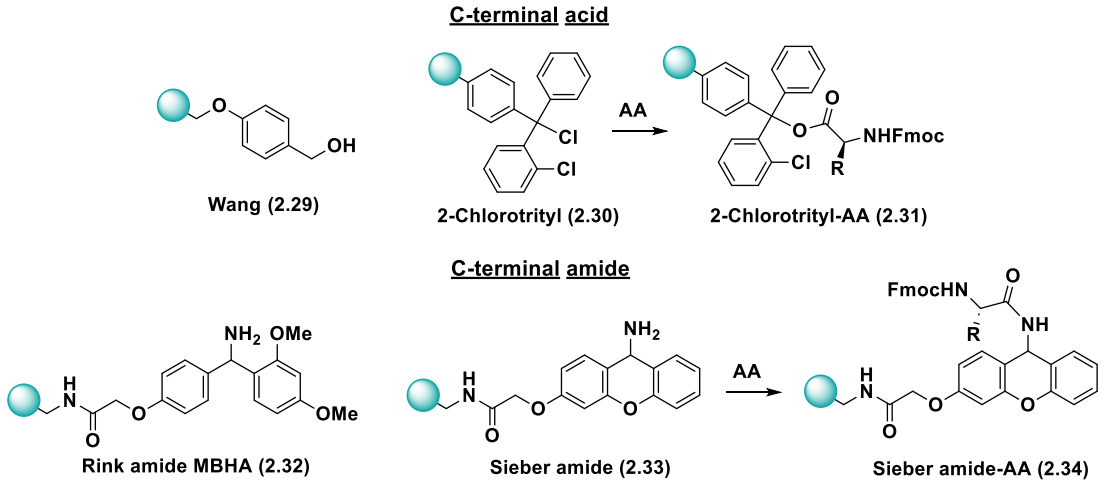


Figure 2.3: *Overview of different linkers used in SPPS*

2.1.1.5 Amino acid protecting groups

In SPPS, amino acid side chains must be protected to stop the formation of side products. The protecting groups used in Fmoc SPPS cannot be base labile to avoid their removal during Fmoc deprotection steps. Instead, acid cleavable side chain protecting groups such as tert-butyl (*t*Bu, **2.35**), trityl (Trt, **2.36**), tert-butyloxycarbonyl (Boc, **2.37**), and 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl (Pbf, **2.38**) are used.

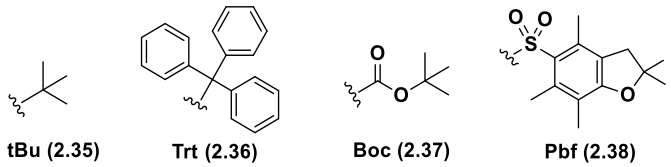


Figure 2.4: *Structures of tBu (**2.25**), Trt (**2.36**), Boc (**2.37**) and Pbf (**2.38**) protecting groups*

For orthogonal reactions on the resin using amino acid side chains, protecting groups are available that can be selectively removed. Lysine can be protected using an alloc group (**2.39**) and aspartic and glutamic acids can be protected using allyl groups (**2.41**). Both alloc and allyl groups can be removed using tetrakis(triphenylphosphine)palladium(0) and phenylsilane. Alternatively, lysine can be protected with 4-methyltrityl (Mtt, **2.40**) and aspartic or glutamic acid can be protected with 2-phenylisopropyl (2-PhiPr, **2.42**), both of which can be removed using a low percentage of TFA, leaving the peptide on resin and keeping other protecting groups

intact. Cysteines can be selectively protected using a 4-methoxytrityl (Mmt, **2.43**) or an acetamidomethyl (Acm, **2.44**) group which can be removed using a low percentage of TFA or iodine respectively.[16]

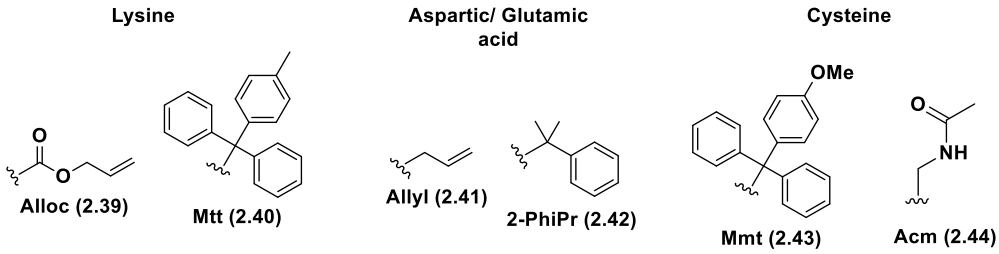


Figure 2.5: *Structure of protecting groups for Lys (Alloc (**2.39**)/Mtt (**2.40**)), Asp or Glu (Allyl (**2.41**)/2-PhiPr (**2.42**)) and Cys (Mmt (**2.43**)/Acm (**2.44**))*

2.1.2 SPPS of difficult sequences

One of the main problems that can occur during SPPS is peptide aggregation, which occurs due to intra- or inter-molecular hydrogen bond formation, leading to the formation of β-sheets (**Figure 2.6**), decreasing the efficiency of any subsequent deprotection and coupling steps.[17,18] The level of aggregation is sequence dependent with an increasing probability found between residues 8-15 and based on the number of hydrophobic residues present.[19,20]

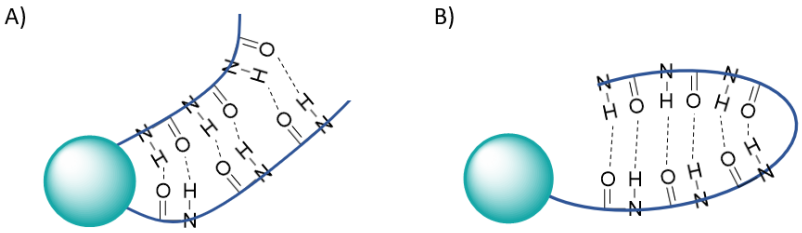


Figure 2.6*: The formation of A) intermolecular and B) intramolecular hydrogen bonding in the growing peptide chain*

Since the invention of SPPS there have been several developments to lower peptide aggregation to make the synthesis of difficult sequences possible. One of the easiest methods is to attempt the synthesis using a resin composed of a different polymer material. The material type can affect the synthesis, largely due to the size of the particles upon swelling, with PEG-based resins having been shown to minimise aggregation, such as in the synthesis of the β-amyloid section 1-42.[21] The development of microwave-assisted SPPS discussed in **Section 2.1.1.3** has also been shown to reduce aggregation.[22]

Low-loading versions of resins may also be available on which there are fewer reactive sites for peptide chains to form with the growing peptide chains therefore spaced further apart on the resin, minimising the intermolecular hydrogen bond formation.[23] Similarly, the addition of a chaotropic salt such as lithium chloride or potassium thiocyanate can also lessen the hydrogen bond interactions between the growing peptide chains.[24]

There have been specific dipeptides developed which can be incorporated in the sequence to stop peptide aggregation. Popular strategies include *N*-protected dipeptides and pseudoproline dipeptides. The amide backbone can be protected with 2,4-dimethoxybenzyl (Dmb) or 2-hydroxy-4-methoxybenzyl (Hmb), reducing the flexibility of the peptide backbone and therefore minimising the ability of the chain to adopt a β-sheet conformation.[25] This can be especially helpful for sequences containing glycine, which has a high flexibility due to its lack of side chain. Since the incorporation of *N*-protected amino acids can be difficult under normal coupling conditions, several dipeptides are available commercially which can then be incorporated into the sequence with conventional coupling methods (**Figure 2.7A**). The Dmb group is removed under the acidic cleavage conditions of the peptide, reforming the desired amino acid.

Alternatively, pseudoproline dipeptides can be inserted into the sequence, similarly reducing the flexibility of the peptide chain.[26,27] These are dipeptides where a threonine, serine or cysteine side chain can be used to form an oxazolidine or thiazolidine derivative with the amide nitrogen, coupled to any other amino acid (**Figure 2.7B**).[27] During peptide cleavage the ring is broken reforming the natural residue.

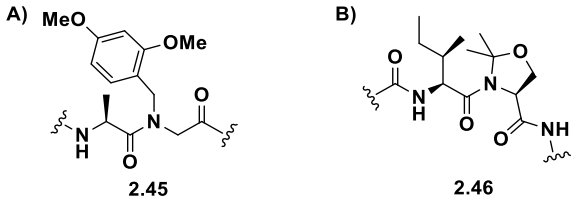


***Figure 2.7**: Structures of A) Ala-(Dmb)Gly (**2.45**) and B) Ile-Ser($\psi^{Me,Me}$pro) (**2.46**) as part of a peptide sequence which become Ala-Gly and Ile-Ser respectively during peptide cleavage with TFA*

The different techniques discussed in this section can all lower peptide aggregation. A combination of different techniques may be necessary and multiple options might have to be tried to optimise the synthesis.

Recently, a pre-print by the Hartrampf group was published in which a SynTag was developed, which can be applied to any sequence to minimise aggregation.[28] For the SynTag six arginine residues are attached to the resin followed by a cleavable Dawson linker on which the desired peptide sequence can be built.

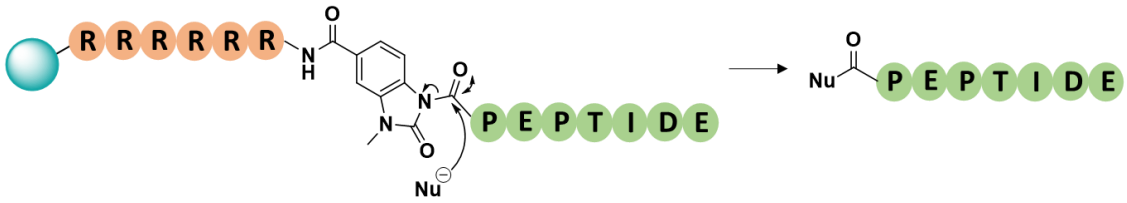


Figure 2.8: *Peptides synthesised using the SynTag consisting of six arginine residues followed by a Dawson linker on which the peptide is built. The peptide can be removed from the SynTag using a nucleophile.*

This arginine sequence minimises aggregation during the synthesis and the Dawson linker allows the cleavage of the peptide from the linker after it is built on resin. This approach was shown to be successful in the synthesis of multiple difficult peptides including amyloid-β (residues 27–42) and MYC (residues 123–143), showing its potential for the application to any peptide sequence.[28,29]

2.1.3 *In vitro* biophysical assays for protein-protein interactions

Once peptides have been synthesised the next step in the drug discovery process is their evaluation against the target protein. *In vitro* biophysical assays are routinely used in early drug discovery to investigate protein-ligand interactions.[30] For the inhibition of PPIs, there is not only a need to examine the interaction of the ligand with the protein, but also the ability of the ligand to disrupt the PPI formation. *In vitro* assays to examine both properties are a useful tool in the development of PPI inhibitors.

2.1.3.1 Direct binding

The binding affinity of the ligand to the target protein can be investigated using different assays depending on the desired information, sample consumption and instrument availability. Fluorescence polarisation (FP), microscale thermophoresis (MST), differential scanning fluorimetry (DSF) and surface plasmon resonance (SPR) are all biophysical techniques with high throughput applications possible. FP, MST and SPR can provide binding affinities, whereas DSF is better used as a relative comparison tool between compounds. SPR has the added benefit of potentially providing information on the kinetics of the interaction. Isothermal titration

calorimetry (ITC) can provide detailed information on both kinetics and binding affinity, but is less suited for high throughput application due to higher sample use. However, it is the only technique that is both label-free and in solution providing the most accurate information about the interaction.[31] FP and MST rely on the use of a fluorescent label on the ligand or protein and DSF requires the addition of a fluorescent dye.[32,33] Surface plasmon resonance (SPR) does not rely on fluorescence but instead requires the protein or peptide to be immobilised, which could also interfere with the binding.[34] In this section a brief discussion of these biophysical techniques is given.

Fluorescence polarisation (FP)

Fluorescence polarisation uses a fluorescently labelled ligand, such as a fluorescein-tagged peptide, to look at the difference in polarisation of the emission light at different protein concentrations.[35] In the absence of protein or when the peptide does not bind the target, the peptide can freely tumble in solution and the polarised excitation light is depolarised upon emission (**Figure 2.9A**). As the peptide starts to bind to the protein target the rotation of the peptide is slowed down and a polarised emission is seen (**Figure 2.9B**).[36] The polarisation can be plotted against the protein concentration to determine a dissociation constant (K_D) value (**Figure 2.9C**).[37]

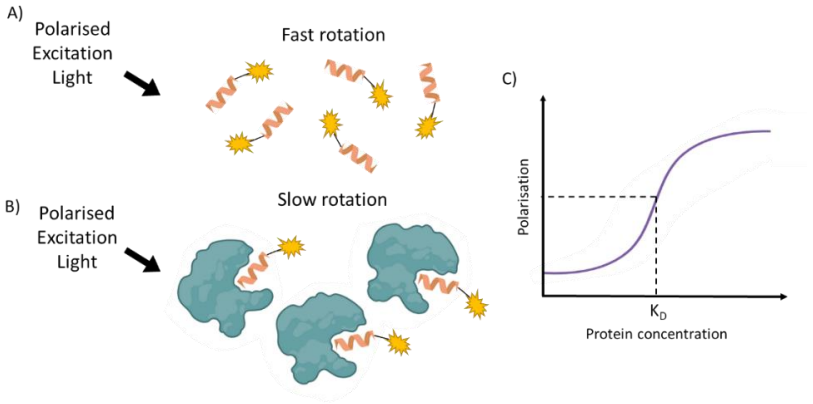


Figure 2.9: *The emitted light is A) depolarised in the absence of protein and B) polarised in the presence of protein to which the ligand binds leading to C) a curve from which the K_D can be determined*

The advantages of FP include that it is useful for high-throughput applications, especially finding its application in investigating peptide-protein interactions due to the ease of synthesis of a fluorescently-tagged peptides.[38,39] The protein consumption depends on the affinity of the interaction between the probe and the protein with a high affinity equating to a low usage of

protein. However, the fluorescent tag may alter the binding properties of the ligand and therefore further confirmation of the K_D determined by FP may be needed.

Microscale thermophoresis (MST)

MST detects the motion of molecules in a microscopic temperature gradient, known as thermophoresis, which changes upon binding.[40] Typically the protein is labelled with a fluorescent dye or the protein is conjugated to green fluorescent protein, but MST can also be used label-free through the intrinsic fluorescence of aromatic residues.[41] Using an infrared laser, the sample is heated in a capillary tube and as the temperature increases, the fluorescence decreases until reaching a steady state due to the movement of the protein away from the heated area (**Figure 2.10A**).[32] As a ligand is added and interacts with the protein, the protein movement is slower with a smaller decrease in fluorescence seen. The ligand can be titrated with a constant protein concentration to determine a dissociation constant (**Figure 2.10B**). MST has a low sample consumption, can be used in any buffer and has the possibility to measure picomolar affinity interactions.[41]

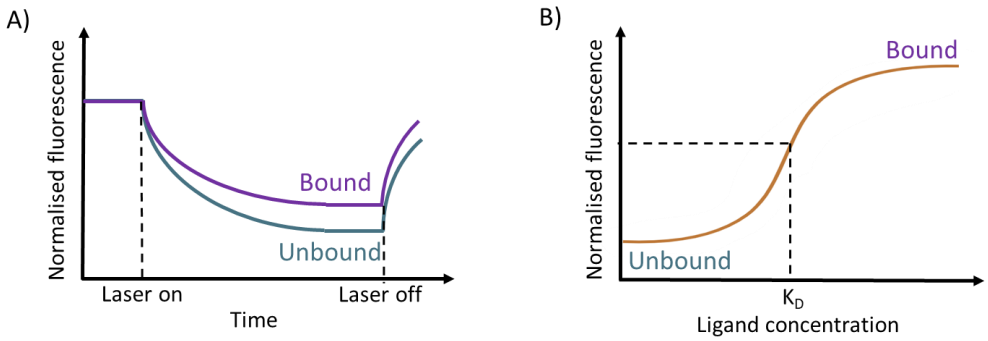


Figure 2.10: *Normalised fluorescence measured in MST showing A) a comparison of the protein unbound (blue) and bound to a ligand (purple) and B) the change in fluorescence with increasing ligand concentration*

Differential scanning fluorimetry (DSF)

DSF uses a fluorescent dye, such as Sypro orange or 1-anilinonaphthalene-8-sulfonic acid (1,8-ANS) which have low fluorescence in high dielectric environments such as buffers. As the sample is heated the protein denatures, revealing buried hydrophobic surfaces (regions with low dielectric environments) to which these dyes bind.[42] The fluorescence is measured over an increase of temperature and the inflection point of the curve can be determined which is the melting temperature (T_m) (**Figure 2.11A**). This point can be measured without ligand present and

with ligand present to measure the change in melting temperature (ΔT_m) (**Figure 2.11B**). The binding of a ligand to the target protein stabilises the protein therefore shifting the T_m to a higher temperature.[43] DSF is a quick and high-throughput assay to determine whether a compound can bind to the target protein with no modifications of the protein or compounds needed. However, it may not be applicable for all proteins, as proteins with hydrophobic patches on their surface would show increased background fluorescence.[44] It also does not allow for the determination of a binding constant or kinetic properties.

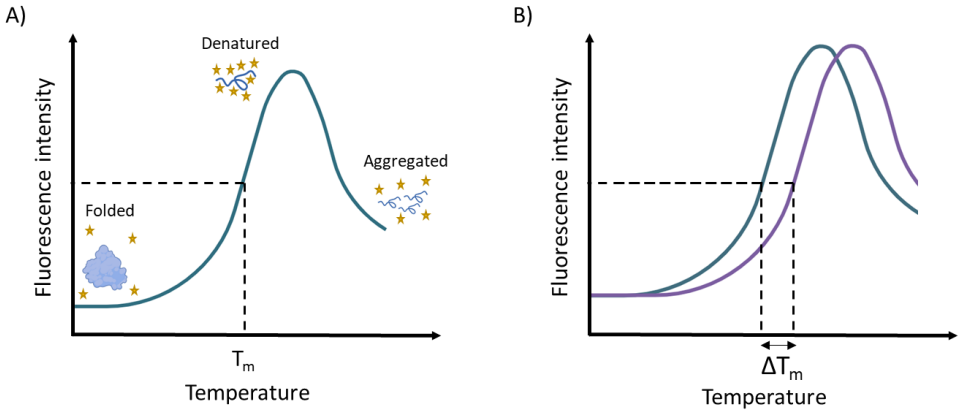


***Figure 2.11**: A schematic overview of DSF showing A) the fluorescence intensity as the folded protein denatures in the presence of a dye and B) difference in melting temperature in the absence (blue) or presence (purple) of ligand*

Surface plasmon resonance (SPR)

In surface plasmon resonance (SPR) a ligand is immobilised on a gold surface over which the analyte is flowed. SPR occurs when the light hits the surface at a certain angle at which the light energy couples to the electrons present in the gold surface.[45,46] Through this excitation the electrons move, with the movement known as plasmon resonance. The angle of the incident light at which plasmon resonance occurs changes with changes on the surface such as the interaction of an analyte with the immobilised ligand (**Figure 2.12A**).[47,48] This change in angle can be followed over an association and dissociation period for a range of analyte concentrations to determine binding affinity and kinetics (**Figure 2.12B**). SPR allows for not only the determination of the binding affinity of the interaction, but can also provide information about the kinetics of the interaction.[49] SPR also does not require a fluorescent tag and therefore no additional compounds have to be synthesised, and has a low protein consumption. However, the immobilisation process may mean that the target site is partially or fully inaccessible and therefore it may not be suitable for every protein.

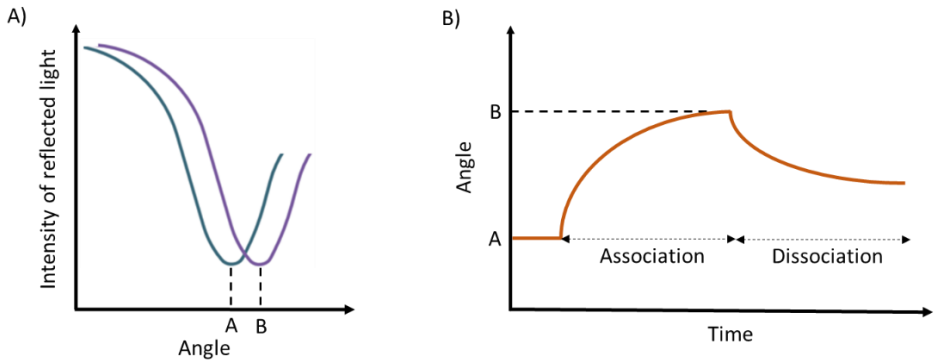


Figure 2.12: *A) The change in the plasmon resonance angle seen without (blue) or with (purple) ligand present. B) This angle change represented as a signal which increases during the association period and then decreases during the dissociation process*

Isothermal titration calorimetry (ITC)

ITC measures the heat change upon binding over the titration of the ligand into the protein solution (**Figure 2.13**).[50] The heat change can be plotted against the ligand concentration from which the association constant (K_a) can be calculated. ITC does not only provide information about the binding affinity of a ligand, but it can also provide the thermodynamic properties of the interaction, such as enthalpy change (ΔH), entropy change (ΔS), free energy change (ΔG) and heat capacity change (ΔC_p) as well as the stoichiometry.[31] Another advantage of ITC is that it does not rely on the use of a fluorescent tag or the immobilisation of protein therefore not interfering with the binding affinity measurements. The disadvantage of the technique is the high protein and ligand consumption which can make other techniques more suitable for the screening of compounds.[51]

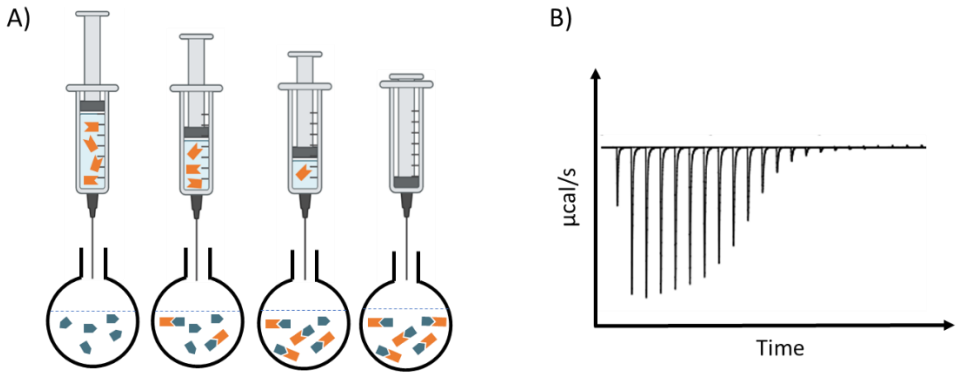


Figure 2.13: *Schematic overview of ITC with A) the injection of the ligand into the protein and B) the heat change measured at every injection.*

2.1.2.2 Inhibition of PPIs

Further assays are needed to investigate the potential of the ligands discovered in the binding assays to inhibit the targeted PPI, to confirm that the ligands bind in the desired region of the protein and to confirm this can outcompete with the binding to the protein partner. To investigate this, assays based on resonance energy transfer such as Förster resonance energy transfer (FRET) and Homogeneous Time Resolved Fluorescence (HTRF) assays, which can take place with all components in solution, or assays relying on immobilisation, such as is Enzyme-linked immunosorbent assay (ELISA) and surface plasmon resonance (SPR) assays, can be used.

Förster resonance energy transfer (FRET)

FRET occurs when a donor and acceptor fluorophore come in close proximity with each other and the excited energy from the donor is transferred to the acceptor from which the emission energy can then be measured. The donor and acceptor fluorophores can be attached to the proteins and the expression of fusion-proteins can be done intracellularly to study protein-protein interactions in cells.[52] If the proteins come in close contact (< 10 nm) with each other a FRET signal will be observed and if treated cells do not give a FRET signal this would show the disruption of the PPI. FRET has also been applied as a microscopic technique in which the FRET and therefore PPI can be observed in real-time.[53] Interference of background fluorescence from buffers, cell lysate and proteins can hamper the use of FRET.[54] Bioluminescence resonance energy transfer (BRET) relies on the same principle as FRET, but uses a bioluminescent (such as luciferase) and an acceptor protein (such as green fluorescent protein). These can be expressed as fusions to the target proteins allowing the examination of interactions in live cells.[55,56]

Homogeneous Time Resolved Fluorescence (HTRF) combines the FRET technology with the time-resolved measurement of fluorescence. This time-resolved fluorescence is possible using a europium or terbium cryptate donors, which have microsecond emission lifetimes compared to nanosecond lifetimes for fluorescein.[57] In these cryptate donors, the lanthanide ion (europium or terbium) is embedded in a trisbipyridine structure allowing for the long-lived emission.[58] Therefore in the HTRF assay the emission is measured with a 50-150 μs delay eliminating any short-lived background fluorescence, allowing for higher sensitivity in the assay.[54] The emission of the donor and acceptor is measured simultaneously and a ratio of the two is reported (**Figure 2.14**), which reduces variations observed well to well. A HTRF assay does not require the

modification of either protein partner or the ligands and with all components in solution, can be used for high throughput applications.[54]

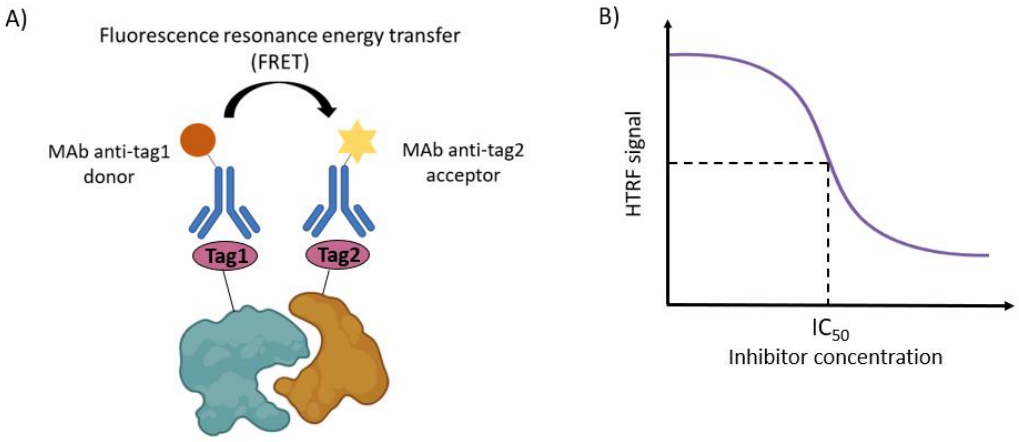


Figure 2.14: *A) The principle of HTRF where if a donor and acceptor antibody come in close proximity a FRET signal is seen and B) with increasing inhibitor concentration the signal is lost*

Enzyme-linked immunosorbent assay (ELISA)

In ELISA based assays, one of the protein partners is immobilised on a microplate through the use of an antibody or covalent bond formation.[59] The plate can then be treated with the inhibitors after which the protein partner is added with washing steps between each step. The protein partner is then detected using an enzyme-tagged antibody specific to the protein partner (**Figure 2.15**). A substrate specific for this enzyme is then added and the signal measured is proportional to the formation of the PPI.[60] A benefit of the immobilisation of the interaction is that it allows for easy washing in between each step. However, compared to HTRF in which everything can be mixed together ELISA is still a longer multi-step process which can limit its high throughput applications.

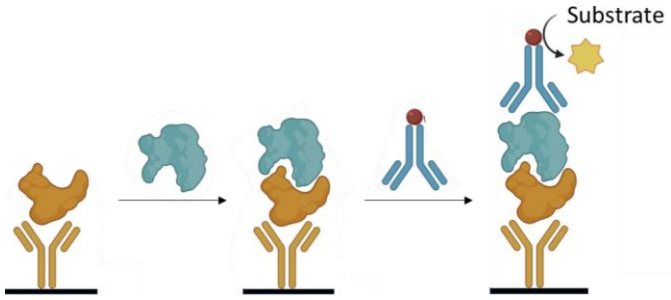


Figure 2.15: *An overview of an ELISA assay in which two proteins need to form an interaction for a signal to occur*

Variations on ELISA assays such as the AlphaLISA assay have been developed for easier high throughput and automation. The AlphaLISA assay uses streptavidin-coated AlphaDonor beads that can interact with a biotin-labelled protein and an anti-tag (e.g. anti-His) conjugated AlphaLISA acceptor bead, which can interact with the protein partner. When the PPI is formed and the beads come into close proximity, the excitation of the donor bead at 680 nm generates singlet oxygen production, triggering the chemiluminescent emission of the acceptor bead at 615 nm.[61] AlphaLISA is a no wash assay, eliminating time-consuming steps found in ELISA assays, allowing for high-throughput applications and lower sample consumption.[61,62]

Surface plasmon resonance (SPR)

SPR has already been discussed as a technique to study the interaction of the ligand with the target protein. It can also be used in a competitive assay to investigate the ability of the ligand to disrupt the PPI. With one protein immobilised on the chip the protein partner is flowed over the chip to form the PPI complex. Increasing concentrations of ligand can then be added to examine the concentration at which the complex can be disrupted.[63]

2.2 Chapter aims

The aims of this chapter are to design and synthesise peptides able to disrupt the dyskerin-dyskerin PPI and to develop the *in vitro* biophysical assays required to assess this. The cryo-EM structure of telomerase will be used to define a starting point for the peptide design, after which it will be explored whether this peptide is able to bind to dyskerin and disrupt the dyskerin-dyskerin PPI. The aim is to both identify the residues within this sequence most important for binding as well as to explore the possibility of modifying the peptide to increase its affinity towards dyskerin. The overarching aim is to develop robust assays to investigate the binding of the synthesised peptides to dyskerin and to investigate their ability to inhibit the dyskerin-dyskerin PPI.

2.3 Results and discussion

2.3.1 Initial peptide design and synthesis

As discuess in **Section 1.2.4** the structure of telomerase revealed two different sections which mediate the PPI between the dyskerin proteins (**Figure 2.16**).[64]

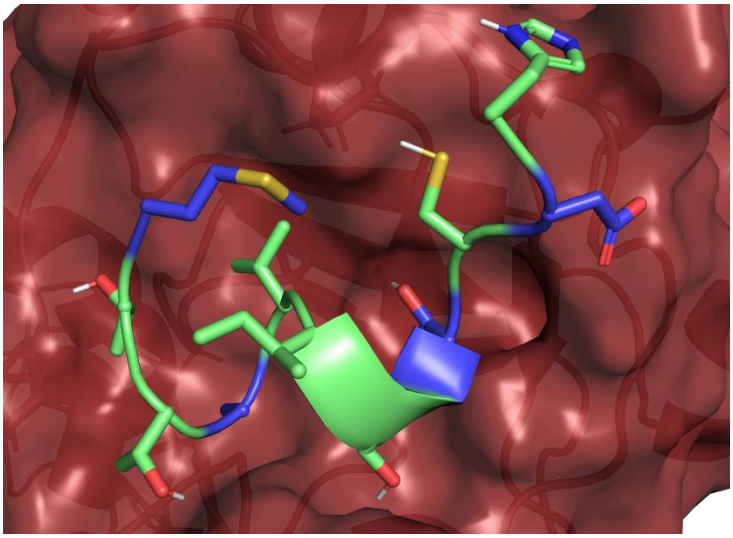


Figure 2.16: *The dyskerin-dyskerin PPI (PDB: 7BGB) with the interacting sequence ($_{350}$MTTAVISTCDH$_{360}$) in green with mutations found in patients with DKC and HH in dark blue.*

As a starting point for the design of peptides targeting the dyskerin-dyskerin PPI, our work began with the section $_{350}$MTTAVISTCDH$_{360}$ from dyskerin due to the A535V mutation being the most common mutation in patients with DKC.[65] Using Fmoc SPPS the synthesis of Ac-MTTAVISTCDH-NH_2 (**2.47**) was first attempted using Rink amide MBHA resin (0.52 mmol/g loading) using room temperature automated peptide synthesis. Rink amide MBHA resin is a polystyrene based resin commonly used in Fmoc SPPS due to its low cost.[66] The peptide was acetylated to form an uncharged N-terminal group and the Rink amide linker provided an uncharged C-terminal amide, mimicking the native protein.[67] This was followed by cleavage from the resin using trifluoroacetic acid (TFA). The scavengers, H_2O, 2,2'-(ethylenedioxy)diethanethiol (DODT) and triisopropylsilane (TIPS) were added to the TFA to prevent reactive moieties from the cleaved protecting groups from reacting with the peptides.[68] After peptide cleavage, the crude material was analysed by RP-HPLC (**Figure 2.17**) which did not show one major peak as desired but instead showed multiple smaller peaks, indicating problems during the synthesis.

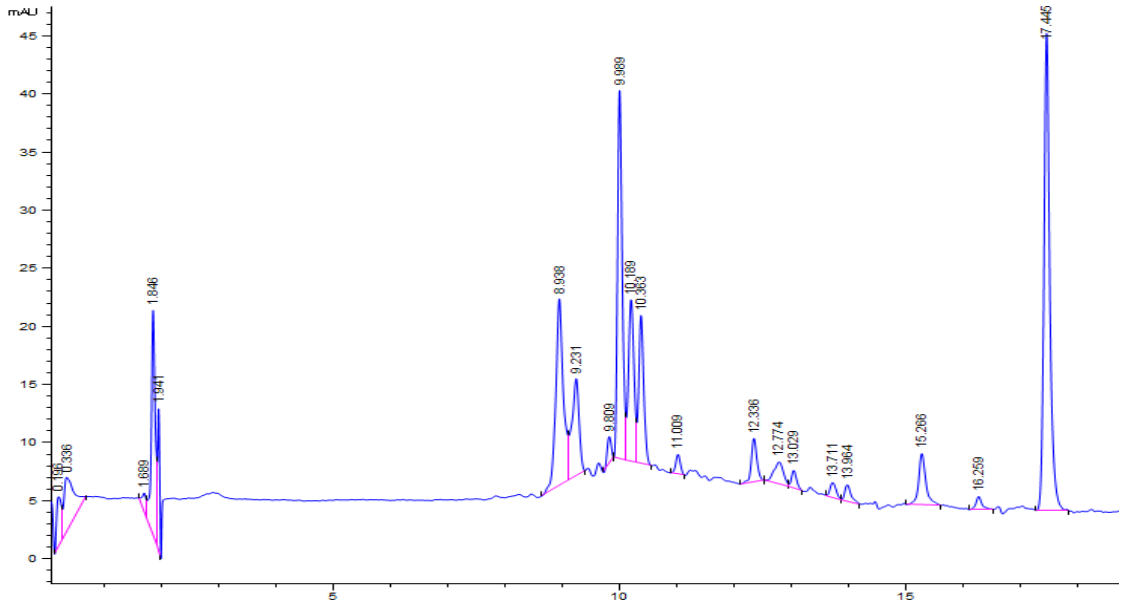
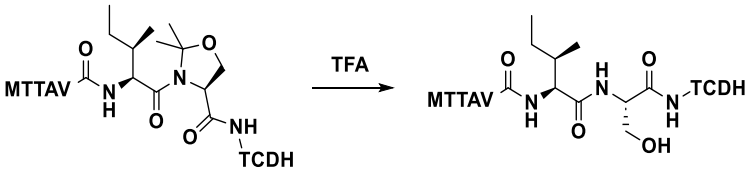


Figure 2.17: *HPLC trace (214 nm, 5-95% MeCN in H_2O + 0.05% TFA over 15 min., then 5 min. at 95%) of crude peptide **2.1** synthesised using Rink amide MBHA resin*

As outlined in section **2.1.1** one of the problems which can occur during SPPS is peptide aggregation. Our initial strategy to reduce aggregation included trial syntheses using different resins and the use of a pseudoproline moiety.

We tested low loading Rink amide MBHA (0.35 mmol/g), NovaPEG Rink amide which is a PEG-based resin, and ProTide Rink amide which is a PEG/polystyrene mixed resin.[23] It was also explored whether an Ile-Ser pseudoproline could be used to improve the synthesis (**Scheme 2.4**).



Scheme 2.4: *The formation of peptide MTTAVISTCDH (**2.47**) from the Ile-Ser pseudoproline version (**2.48**) using TFA*

All four methods were tried using room temperature automated SPPS and after cleavage the crude mixture from each method was analysed by RP-HPLC (**Figure 2.18**).

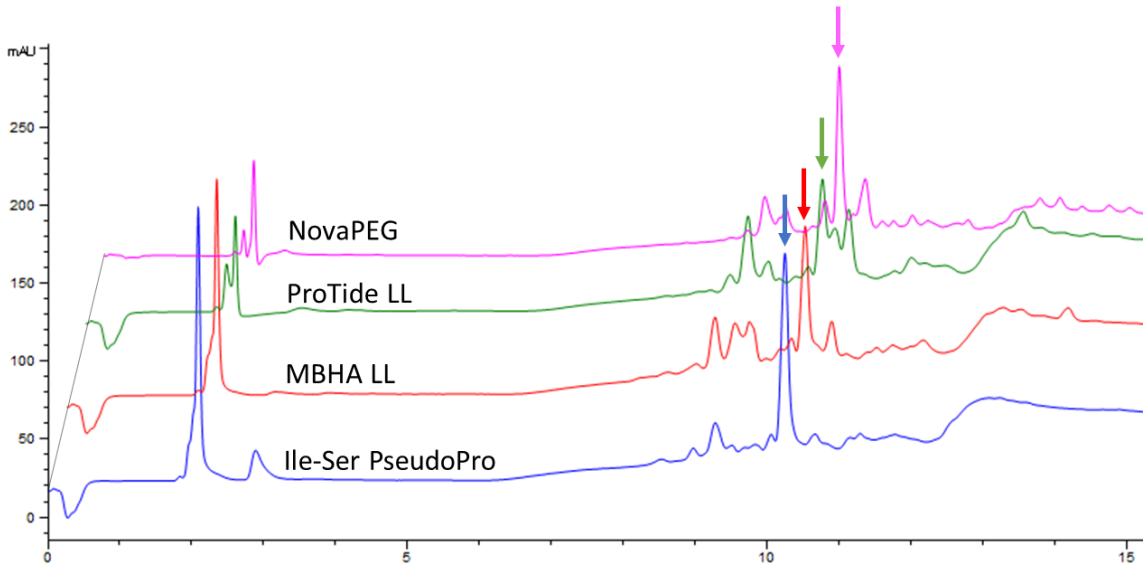


Figure 2.18: *HPLC traces (214 nm, 5-95% MeCN in H_2O + 0.05% TFA over 15 min) of crude peptide **2.1** using an Ile-Ser pseudoproline (blue), MBHA LL (red), ProTide LL (green) and NovaPEG (pink). The traces are offset 1% in time and 40% in absorbance from each other.*

The crude purity levels of the different methods have been summarised in **Table 2.1**. The ProTide low loading resin (0.35 mmol/g) (**Figure 2.18 green**) showed little improvement on the Rink amide MBHA resin (21% compared to 14% crude purity). The Rink amide MBHA low loading resin (**Figure 2.18 red**) and NovaPEG resin (**Figure 2.18 pink**) showed some improvement (30% and 29% crude purity respectively). Although for NovaPEG the impurities were still present, the product peak showed larger separation from the impurity peaks, which would make purification more easily achievable. The use of pseudoproline (**Figure 2.18 blue**) showed the most improvement (54% crude purity), with one major product peak (10.2 min.).

Table 2.1: *The different synthetic methods attempted in the synthesis of peptide **2.1** and their crude purity as measured by HPLC through the integration of the 5-15 min. section of the spectrum*

Method	Crude purity (%)
Rink amide MBHA resin (0.52 mmol/g)	14
Ile-Ser pseudoproline with Rink amide MBHA resin (0.52 mmol/g)	54
Rink amide MBHA low loading resin (0.35 mmol/g)	30
ProTide Rink amide low loading resin (0.20 mmol/g)	21
NovaPEG Rink amide (0.57 mmol/g)	29

The product peak was successfully isolated from the crude mixture of the Ile-Ser pseudoproline method with the peak at 10.2 min. confirmed as the target peptide **2.47** by LCMS (m/z calc. = 1219.55, m/z obs. = 1219.5)

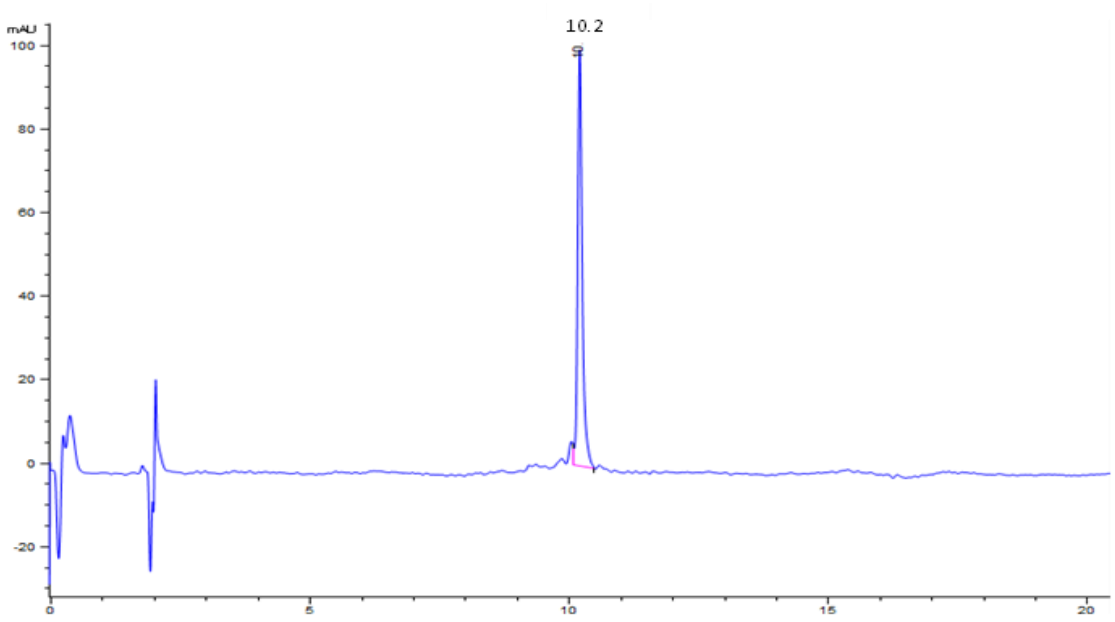


Figure 2.19: *HPLC trace of purified peptide* ***2.47*** *(214 nm, 5-95% MeCN in H_2O + 0.05% TFA over 15 min, then 95% MeCN in H_2O + 0.05% TFA for 5 min.)*

For the peptide design, the main aim of the peptides synthesised was to define a sequence that may be able to bind to dyskerin and to investigate the residues most important for binding. To do so, a set of longer peptides, where residues extending along the native protein sequence were added to each side of the peptide, were designed (**Table 2.2**). From the telomerase structure these residues do not seem to be directly interacting at the PPI, but they could have an impact on stabilising the α-helical structure of the peptide.

Table 2.2: *The sequences, yields, retention time (t_R) and m/z calculated and observed of peptides* ***2.47, 2.49*** *and* ***2.50***

Nr.	Sequence	Yield (%)	Purity (%)	t_R (min.)	m/z calc.	m/z obs.
2.47	Ac-MTTAVISTCDH-NH_2	14	96	10.2	1219.6	1219.5
2.49	Ac-LMTTAVISTCDHG-NH_2	5	90	10.9	695.3	695.6
2.50	Ac-ALMTTAVISTCDHGI-NH_2	10	93	11.8	787.4	787.8

Peptides **2.49** and **2.50** were synthesised using a combination of NovaPEG Rink amide resin and pseudoproline Ile-Ser to minimise aggregation, as aggregation can increase with peptide length. This proved successful with both peptides acquired as a single peak with the correct mass for

the $[M+2H]^{2+}$ observed by LCMS (for **2.49**: m/z calc. = 695.32 and m/z obs. = 695.6, for **2.50**: m/z calc. = 787.38 and m/z obs. = 787.8).

Additionally, an alanine scan of peptide **2.47**, where each residue was sequentially replaced with alanine, was designed to investigate the importance of each residue in the sequence (**Table 2.3**).

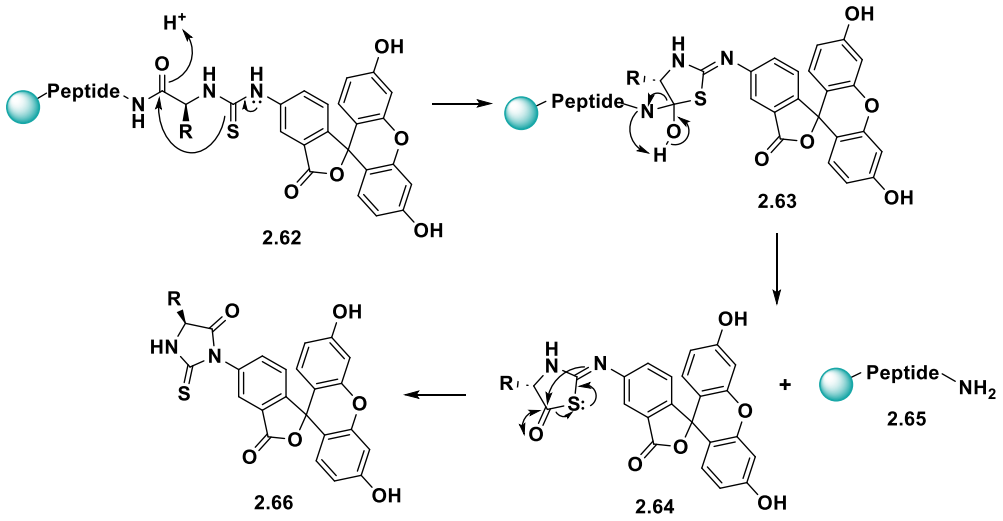
Table 2.3: *The sequences of the alanine scan peptides* ***2.51-2.61***

Nr.	Sequence	Yield (%)	Purity (%)	t_R (min.)	m/z calc.	m/z obs.
2.51	Ac-ATTAVISTCDH-NH_2	3	97	9.5	1159.5	1159.9
2.52	Ac-MATAVISTCDH-NH_2	6	90	10.7	1189.5	1189.9
2.53	Ac-MTAAVISTCDH-NH_2	-	-	-	-	-
2.54	Ac-MTTAAISTCDH-NH_2	3	96	9.7	1191.5	1191.9
2.55	Ac-MTTAVASTCDH-NH_2	12	91	10.2	1177.5	1178.1
2.56	Ac-MTTAVIATCDH-NH_2	3	96	10.5	1203.6	1204.1
2.57	Ac-MTTAVISACDH-NH_2	2	95	10.6	1189.5	1190.1
2.58	Ac-MTTAVISTADH-NH_2	3	95	9.6	1187.6	1188.2
2.59	Ac-MTTAVISTCAH-NH_2	3	94	10.2	1175.6	1176.2
2.60	Ac-MTTAVISTCDA-NH_2	-	-	-	-	-
2.61	FITC-βAla-MTTAVISTCDH-NH_2	1	97	10.7	819.3	819.6

The synthesis of the alanine scan peptides was first attempted using NovaPEG Rink amide resin due to the increased cost of the Ile-Ser pseudoproline dipeptide. This was successful for peptides **2.55**-**2.59**, for which pure peptide was collected after RP-HPLC purification upon the first attempt. For peptides **2.51**-**2.54**, the HPLC traces of the crude peptides mostly showed the desired product peak with minimal impurities. However, degradation occurred after purification with the HPLC trace of the isolated peaks either no longer showing the product peak at all or showing multiple impurities had formed. This may be due to prolonged time (up to 18 h) at slightly elevated temperatures (30 °C) during solvent evaporation after RP-HPLC purification. These peptides were resynthesised, which was attempted using both NovaPEG Rink amide resin and the pseudoproline Ile-Ser moiety to simplify the purification. The solvent was then removed at 20 °C during a shorter period of time (up to 4 h), which was successful for all except for peptide **2.53**, which still showed degradation. The synthesis of peptide **2.60** was difficult even with both techniques used to minimise aggregation and the crude material was insoluble in MeCN and MeOH or either in combination with H_2O which was used to purify the other peptides. In this peptide the Ala replaces a His, a charged amino acid, perhaps lowering the solubility of the

peptide. All purified peptides were confirmed by a single peak in the analytical HPLC and LCMS to confirm the desired m/z with the data shown in **Table 2.3**.

To investigate the binding affinity of these peptides to dyskerin a fluorescent probe of peptide **2.47** was designed for the use in a fluorescence polarisation assay. Fluorescein isothiocyanate (FITC) was conjugated to the peptide through a β-alanine linker giving peptide **2.61** (**Table 2.3**). The β-alanine linker is used as a spacer between the peptide and FITC stopping any interference of fluorescein on the binding of the peptide to dyskerin. It is also needed to prevent the Edman degradation of the fluorescein thiourea (**2.62**) eliminating the last amino acid in the sequence (**2.65**) forming fluorescein thiazolinone (**2.64**) which rearranges to form a fluorescein thiohydantoin (FTH, **2.66**) (**Scheme 2.5**).[69]



Scheme 2.5: *The mechanism for the Edman degradation of a fluorescein thiourea (**2.62**) leading to the elimination of the last amino acid in the sequence to form FTH (**2.66**).*

Peptide **2.61** was synthesised using NovaPEG Rink amide resin using automated SPPS until after the coupling and deprotection of β-alanine. The fluorescein tag was conjugated to the peptide using FITC and DIPEA over 18 h and kept in the dark to prevent any quenching of the fluorophore. After this the peptide was cleaved and purified by RP-HPLC with a single peak isolated (t_R = 10.7 min.) with the correct mass observed by LCMS for the $[M+2H]^{2+}$ (m/z calc. = 819.31, m/z obs. = 819.6).

2.3.2 The development of a fluorescence polarisation binding assay

We first used fluorescence polarisation (FP) to investigate the binding affinity of the synthesised peptides to dyskerin due to ease of synthesis of a fluorescent probe as well as the ease of screening a large number of peptides with low protein and peptide consumption. Fluorescence polarisation provides a high throughput method to evaluate the prepared peptides in a competition assay with the fluorescent probe. The fluorescence polarization assay was tested using MBP-dyskerin (see **Chapter 5** for the expression of MBP-dyskerin) and FITC-βAla-MTTAVISTCDH-NH_2 (**2.61**) in 10 mM HEPES, 150 mM NaCl, 0.05% Tween-20, pH 7.4 buffer. The aim for the probe concentration is to keep it as low as possible for higher sensitivity and it cannot be larger than two times the K_D.[35] Since the K_D of this interaction was not known, the assay was tested with both a 10 and 5 nM concentration of the probe with both concentrations above the limit of detection and therefore the lower concentration of 5 nM of (**2.61**) was chosen.[70] For FP assays the protein concentration is plotted against anisotropy, which is derived from the intensity of the parallel and perpendicular light through **Equation 2.1**.[71]

$$r = \frac{I_a - I_b}{I_a + 2I_b}$$

Equation 2.1*: The calculation of anisotropy (r) from the intensity of parallel emission light (I_a) and perpendicular emission light (I_b)*

The assay was first tested with a dyskerin protein concentration up to 11.8 µM. Measurements were taken at different time points from 15 min up to 18 h but over time there was only a decrease in signal and no upper plateau was reached.

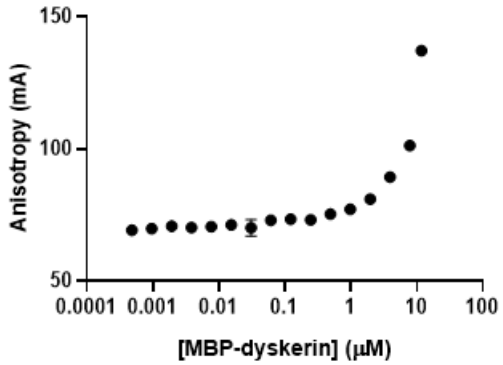


Figure 2.20*: Anisotropy measured of 5 nM peptide **2.61** with MBP-dyskerin after 15 min incubation*

Binding of **2.61** to MBP-dyskerin was seen from approximately 1 µM, but since no upper plateau was reached going up 11.8 µM a K_D determination was not possible. To reach an upper plateau the MBP-dyskerin concentration would need to be much higher, which would require a large uneconomical amount of protein. Instead, the FP assay could be improved by finding a higher affinity probe that would reach an upper plateau at the tested concentrations. We hypothesised that this probe may have a low binding affinity due to the loss of secondary structure of the linear peptide. As such, we used circular dichroism to investigate the secondary structure of peptide **2.47**.

2.3.3 Investigating the α-helicity of peptide 2.47

Circular dichroism was used to investigate the secondary structure of the peptides synthesised. Circular dichroism measures the difference in left- and right-handed circularly polarised light over a range of wavelengths, for peptides the region from 185 to 260 nm is of interest. Within $_{350}$MTTAVISTCDH$_{360}$ the section $_{352}$TTAVIST$_{357}$ is an α-helix when part of the dyskerin protein. For an α-helix minima are expected at 208 and 222 nm and a maximum at 193 nm.[72] The data collected was analysed by Dichroweb using the method CDSSTR.[73, 74] The spectra was recorded at 167 µM (0.20 mg/mL) in 50/50 10 mM phosphate buffer pH 7.4/trifluoroethanol. Trifluoroethanol (TFE) has been shown to enhance the α-helical structure of peptides in solution.[75] Peptide **2.47** was tested to investigate whether it would form an α-helix outside the rest of the protein structure.

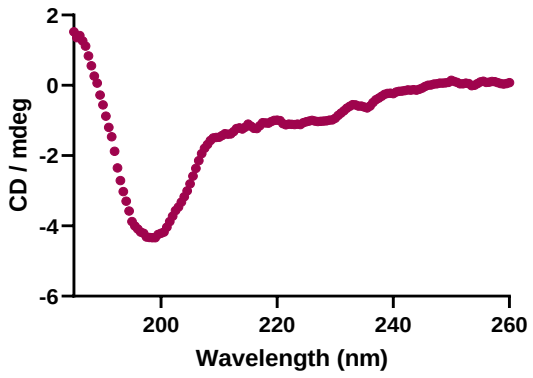


Figure 2.21: *Circular dichroism spectra of peptide **2.47** at 167 µM in 50/50 10 mM phosphate buffer pH 7.4/TFE*

The linear peptide **2.47** showed low α-helicity (3%) using circular dichroism. The lack of structure of the peptide could be lowering its binding affinity to dyskerin through the energy requirement

of folding into a helical structure before binding. One way of overcoming this and therefore potentially increasing the binding affinity of the peptide is through peptide stapling which can preorganise the peptide into an α-helix through the connection of two amino acids on the same side of the helix.

2.3.4 The design, synthesis and evaluation of hydrocarbon stapled peptides

Peptide stapling can be applied to α-helical peptides to improve their binding affinity, cell permeability and proteolytic stability.[76] In peptide stapling two residues are covalently bonded together through the reaction of their side chains. Different types of staples have been developed, including lactam, triazole and hydrocarbon staples as well as a range of different thioether staples (**Figure 2.22**).[77]

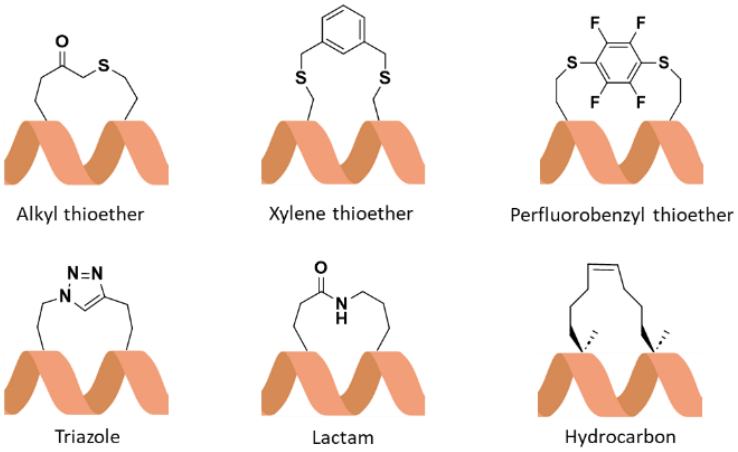


***Figure 2.22**: Different peptide stapling methods used to enhance the α-helicity of the peptides*

Peptide stapling using natural amino acids, such as through cysteines in thioether staples and through aspartic or glutamic acid and lysine in lactam staples provides the advantage of using amino acids that do not require any different coupling conditions. Modifications of these conditions may be necessary for the incorporation of the unnatural amino acids such as the ones used in hydrocarbon stapling.[78,79] Although hydrocarbon stapling requires unnatural amino acids, it has become one of the most frequently used stapling methods due to its success in increasing α-helicity, binding affinity and cell permeability.[80,81] It was first successfully applied to BH3 derived peptides targeting Bcl-2, and has since been used to improve peptide inhibitors for many targets, including the cancer targets β-catenin, MDM2/X and the Nac1 POZ domain.[81–84] The staple is formed through the reaction of two alkene groups in a Grubbs ring-closing

metathesis reaction. Depending on the stereochemistry of the alkenyl amino acids different length staples can be formed, with *i*, *i*+4 stapling using the S-enantiomer, and *i*, *i*+7 stapling using a combination of the S- and R-enantiomers as illustrated in **Figure 2.23**.[80]

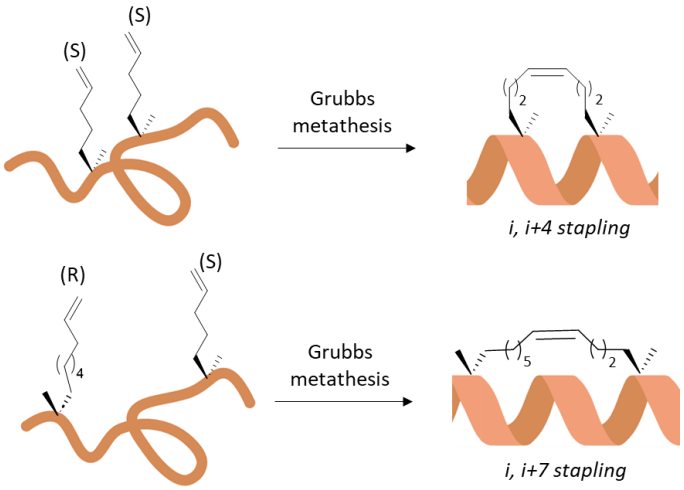
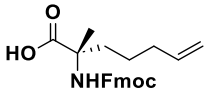


***Figure 2.23**: (S)- and (R)-enantiomers of alkenyl amino acids for i, i+4 and i, i+7 hydrocarbon stapling using a ring-closing Grubbs metathesis reaction*

2.3.4.1 Hydrocarbon stapled fluorescent peptide design, synthesis and evaluation

The first stapling method chosen to apply to the peptides derived from dyskerin was hydrocarbon stapling. The amino acid building blocks are commercially available due to their widespread use. The only concern with this method was the potential impact on solubility as this had been a factor in other stapled peptides in the literature.[85] Hydrocarbon stapling can be used for *i*, *i*+4 and for *i*, *i*+7 stapling, but since the peptide derived from dyskerin is only α-helical over the six residue sequence $_{351}$TTAVIST$_{357}$ *i*, *i*+4 stapling was applied. For *i*, *i*+4 stapling Fmoc-(S)-2-(4-pentenyl)Ala-OH (S_5, **2.67**) is used as the building block and introduced at both positions in the sequence (**Figure 2.24**).



2.67

***Figure 2.24:** Structure of Fmoc-(S)-2-(4-pentenyl)Ala-OH (S_5, **2.67**)*

Since the $_{351}$TTAVIST$_{357}$ section of dyskerin in telomerase forms an α-helical structure initially the insertion of the staple within this section was explored. However, we chose to not insert the staple in the A353 position as this alanine residue faces the PPI and inserting it in this position

would most likely interfere with the peptide binding to dyskerin. As well as this the A353V mutation is found in patients with dyskeratosis congenita, which could indicate that a larger side chain in this position may not be tolerated. Within $_{351}$TTAVIST$_{357}$ this gives two different stapling positions: $_{351}$**S$_5$**TAV**S$_5$**ST$_{357}$ and $_{351}$T**S$_5$**AVI**S$_5$**T$_{357}$.

Work began with the fluorescently labelled FITC-βAla-Nle**S$_5$**TAV**S$_5$**STCDH-NH_2 (**2.68**) to test as a probe in a FP assay, as well as the unlabelled version of the peptide Ac-Nle**S$_5$**TAV **S$_5$**STCDH-NH_2 (**2.69**). For compatibility with the synthesis of hydrocarbon stapled peptides the Met (**2.70**) residue in the sequence was replaced with a norleucine (Nle, **2.71**). The thioether in methionine can reduce the efficiency of the ruthenium catalyst which can impact the staple formation.[86] Nle mimics the methionine side chain but instead of a sulfur residue contains a carbon (**Figure 2.25**).

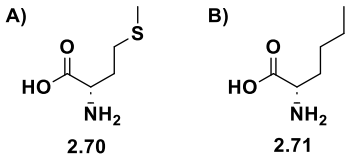


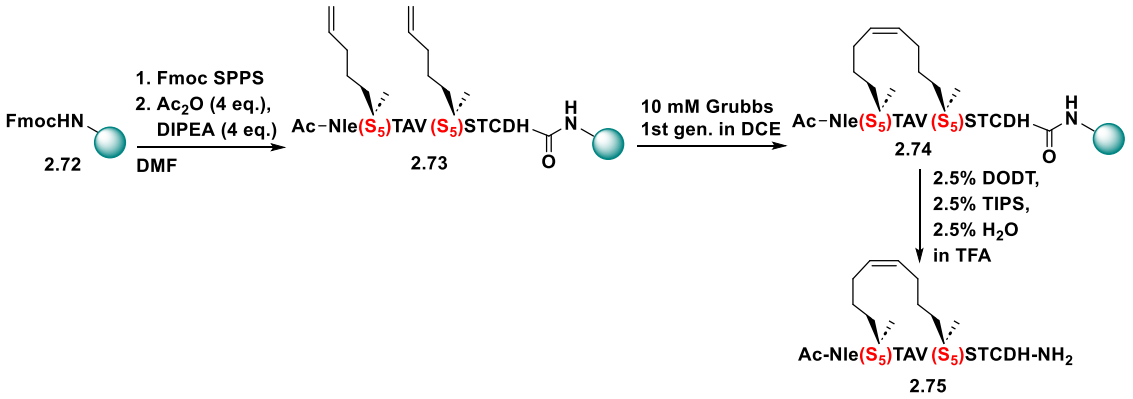
Figure 2.25: *Structures of A) L-Methionine (**2.70**) and B) L-Norleucine (**2.71**)*

In peptide **2.47** (Ac-MTTAVISTCDH-NH_2) methionine is the first residue in the sequence and as part of the native protein it is not a residue on the interacting face, therefore making it unlikely that the replacement affects the binding affinity.

Table 2.4: *Hydrocarbon stapled i, i+4 peptides **2.68** and **2.69***

Nr.	Sequence	Yield (%)	Purity (%)	t_R (min.)	m/z calc.	m/z obs.	Helicity (%)
2.68	FITC-βAla-Nle**S$_5$**TAV**S$_5$**STCDH-NH_2	5	92	14.0	828.3	828.6	n/a
2.69	Ac-Nle**S$_5$**TAV**S$_5$**STCDH-NH_2	5	91	13.8	1237.6	1237.7	34

Peptides **2.68** and **2.69** were synthesised using microwave-assisted SPPS using oxyma and DIC as described in section **2.1.2**. For peptide **2.69** the peptide was fully synthesised and acetylated before the Grubbs metathesis reaction to form the staple was performed (**Scheme 2.6**). The metathesis reaction took place using a solution of 10 mM Grubbs 1st generation catalyst in dichloroethane (DCE) with microwave heating to 40 °C for 30 min which was repeated twice.[86] In this process the two alkene residues react with each other to form a double bond releasing ethene.



Scheme 2.6: *Overview of the SPPS of hydrocarbon stapled peptides*

For the FITC-tagged peptide **2.68** the staple formation took place after the incorporation of Fmoc-β-alanine with the Fmoc protecting group in place. After the staple formation, the peptide was deprotected and FITC was added as for peptide **2.61**. NovaPEG Rink amide resin was first chosen for the synthesis of peptides **2.68** and **2.69** as this was shown to be successful in the synthesis of the linear peptides. Tentagel S RAM resin was later used as an alternative PEG based resin due to supply issues of the NovaPEG Rink amide resin. Both resins gave consistent results with the crude peptide showing high purity and both peptides isolated as a single peak with the correct mass as outlined in **Table 2.4**.

Peptide **2.68** was then tested in a FP assay to explore whether the insertion of the staple improved the binding affinity of the peptide towards dyskerin. The peptide was tested at 5 nM in 10 mM HEPES, 150 mM NaCl, 0.05% Tween-20, pH 7.4 buffer as used for peptide **2.14** (**Figure 2.26**).

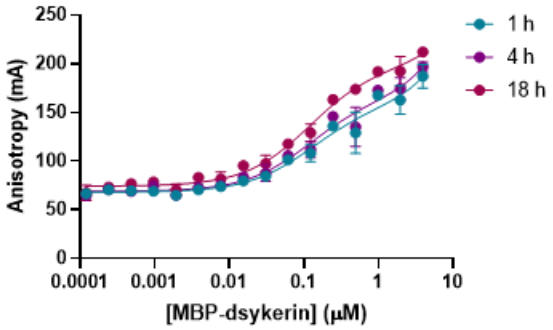


Figure 2.26: *The fluorescence anisotropy measured for peptide* ***2.68*** *(5 nM) in 10 mM HEPES, 150 mM NaCl, 0.05% Tween-20, pH 7.4 recorded at 1 h, 4 h and 18 h timepoints*

The fluorescence polarisation was recorded at 1 h, 4 h and 18 h time points to determine the optimal incubation time for the equilibration of the assay. The hydrocarbon stapled version showed an increased binding affinity towards dyskerin compared to its linear counterpart. The signal size increased with a longer incubation period, but the binding affinity remained similar (**Table 2.5**). Therefore, the impact of the incubation time was minimal and not critical for the experiment.

Table 2.5: *The K_D values measured by FP of peptide **2.68** (5 nM) binding to MBP-dyskerin in 10 mM HEPES, 150 mM NaCl, 0.05% Tween-20 at 1 h, 4 h and 18 h time points*

Time point	K_D (nM) [95% CI]
1 h	125 [73.3 to 214]
4 h	115 [70.9 to 188]
18 h	118 [86.8 to 161]

This K_D value could then be used to develop a competition FP assay in which unlabelled peptides can be titrated into the assay to determine whether they can compete with peptide **2.68**. A value between the K_D and 80% Bmax was chosen, with Bmax being 197 mA at 18 h, giving a range from 118 to 433 nM.[35]

The MBP-dyskerin concentration was set to 300 nM in the competition assay and peptide **2.69** was tested from a top concentration of 200 µM. Peptide **2.69** showed competition with the fluorescent probe with an IC_{50} = 2.06 µM (95% CI = 1.62 to 2.64 µM) (**Figure 2.27**).

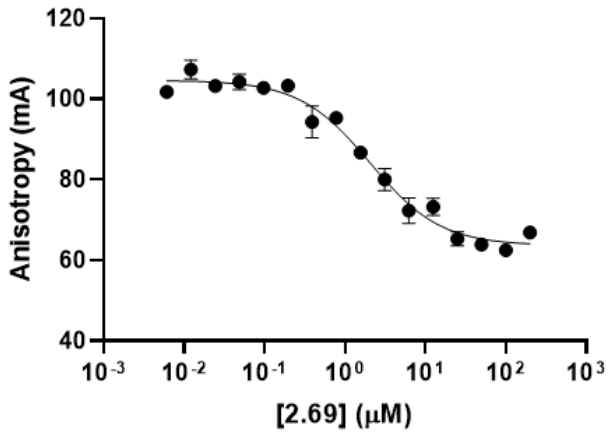


Figure 2.27*: Anisotropy measured over a range of peptide **2.69** concentrations with MBP-dyskerin (300 nM), peptide **2.68** (5 nM) in 10 mM HEPES, 150 mM NaCl, 0.05% Tween-20, pH 7.4*

Peptide **2.69** was also tested using circular dichroism to investigate the effect of the staple insertion on the α-helical structure of the peptide (**Figure 2.28**).

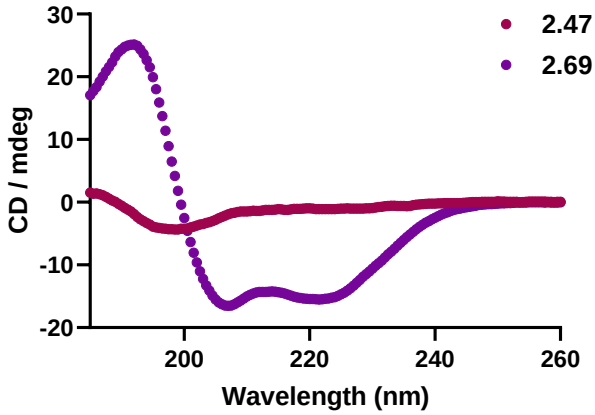


Figure 2.28: *Circular dichroism spectra of peptides **2.47** and **2.69** in 50/50 10 mM phosphate buffer pH 7.4/TFE*

The insertion of the hydrocarbon staple in peptide **2.69** increased the helicity of the peptide from 3% to 34%. Although not a completely helical peptide, this did show a significant increase and now with the competition FP assay in place, it was explored whether the movement of the staple along the peptide could further improve the binding affinity towards MBP-dyskerin.

2.3.4.2 Probing other stapling positions

The data collected for peptide **2.69** showed the potential of using hydrocarbon stapling to increase the binding affinity of the peptide towards dyskerin. Therefore, the staple was introduced into different positions within the sequence creating a staple scan (**Table 2.6**). The residues A353 and T357 were avoided as a stapling position as these residues are on the interacting face of the peptide.

Table 2.6: *The sequences and IC_{50} values recorded in the FP assay for peptides **2.69, 2.76-2.79***

Nr.	Sequence	Yield (%)	Purity (%)	t_R (min.)	m/z calc.	m/z obs.	FP IC_{50} (μM)	Helicity (%)
2.69	Ac-Nle**S$_5$**TAV**S$_5$**STCDH-NH_2	5	91	13.8	1237.6	1237.7	2.06	34
2.76	Ac-NleT**S$_5$**AVI**S$_5$**TCDH-NH_2	12	97	12.4	1263.7	1263.7	5.18	35
2.77	Ac-NleTTA**S$_5$**IST**S$_5$**DH-NH_2	12	98	12.7	1249.7	1249.8	>100	51
2.78	Ac-NleTTAV**S$_5$**STC**S$_5$**H-NH_2	13	65	13.2	1223.7	1224.1	>100	-
2.79	Ac-NleTTAVI**S$_5$**TCD**S$_5$**-NH_2	2	86	13.6	1225.7	1225.6	49.9	-

These peptides were synthesised in the same manner as peptide **2.69** with HPLC and LCMS data confirming the product formation (**Table 2.6**). Peptides **2.76**-**2.79** were tested in the competition FP assay to determine whether their binding was improved compared to **2.69** (**Figure 2.29**).

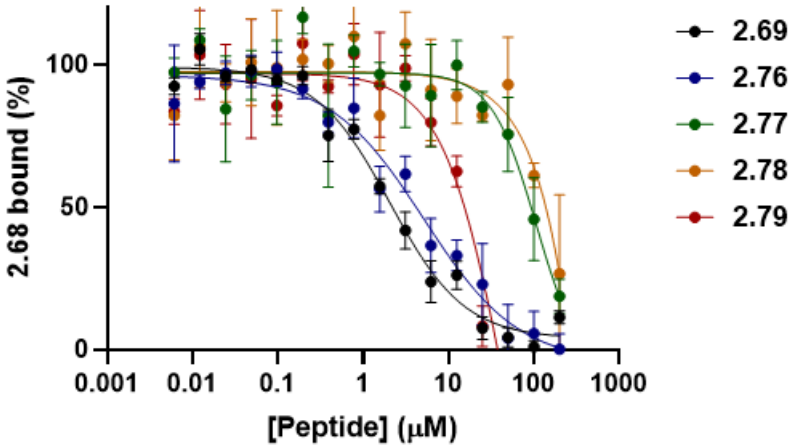


Figure 2.29: *FP competition assay with MBP-dyskerin (300 nM),* ***2.68*** *(5 nM) and* ***2.69*** *(black),* ***2.76*** *(blue),* ***2.77*** *(green)* ***2.78*** *(orange) and* ***2.79*** *(red)*

Peptide **2.76** with the staple within the α-helix showed the highest binding affinity after **2.69** with an IC_{50} of 5.18 μM (95% CI = 3.14 to 10.6 μM). The staple was not tolerated well in peptides **2.77** and **2.78** interfering with the binding. For peptide **2.79** the hydrocarbon staple showed some binding to dyskerin still but weaker than peptide **2.69**. Interestingly the anisotropy signal for **2.79** went lower than the bottom plateau for **2.69** (≈ 60 mA),. to ≈20 mA at 200 μM. This may be due to aggregation or precipitation occurring which lowers the anisotropy.[87] Another question posed around the FP assay was its applicability to study the interaction of a homodimer. Since the dyskerin-dyskerin PPI is an interaction between two of the same proteins, these could be interacting with thesmelves in solution. If there is dimer formation this may then affect the results obtained in the FP assay and therefore it would be important to validate the results in a secondary assay.

The helicity of **2.76** and **2.77** was also explored to understand the relationship between the helical structure of the peptides and their binding affinity.

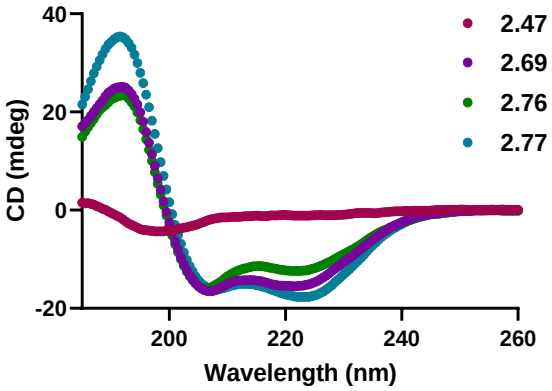


Figure 2.30: *Circular dichroism spectra of peptides **2.47, 2.69, 2.76** and **2.77** in 50/50 10 mM phosphate buffer pH 7.4/TFE*

The introduction of a hydrocarbon staple into the sequence increased the α-helicity to 34% for peptide **2.69**. The movement of the staple toward the C-terminus in peptides **2.76** and **2.77** further increased this to 35% and 51% respectively. **2.77** showed the highest helicity, which could be due to the staple being placed more towards the C-terminus which may have allowed the N-terminus to regain some of its natural helical structure.

2.3.4.3 Introducing mutations into the stapled sequences

The introduction of the hydrocarbon staple into the sequence improved the binding affinity to low micromolar levels, but ideally for therapeutic purposes the binding affinity would need to be improved to nanomolar levels. We applied structure-based drug design and the introduction of mutations into the sequence based on the telomerase cryo-EM structure. Four new sequences were designed and synthesised (peptides **2.80**-**2.83**). A tryptophan residue in place of T352 was explored for potential π-π stacking with W52 (**Figure 2.31A**). The introduction of tryptophan at the end of the sequence (H360W) in place of the histidine could extend the π-π stacking interactions with H68 (**Figure 2.31B**). A T357D mutation could potentially allow for new H-bonding interactions to T67 (**Figure 2.31C**). An A353L mutation was also tried as the alanine side chain sits in a hydrophobic pocket created by W52, P53 and L56 of dyskerin which the leucine side chain could further occupy (**Figure 2.31D**).

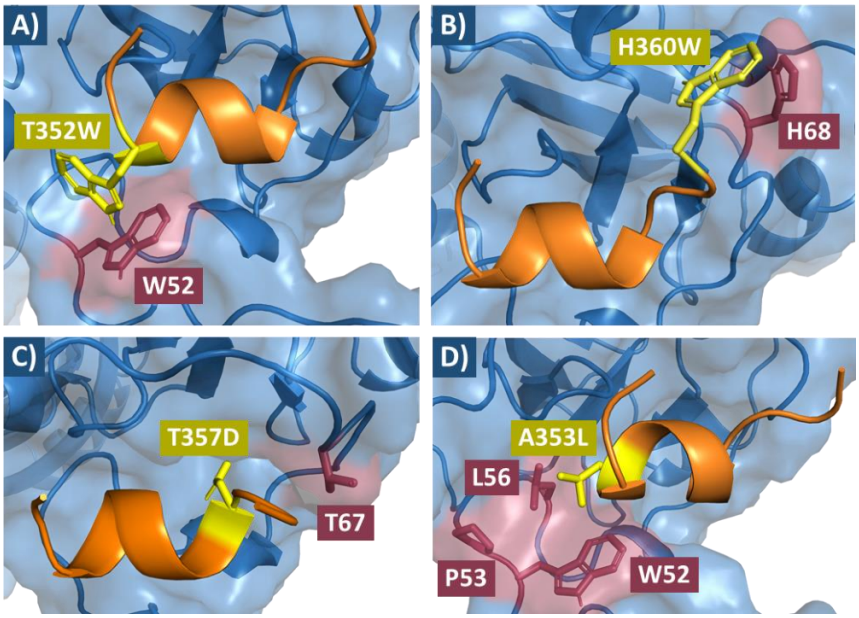


Figure 2.31: *Maestro generated structures of mutant peptides of $_{350}$MTTAVISTCDH$_{360}$ containing A) T352W, B) H360W, C) T357D and D) A353L mutations highlighted in yellow*

We chose to insert the A353L, T357D and H360W mutations into the stapled peptide **2.76**. The difference in binding affinity between **2.76** (IC_{50} = 5.18 µM) and **2.69** (IC_{50} = 2.06 µM) had been small and peptide **2.76** had a higher crude purity leading to a better overall yield, potentially due to the staple position lowering the aggregation in **2.76**. The T352W mutation was inserted into peptide **2.69** since it is in the position of the staple in peptide **2.76**. All peptides were synthesised as the previously discussed for the hydrocarbon stapled peptides leading to the isolation of the product as confirmed by HPLC and LCMS as outlined in **Table 2.7**.

Table 2.7: *Sequences, yields, HPLC retention time and LCMS m/z calculated and observed of peptides **2.80**-**2.83** with their IC_{50} as measured in the FP assay with peptide **2.68** as the fluorescent probe*

Nr.	Sequence	Yield (%)	Purity (%)	t_R (min.)	m/z calc.	m/z obs.	FP IC_{50} (µM)
2.80	Ac-Nle**S**$_5$WAV**S**$_5$STCDH-NH_2	6	90	14.6	1322.7	1322.7	8.38
2.81	Ac-NleT**S**$_5$LVI**S**$_5$TCDH-NH_2	2	95	14.4	1305.7	1305.3	40.9
2.82	Ac-NleT**S**$_5$AVI**S**$_5$DCDH-NH_2	7	97	12.9	1277.7	1277.8	34.9
2.83	Ac-NleT**S**$_5$AVI**S**$_5$TCDW-NH_2	4	95	14.2	656.9	657.4	78.4

These peptides were then tested in the competition FP assay to investigate the effect of the introduced mutation on their binding affinity (**Figure 2.32**).

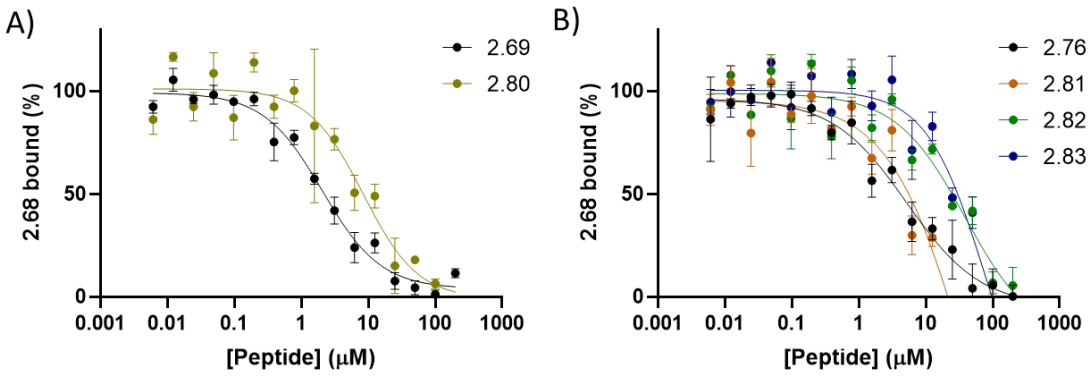


Figure 2.32: *FP spectra of A) peptides **2.69** and **2.80, a**nd B) **2.76, 2.81**-**2.83** in competition with peptide **2.68** (5 nM) and MBP-dyskerin (300 nM) in 10 mM HEPES, 150 mM NaCl, 0.05% Tween-20, pH 7.4 buffer*

The introduction of mutations into the hydrocarbon staple was not successful in further increasing the binding affinity. Peptide **2.80** with a T352W mutation retained activity but did not further improve it. Peptides **2.82** and **2.83** showed a decrease in binding affinity. Although for **2.81** the start of the FP curve looked similar as to **2.80** the anisotropy measured at higher concentration was again lower than **2.69**.

2.3.5 The development of a surface plasmon resonance assay for MBP-dyskerin

As an alternative to the FP assay we explored SPR for measuring the binding affinity of the synthesised peptides to MBP-dyskerin. In the FP assay the MBP-dyskerin protein is free in solution which means that dyskerin could interact with the peptides and with itself, which could interfere with the measurements taken. Since the protein is immobilised in the SPR assay this is less of a concern and therefore SPR was explored as a complimentary assay to the FP assay.

A CM5 chip which contains a matrix of carboxymethylated dextran attached to a gold surface was selected for the SPR assay. This chip type is commonly used for the covalent immobilisation of proteins to the chip using amide coupling, which allows for the conjugation of any proteins through lysine residues. Dyskerin has multiple lysine residues throughout the protein which can be used for this conjugation. Although some of the protein may conjugate to the chip in an orientation in which the binding site is inaccessible, the spread of lysine residues throughout the protein should allow for the protein to also be conjugated with the binding site accessible. This method was chosen over the use of an NTA chip on which the protein can be non-covalently immobilised using its His-tag. When using an NTA chip the interaction between the His-tag on

the protein with the nickel on the chip needs to be strong. If this interaction is weak the protein may dissociate from the chip during the experiment invalidating the results. Since this was unknown for dyskerin and the cost of the chip was high, the use of the CM5 chip was preferred. An additional benefit is that only a very small amount of protein is used with the CM5 chip as the protein only has to be immobilised one time for all experiments, whereas new protein would have to be used for every peptide tested in the case of the NTA chip.

The carboxylic acid groups on the chip surface are first activated using 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) and *N*-hydroxysuccinimide (NHS) after which lysine residues on the protein surface react with the activated ester to form a covalent amide bond to the surface of the chip (**Figure 2.33**).

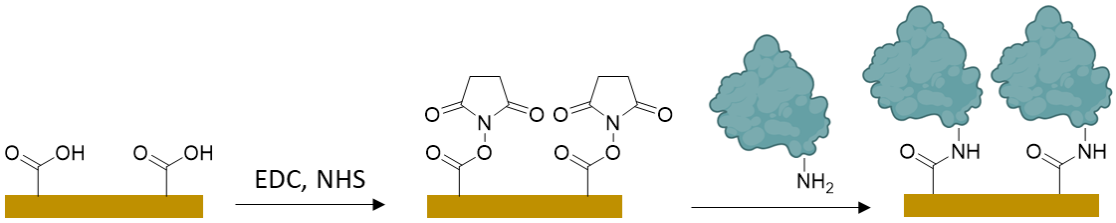


Figure 2.33: *The activation of the SPR chip surface using EDC and NHS after which the protein is immobilised*

The immobilisation level depends on the signal size required and the difference in molecular weight between the ligand (MBP-dyskerin) and the analyte (peptides) through **Equation 2.2**.

$$R_{max} = \frac{MW_A}{MW_L}\ x\ R_L\ x\ S_M$$

Equation 2.2: *The calculation of R_{max} for surface plasmon resonance, R_{max} = the maximum signal, MW_A = Molecular weight of the analyte, MW_L = Molecular weight of the ligand, R_L = Level of immobilised ligand in RU, S_M = Stoichiometry of the analyte-ligand interaction*

The R_{max} is the maximum signal seen when the analyte has fully occupied any binding spaces on the protein (the ligand) and is dependent on the difference in molecular weight of the analyte and the ligand as well as the stoichiometry of their interaction and the immobilisation level of the ligand (MBP-dyskerin in this case). Depending on the purpose of the SPR measurements, i.e. kinetics or affinity the R_{max} can be chosen. For this thesis, the purpose of the measurements was to determine the binding affinity as a complementary assay to the FP assay. The R_{max} desired for affinity measurements tends to be between 30-100 RU.[88] Therefore, a value in the middle of this (≈50 RU) was chosen for the calculation of the immobilisation level. Using the equation with the

molecular weight of peptide **2.47** as an example this gives R_L = 4430 RU and therefore to reach that level the immobilisation level to be aimed for was set to 5000 RU.

Before the immobilisation, pH scouting was used to determine the optimal pH value for a favourable electrostatic interaction between the negatively charged carboxylate groups on the chip surface and the positive charged lysine residues in MBP-dyskerin, which depends on the pI of the protein. The pI is the isoelectric point of the protein, which is the pH value at which the net charge of the protein is zero. For MBP-dyskerin, the pI is 8.0 as calculated through Expasy, an online calculator.[89] In this process the MBP-dyskerin protein was diluted 1 in 10 from its stock solution at 0.85 mg/mL in 10 mM HEPES, 150 mM NaCl, 0.05% Tween-20, pH 7.4 into 10 mM sodium acetate buffer at the following pH levels: 5.5, 5.0, 4.5 and 4.0. These four solutions were then flowed over the chip to assess the pH level most suitable for the immobilisation of the protein to the chip (**Figure 2.34**). In this process the pH level at which there is a positive response larger than the required immobilisation level is wanted.

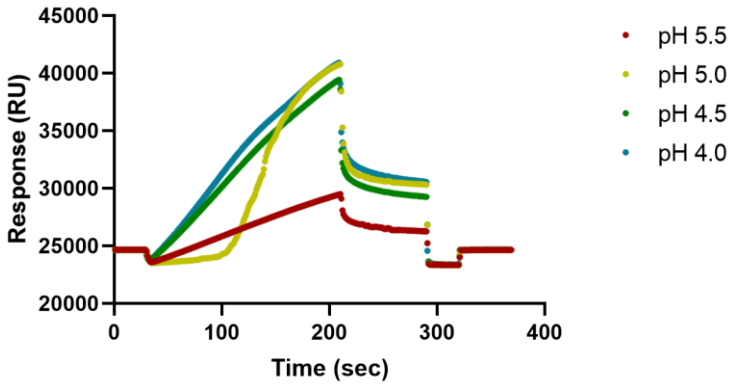


Figure 2.34: *pH scouting of MBP-dyskerin in sodium acetate buffer pH 5.5, 5.0, 4.5 and 4.0*

The sodium acetate buffer at pH 5.0 was chosen as this was the least acidic buffer at which a signal greater than the desired immobilisation level of ≈5000 RU was observed.

After the activation of the chip surface the protein was added in small injections and the chip was washed with buffer in between each injection until the desired immobilisation level was reached. Any remaining carboxylic acid groups on the chip were blocked with ethanolamine, to prevent non-specific binding of the analytes to the chip. A reference surface was also prepared which was activated similarly and then blocked with ethanolamine without the addition of the protein. Both the reference surface and a blank sample are subtracted from the signal seen to

give accurate data. Although the protein immobilisation level had been set to 5000 RU, in the immobilisation process nearly 10000 RU of protein appeared to have been immobilised. The chip was washed with 10 mM sodium hydroxide (NaOH) (5 x 30 sec.) to remove any unbound protein which may only be interacting with the chip electrostatically. In this process a large amount of protein was lost from the chip with the eventual baseline settling back down to an immobilisation level of around 5000 RU. This loss in protein could be due to dyskerin forming a dimer in solution with the dimer disrupted during the NaOH washes.

In the FP assay the best binding peptide had been peptide **2.69**. Therefore, this peptide was chosen as a starting point for the SPR assay. A 1 in 2 dilution series (200 µM to 6.25 µM) was made in a hepes-buffered saline buffer, HBS-EP+ (10 mM HEPES, 150 mM NaCl, 3 mM EDTA, 0.05% Tween-20, pH 7.4) containing 1% DMSO for peptide solubility (**Figure 2.35**).

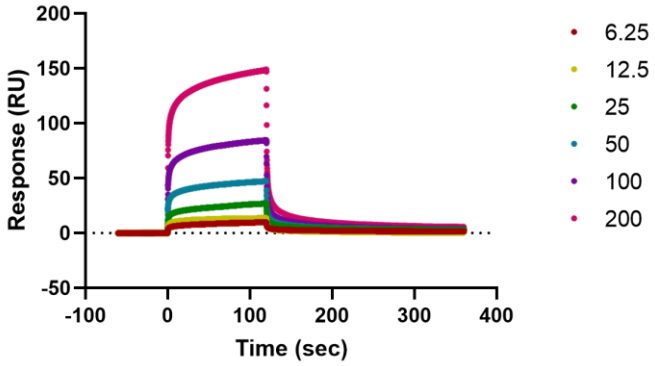


***Figure 2.35**: SPR spectrum of peptide **2.69** in HBS-EP+ with 1% DMSO*

The K_D as determined by the steady state analysis was 712 µM, which falls outside of the measured concentration range. The signal size observed at the higher concentrations (i.e. 140 RU at 200 µM) was also above the R_{max} value of 50 RU. This was concerning as although binding was seen this may mean that there is more than one peptide molecule binding per protein changing the stoichiometry of the interaction and therefore a higher signal could be observed. Alternatively it has also been noted in literature that a change in conformation of the protein can change the response.[90]

To explore whether this potential non-specific binding to the protein could be stopped, different buffer compositions were tried (**Table 2.8**). As an alternative to HBS-EP+, a phosphate-buffered

saline buffer, PBS-P+ (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 0.05% Tween-20, pH 7.4) is commonly used as the running buffer for small molecule analysis.

Table 2.8: *Buffer compositions of buffers 1-5*

Number	Buffer
1	HBS-EP+ + 1% DMSO, pH 7.4
2	HBS-EP+ + 5% DMSO, pH 7.4
3	PBS-P+ + 1% DMSO, pH 7.4
4	PBS-P+ + 5% DMSO, pH 7.4
5	PBS-P+ + 5% DMSO, pH 8.0

HBS-EP+ was tried with a higher percentage DMSO in case these issues are due to limited peptide solubility. PBS-P+ buffer was tried with 1% and 5% DMSO as well as at a higher pH of 8.0. The pI of MBP-dyskerin is 8.0 but dyskerin by itself has a higher pI of 8.6 and therefore a higher pH buffer may be more suitable. The spectra were recorded in each buffer composition with peptide **2.69** (**Figure 2.36**).

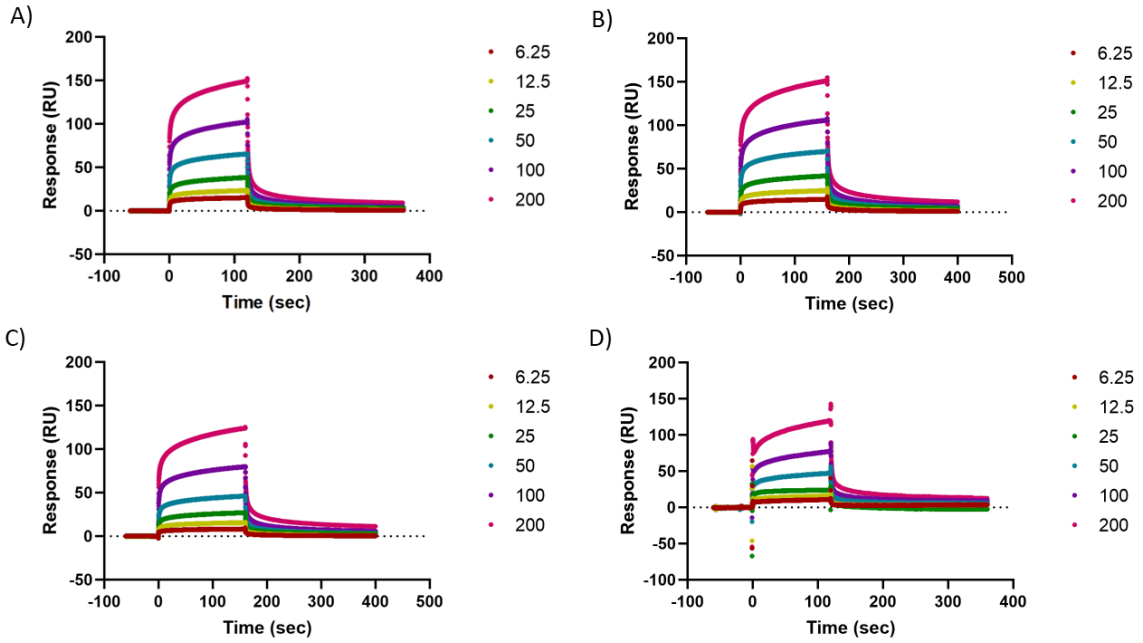


Figure 2.36: *SPR spectra recorded for peptide* ***2.69*** *in A) buffer 2, B) buffer 3, C) buffer 4 and D) buffer 5*

Regardless of the buffer composition the R_{max} value was higher than expected for peptide **2.69** (**Table 2.9**). The best binding affinity (K_D = 149 µM) was found in PBS-P+ with 5% DMSO at pH 7.4.

Table 2.9: *The K_D and R_{max} values for peptide **2.69** in buffers 1-5*

Buffer	K_D (µM)	R_{max} (RU)
1	712	659
2	276	289
3	184	274
4	149	254
5	278	276

The other hydrocarbon stapled peptides were also screened in the different buffers to determine their binding affinity (**Table 2.10**) and to investigate whether there was any trend between their binding affinity in the FP and SPR assays.

Table 2.10: *The FP IC_{50} values and the K_D values measured by SPR for peptides **2.69, 2.76-2.83** in buffers 1-5 (**Table 2.8**) with K_D values below 100 µM in green and between 100-500 µM in orange*

Peptide	FP IC_{50} (µM)	Buffer 1 K_D (µM)	Buffer 2 K_D (µM)	Buffer 3 K_D (µM)	Buffer 4 K_D (µM)	Buffer 5 K_D (µM)
2.69	2.06	712	276	184	149	306
2.76	5.18	>1000	611	>1000	959	>1000
2.77	107	262	>1000	>1000	457	>1000
2.78	>100	-	525	905	330	>1000
2.79	56	-	345	274	310	>1000
2.80	8.38	52	>1000	-	23	>1000
2.81	40.9	64	40	-	82	-
2.82	34.9	267	362	-	177	432
2.83	78.4	>1000	661	-	290	356

From the hydrocarbon stapled peptide series (**2.69, 2.76**-**2.79**) peptide **2.69** still showed the highest binding affinity as found in the FP assay. However, the binding affinity of peptide **2.76** was the lowest in the SPR assay even though it was the second-best peptide in the FP assay. For both **2.69** and **2.76**, the signals recorded extended above the R_{max} therefore indicating there may

be an element of non-specific binding. For peptides **2.77**, **2.78** and **2.79** this was not the case with an expected R_{max} value recorded, with all three peptides recording similar binding affinities.

Although in the FP assay the introduction of mutations into the sequence did not seem to improve the binding affinity, the opposite is true in the SPR assay where especially peptides **2.80** and **2.81** show a great increase in binding affinity in at least some of the buffers. However, the same issue of the R_{max} signal being larger than expected is still observed regardless of buffer composition. For example, for peptide **2.80** the highest binding affinity was seen for buffer 4 (K_D of 23 µM), but the R_{max} was 608 RU. For peptide **2.81** the highest binding affinity was seen in buffer 2 (K_D of 40 µM), but the R_{max} value was 258 RU. For **2.82** the highest binding affinity was observed in buffer 4 and the R_{max} is closer to what is expected (R_{max} = 77 RU). Interestingly this is the peptide with a mutation to a charged residue lowering the overall hydrophobicity of the peptide, which could be contributing to the non-specific binding of some of the other peptides.

Another explanation of the high signal sizes could be that the reference surface was not properly accounting for the effect of the peptide on the surface. The reference surface does not contain a protein but instead has ethanolamine on the surface which is a small molecule. At the high peptide concentrations needed in the assay due to the weak binding of the peptides the bulk effect of the peptide solution on the protein may be different from the reference surface potentially leading to a bulk contribution seen in the spectrum. This was thought to be potentially occurring due to the shape of the spectra with a bulk contribution to the spectrum leading to a squarer shaped spectrum (**Figure 2.37**).

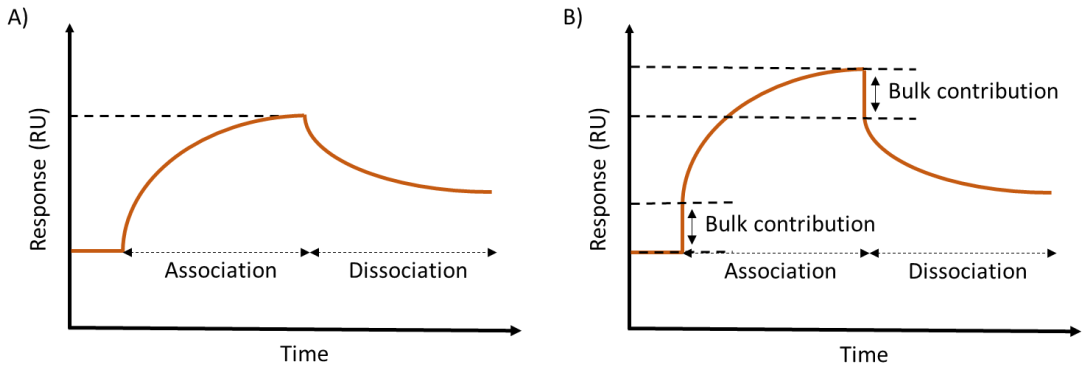


Figure 2.37: *An SPR spectrum A) without bulk contribution and B) with bulk contribution*

2.3.6 The development of an HTRF assay to investigate the inhibition of the dyskerin-dyskerin PPI

The FP and SPR assays investigate the binding of the peptides to dyskerin, but to further explore their ability to disrupt the dyskerin-dyskerin protein-protein interaction a HTRF assay was developed. This assay makes use of two anti-6His antibodies, one conjugated to a Tb cryptate donor and the other conjugated to a d2 acceptor, a small molecule of ≈1000 Da.[91,92] When the dyskerin-dyskerin dimer forms, these antibodies come into close enough proximity for Förster resonance energy transfer (FRET) to occur. As the dimer formation is inihibited by the peptides, the antibodies are no longer in close proximity and the FRET signal is lost (**Figure 2.38**).

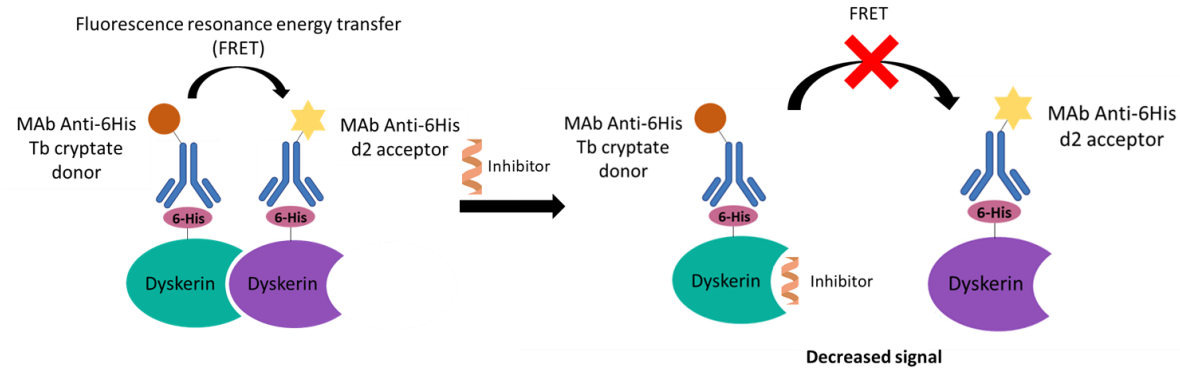


Figure 2.38: *Schematic overview of the HTRF assay developed for the dyskerin-dyskerin PPI*

The dyskerin-dyskerin interaction is a homodimer interaction and two anti-his antibodies were used. This may mean that for some of the dimers this will bring two donor or two acceptor antibodies together instead of a donor with an acceptor. However, this was not thought to be a problem for the assay as it was thought they would all be inhibited equally. The use of two of the same anti-tag antibodies for the use in HTRF has also been reported by Revvity.[93,94]

Work began on the HTRF assay with the determination of the optimal protein concentration for the largest assay window. In a HTRF assay the ratio of the acceptor emission over the donor emission is taken and reported according to **Equation 2.3**.

$$HTRF\ ratio = \frac{Signal\ at\ 665\ nm}{Signal\ at\ 620\ nm}\ x\ 10^4$$

Equation 2.3: *Calculation of the HTRF ratio*

From a top dyskerin protein concentration of 5.7 µM a 1 in 2 series dilution was prepared and tested in the assay (**Figure 2.39**).

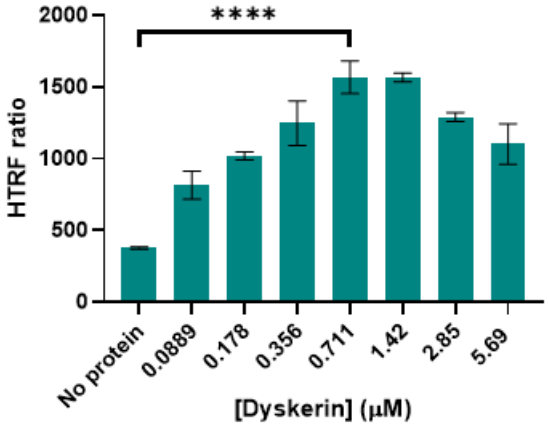


Figure 2.39: *HTRF ratio measured at different dyskerin protein concentrations after 18 h incubation at 4 °C in 10 mM HEPES, 150 mM NaCl, 0.05% Tween-20, pH 7.4*

Measurements were taken at 1 h, 2 h and 3 h with the plate kept at 4 °C between measurements. For the use of the Tb cryptate donor antibody it is key to keep the plate at 4 °C for the FRET signal and therefore the plate was kept on ice until just before the plate reading. After leaving the plate at room temperature for 10 min, the signal has decreased nearly back down to the baseline level of no protein present (**Figure 2.40B**). From this data the largest assay window was observed at both 1.42 µM and 0.71 µM (**Figure 2.39**). At higher protein concentrations the signal went down again, known as the Hook effect. This effect occurs when the concentration of protein is much higher than the antibody concentration therefore saturating both antibodies, lowering the signal seen.[95] The signal was also shown to be stable between 1 and 3 hours (**Figure 2.40A**).

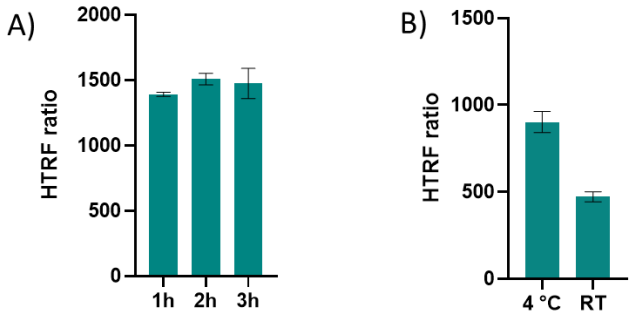


Figure 2.40: *HTRF Signal size measured at A) different time points of an MBP-dyskerin concentration of 0.711 µM and B) different temperatures of an MBP-dyskerin concentration of 0.711 µM at the 3 h time point*

To minimise the protein used a concentration of 0.711 µM was chosen for the screening of the peptides synthesised. The linear peptides **2.47**, **2.49** and **2.50** as well as the alanine scan peptides **2.51**-**2.60** were tested in the assay at 200 µM with an incubation time of 1 h.

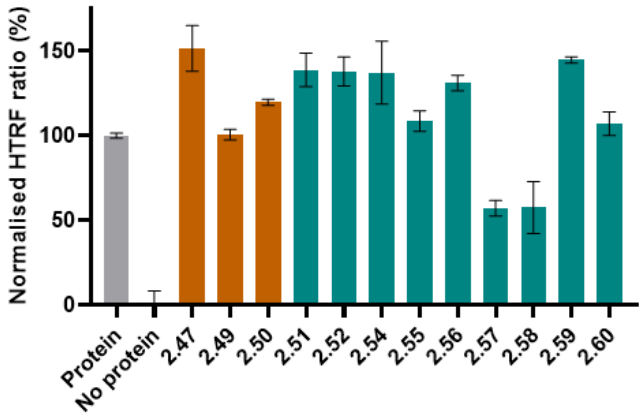


Figure 2.41: *Normalised HTRF ration in the presence of 200 µM of linear peptides **2.47, 2.49** and **2.50** (orange) and alanine scan peptides **2.51**-**2.60** (blue) at 0.711 µM dyskerin concentration after 1 h incubation at 4 °C in 10 mM HEPES, 150 mM NaCl, 0.05% Tween-20, pH 7.4*

Peptides **2.47**, **2.49** and **2.50** showed any effect on the dyskerin-dyskerin dimer formation. This was not unexpected as the binding affinity of the FITC labelled linear peptide (**2.61**) had been low in the FP assay. The alanine scan synthesised in section **2.3.2** was also evaluated in the HTRF assay. The parent peptide **2.47** had not shown any effect on the dimer formation, but there are some examples in literature of alanine mutations shown to increase the binding of the peptide.[96,97] The peptides were screened in the assay at 200 µM to assess their ability to disrupt the dimer formation (**Figure 2.41**). Interestingly in the assay two of the alanine scan peptides did show an effect on dimer formation with a 43 ± 5% and 42 ± 15% reduction in dimer formation at 200 µM seen for peptides **2.57** and **2.58** respectively. This could potentially mean that the

introduction of an alanine residue in these sequences increased the affinity of the peptides towards dyskerin which can now partially inhibit the interaction. For some of the peptides (e.g **2.47**) the signal observed was larger than the protein control signal. This could potentially indicate the peptide stabilising the dimer formation, but this experiment was only repeated once in triplicate and further repetitions would have to be done to confirm this effect.

The hydrocarbon stapled peptide series had shown an increase in binding affinity in the FP assay and therefore were thought to maybe be able to inhibit the PPI. The staple scan (**2.69, 2.76**-**2.79**) was tested in the HTRF assay at 200 µM (**Figure 2.42**).

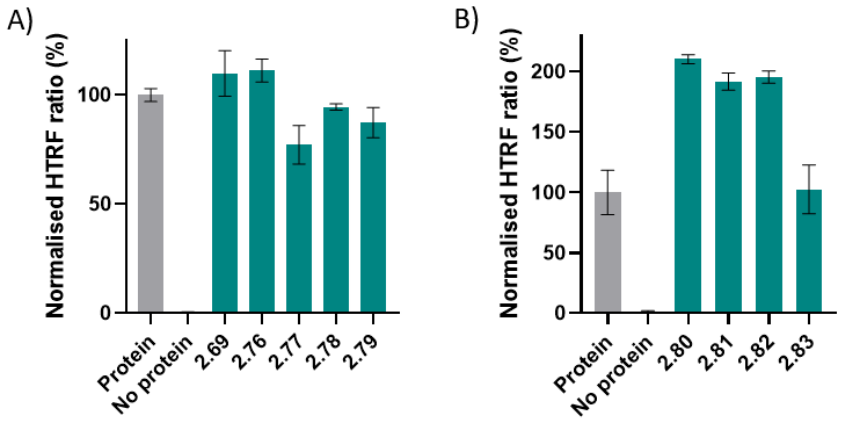


Figure 2.42: *Normalised HTRF ratio in the presence of A) 200 µM of peptides **2.69, 2.76**-**2.79** and B) 100 µM of peptides **2.80**-**2.83** at 0.711 µM dyskerin concentration after 1 h incubation at 4 °C in 10 mM HEPES, 150 mM NaCl, 0.05% Tween-20, pH 7.4*

Interestingly the only inhibition of the dyskerin-dyskerin dimer was seen for peptide **2.77** (Ac-NleTTA**S$_5$**IST**S$_5$**DH-NH_2) with a slight inhibition (23 ± 9%) at 200 µM. **2.77** was not the peptide that had shown the best binding affinity in either the FP or SPR assay, which could further confirm that the other hydrocarbon stapled peptides are binding to MBP-dyskerin non-specifically instead of at the dyskerin-dyskerin dimer interface. The hydrocarbon stapled peptides with mutations (**2.80**-**2.83**) were screened in the assay at 100 µM due to limited peptide availability. Peptides **2.80** and **2.81** had shown the best binding affinity in the SPR assay and therefore it was of interest to see whether this would translate to the HTRF assay. However, none of these peptides showed inhibition of the dyskerin-dyskerin PPI (**Figure 2.42B**).

Out of all the linear and hydrocarbon stapled peptides two peptides from the alanine scan (**2.57** and **2.58**) and one hydrocarbon stapled peptides (**2.77**) showed the ability to partially inhibit the

dyskerin-dyskerin PPI. However, this was not in line with the fluorescence polarisation assay data collected. Due to the hydrophobic nature of the hydrocarbon stapling, the peptides were thought to bind dyskerin non-specifically. To investigate this further we explored a less hydrophobic stapling method, lactam stapling.

2.3.7 The design, synthesis, and evaluation of lactam stapled peptides

A lactam staple is formed through the reaction of a lysine side chain with an aspartic or glutamic acid to form an amide bond. Amide bonds are more hydrophilic in character due to their ability to form hydrogen bonds compared to the hydrocarbon staple which cannot do so.

Lactam stapled peptides can be synthesised on resin through the introduction of a specific side chain-protected aspartic acid (Asp) or glutamic acid (Glu) and a lysine (Lys). The side chain of Asp/Glu can be protected as a 2-phenylisopropyl ester (O-2-PhiPr) group and Lys with a 4-methyltrityl (Mtt) group (**Figure 2.43**). These amino acids are commercially available and both side chain protecting groups can be removed with 2.25% TFA and 5% TIPS in DCM. This low percentage of TFA does not remove any other protecting groups (*t*Bu, Trt, Boc and Pbf) and does not cleave the peptide from the Tentagel S Ram resin used.

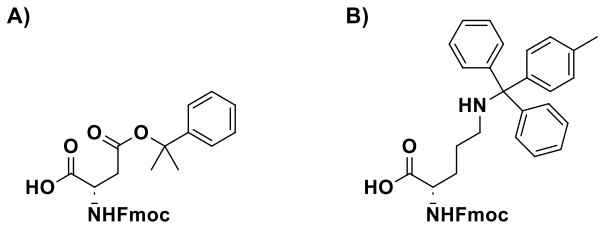
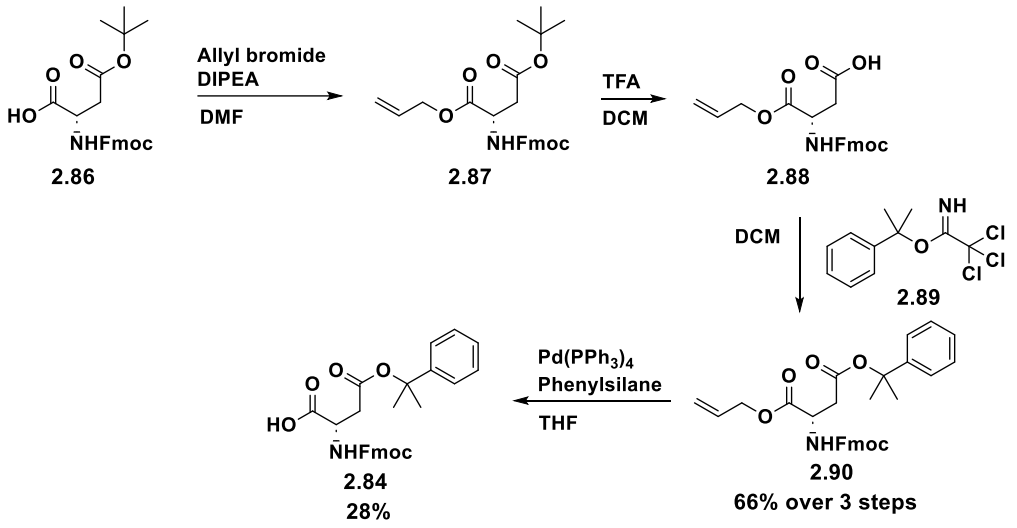


Figure 2.43: *Structures of A) Fmoc-Asp(O-2-PhiPr)-OH (**2.84**) and B) Fmoc-Lys(Mtt)-OH (**2.85**)*

Fmoc-Asp(O-2PhiPr)-OH (**2.84**) is commercially available but expensive and so we synthesised this amino acid in house as outlined in **Scheme 2.7**. The amino acid can be prepared from Fmoc-Asp(O*t*Bu)-OH (**2.86**). The main chain carboxylic acid is protected with an allyl group (**2.87**), which is followed by a deprotection of the side chain *t*Bu group (**2.88**). This allows for the protection with the 2-phenylisopropanol group (**2.90**). Finally, the main chain carboxylic acid is deprotected to form the final amino acid (**2.84**).



Scheme 2.7: *Synthesis of Fmoc-L-Asp(O-2PhiPr)-OH (**2.84**) from Fmoc-L-Asp(O^{t}Bu)-OH (**2.86**)*

The synthesis followed was adapted from the literature synthesis of Fmoc-β-Me-Asp(PhOiPr)-OH.[98] In this synthesis the product is not purified until after the incorporation of the O-2PhiPr group with only aqueous work up steps before then. In the first step excess DIPEA is removed from the reaction mixture through washes with 0.1 M HCl, followed by ^{t}Bu removal using TFA. Excess TFA is removed *in vacuo* after which the O-2PhiPr protecting group is installed with 1-methyl-1-phenylethyl-2,2,2-trichloroethanimidate (**2.89**). **2.89** was prepared from 2-phenylisopropanol (**2.91**) with trichloroacetonitrile using sodium hydride, and used directly without purification following literature procedures.[99]

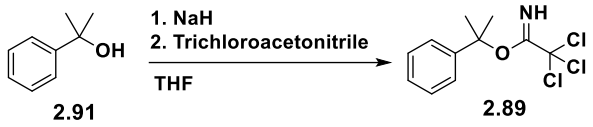


Figure 2.44: *Synthesis of 1-methyl-1-phenylethyl-2,2,2-trichloroethanimidate (**2.89**) from 2-phenylisopropanol (**2.91**)*

Fmoc-L-Asp(O-2PhiPr)-OAllyl was purified by normal phase flash column chromatography to give a 66% yield over three steps as confirmed by ^{1}H NMR which showed the presence of both the allyl (three protons in the 5.83-5.22 ppm region) and O-2PhiPr protecting groups (six protons at 1.77 ppm). The allyl group was then removed using $Pd(PPh_3)_4$ and phenylsilane giving Fmoc-L-Asp(O-2PhiPr)-OH in 28% yield with the O-2PhiPr peaks observed in the ^{1}H NMR but no allyl peaks found (no peaks seen in 5-6 ppm region other than one proton at 5.78 ppm for the NH). A major side product was the product without the O-2PhiPr protecting group, potentially having

been cleaved during purification in acidic silica.

Now having access to the amino acids building blocks, a FITC tagged lactam stapled peptide (**2.92**), FITC-βAla-NleDTAVKSTCDH-NH_2, was designed. Three unlabelled lactam stapled peptides were also designed with the staple inserted in three different positions within the sequence (**Table 2.11**). Unlike the hydrocarbon staple the lactam staple is not symmetrical and the peptide could be synthesised with a DXXXK or KXXXD arrangement. This was briefly explored, but the synthesis of the KXXXD sequence was less successful with a high levels of impurities seen, perhaps due to increased aggregation.

Table 2.11: *The sequence of lactam stapled peptides* ***2.92-2.95***

Nr.	Sequence	Yield (%)	Purity (%)	t_R (min.)	m/z calc.	m/z obs.
2.92	FITC-βAla-Nle**D**TAV**K**STCDH-NH_2	7	87	11.5	816.3	816.1
2.93	Ac-Nle**D**TAV**K**STCDH-NH_2	8	91	10.9	1212.6	1212.6
2.94	Ac-NleT**D**AVI**K**TCDH-NH_2	2	86	12.0	1236.6	1237.0
2.95	Ac-NleTTA**D**IST**K**DH-NH_2	13	99	9.8	1224.6	1224.7

The peptides were synthesised using automated peptide synthesis until the Lys and Asp residues were incorporated. The first step in staple formation was the selective deprotection of the Lys and Asp protecting groups using 2.25% TFA and 5% Tips in DCM. The staple was then formed through an amide bond formation using benzotriazol-1-yloxytripyrrolidinophosphonium hexafluorophosphate (PyBOP) (**Figure 2.45A**) and DIPEA. PyBOP is the choice for the cyclisation over the use of HBTU due to the possible guanidinium side product (**Figure 2.45B**) which can form on the lysine if HBTU is used in excess of the acid it activates. PyBOP does not have this problem and therefore an excess of the reagent can be used driving the staple formation to completion.

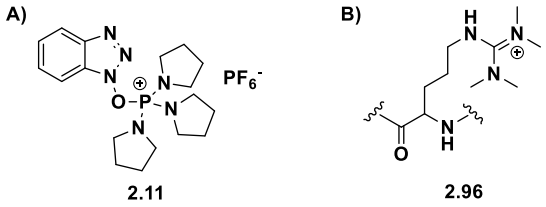
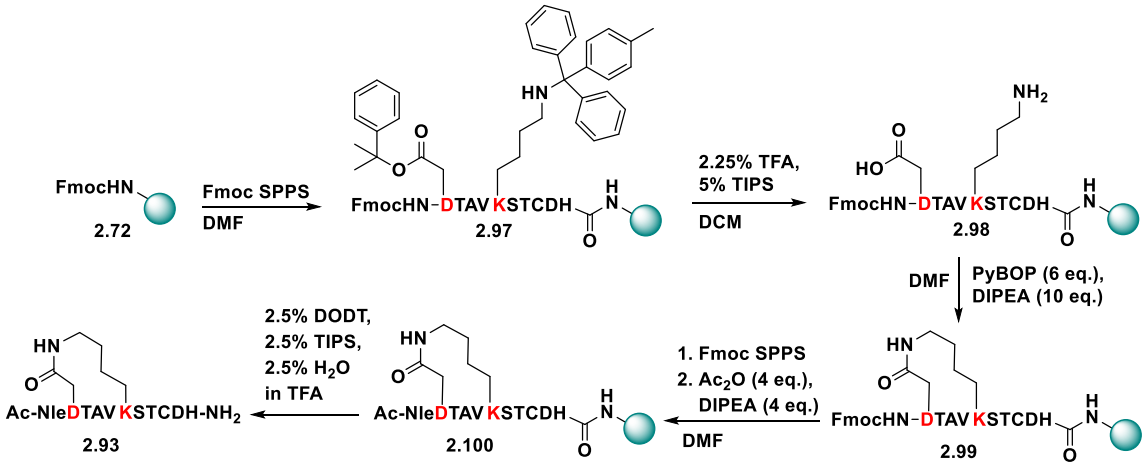


Figure 2.45: *Structures of A) PyBOP (**2.11**) and B) the Lys guanidinium side product (**2.96**)*

The synthesis of the peptide was then continued starting with an Fmoc deprotection followed by the coupling of any remaining amino acids. The peptide was cleaved and purified, and confirmed by observation of a single peak in the HPLC spectrum (t_R = 10.9 min.) and the correct mass by LCMS (m/z calc. = 816.8 $[M+2H]^{2+}$, m/z obs. = 816.1).



Scheme 2.8: *Overview of the SPPS of lactam stapled peptide* ***2.93***

The fluorescent probe, **2.92**, was then tested in the FP assay to investigate its binding affinity towards dyskerin (**Figure 2.46**).

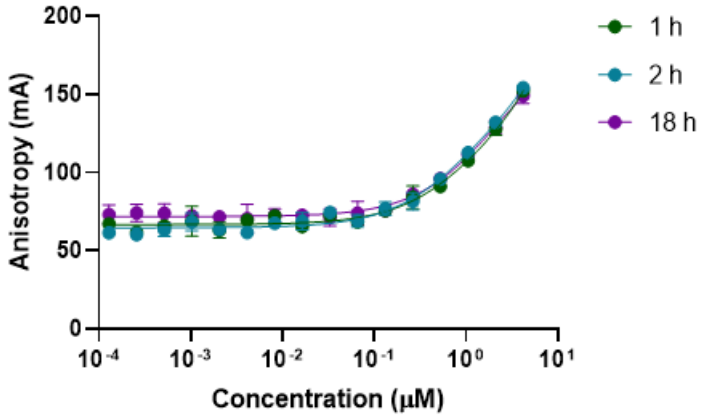


Figure 2.46: *The fluorescence anisotropy measured for peptide* ***2.92*** *(5 nM) in 10 mM HEPES, 150 mM NaCl, 0.05% Tween-20, pH 7.4 recorded at 1 h, 2 h and 18 h timepoints*

This peptide showed an increased binding affinity towards MBP-dyskerin compared to the linear peptide, but the binding affinity was lower than the hydrocarbon stapled peptide with the determination of a K_D value not possible. The anisotropy was measured after different incubation

times of 1 h, 2 h and 18 h with minimal change in the signal observed. Due to the relatively low binding affinity of this fluorescent peptide, this was still not an ideal probe for an FP assay.

The peptides were then screened at 200 µM in the HTRF assay to determine whether they were able to inhibit the dyskerin-dyskerin PPI (**Figure 2.47**).

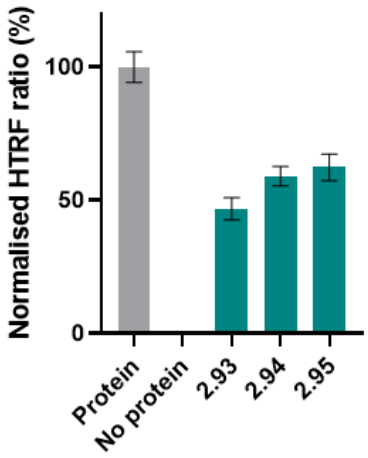


Figure 2.47: *Normalised HTRF ratio in the presence of 200 µM of peptides* ***2.93-2.95*** *at 0.711 µM dyskerin concentration after 18 h incubation at 4 °C in 10 mM HEPES, 150 mM NaCl, 0.05% Tween-20, pH 7.4*

The lactam stapled peptides were better inhibitors of the dyskerin-dyskerin PPI than the hydrocarbon stapled versions with activity seen in the HTRF assay for all three peptides, with peptide **2.93** having the greatest effect (53 ± 4% inhibition at 200 µM) over **2.94** (41 ± 4%) and **2.95** (38 ± 5%). This showed that it is possible to inhibit the dyskerin-dyskerin PPI with lactam stapled peptides with higher activity observed in this series compared to the hydrocarbon stapled peptides.

Peptides **2.93**-**2.95** were further explored in the SPR assay to see whether it was possible to determine a K_D value for these peptides. Two buffers were explored, buffer 2 (HBS-EP+ + 5% DMSO, pH 7.4) and buffer 4 (PBS-P+ + 5% DMSO, pH 7.4) which had been found to give better results than their counterparts with 1% DMSO for the hydrocarbon stapled peptides in **Section 2.3.5**. However, for the lactam stapled peptides no good curves were observed in the PBS-P+ buffer and therefore only the results in HBS-EP+ are discussed here (**Table 2.12**).

Table 2.12: *SPR K_D values recorded in HBS-EP+ + 5% DMSO, pH 7.4 buffer and their inhibition of the dyskerin-dyskerin PPI as measured by HTRF assay at 200 µM*

Nr.	HBS-EP+ pH 7.4 + 5% DMSO, K_D (µM)	R_{max} (RU)	HTRF inhibition at 200 µM (%)
2.93	142 ± 12	147 ± 5.2	53 ± 4
2.94	35.5 ± 11	91.7 ± 7.5	41 ± 4
2.95	44.1 ± 9.0	89.2 ± 5.5	38 ± 5

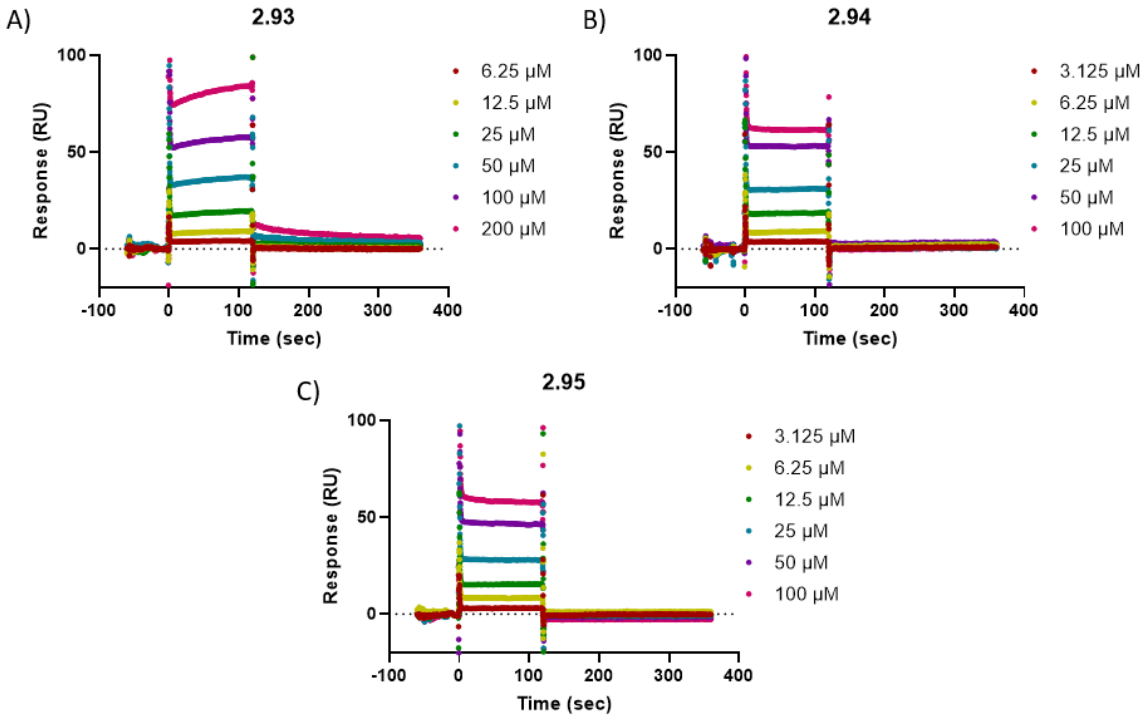


Figure 2.48: *SPR spectra of peptides A)* ***2.93****, B)* ***2.94*** *and C)* ***2.95*** *in HBS-EP+ pH 7.4 + 5% DMSO*

From the SPR results (**Figure 2.48**), peptide **2.94** displayed the strongest binding affinity towards MBP-dyskerin (K_D = 35.5 µM) followed by peptide **2.95** (K_D = 41.1 µM) and finally **2.93** (K_D = 142 µM). These peptides all presented similar activity in the HTRF assay with **2.93** showing a slightly higher activity, but this was not reflected in the SPR data. However, the difference in K_D between the compounds was small and the Rmax values were larger than expected as was observed for the hydrocarbon stapled peptides discussed in **Section 2.3.5** and therefore further optimisation of the SPR assay may still be needed.

2.4 Conclusion and Future work

In this chapter the first known peptide inhibitors of the dyskerin-dyskerin PPI have been reported. The Cryo-EM structure of dyskerin was analysed and the section $_{350}$MTTAVISTCDH$_{360}$ was selected as a starting point for the design of peptides targeting this PPI and the SPPS of this peptide (**2.47**) was optimised with longer peptides (**2.49** and **2.50**) and an alanine scan (**2.51-2.60**) synthesised. Using a fluorescence polarisation assay it was shown that peptide (**2.47**) weakly bound to dyskerin and had low α-helicity (3%). Peptide stapling was explored to improve these properties with hydrocarbon stapling first attempted. A hydrocarbon stapled fluorescent probe (**2.68**) showed a significant increase in binding to dyskerin (K_D = 118 nM). Alternative stapling positions were explored (**2.76-2.79**) as well as the introduction of mutations into the sequence (**2.80-2.83**), but the binding affinity was not improved. A HTRF assay was developed to investigate the ability of the peptides to inhibit the dyskerin-dyskerin PPI. At 200 μM, two alanine scan peptides (**2.57** and **2.58**) and one hydrocarbon stapled peptides (**2.77**) partially inhibited the dyskerin-dyskerin PPI. Hydrocarbon stapling was also shown to increase the α-helicity of the peptides from 3% for the linear peptide **2.47** to 34%, 35% and 51% for hydrocarbon stapled peptide **2.69**, **2.76** and **2.77** respectively. However, the best binding peptide (**2.69**) in the FP assay of the hydrocarbon stapled series did not have an effect in the HTRF assay and this was thought to be due to non-specific binding of the peptide to dyskerin because of its hydrophobic character. Therefore, lactam stapling was used as an alternative to hydrocarbon stapling with three stapled peptides synthesised (**2.93**-**2.95**). All three peptides were able to disrupt the dimer formation with **2.93** showing the greatest inhibition (53% inhibition at 200 μM), giving a total of six peptides capable of disrupting the dyskerin-dyskerin PPI.

Future work will focus on the development of the lactam stapled series including the introduction of mutations in the sequence to further improve the binding affinity, such as the combination of the active alanine scan peptides with a lactam staple. Additionally, the development of peptides able to control the intrinsically disordered section of the dyskerin-dyskerin interaction are to be explored. It may be that a synergistic affect is observed for molecules able to control both of these interaction sites.

2.5 References

1 R. B. Merrifield, Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide, *J. Am. Chem. Soc.*, 1963, **85**, 2149–2154.

2 R. Sheppard, The fluorenylmethoxycarbonyl group in solid phase synthesis, *J. Pept. Sci.*, 2003, **9**, 545–552.

3 K. B. Jadhav, K. J. Woolcock and M. Muttenthaler, Anhydrous hydrogen fluoride cleavage in Boc solid phase peptide synthesis, *Methods Mol. Biol. Clifton NJ*, 2020, **2103**, 41–57.

4 C.-D. Chang and J. Meienhofer, Solid-phase peptide synthesis using mild base cleavage of Nαfluorenylmethyloxycarbonyl amino acids, exempllified by a synthesis of dihydrosomatostatin, *Int. J. Pept. Protein Res.*, 1978, **11**, 246–249.

5 L. A. Carpino, The 9-fluorenylmethyloxycarbonyl family of base-sensitive amino-protecting groups, 1987, **20**, 401–407.

6 T. B. McMahon and P. Kebarle, Acidities of carbon and nitrogen acids: the aromaticity of the cyclopentadienyl anion, 1976, **98**, 3399–3406.

7 D. A. T. Pires, M. P. Bemquerer and C. J. do Nascimento, Some Mechanistic Aspects on Fmoc Solid Phase Peptide Synthesis, *Int. J. Pept. Res. Ther.*, 2014, **20**, 53–69.

8 O. F. Luna, J. Gomez, C. Cárdenas, F. Albericio, S. H. Marshall and F. Guzmán, Deprotection Reagents in Fmoc Solid Phase Peptide Synthesis: Moving Away from Piperidine?, *Molecules*, 2016, **21**, 1542.

9 W. König and R. Geiger, A new method for synthesis of peptides: activation of the carboxyl group with dicyclohexylcarbodiimide using 1-hydroxybenzotriazoles as additives, *Chem. Ber.*, 1970, **103**, 788–798.

10 J. M. Collins, K. A. Porter, S. K. Singh and G. S. Vanier, High-Efficiency Solid Phase Peptide Synthesis (*HE* -SPPS), *Org. Lett.*, 2014, **16**, 940–943.

11 S. L. Pedersen, A. Pernille Tofteng, L. Malik and K. J. Jensen, Microwave heating in solid-phase peptide synthesis, *Chem. Soc. Rev.*, 2012, **41**, 1826–1844.

12 K. D. Wehrstedt, P. A. Wandrey and D. Heitkamp, Explosive properties of 1-hydroxybenzotriazoles, *J. Hazard. Mater.*, 2005, **126**, 1–7.

13 J. B. Sperry, C. J. Minteer, J. Tao, R. Johnson, R. Duzguner, M. Hawksworth, S. Oke, P. F. Richardson, R. Barnhart, D. R. Bill, R. A. Giusto and J. D. I. Weaver, Thermal Stability Assessment of Peptide Coupling Reagents Commonly Used in Pharmaceutical Manufacturing, *Org. Process Res. Dev.*, 2018, **22**, 1262–1275.

14 S. A. Palasek, Z. J. Cox and J. M. Collins, Limiting racemization and aspartimide formation in microwave-enhanced Fmoc solid phase peptide synthesis, *J. Pept. Sci.*, 2007, **13**, 143–148.

15 J. A. Moss, Guide for Resin and Linker Selection in Solid-Phase Peptide Synthesis, *Curr. Protoc. Protein Sci.*, 2005, **40**, 1–19.

16 M. Conda-Sheridan and M. Krishnaiah, in *Peptide Synthesis: Methods and Protocols*, eds. W. M. Hussein, M. Skwarczynski and I. Toth, Springer US, New York, NY, 2020, pp. 111–128.

17 M. Paradís-Bas, J. Tulla-Puche and F. Albericio, The road to the synthesis of "difficult peptides", *Chem. Soc. Rev.*, 2016, **45**, 631–654.

18 S. Srisailam, T. K. S. Kumar, T. Srimathi and C. Yu, Influence of Backbone Conformation on Protein Aggregation, *J. Am. Chem. Soc.*, 2002, **124**, 1884–1888.

19 A.-M. Fernandez-Escamilla, F. Rousseau, J. Schymkowitz and L. Serrano, Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins, *Nat. Biotechnol.*, 2004, **22**, 1302–1306.

20 S. Mohapatra, N. Hartrampf, M. Poskus, A. Loas, R. Gómez-Bombarelli and B. L. Pentelute, Deep Learning for Prediction and Optimization of Fast-Flow Peptide Synthesis, *ACS Cent. Sci.*, 2020, **6**, 2277–2286.

21 F. García-Martín, M. Quintanar-Audelo, Y. García-Ramos, L. J. Cruz, C. Gravel, R. Furic, S. Côté, J. Tulla-Puche and F. Albericio, ChemMatrix, a poly(ethylene glycol)-based support for the solid-phase synthesis of complex peptides, *J. Comb. Chem.*, 2006, **8**, 213–220.

22 M. Erdélyi and A. Gogoll, Rapid Microwave-Assisted Solid Phase Peptide Synthesis, *Synthesis*, 2002, 1592–1596.

23 J. P. Tam and Y. A. Lu, Coupling Difficulty Associated with Interchain Clustering and Phase Transition in Solid Phase Peptide Synthesis, *J. Am. Chem. Soc.*, 1995, **117**, 12058–12063.

24 L. K. Mueller, A. C. Baumruck, H. Zhdanova and A. A. Tietze, Challenges and Perspectives in Chemical Synthesis of Highly Hydrophobic Peptides, *Front. Bioeng. Biotechnol.*, 2020, **8**, 1–17.

25 V. Cardona, I. Eberle, S. Barthélémy, J. Beythien, B. Doerner, P. Schneeberger, J. Keyte and P. D. White, Application of Dmb-Dipeptides in the Fmoc SPPS of Difficult and Aspartimide-Prone Sequences, *Int. J. Pept. Res. Ther.*, 2008, **14**, 285–292.

26 M. Mutter, A. Nefzi, T. Sato, X. Sun, F. Wahl and T. Wöhr, Pseudo-prolines (psi Pro) for accessing 'inaccessible' peptides., *Pept. Res.*, 1995, **8**, 145–153.

27 T. Wöhr and M. Mutter, Pseudo-prolines in peptide synthesis: Direct insertion of serine and threonine derived oxazolidines in dipeptides, *Tetrahedron Lett.*, 1995, **36**, 3847–3848.

28 H. Bürgisser, E. T. Williams, R. Lescure, A. Premanand, A. Jeandin and N. Hartrampf, *A versatile "Synthesis Tag" (SynTag) for the chemical synthesis of aggregating peptides and proteins*, Chemistry, 2023.

29 J. B. Blanco-Canosa, B. Nardone, F. Albericio and P. E. Dawson, Chemical Protein Synthesis Using a Second-Generation N-Acylurea Linker for the Preparation of Peptide-Thioester Precursors, *J. Am. Chem. Soc.*, 2015, **137**, 7197–7209.

30 G. Holdgate, K. Embrey, A. Milbradt and G. Davies, Biophysical methods in early drug discovery, *ADMET DMPK*, 2019, **7**, 222–241.

31 S. Leavitt and E. Freire, Direct measurement of protein binding energetics by isothermal titration calorimetry, *Curr. Opin. Struct. Biol.*, 2001, **11**, 560–566.

32 M. Jerabek-Willemsen, C. J. Wienken, D. Braun, P. Baaske and S. Duhr, Molecular Interaction Studies Using Microscale Thermophoresis, *ASSAY Drug Dev. Technol.*, 2011, **9**, 342–353.

33 T. Heyduk, Y. Ma, H. Tang and R. H. Ebright, in *Methods in Enzymology*, Academic Press, 1996, vol. 274, pp. 492–503.

34 R. B. M. Schasfoort, *Introduction to Surface Plasmon Resonance*, 2017.

35 N. J. Moerke, Fluorescence Polarization (FP) Assays for Monitoring Peptide-Protein or Nuleic Acid-Protein Binding, *Curr. Protoc. Chem. Biol.*, 2009, **1**, 1–15.

36 D. M. Jameson and S. E. Seifried, Quantification of Protein–Protein Interactions Using Fluorescence Polarization, *Methods*, 1999, **19**, 222–233.

37 A. M. Rossi and C. W. Taylor, Analysis of protein-ligand interactions by fluorescence polarization, *Nat. Protoc.*, 2011, **6**, 365–387.

38 J. C. Owicki, Fluorescence Polarization and Anisotropy in High Throughput Screening: Perspectives and Primer, *J. Biomol. Screen.*, 2000, **5**, 297–306.

39 Y. Du, in *Protein-Protein Interactions: Methods and Applications*, eds. C. L. Meyerkord and H. Fu, Springer, New York, NY, 2015, pp. 529–544.

40 C. J. Wienken, P. Baaske, U. Rothbauer, D. Braun and S. Duhr, Protein-binding assays in biological liquids using microscale thermophoresis, *Nat. Commun.*, 2010, **1**, 100.

41 S. A. I. Seidel, P. M. Dijkman, W. A. Lea, G. van den Bogaart, M. Jerabek-Willemsen, A. Lazic, J. S. Joseph, P. Srinivasan, P. Baaske, A. Simeonov, I. Katritch, F. A. Melo, J. E. Ladbury, G. Schreiber, A. Watts, D. Braun and S. Duhr, Microscale thermophoresis quantifies biomolecular interactions under previously challenging conditions, *Methods*, 2013, **59**, 301–315.

42 C. Bergsdorf and S. K. Wright, in *Methods in Enzymology*, ed. C. A. Lesburg, Academic Press, 2018, vol. 610, pp. 135–165.

43 F. H. Niesen, H. Berglund and M. Vedadi, The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability, *Nat. Protoc.*, 2007, **2**, 2212–2221.

44 K. Gao, R. Oerlemans and M. R. Groves, Theory and applications of differential scanning fluorimetry in early-stage drug discovery, *Biophys. Rev.*, 2020, **12**, 85–104.

45 H. Raether, *Surface Plasmons on Smooth and Rough Surfaces and on Gratings*, Springer, Berlin, Heidelberg, 1988, vol. 111.

46 H. P. Ho and W. W. Lam, Application of differential phase measurement technique to surface plasmon resonance sensors, *Sens. Actuators B Chem.*, 2003, **96**, 554–559.

47 K. Matsubara, S. Kawata and S. Minami, Optical chemical sensor based on surface plasmon measurement, *Appl. Opt.*, 1988, **27**, 1160–1163.

48 P. Englebienne, A. V. Hoonacker and M. Verhas, Surface plasmon resonance: principles, methods and applications in biomedical sciences, *Spectroscopy*, 2003, **17**, 255–273.

49 M. M. Morelock, R. H. Ingraham, R. Betageri and S. Jakes, Determination of Receptor-Ligand Kinetic and Equilibrium Binding Constants using Surface Plasmon Resonance: Application to the lck SH2 Domain and Phosphotyrosyl Peptides, *J. Med. Chem.*, 1995, **38**, 1309–1318.

50 E. Freire, O. L. Mayorga and M. Straume, Isothermal titration calorimetry, *Anal. Chem.*, 1990, **62**, 950A-959A.

51 A. Ababou and J. E. Ladbury, Survey of the year 2005: literature on applications of isothermal titration calorimetry, *J. Mol. Recognit.*, 2007, **20**, 4–14.

52 M. Wade, J. Méndez, N. P. Coussens, M. R. Arkin and M. A. Glicksman, in *Assay Guidance Manual*, eds. S. Markossian, A. Grossman, K. Brimacombe, M. Arkin, D. Auld, C. Austin, J. Baell, T. D. Y. Chung, N. P. Coussens, J. L. Dahlin, V. Devanarayan, T. L. Foley, M. Glicksman, K. Gorshkov, J. V. Haas, M. D. Hall, S. Hoare, J. Inglese, P. W. Iversen, S. C. Kales, M. Lal-Nag, Z. Li, J. McGee, O. McManus, T. Riss, P. Saradjian, G. S. Sittampalam, M. Tarselli, O. J. Trask, Y. Wang, J. R. Weidner, M. J. Wildey, K. Wilson, M. Xia and X. Xu, Eli Lilly & Company and the National Center for Advancing Translational Sciences, Bethesda (MD), 2004.

53 R. B. Sekar and A. Periasamy, Fluorescence resonance energy transfer (FRET) microscopy imaging of live cell protein localizations, *J. Cell Biol.*, 2003, **160**, 629–633.

54 F. Degorce, A. Card, S. Soh, E. Trinquet, G. P. Knapik and B. Xie, HTRF: A Technology Tailored for Drug Discovery –A Review of Theoretical Aspects and Recent Applications, *Curr. Chem. Genomics*, 2009, **3**, 22–32.

55 K. G. Harikumar, Y. Yan, T.-H. Xu, K. Melcher, H. E. Xu and L. J. Miller, Bioluminescence Resonance Energy Transfer (BRET) Assay for Determination of Molecular Interactions in Living Cells, *Bio-Protoc.*, 2017, **7**, e2904.

56 Y. Xu, D. W. Piston and C. H. Johnson, A bioluminescence resonance energy transfer (BRET) system: Application to interacting circadian clock proteins, *Proc. Natl. Acad. Sci.*, 1999, **96**, 151–156.

57 M. Y. Berezin and S. Achilefu, Fluorescence Lifetime Measurements and Biological Imaging, *Chem. Rev.*, 2010, **110**, 2641–2684.

58 B. Alpha, J.-M. Lehn and G. Mathis, Energy Transfer Luminescence of Europium(III) and Terbium(III) Cryptates of Macrobicyclic Polypyridine Ligands, *Angew. Chem. Int. Ed. Engl.*, 1987, **26**, 266–267.

59 E. Engvall, The ELISA, Enzyme-Linked Immunosorbent Assay, *Clin. Chem.*, 2010, **56**, 319–320.

60 S. Aydin, A short history, principles, and types of ELISA, and our laboratory experience with peptide/protein analyses using ELISA, *Peptides*, 2015, **72**, 4–15.

61 L. Beaudet, R. Rodriguez-Suarez, M.-H. Venne, M. Caron, J. Bédard, V. Brechler, S. Parent and M. Bielefeld-Sévigny, AlphaLISA immunoassays: the no-wash alternative to ELISAs for research and drug discovery, *Nat. Methods*, 2008, **5**, an8–an9.

62 Immunoassays Alpha |PerkinElmer, https://www.perkinelmer.com/nl/category/immunoassays-alpha, (accessed 17 October 2023).

63 H. H. Nguyen, J. Park, S. Kang and M. Kim, Surface Plasmon Resonance: A Versatile Technique for Biosensor Applications, *Sensors*, 2015, **15**, 10481–10510.

64 G. E. Ghanim, A. J. Fountain, A. M. M. van Roon, R. Rangan, R. Das, K. Collins and T. H. D. Nguyen, Structure of human telomerase holoenzyme with bound telomeric DNA, *Nature*, 2021, **593**, 449–453.

65 G. Sarek, P. Marzec, P. Margalef and S. J. Boulton, Molecular basis of telomere dysfunction in human genetic diseases, *Nat. Struct. Mol. Biol.*, 2015, **22**, 867–874.

66 G. R. Matsueda and J. M. Stewart, A p-methylbenzhydrylamine resin for improved solid-phase synthesis of peptide amides, *Peptides*, 1981, **2**, 45–50.

67 L. H. Brinckerhoff, V. V. Kalashnikov, L. W. Thompson, G. V. Yamshchikov, R. A. Pierce, H. S. Galavotti, V. H. Engelhard and C. L. Slingluff, Terminal modifications inhibit proteolytic degradation of an immunogenic MART-127-35 peptide: Implications for peptide vaccines, *Int. J. Cancer*, 1999, **83**, 326–334.

68 M. Gairí, P. Lloyd-Williams, F. Albericio and E. Giralt, Severe side-reaction in the acidolytic cleavage of a C-terminal Met-containing peptide from the solid support. Formation of the homoserine lactone peptide, *Tetrahedron Lett.*, 1994, **35**, 175–178.

69 M. Jullian, A. Hernandez, A. Maurras, K. Puget, M. Amblard, J. Martinez and G. Subra, N-terminus FITC labeling of peptides on solid support: the truth behind the spacer, *Tetrahedron Lett.*, 2009, **50**, 260–263.

70 J. Kim, S. Felts, L. Llauger, H. He, H. Huezo, N. Rosen and G. Chiosis, Development of a Fluorescence Polarization Assay for the Molecular Chaperone Hsp90, *SLAS Discov.*, 2004, **9**, 375–381.

71 G. Mocz, Information Content of Fluorescence Polarization and Anisotropy, *J. Fluoresc.*, 2006, **16**, 511–524.

72 G. Holzwarth and P. Doty, The Ultraviolet Circular Dichroism of Polypeptides, *J. Am. Chem. Soc.*, 1965, **87**, 218–228.

73 L. Whitmore and B. A. Wallace, Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases, *Biopolymers*, 2008, **89**, 392–400.

74 N. Sreerama and R. W. Woody, Estimation of protein secondary structure from circular dichroism spectra: Comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set, *Anal. Biochem.*, 2000, **287**, 252–260.

75 D. Roccatano, G. Colombo, M. Fioroni and A. E. Mark, Mechanism by which 2,2,2-trifluoroethanol/water mixtures stabilize secondary-structure formation in peptides: A molecular dynamics study, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 12179–12184.

76 M. Moiola, M. G. Memeo and P. Quadrelli, Stapled Peptides—A Useful Improvement for Peptide-Based Drugs, *Molecules*, 2019, **24**, 3654.

77 Y. H. Lau, P. De Andrade, Y. Wu and D. R. Spring, Peptide stapling techniques based on different macrocyclisation chemistries, *Chem. Soc. Rev.*, 2015, **44**, 91–102.

78 X. Li, S. Chen, W.-D. Zhang and H.-G. Hu, Stapled Helical Peptides Bearing Different Anchoring Residues, *Chem. Rev.*, 2020, **120**, 10079–10144.

79 D. P. Fairlie and A. Dantas de Araujo, Stapling peptides using cysteine crosslinking, *Pept. Sci.*, 2016, **106**, 843–852.

80 C. E. Schafmeister, J. Po and G. L. Verdine, An All-Hydrocarbon Cross-Linking System for Enhancing the Helicity and Metabolic Stability of Peptides, *J Am Chem Soc*, 2000, **122**, 5891–5892.

81 L. D. Walensky, A. L. Kung, I. Escher, T. J. Malia, S. Barbuto, R. D. Wright, G. Wagner, G. L. Verdine and S. J. Korsmeyer, Activation of apoptosis in vivo by a hydrocarbon-stapled BH3 helix, *Science*, 2004, **305**, 1466–1470.

82 Y. S. Chang, B. Graves, V. Guerlavais, C. Tovar, K. Packman, K. H. To, K. A. Olson, K. Kesavan, P. Gangurde, A. Mukherjee, T. Baker, K. Darlak, C. Elkin, Z. Filipovic, F. Z. Qureshi, H. Cai, P. Berry, E. Feyfant, X. E. Shi, J. Horstick, D. A. Annis, A. M. Manning, N. Fotouhi, H. Nash, L. T. Vassilev and T. K. Sawyer, Stapled α-helical peptide drug development: A potent dual inhibitor of MDM2 and MDMX for p53-dependent cancer therapy, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 3445–3454.

83 K. Takada, D. Zhu, G. H. Bird, K. Sukhdeo, J. Zhao, M. Lemieux, D. E. Carrasco, J. Ryan, D. Horst, N. C. Munshi, W. Xu, A. L. Kung, R. A. Shivdasani, L. D. Walensky and D. R. Carrasco, Targeted Disruption of the BCL9/β-catenin Complex Inhibits Oncogenic Wnt Signaling, *Sci. Transl. Med.*, 2012, **4**, 117–136.

84 T. Wu, P. He, W. Wu, Y. Chen and F. Lv, Targeting oncogenic transcriptional corepressor Nac1 POZ domain with conformationally constrained peptides by cyclization and stapling, *Bioorganic Chem.*, 2018, **80**, 1–10.

85 X. Ran, L. Liu, C. Y. Yang, J. Lu, Y. Chen, M. Lei and S. Wang, Design of High-Affinity Stapled Peptides to Target the Repressor Activator Protein 1 (RAP1)/Telomeric Repeat-Binding Factor 2 (TRF2) Protein-Protein Interaction in the Shelterin Complex, *J. Med. Chem.*, 2016, **59**, 328–334.

86 G. H. Bird, W. Christian Crannell and L. D. Walensky, Chemical Synthesis of Hydrocarbon-Stapled Peptides for Protein Interaction Research and Therapeutic Targeting, *Curr. Protoc. Chem. Biol.*, 2011, **3**, 99–117.

87 M. A. Simon, P. Ecsédi, G. M. Kovács, Á. L. Póti, A. Reményi, J. Kardos, G. Gógl and L. Nyitray, High-throughput competitive fluorescence polarization assay reveals functional redundancy in the S100 protein family, *FEBS J.*, 2020, **287**, 2834–2846.

88 A. M. Giannetti, in *Methods in Enzymology*, ed. L. C. Kuo, Academic Press, 2011, vol. 493, pp. 169–218.

89 Expasy - Compute pI/Mw tool, https://web.expasy.org/compute_pi/, (accessed 3 October 2023).

90 H. Sota, Y. Hasegawa and M. Iwakura, Detection of Conformational Changes in an Immobilized Protein Using Surface Plasmon Resonance, *Anal. Chem.*, 1998, **70**, 2019–2024.

91 MAb Anti 6HIS-Tb cryptate HTRF reagent | Cisbio, https://uk.cisbio.eu/mab-anti-6his-tb-cryptate-40045, (accessed 14 October 2023).

92 MAb Anti 6HIS-d2 HTRF reagent | Cisbio, https://uk.cisbio.eu/mab-anti-6his-d2-40043, (accessed 14 October 2023).

93 Tau aggregation HTRF kit | Cisbio, https://uk.cisbio.eu/tau-aggregation-kit-40658, (accessed 25 October 2023).

94 Alpha-Synuclein aggregation HTRF kit | Cisbio, https://uk.cisbio.eu/alpha-synuclein-aggregation-kit-40643, (accessed 25 October 2023).

95 L. Einhorn and K. Krapfenbauer, HTRF: a technology tailored for biomarker determination—novel analytical detection system suitable for detection of specific autoimmune antibodies as biomarkers in nanogram level in different body fluids, *EPMA J.*, 2015, **6**, 23.

96 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, Structure-based design and synthesis of potent cyclic peptides inhibiting the YAP-TEAD protein-protein interaction, *ACS Med. Chem. Lett.*, 2014, **5**, 993–998.

97 W. Song, K. Wang, W. Wang, P. Yang and X. Dang, Grafting, Stripping and Stapling of Helical Peptides from the Dimerization Interface of ONFH-Related Bone Morphogenetic Protein-2, *Protein J.*, 2019, **38**, 12–22.

98 A. M. Giltrap, F. P. J. Haeckl, K. L. Kurita, R. G. Linington and R. J. Payne, Synthetic Studies Toward the Skyllamycins: Total Synthesis and Generation of Simplified Analogues, *J. Org. Chem.*, 2018, **83**, 7250–7270.

99 H. Martin-Gómez, F. Albericio and J. Tulla-Puche, A Lasso-Inspired Bicyclic Peptide: Synthesis, Structure and Properties, *Chem. – Eur. J.*, 2018, **24**, 19250–19257.

Chapter 3

Design, synthesis, and *in vitro* biological evaluation of small molecules targeting the dyskerin-dyskerin PPI

3.1 Introduction

3.1.1 Targeting PPIs with small molecules

As described in **Section 1.2.2,** previous work on targeting telomerase with small molecules has focused on targeting the G-quadruplex DNA structures found in telomeres and on targeting telomerase reverse transcriptase (TERT) using an oligonucleotide inhibitor mimicking hTR, the mRNA template used for telomere elongation. A large number of natural products have also been found to inhibit telomerase activity but often with unknown mechanisms. In targeting telomerase there has been a focus on targeting the catalytic core of telomerase rather than the H/ACA RNP. Now with a high resolution cryo-EM structure of telomerase available as described in **Section 1.2.1** the PPIs found within the H/ACA RNP have emerged as a new target to inhibit telomerase assembly.

Small molecule drug leads provide an advantage over peptides due to their relatively easy modification enabling the optimisation of affinity, pharmacokinetic, pharmacodynamic and ADME properties, supported by decades of research on how to do so.[1,2] This allows for some of the problems faced by peptide therapeutics, such as limited proteolytic stability and a lack of cell permeability, to be overcome.[3]

Compared to classical drug targets such as ion channels and enzymes, the discovery of small molecules targeting PPIs can be difficult due to the large flat surface of the interaction between two proteins.[4] However, small molecule inhibitors do not necessarily have to cover the entire binding interface of the PPI. Instead they can target the parts of the PPI that are the most important for binding, deemed hot spots.[5] These hot spots can be discovered through site directed-mutagenesis, in which alanine point mutations are introduced at residues across the interacting surface and the change in binding affinity is measured.[6] However, the expression of multiple mutated proteins can be lengthy and therefore computational alanine scanning methods have become popular, with different web applications such as Robetta and BAlaS available.[7–9] Alternatively, with a known peptide binder, an alanine scan of the peptide can also be performed with a loss of affinity or activity highlighting the residues most important for the interaction.[10,11] The discovery of small molecules can then focus on targeting these hot spots.[12]

Methods used for traditional drug targets, such as high-throughput screening, are less successful for identifying PPI inhibitors because PPIs lack good binding sites for drug-like small molecules.[13]

Instead there has been a focus on other techniques, such as fragment-based screening and *in silico* screening.[14–16] A fragment-based approach screens a library of fragments with NMR, surface plasmon resonance (SPR), thermal shift assays, or isothermal titration calorimetry (ITC).[17] The fragments found to bind can then be then soaked into protein crystals and analysed with X-ray crystallography.[18] These fragments can then be developed into a lead compound with fragment growth or combining fragments (**Figure 3.1**), successfully applied to the discovery of Venetoclax, a Bcl-2 inhibitor.[19]

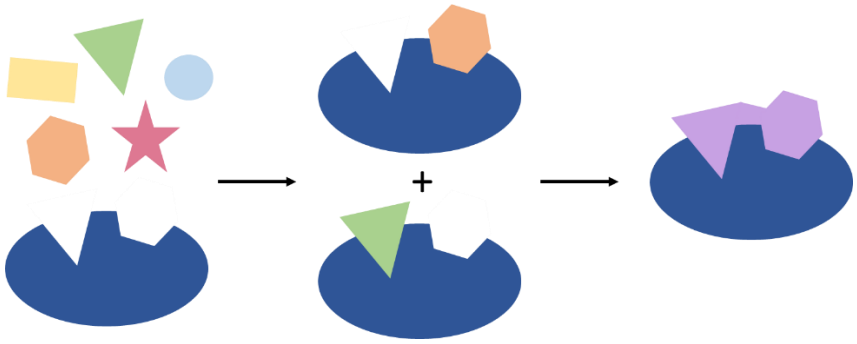


Figure 3.1: *Schematic overview of fragment-based drug design*

In silico screening methods have been successful at identifying small molecule inhibitors of PPIs, such as hot spot-based virtual screening in which fragments are screened to bind to a specific hot spot. This method was successfully applied to identify small molecule inhibitors of the PPI between interferon-α and its receptor.[20]

3.1.2 Peptide-directed binding

A new method to discover small-molecule inhibitors of PPIs, called peptide-directed binding, utilizes *in silico* screening and computational modelling to convert α-helical peptide inhibitors into small-molecule inhibitors.[21] Outlined in **Figure 3.2**, NoxaB, a peptide targeting Mcl-1, was split into two halves and a reactive end, an alkyne or azide, was installed into each half. These can be reacted with small molecule fragments containing azides or alkynes to form a triazole through a copper-catalysed azide-alkyne cycloaddition (CuAAC). A library of azides and alkynes were screened virtually to replace half of the peptide creating small-molecule peptide hybrids. The top candidates were synthesised and their binding was evaluated by fluorescence polarisation. The small molecule parts of the hybrids that demonstrated binding to Mcl-1 were then combined to form small-molecule inhibitors.

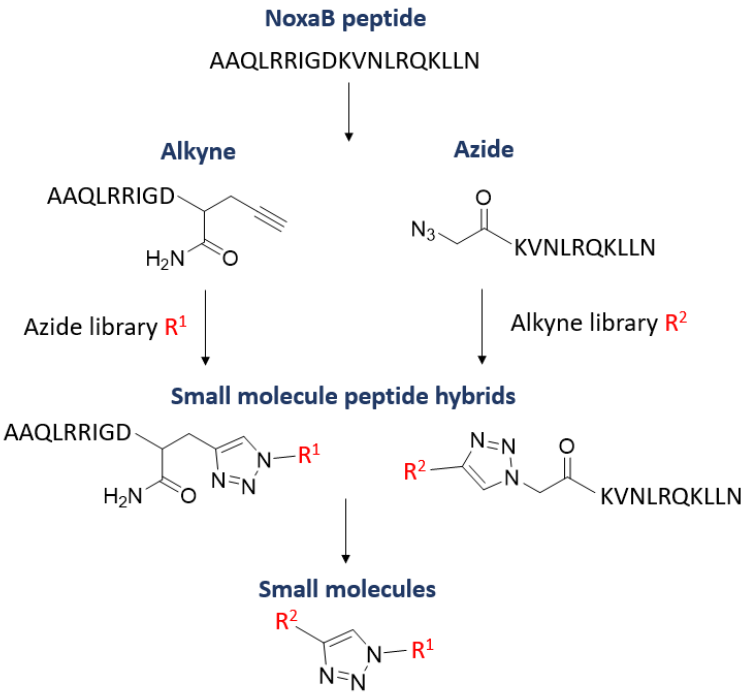


Figure 3.2: *Peptide-directed binding applied to the NoxaB peptide targeting Mcl-1*

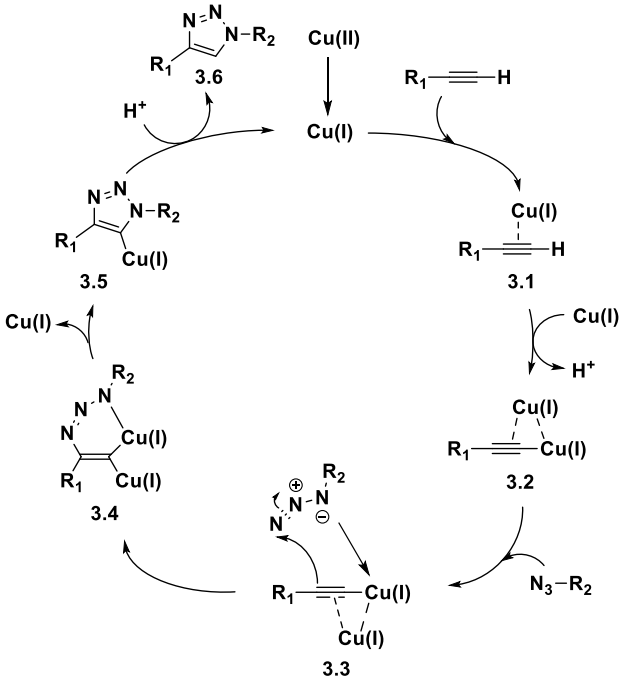
This gave 104 potential small molecules of which 35 were synthesised. Out of these 35, more than 50% showed an IC_{50} <100 µM, a very high hit rate compared to other methods such as high-throughput screening with hit rates of 0.1-1% depending on the target.[21,22] This technique was then further explored as a complete *in silico* process to design the small molecules without the need to synthesise the small-molecules peptide hybrids. Here a p53 derived peptide was used to discover small molecule inhibitors of hDM2/hDMX. The hit rate was, again, remarkably high with 10 of the 20 compounds synthesised able to bind to the target. This shows the promise of this technique, which requires minimal synthesis to find new hit compounds.[23] Although peptide-directed binding has shown excellent results against well studied targets, it is yet to be used for PPIs without known small molecule inhibitors and for lesser explored targets. Click chemistry is ideal for peptide-directed binding due the selective nature of the bond formed, and tolerance of a large range of functional groups therefore not needing protection/deprotection steps that add to the length of the synthesis.

3.1.3 Triazole click chemistry

In 2022 Carolyn Bertozzi, Barry Sharpless and Morten Meldal won the Nobel prize for the development of click and bioorthogonal chemistry. The principle feature of click chemistry is two

components reacting together in a selective matter and was first published in 2001 for the formation of a triazole.[24] Since then other types of click chemistry have also been developed using the inverse electron demand Diels-Alder reaction of tetrazines and sulfur fluoride exchange (SuFEx) chemistry.[25,26]

Click chemistry is synonymous with two different methods of 1,4-disubstituted 1,2,3-triazole (subsequently referred to as triazole) formation, copper(I)-catalyzed azide-alkyne cycloaddition (CuAAC) or strain-promoted azide-alkyne cycloaddition (SPAAC). CuAAC is a Huisgen 1,3-dipolar cycloaddition reaction between a terminal alkyne and an aliphatic azide. The reaction utilises a copper catalyst, such as copper sulfate, which is reduced by sodium ascorbate to form a Cu(I) species that activates the alkyne.[27,28] Fokin *et al.* first proposed a concerted mechanism utilising two Cu(I) species (**Scheme 3.1**).[29] After the activation of the alkyne with the first Cu(I) species (**3.1**) a second Cu(I) species then substitutes the proton on the terminal alkyne (**3.2**). The azide then reacts with the alkyne whilst coordinating to the Cu(I) (**3.3**) forming the triazole (**3.4**) and releasing one of the Cu(I) species (**3.5**). A proton then substitutes the other Cu(I) species forming the final triazole product (**3.6**). The Cu(I) can then go on catalytically to activate the alkyne restarting the cycle once again.



Scheme 3.1: *Mechanism of CuAAC as described by Fokin et al[29]*

CuAAC works effectively in combination with solid phase synthesis, utilised in peptide stapling and the synthesis of macrocyclic peptides.[30–33] The uses of CuAAC on resin, including its use in peptides, peptide nucleic acids (PNAs), small molecules and macrocyclic structures, were reviewed by Albericio *et al* in 2016.[34] Although CuAAC has proven useful in different synthetic applications, the use of the Cu(I) species limits its use in living systems, such as in cells, due to the formation of reactive oxygen species (ROS).[35] To overcome this, the strain-promoted azide-alkyne cycloaddition (SPAAC) was developed.[36] In SPAAC the alkyne is part of a strained ring system (such as a cyclooctyne, **3.7** in **Figure 3.3**) making it more reactive due to the favourable energetics of strain release, which therefore make the use of Cu(I) not necessary.

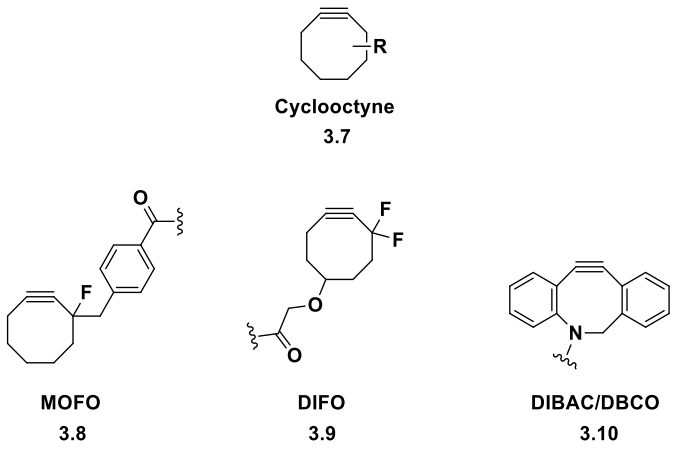


Figure 3.3: *Structure of cyclooctyne (**3.7**), monofluorinated cyclooctyne (MOFO) (**3.8**), difluorinated cyclooctyne (DIFO) (**3.9**) and DIBAC/DBCO (**3.10**)*

The main drawbacks with SPAAC are the lower reaction kinetics compared to CuAAC, poor water solubility and side reactions with cysteines present in proteins.[37,38] To overcome these issues cyclooctyne derivatives have been developed. Monofluorinated cyclooctyne (MOFO) increased the reaction rate of the SPAAC due to its electron-withdrawing nature.[37] This was further increased through the introduction of a second fluoro-substituent creating a difluorinated cyclooctyne (DIFO).[39] The electron-withdrawing nature of the fluorine atoms increases the electrophilicity of the alkyne. DIFO is stable in aqueous conditions and in the presence of β-mercaptoethanol, a cysteine residue mimic.[39] Alternatively, dibenzyl derivates activate the alkyne further through increased ring strain and the incorporation of a nitrogen in the ring, such as in dibenzoazacyclooctyne (DIBAC/DBCO), improves solubility.[40,41]

3.2 Chapter aims

In this chapter we explore whether it is possible to use the cryo-EM structure of telomerase to discover small molecule inhibitors of the dyskerin-dyskerin PPI *in vitro* using peptide-directed binding and triazole forming click chemistry. The cryo-EM structure of telomerase will be evaluated at the dyskerin-dyskerin binding interface with the interacting sequence $_{350}$MTTAVISTCDH$_{360}$ of dyskerin split into two halves, replacing the other half with a small molecule fragment *in silico*. These semi-peptides will be evaluated to rank the fragments, with the top fragments of each half combined to create virtual small molecules. The top small molecules from this *in silico* screen will be evaluated experimentally in the HTRF assay developed in **Chapter 2** to determine their ability to disrupt the PPI formation and a structure-activity relationship established. The best compounds will be evaluated in our SPR assay to determine their binding affinity towards dyskerin, and determine a relationship between their binding affinity and their ability to disrupt the protein dimer formation.

3.3 Results and discussion

3.3.1 *In silico* design of semi-peptides

In the process of peptide-directed binding, the peptide is split into two halves with each half replaced with either an azide or an alkyne fragment. Looking at the dyskerin-dyskerin PPI, the section $_{350}$MTTAVISTCDH$_{360}$ of the protein was taken as the initial starting peptide. Even though this section as a peptide (**2.47**) was not successful at the inhibition of this interaction, it is the part of the protein at the interacting site in the structure. Previously peptide-directed binding has been applied to PPIs with a known structure of a protein with a peptide binder.[21,23] In this work we will explore whether it is possible to develop small molecules directly from the PPI structure, expanding the scope of peptide-directed binding. Although no structure is available of peptide **2.47** with dyskerin, the protein around dyskerin section $_{350}$MTTAVISTCDH$_{360}$ was removed leaving it as a virtually generated peptide and in subsequent interactions this section will be referred to as a peptide for simplicity. It was chosen to split the peptide in half between the V354 and I355 residues in the middle of the peptide (**Figure 3.4**). This equally split the M350T, A353V, T357A and D359N mutations found in patients with DKC, which could indicate the residues most important for the interaction and with I355 on the non-interacting face of the peptide this also left 5 residues on each side able to interact with dyskerin.

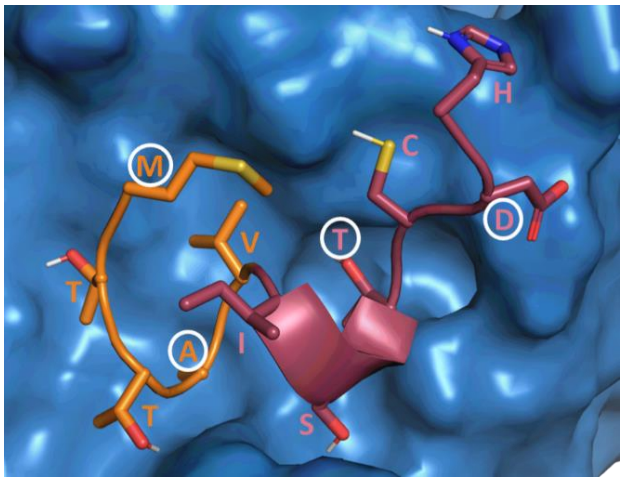
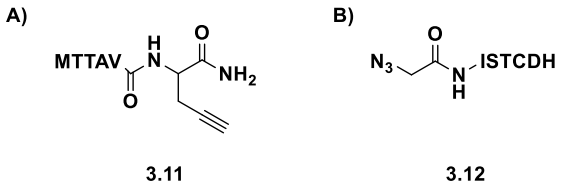


Figure 3.4: *Structure of dyskerin-dyskerin interface in telomerase[42](pdb: 7BGB) with dyskerin (blue) and the semi-peptides of dyskerin $_{350}$MTTAV$_{354}$ (orange) and $_{355}$ISTCDH$_{360}$ (burgundy). The residues with mutations found in DKC are highlighted with a white circle*

Virtually, propargylglycine was added to the C-terminus of V354 and azidoacetic acid was added to the N-terminus of I355 (**Figure 3.5**). These termini could then be used in the covalent docking module to screen azide and alkyne containing fragments.



3.11 **3.12**

Figure 3.5: *Structures of A) MTTAV-propargylglycine (**3.11**) and B) azidoacetic acid-ISTCDH (**3.12**)*

3.3.1.1 Library selection

Peptide-directed binding requires libraries of azide and alkyne fragments that cover a wide chemical space but are ideally immediately available or rapidly synthetically achievable. CuAAC is tolerant of the presence of other functional groups and so there was no need for the exclusion of specific groups from the library.[28]

For the azide and alkyne selection different libraries are available online such as the Enamine and Zinc libraries.[43,44] The Zinc library is an online compound library which combines a number of libraries from different suppliers leading to greater structural diversity and this library was therefore selected for the *in silico* screen.[44] The selection of fragments was refined to compounds which were commercially available, allowing us to purchase the top fragments for the synthesis of the small molecules. The resulting azide and alkyne library of fragments (3347

and 33418 fragments respectively) were then prepared for *in silico* screening using the LigPrep module. During LigPrep the fragments are converted into 3D-structures with different files generated for different, stable conformers, tautomers, or ionisation states.[45]

3.3.2 Covalent docking of azide fragments

Using the half-peptide model the azide fragments were then screened using covalent docking. Covalent docking is normally applied to the investigation of covalent drugs, but the workflow can in this case be applied for peptide-directed binding. The small molecule fragments are first screened non-covalently with a positional constraint on the reactive termini (which have been mutated to alanine to allow for a close positioning of both termini) with a 5 Å cut-off distance between the termini to select for the most appropriate binding poses.[46,47] The termini are then mutated back to the alkyne and azide allowing for the formation and docking of the triazole using the fragment's top binding poses from the non-covalent screen (**Figure 3.6**). The number of azide fragments (3347) is a relatively large number of fragments to screen taking over three months, but it was chosen to include all the fragments to gain the greatest structural diversity.

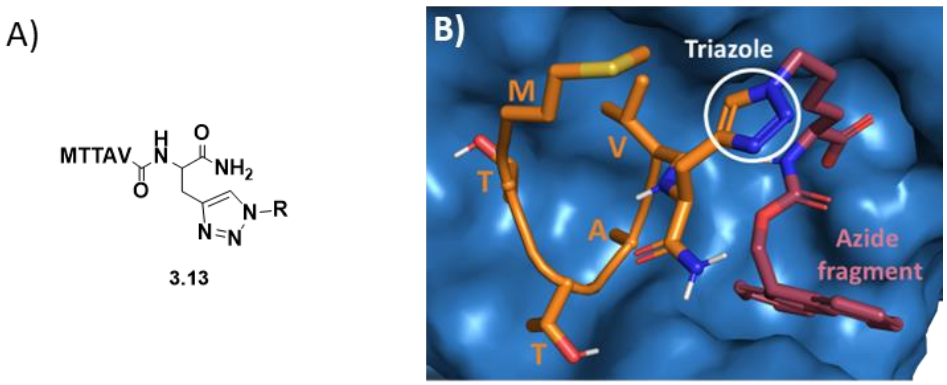
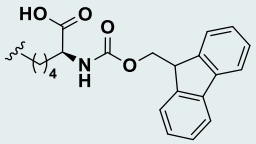
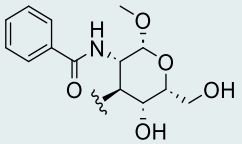
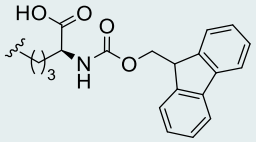
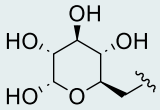
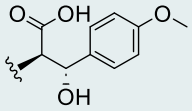
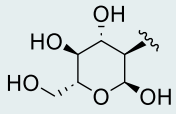
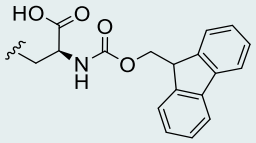
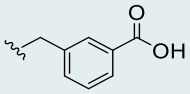
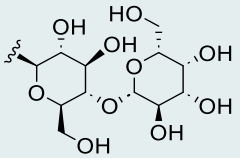
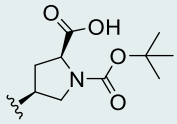
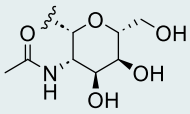


***Figure 3.6**: A) General structure of semi-peptides (**3.13**) formed with azide fragments and B) Example structure of semi-peptide highlighting the peptide section $_{350}$MTTAV$_{354}$ (orange), the triazole linker (white circle) and the small molecule fragment (burgundy)*

The semi-peptides formed were ranked according to their docking score, which considers both the ligand-receptor interactions and the ligand strain energy in adopting the right binding pose, with the more negative the docking score the more energetically favourable the interaction.[48] For the peptide-directed binding process the docking score is used for the relative ranking of the compounds to determine which fragment to take forward into the triazole formation. The semi-peptide structures were manually evaluated to select for immediate commercial availability , or

compounds that could be synthesised readily. After this evaluation the top 11 azide fragments were selected.

Table 3.1: *The top 11 azide fragments in the covalent docking process as ranked by their covalent docking score.*

Nr.	Structure	Docking score	Nr.	Structure	Docking score
3.14		-5.699	**3.20**		-4.848
3.15		-5.521	**3.21**		-4.845
3.16		-5.458	**3.22**		-4.844
3.17		-5.056	**3.23**		-4.834
3.18		-5.007	**3.24**		-4.757
3.19		-4.915			

In this top 11, three of the semi-peptides (**3.14**, **3.15** and **3.17**) had high similarity, all having a Fmoc-protected amino acid with different length side chains containing the azide. A further five of the suggested fragments (**3.18**, **3.19**, **3.20**, **3.21** and **3.22**) were based around a carbohydrate ring. With a high recurrence of similar fragments it seemed as though these two motifs would be best to fit the binding area of dyskerin. **3.16**, **3.23** and **3.24** have unique structural motifs, although **3.16** and **3.23** still contain an aromatic ring and carboxylic acid as found for **3.14**, **3.15** and **3.17**. Initially it was set out to select ten azide fragments, but due to its unique structural motif **3.24** was included in the selection to better understand the translation of the semi-peptide docking into the small molecule docking.

The cryo-EM structure suggests both hydrophobic and hydrophilic interactions for the peptide segment being replaced in this modelling step ($_{355}$ISTCDH$_{360}$). The aspartic acid residue of the peptide can form a H-bonding interaction with T67 in the protein and the histidine residue of the peptide forms a π-π stacking interaction with H68 in the protein (**Figure 3.7A**). The other residues in the peptide mainly participate in Van der Waals interactions with hydrophobic residues in the protein, especially around the serine and threonine where a pocket is found which the fragment can occupy. Interestingly nine semi-peptides (**3.14**, **3.15**, **3.16**, **3.18**, **3.20**, **3.21**, **3.22**, **3.23**, **3.24**) formed a new H-bonding interaction with R322 in the protein. The peptide does not interact with this residue, providing an opportunity for the small molecules to form interactions the peptide is not able to.

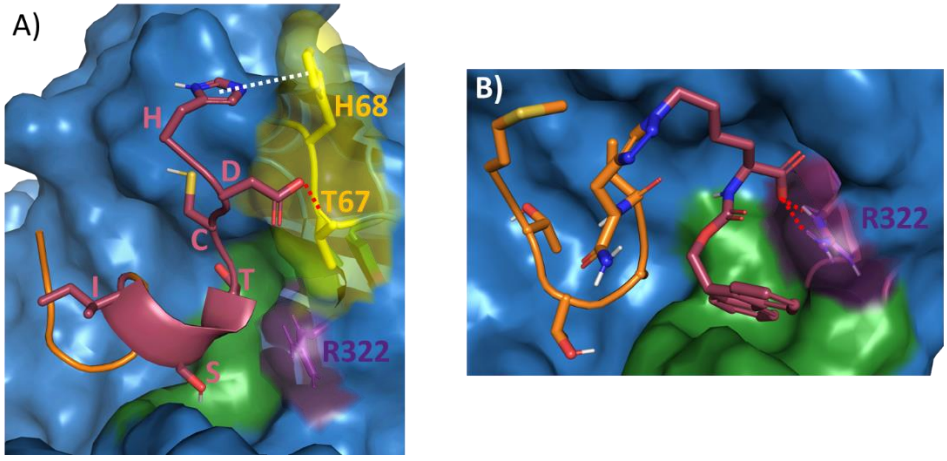


Figure 3.7: *A) Cryo-EM structure of dyskerin-dyskerin PPI with dyskerin (blue) and $_{350}$MTTAV$_{354}$ (orange) and $_{355}$ISTCDH$_{360}$ (burgundy). π-π stacking interaction between H360 and H58 (white dotted line) and a H-bond interaction between D359 and T67 (red dotted line) highlighted. S356 and T357 sit in a hydrophobic pocket (green). R322 is highlighted in purple. B) Structure of dyskerin (blue) with peptide-small molecule hybrid* ***3.14,*** *$_{350}$MTTAV$_{354}$ (orange) and the small molecule fragment (burgundy) highlighting H-bonding from R322 to the carboxylic acid of the fragment (red dotted line) and the Fmoc group sitting in a hydrophobic pocket (green).*

For example, in semi-peptide **3.14**, (**Figure 3.7B**) the carboxylic acid in the fragment forms H-bonding interactions with R322. The fragment does not follow the positioning of the Asp and His residues in the peptide and the aromatic ring from the Fmoc group of the fragment extends in the direction of the serine side chain of the peptide. This aromatic group sits in a hydrophobic pocket created by hydrophobic residues L47, T49, W52 and F59.

3.3.3 Covalent docking of alkyne fragments

Due to the large number of terminal alkyne fragments available in the Zinc database (33418) it would not be possible to apply covalent docking to the fragments as was done for the azide

fragments. Therefore, we examined two different refinement methods, high-throughput non-covalent molecular docking (**section 3.3.3.1**) and molecular weight limitation (**section 3.3.3.2**).

3.3.3.1 Non-covalent docking refinement of alkyne library

We first explored ranking the entire alkyne fragment library using non-covalent docking. In this process the fragments were screened without the covalent attachment, identifying potential fragments which can replace the $_{350}$MTTAV$_{354}$ section of the peptide. This method is faster allowing for the high-throughput of fragments, but the compounds are not restrained by their attachment to the peptide so may adopt any binding pose. The fragments were then ranked according to their docking score and the top 1000 compounds (limiting the covalent docking process to a month) were taken forward into the covalent docking process generating semi-peptides (**Figure 3.8**).

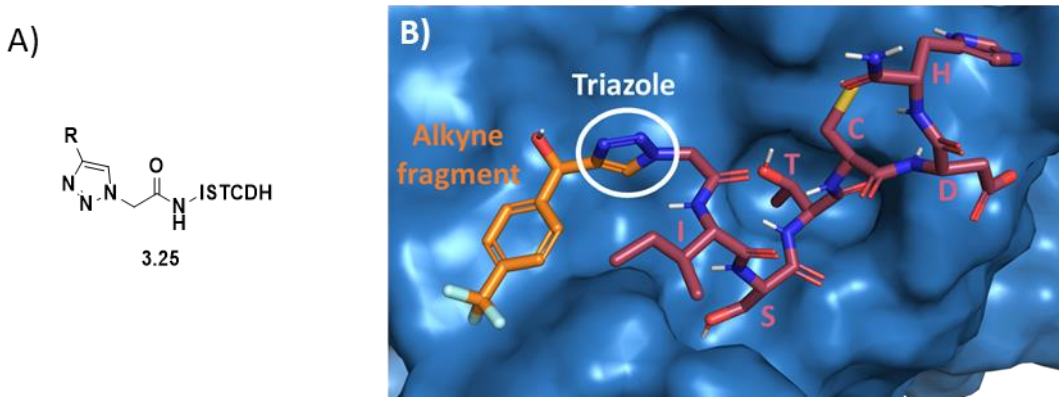
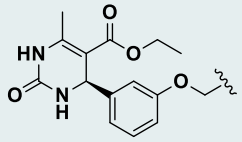
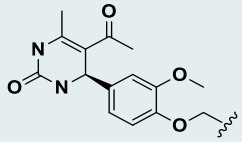
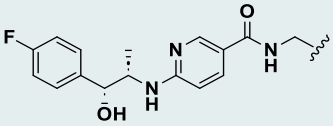
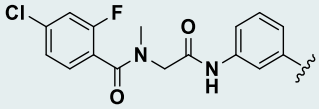
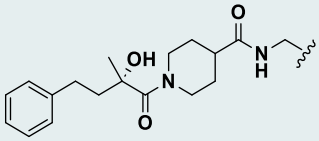
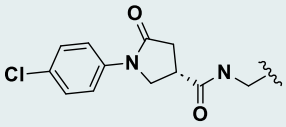
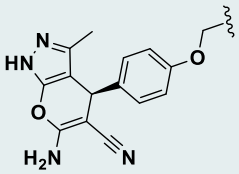
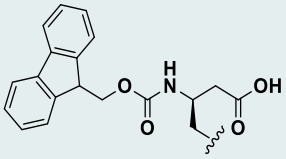
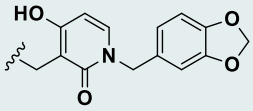
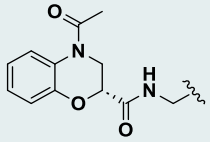


Figure 3.8: *A) General structure of semi-peptides (**3.25**) formed with alkyne small molecule fragments. B) Example structure of peptide-small molecule hybrid highlighting the peptide section $_{355}$ISTCDH$_{360}$ (burgundy), the triazole linker (white circle) and the small molecule fragment (orange).*

The semi-peptides were then evaluated to select the top compounds that were immediately commercially available or could be synthesised readily (**Table 3.2**).

Table 3.2: *Top 10 alkynes from the covalent docking process of the top 1000 compounds determined through Glide docking.*

Nr.	Structure	Docking score	Nr.	Structure	Docking score
3.26		-4.705	**3.31**		-4.472
3.27		-4.556	**3.32**		-4.455
3.28		-4.548	**3.33**		-4.369
3.29		-4.520	**3.34**		-4.328
3.30		-4.487	**3.35**		-4.321

Analysing the protein structure with the peptide, the $_{350}$MTTAV$_{354}$ section of the peptide does not form any H-bonding or π-π stacking interaction as seen for $_{355}$ISTCDH$_{360}$ (**Figure 3.9**).

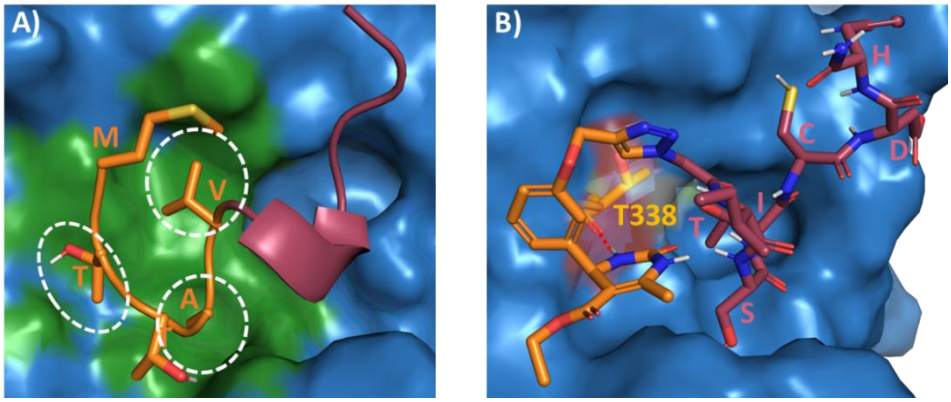


Figure 3.9: *Cryo-EM structure of dyskerin-dyskerin PPI with A) dyskerin (blue), $_{350}$MTTAV$_{354}$ (orange) and $_{355}$ISTCDH$_{360}$ (burgundy) highlighting the binding pockets of T351, A353 and V354 (white dotted lines) with hydrophobic surfaces highlighted in green. B) Peptide-small molecule hybrid **3.26** with the peptide section $_{354}$ISTCDH$_{360}$ (burgundy) and the alkyne fragment (orange) highlighting the H-bond between the urea NH of the fragment and the backbone of T338 of dyskerin (red dotted line)*

Within this section the side chains of T351, A353 and V354 sit in shallow hydrophobic pockets through which they interact with dyskerin (**Figure 3.9A**). The top semi-peptide (**3.26**) sits in the spaces normally occupied by the $_{351}$TTAV$_{354}$ section of the peptide with the triazole in the pocket of the V354 side chain (**Figure 3.9B**), the phenyl ring in the space of the T351 side chain, the ethyl ester in the place of the T352 side chain and the urea ring in the pocket of the A353 side chain. The presence of the urea moiety in the small molecule fragment allows for the H-bonding interaction between an NH and the backbone of T338 of the protein (**Figure 3.9B**) which the A353 side chain cannot form.

3.3.3.2 Molecular weight refinement of alkyne library

We also examined refining the molecular weight of the compounds to keep the overall molecular weight of the triazole compounds low. The azide fragments identified through the covalent docking differ in size, but when the alkyne fragments are added many compounds with high molecular weights (>500) will be identified, and are often poor hits.[49] PPI inhibitors commonly fall outside of Lipinski's rules due to the large binding surface they need to span, such as Venetoclax.[19] However, limiting the molecular weight of the library could still be advantageous as the process of peptide-directed binding is used to find hit compounds which may gain in size in further development. **Table 3.3** describes the number of alkyne fragments in the zinc library under a specified molecular weight.

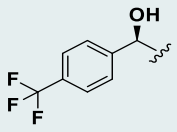
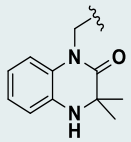
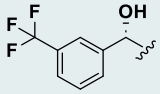
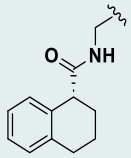
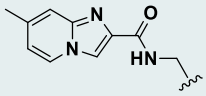
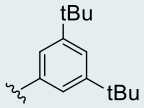
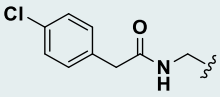
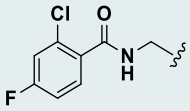
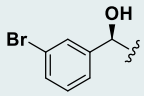
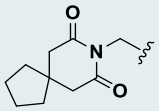
Table 3.3: *The molecular weight limit set and the number of alkyne fragments within that limit.*

Molecular weight limit (g mol^{-1})	Number of alkyne fragments
250	3265
225	1477
220	1173
215	846

Initially a molecular weight limit of 250 was explored, but this still included 3265 fragments. Therefore the molecular weight was further limited with a limit of 215 chosen generating 846 fragments keeping the number of fragments below 1000 to limit the time span of the covalent docking process. These fragments were taken forward into covalent docking to form semi-peptides (**Figure 3.8**). The semi-peptides were then ranked according to their docking score. The compounds were sifted to select compounds that were commercially available or readily synthesised. Since the fragments in this sift were all simpler, a consequence of their lower

molecular weight, the percentage of readily synthesised molecules was higher. The top 10 fragments are listed in **Table 3.4**.

Table 3.4: *Top 10 semi-peptides from the covalent docking process of alkynes with a MW < 215*

Nr.	Structure of R	Docking score	Nr.	Structure of R	Docking score
3.36		-4.036	**3.41**		-3.846
3.37		-3.963	**3.42**		-3.789
3.38		-3.913	**3.43**		-3.763
3.39		-3.876	**3.44**		-3.731
3.40		-3.857	**3.45**		-3.700

The top compounds from this covalent docking screen, except for **3.45,** all showed an aromatic ring occupying the hydrophobic pocket of T351. In replacing the MTTAV section with a small molecule fragment, the space occupied by the Ala is not occupied in most structures. Exceptionally, **3.37** has a hydroxy group in this Ala pocket, allowing for H-bonding to T338, also seen for fragment **3.26** in Section **3.3.3.1**. There is a shift seen in the position of I355 of the peptide which forms hydrophobic interactions with dyskerin, not seen in the cryo-EM structure. In two of the semi-peptides (**3.38** and **3.39**) the triazole ring participates in binding through a cation-π interaction with K339, showing the possibility for the triazole to not only act as a linker between the fragments but to further enhance affinity.

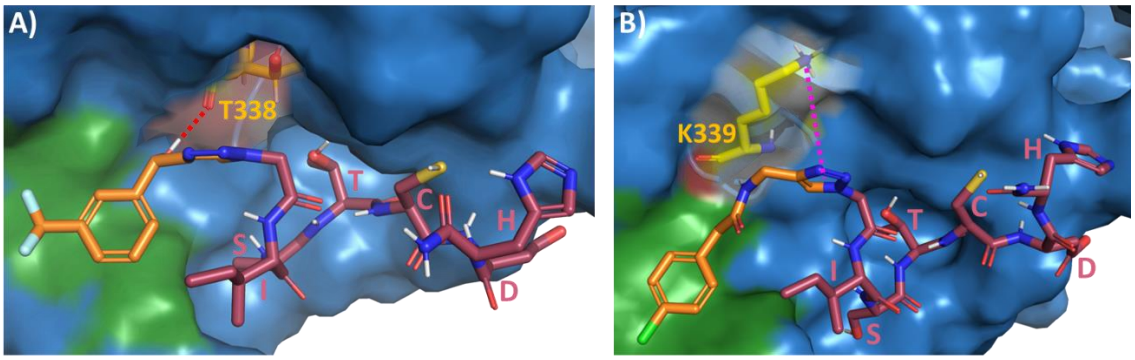


Figure 3.10: *Covalent docking structures with dyskerin (blue) with hydrophobic surfaces (green) and peptide section* $_{355}$*ISTCDH*$_{360}$ *(burgundy), alkyne fragment (orange) of A)* ***3.37*** *highlighting the H-bond from the hydroxy to T338 (red dotted line) and B)* ***3.39*** *highlighting the cation-π interaction between the triazole and K339 (pink dotted line)*

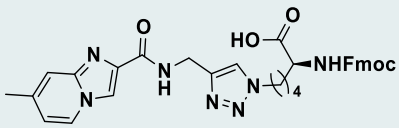
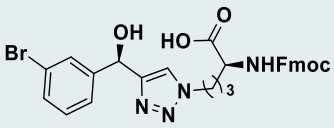
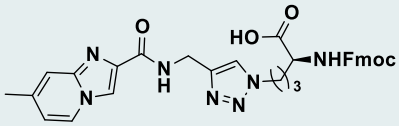
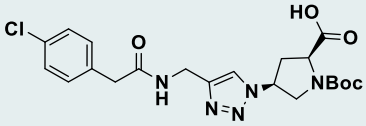
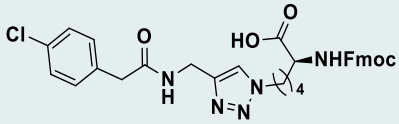
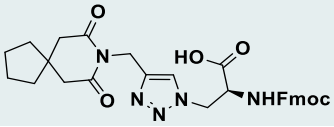
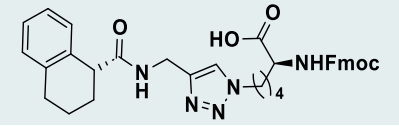
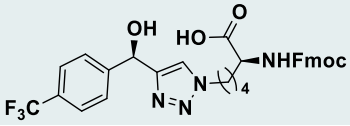
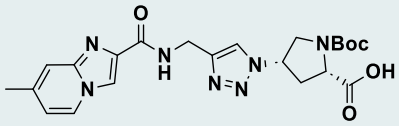
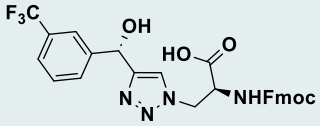
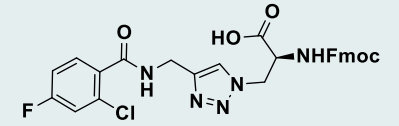
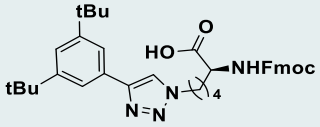
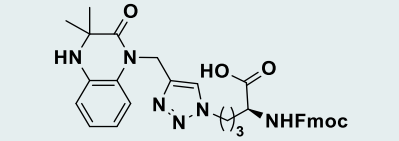
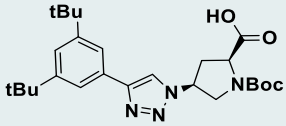
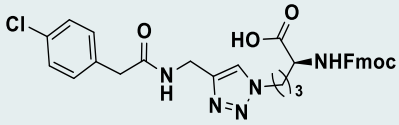
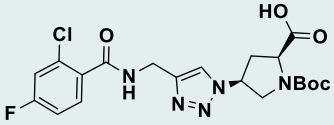
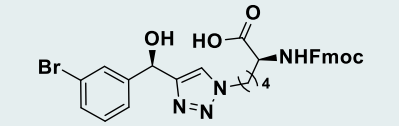
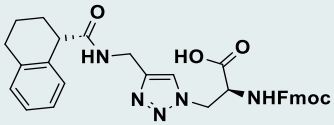
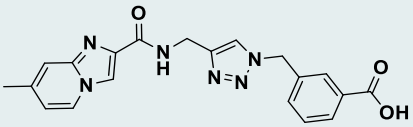
Comparing the two approaches to refining the alkyne fragment library, there was no overlap in compounds. The docking scores for the fragments found through non-covalent docking were better than the size refined fragments, due to their larger size presenting more interactions between the fragment and protein. We chose to take the fragments listed in both **Table 3.2** and **Table 3.4** forward into the *in silico* formation of the triazole compounds to assess which were the most promising compounds to synthesise.

3.3.4 Triazole induced-fit docking

From the covalent docking of the azides, the 11 top fragments as listed in **Table 3.1** were combined with the alkyne fragments to generate virtual triazoles. 11 azides and 20 alkynes generate 220 triazoles which were docked using the induced-fit docking module in Schrodinger. In induced-fit docking the flexibility of the receptor is accounted for, since proteins are dynamic and conformationally flexible in the presence of a ligand.[50] The compounds were ranked based on their induced-fit docking score, which combines the docking score with 5% of the energy change associated with protein movement (known as Prime energy in Schrodinger).[51] Unexpectedly, the induced-fit docking scores of the triazoles formed with the smaller alkyne fragments were better than for the larger fragments. The best docking score for the larger fragment library was -726.86, which was worse than the lowest scoring alkyne from the small alkyne library. This was the opposite of what was seen for the semi-peptides perhaps reflecting the larger fragments are potentially inflexible and unable to fit in the predicted binding site. The small molecules were ranked by their induced-fit docking scores with the aim of exploring approximately twenty compounds experimentally to determine whether the peptide-directed binding process had been successful. In the end 19 compounds were selected for synthesis since

there was a large gap in the induced-fit docking score to the twentieth compound. These 19 compounds are highlighted in **Table 3.5**.

Table 3.5: *Top 19 small molecules as ranked by their induced-fit docking (IFD) score.*

Compound	Nr. and IFD score	Compound	Nr. and IFD score
	**3.47** -738.39		**3.57** -734.68
	**3.48** -737.88		**3.58** -734.55
	**3.49** -736.84		**3.59** -734.51
	**3.50** -736.64		**3.60** -734.50
	**3.51** -736.56		**3.61** -734.43
	**3.52** -735.72		**3.62** -734.31
	**3.53** -735.66		**3.63** -734.30
	**3.54** -735.52		**3.64** -734.20
	**3.55** -735.18		**3.65** -734.14
	**3.56** -734.87		

For the azide fragments in this top 22 there was a high recurrence of Fmoc-protected amino acid azide derivatives. The one carbon, three carbon and four carbon length side chain versions are found in the small molecules with the four carbon linker found in three of the top four compounds. In the covalent docking process several carbohydrate-based fragments were identified as potential binders, but interestingly in the induced-fit docking process these fragments were not highly ranked.

Interestingly, different binding poses for the same compound were found near the top with, for example, **3.47** having the top two poses overall in the induced-fit docking ranking with the aromatic rings of the Fmoc-group flipped (**Figure 3.11**). However, in both binding poses there are multiple H-bonding interactions from the carboxylic acid to R322 and K46 of dyskerin anchoring the small molecule in place. The semi-peptide containing this azide fragment (**3.14**) formed a H-bond interaction with R322 but not with K46 due to the carboxylic acid of the semi-peptide being positioned on the opposite side of R322, not near K46. In the docking of the small molecules the structures have moved to position the carboxylic acid in between R322 and K46 presumably due to the extra H-bonding interaction favouring this position.

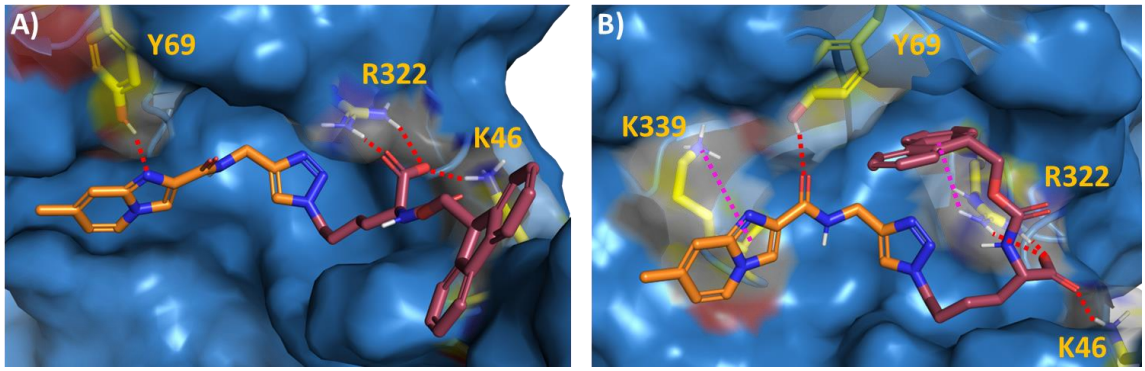
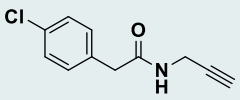
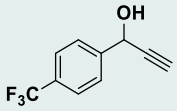
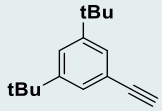
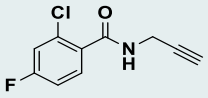
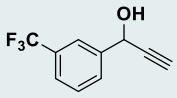
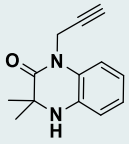
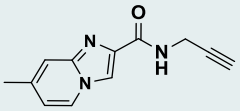
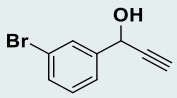
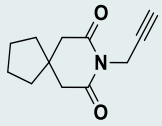
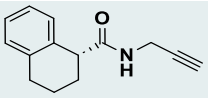


Figure 3.11: *A) Top binding pose of* ***3.47****, B) Second binding pose of* ***3.47***

3.3.5 Synthesis of alkyne fragments

For the synthesis of the top 19 small molecules as identified by induced-fit docking a total of 10 different alkyne fragments were required as listed in **Table 3.6**.

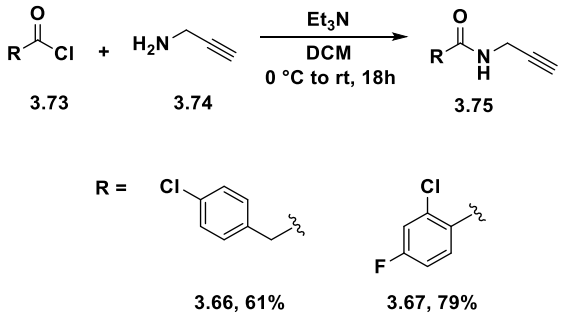
Table 3.6: *Alkyne fragments **3.66**-**3.75** required for the synthesis of compounds **3.47**-**3.65***

Nr.	Structure	Nr.	Structure	Nr.	Structure
3.66		**3.70**		**3.73**	
3.67		**3.71**		**3.74**	
3.68		**3.72**		**3.75**	
3.69					

Several of the alkyne fragments could be synthesised using a one-step procedure from cheap starting materials with the other three alkynes (**3.73**-**3.75**) purchased.

Fragments **3.66**-**3.69** can be synthesised through an amide bond formation between an acid or acid chloride and propargylamine to introduce the alkyne group into the fragment.

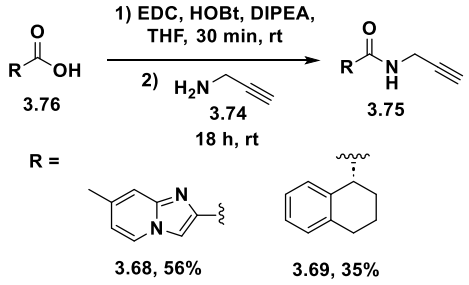
For the synthesis of **3.66** and **3.67** the precursors were commercially available as the acid chloride and were reacted with propargylamine in the presence of triethylamine to negate the formation of hydrochloric acid (**Scheme 3.2**). In the presence of hydrochloric acid the propargylamine can be protonated and it would therefore no longer be able to act as a nucleophile in this reaction.



Scheme 3.2: *Reaction of an acid chloride with propargylamine to form alkyne fragments **3.66** and **3.67***

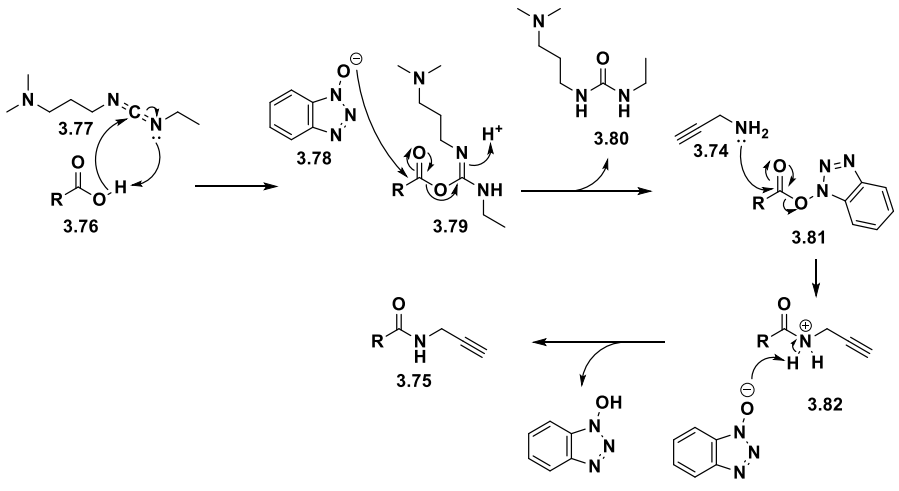
Both compounds were purified by flash column chromatography to give **3.66** and **3.67** in 61% and 79% yield respectively. The formation of the compounds were evidenced by ^{1}H NMR through a shift in the methylene group of the amide (4.02 ppm for **3.66** and 4.26 ppm for **3.67**) compared to the propargylamine (3.45 ppm). The correct mass was also observed by LCMS for both **3.66** (m/z calc = 208.1, m/z obs. = 208.1) and **3.67** (m/z calc = 212.0, m/z obs. = 212.1).

Alkyne fragments **3.68** and **3.69** were formed through amide bond formation, but without the acid chloride commercially available we chose to purchase the compounds as the carboxylic acid. The amide bond was formed with the coupling reagent (EDC) (**3.77**), HOBt (**3.78**) and DIPEA in THF (**Scheme 3.3**).



Scheme 3.3: *Synthesis of alkyne fragments **3.58** and **3.59***

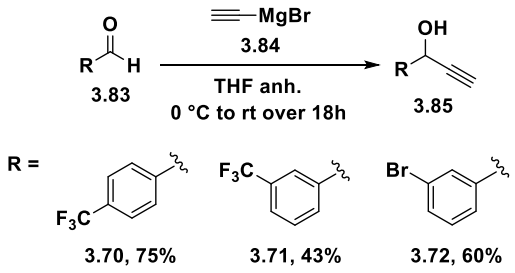
The mechanism of the amide bond formation using HOBt (**3.78**) and EDC (**3.77**) is outlined in **Scheme 3.4**. The carboxylic acid (**3.76**) is first activated by EDC (**3.79**) after which it reacts with HOBt to form an activated ester (**3.81**). The propargylamine reacts with **3.81** to form an amide bond (**3.82**) leading to the product (**3.75**) through a proton transfer.



Scheme 3.4: *Mechanism of amide bond formation between a carboxylic acid (**3.76**) and propargylamine (**3.74**) using EDC (**3.77**) and HOBt (**3.78**)*

EDC is commonly used in the solution phase synthesis of amide bonds due to the tertiary amine present in the molecule (**3.77**), making the removal of the urea byproduct (**3.80**) possible during the workup through acidic washes. Purification was achieved using flash column chromatography, providing **3.68** and **3.69** in a 56% and 35% yield respectively.

Alkyne fragments **3.70**-**3.72** were formed through a Grignard reaction of the corresponding aldehyde with ethynylmagnesium bromide. Grignard reagents are organometallic reagents that allow an organic species, in this case the alkyne, to act as a nucleophile reacting with the aldehyde and creating a new carbon-carbon bond.



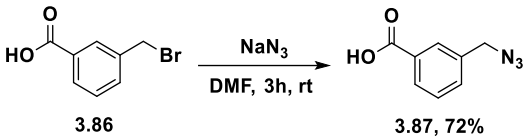
Scheme 3.5**: Synthesis of alkyne fragments **3.70-3.72

During the synthesis of these fragments using the method outlined in **Scheme 3.5** a racemic mixture is formed. From the covalent docking results single enantiomers were highlighted as potential binders. However, the compounds were not available as a single enantiomer. For the synthesis we chose not to pursue an enantioselective synthesis of these compounds because these alkynes were only used to find a potential hit compound. If this alkyne was present in any of the hits found and were to be taken forward into further optimisation, then the problem of the stereocentre would have to be addressed. The formation of **3.70** was evidenced by ^{1}H NMR through the presence of the proton on the carbon attached to the hydroxy at 5.53 ppm with coupling to the terminal alkyne proton (2.71 ppm) observed (J = 2.4 Hz). For **3.71** the same proton was observed at 5.53 ppm with coupling to the alkyne proton (2.72 ppm) observed (J = 2.2Hz). Similarly, for **3.72** the methine proton was observed at 5.42 ppm coupling to the alkyne proton (2.68 ppm) (J = 2.2 Hz).

3.3.6 Synthesis of azide fragment

The azide fragments required for compounds **3.47**-**3.55** and **3.57**-**3.65** were purchased. Most synthetic routes to these target azides were multi-step making their purchase more economical.

However, azide fragment **3.87** was synthesised in a one-step substitution reaction of 3-bromomethyl benzoic acid with sodium azide in DMF.



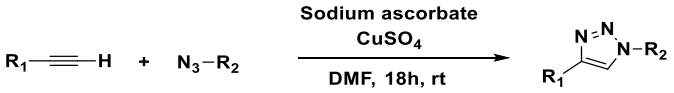
3.86 3.87, 72%

Scheme 3.6: *Synthesis of azide fragment* ***3.64***

The azide formation of **3.87** was evidenced by its IR spectrum, with a characteristic peak seen at 2090 cm^{-1} for the azide and the correct mass found by LCMS (m/z calc. = 176.1, m/z obs. = 176.0).

3.3.7 Triazole synthesis

The triazoles were synthesized through a CuAAC reaction as described in **Section 3.1.2**. The CuAAC reactions were all carried out using copper (II) sulfate and sodium ascorbate in DMF at room temperature for 18 h.



Scheme 3.7: *General triazole synthesis*

For the synthesis of **3.47**, **3.48**, **3.49**, **3.50**, **3.52**, **3.59** and **3.64** a literature procedure was followed for the workup of the reaction.[23] The reaction mixture was extracted with DCM after which the organic layers were washed with water to remove any residual DMF as well as the copper sulfate and sodium ascorbate, which are both water soluble. The compound obtained required purification to remove any traces of either alkyne or azide starting material. Some purification was possible with flash chromatography using MeOH in DCM (0-40%), but a mixture of the product and azide starting material was found. Further purification was attempted using reverse-phase column chromatography yielding pure product. This provided the desired compounds in poor to moderate yields (10%-38%), but enough compound was isolated for the desired use in the HTRF and SPR assays. After the isolation of the compounds, limited solubility was seen in DCM, MeOH and EtOAc. This may be contributing to the low yield of the reactions due to low mass recovery from the aqueous work-up. In addition to this, for **3.51** no compound was recovered in the work-up and therefore the DMF was removed *in vacuo* before purification. This yielded **3.51**, **3.55**, **3.57** and **3.58**, but still in low yield (11%-19%). The remaining compounds (**3.53**, **3.54**, **3.56**, **3.60**, **3.61**, **3.62**, **3.63** and **3.65**) were purified by reverse-phase column

chromatography (0-100% MeOH in H_2O + 0.05% TFA) by directly loading the DMF reaction mixture onto the column. The compounds elute at high percentages of MeOH (80-100%) which allowed for the DMF to be washed away in the initial few column volumes without MeOH present. This method skipped time-consuming workup steps and the compounds could be purified by a single chromatography step. The formation of the desired products was confirmed by LCMS and ^{1}H NMR specifically looking for the triazole proton which is not present in either starting material. This was challenging for the compounds with Fmoc groups or other aromatic moieties as the triazole peak can be obscured, but the presence of the proton could be noted through the integration of the spectrum as well as by 2D-NMR. For example, for **3.58,** the triazole proton was observed at 8.04 ppm and the correct mass was observed by LCMS (m/z calc. = 464.2, m/z obs. = 464.3). The data for the synthesised triazoles can be found in **Table 3.7** with full ^{1}H and ^{13}C NMR data listed in **chapter 6**.

Table 3.7: *Overview of compounds* ***3.34-3.55*** *giving yield, calculated m/z and observed m/z*

Nr.	Yield / %	m/z calc.	m/z obs.	Nr.	Yield / %	m/z calc.	m/z obs.
3.47	10	608.3	608.4	**3.57**	15	591.1	591.3
3.48	31	594.2	594.4	**3.58**	19	464.2	464.3
3.49	24	602.2	602.2	**3.59**	35	558.2	558.3
3.50	18	608.3	608.4	**3.60**	58	595.2	595.3
3.51	11	470.2	470.3	**3.61**	38	553.2	553.1
3.52	48	564.1	564.2	**3.62**	73	609.3	609.4
3.53	79	595.3	595.3	**3.63**	63	471.3	471.4
3.54	41	588.2	588.3	**3.64**	30	468.1	468.3
3.55	13	605.1	605.3	**3.65**	50	566.2	566.3
3.56	41	391.1	391.2				

3.3.8 Effect of triazoles on the dyskerin-dyskerin PPI

The triazole compounds were screened at 100 µM in the HTRF assay described in **Section 2.3.6** to determine whether the synthesized compounds would be able to inhibit the dyskerin-dyskerin PPI. The assay was measured after a 1 h incubation period at 4 °C.

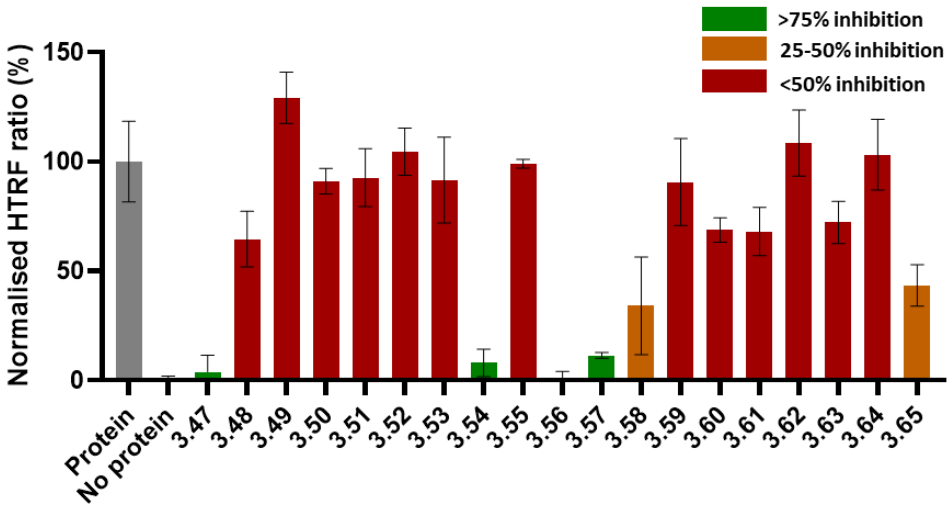


Figure 3.12: *Normalised HTRF ratio (%) of all small molecules tested at 100 µM recorded after 1 h incubation with >75% inhibition highlighted in green, 25-50% inhibition in orange and <50% inhibition in red*

Excitingly, compounds **3.47**, **3.54**, **3.56** and **3.57** showed high activity with 96%, 92%, 99% and 89% inhibition respectively at 100 µM. A further two compounds, **3.58** and **3.65** inhibited 66% and 57% of the PPI formation at 100 µM respectively. The activity of the four most active compounds (**3.47**, **3.54**, **3.56**, and **3.57**) was investigated through a 1 in 2 series dilution from 100 µM (**Figure 3.13**). Compound **3.56** was the most active compound (IC_{50} = 13 µM) over **3.47** (IC_{50} = 43 µM), **3.54** (IC_{50} = 59 µM) and **3.55** (IC_{50} = 74 µM) in the HTRF assay.

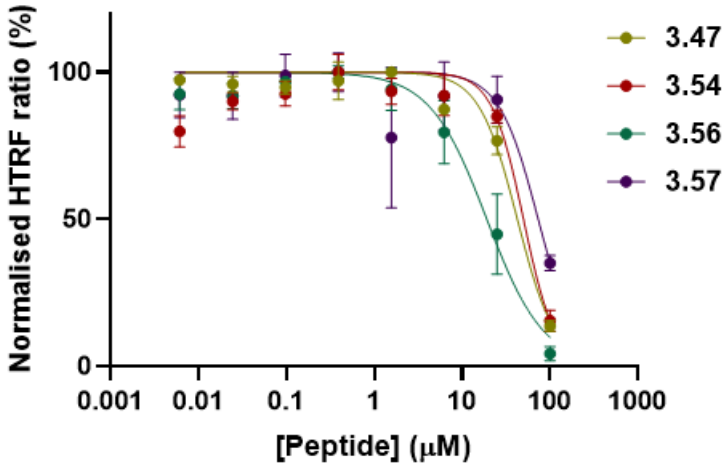
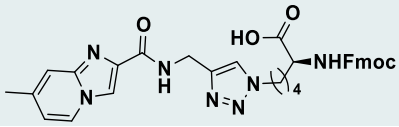
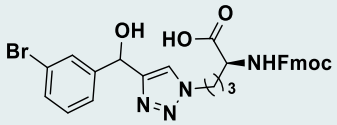
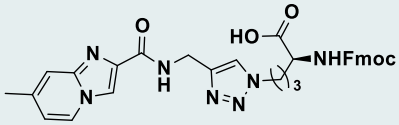
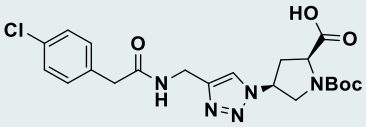
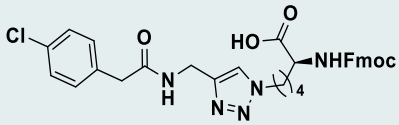
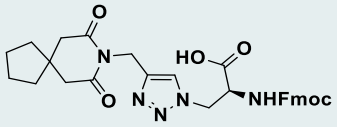
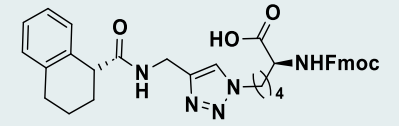
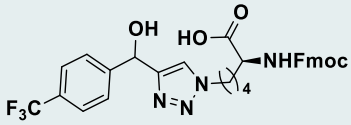
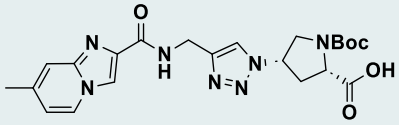
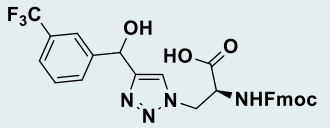
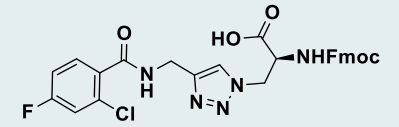
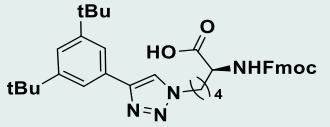
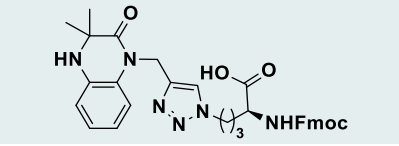
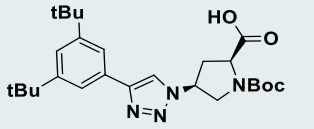
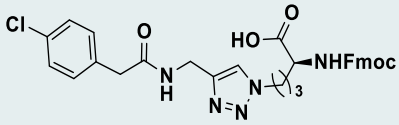
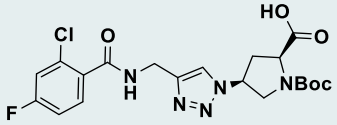
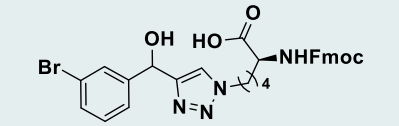
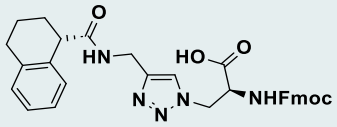
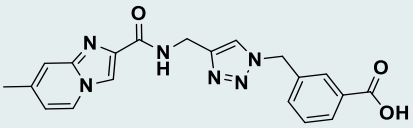


Figure 3.13: *Normalised HTRF ratio (%) of compounds **3.47**, **3.54**, **3.56** and **3.57** after 1h incubation*

Table 3.8: *Top 19 small molecules with their activity in the HTRF assay.*

Compound	Nr. and HTRF IC_{50} [95% CI]	Compound	Nr. and HTRF IC_{50} [95% CI]
	**3.47** 43 µM [37-49]		**3.57** 74 µM [54-96]
	**3.48** N/A		**3.58** ≈100 µM
	**3.49** N/A		**3.59** N/A
	**3.50** N/A		**3.60** N/A
	**3.51** N/A		**3.61** N/A
	**3.52** N/A		**3.62** N/A
	**3.53** N/A		**3.63** N/A
	**3.54** 59 µM [38-66]		**3.64** N/A
	**3.55** N/A		**3.65** ≈100 µM
	**3.56** 19 µM [15-24]		

It was then explored whether the structure of the compounds could be related to their activity (**Table 3.8**). The induced-fit docking poses of **3.56** were analysed to investigate the structure-activity relationship of the compound. The carboxylic acid in **3.56** forms a H-bonding interaction

with both R322 and K46 of dyskerin. The triazole is also placed in a way that it can enhance the binding of the compound through π-π stacking with F59 and a H-bond with Y69.

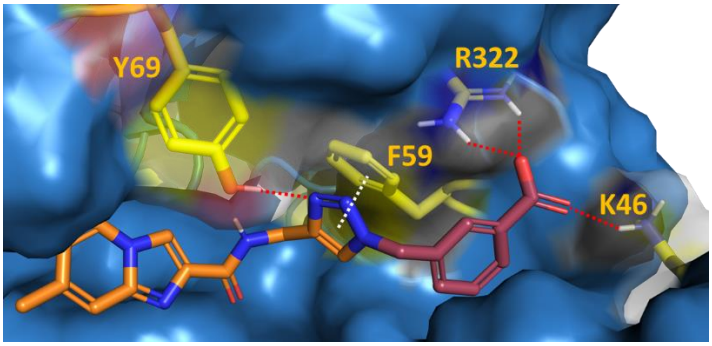


Figure 3.14: *Top binding pose of* ***3.56*** *with dyskerin (blue) as found through induced-fit docking highlighting π-π stacking with white dotted line and H-bonding with red dotted lines.*

The azide fragment (**3.10**) used for compound **3.56** was not used in any of the other triazoles synthesised. However, the alkyne fragment (**3.38**) appeared in triazoles **3.47**, **3.48** and **3.51** (**Figure 3.15**).

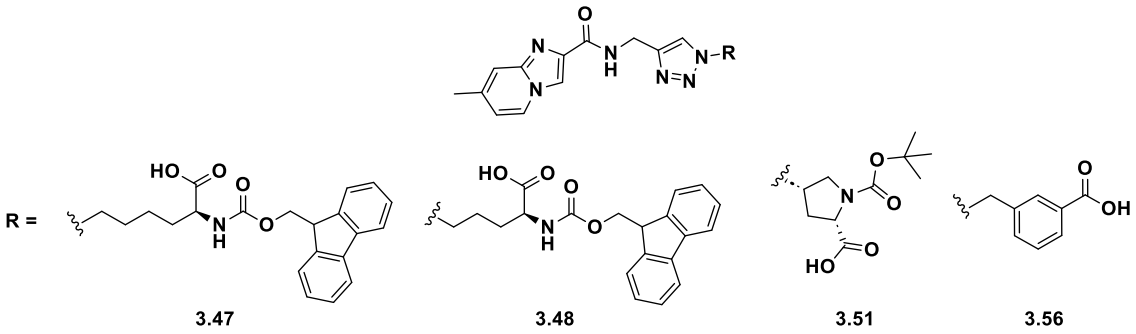


Figure 3.15: *Structure of compounds* ***3.47, 3.48, 3.51*** *and* ***3.56***

Compound **3.47** is structurally very similar to **3.48** with only a difference of one carbon in the chain on the azide side. Interestingly this changed the activity of the compounds significantly with at 100 µM 96% inhibition of the PPI seen for **3.47** but only 35% inhibition seen for **3.48.** The longer carbon chain of **3.47** potentially allowed for a better position of the carboxylic acid or aromatic rings in the azide fragment. The structure of **3.51** is very different from **3.47** and **3.48**, although this also has the same carboxylic acid with can interact with R322, but no activity was seen at 100 µM. This loss in affinity may be due to the lack of aromatic ring which could be forming favoured hydrophobic interactions. The higher activity of **3.56** compared to **3.47** could be due to the additional interactions (H-bond and π-π stacking) **3.56** could form from its triazole to dyskerin (**Figure 3.14**).

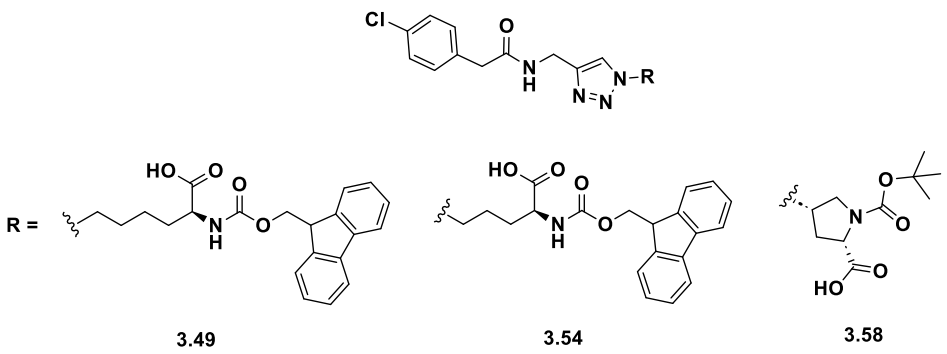


Figure 3.16**: Structures of compounds **3.49**, **3.54** and **3.58

The structure of the third most active compound **3.54** (IC_{50} = 59 µM) on the azide side also has a carboxylic acid which can interact with R322, as well as an aromatic ring on the alkyne side of the molecule. Interestingly, for **3.49**, which contains the same alkyne side as triazole **3.54**, but with a one-carbon longer chain, no inhibition of the dyskerin-dyskerin PPI was observed at 100 µM, suggesting the longer chain is too large for the binding site. Compound **3.58**, containing the same alkyne fragment as **3.54** maintained some activity (66% inhibition at 100 µM) indicating the non-aromatic proline derivative as the azide fragment is tolerated but not as active.

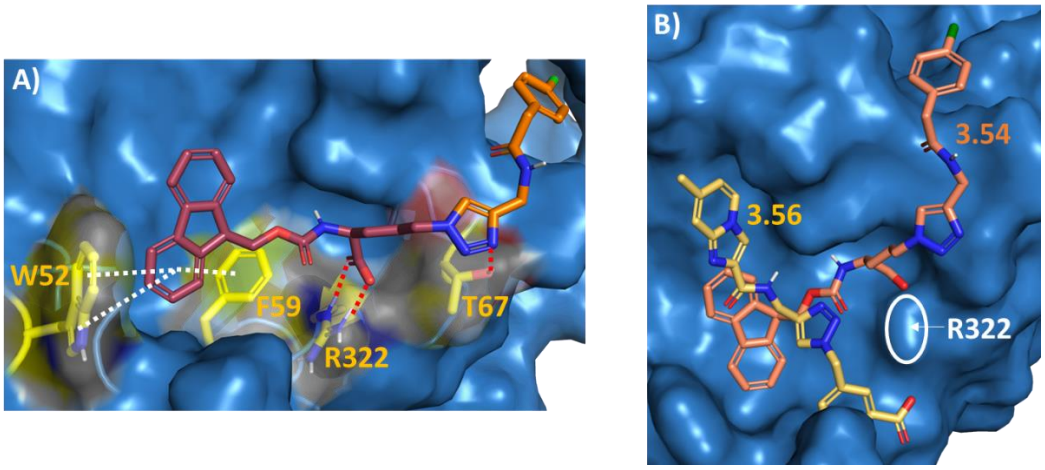


***Figure 3.17:** A) Top binding pose for **3.54** as found through induced-fit docking highlighting π-π stacking interaction (white dotted lines) with W52 and F59 and H-bond interactions (red dotted lines) with R322 and T6. B) Top binding poses of **3.54** (orange) and **3.56** (yellow) as found through induced-fit binding*

For **3.54** in the induced-fit docking model, in the top 5 binding poses the compound is bound in the opposite direction as was planned from the covalent docking with the azide and alkyne fragments swapped from the side they were docked to. This could be due to the multiple π-π stacking interactions from the Fmoc group of the fragment to W52 and F59 as well as the H-bonding from R322 to the carboxylic acid of the fragment anchoring this compound in this

conformation (**Figure 3.17A**). Although both **3.54** and **3.56** contain carboxylic acid groups which can form H-bonding with R322, they are found to bind from opposite directions in the induced-fit docking model (**Figure 3.17B**).

3.3.9 Determination of binding affinity through SPR

Compounds **3.47**, **3.54** and **3.56** were explored in the SPR assay to determine their binding affinities. They were first explored in HBS-EP+ at pH 7.4 with 5% DMSO. The binding affinity for **3.54** and **3.56** were inside the measured range (<200 µM) and these two compounds were explored in different buffers to assess the optimal conditions for the interaction with PBS-P+ at pH 7.4 with 5% DMSO chosen as well as PBS-P+ with 5% DMSO at pH 8.0 which had previously been explored for the peptides in **Section 2.3.5** (**Table 3.9**).

Table 3.9: *K_D values measured by SPR for compounds **3.47**, **3.54** and **3.56** in buffers 1-3.*

Nr.	1: K_D HBS-EP+ with 5% DMSO pH 7.4 (µM)	2: K_D PBS-P+ with 5% DMSO pH 7.4 (µM)	3: K_D PBS-P+ with 5% DMSO pH 8.0 (µM)
3.47	309 ± 40	-	-
3.54	190 ± 18	220 ± 61	226 ± 83
3.56	89 ± 15	28.7 ± 6.1	38.6 ± 5.7

In the SPR assay compound **3.56** recorded a higher binding affinity than **3.54** which was in line with the activity seen in the HTRF assay in **Section 3.3.4**.

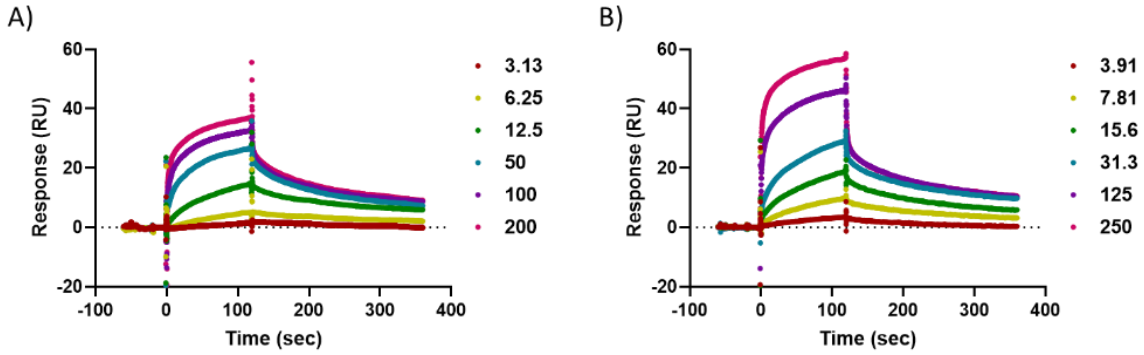


Figure 3.18: *SPR sensorgrams of **3.56** in A) PBS-P+, pH 7.4 with 5% DMSO and B) PBS-P+, pH 8.0 with 5% DMSO*

Compound **3.56** showed the best binding affinities in the PBS-P+ buffers with curvature seen in the sensorgrams indicating slower on and off rates which could be due to its higher binding affinity (**Figure 3.18**). As seen for some of the peptides, there was issues with the R_{max} value especially for **3.54** (**Table 3.10**) with the signal size larger than expected based on the

immobilisation of 5000 RU MBP-dyskerin and a 1:1 binding stoichiometry. Looking at the compounds in buffer 2, for which **2.56** recorded the best K_D, the R_{max} for **3.54** recorded of 322 is over 10-fold higher than the theoretical R_{max} (27.2 RU) (**Table 3.10**). For **3.56** the recorded R_{max} of 43.7 is just over 2-fold higher than the theoretical R_{max} (18.1 RU) (**Table 3.10**), which although higher is much closer to what would be expected. Looking at the difference of the experimental R_{max} of **3.56** in buffer 2 and 3 (43.7 RU and 64.7 RU respectively), this could potentially be solved by further buffer investigations.

Table 3.10: *K_D, theoretical R_{max} and experimental R_{max} of compounds **3.54** and **3.56** in buffer 2*

Nr.	K_D (µM)	R_{max} theoretical (RU)	R_{max} experimental (RU)
3.54	220	27.2	322 ± 50
3.56	28.7	18.1	43.7 ± 1.6

It was also observed that the MBP-dyskerin protein can lose the MBP-tag when stored for longer time periods at 4 °C and therefore it may be possible that the protein on the chip has over time lost its MBP-tag, which would almost half the MW of the protein from 108 kDa down to 58 kDa and therefore would double the theoretical R_{max} (36.2 RU) which is very close to the experimental value for **3.56** recorded in buffer 2. This could not explain the experimental R_{max} for **3.54** which may indicate the non-specific binding of this compound, which could indicate it may not be as useful for further drug development.

3.4 Conclusion and future work

Using the process of peptide-directed binding, small molecules targeting the dyskerin-dyskerin PPI were designed. The peptide was split into two halves in between V354 and I355 with each half of the peptide replaced with an azide or alkyne fragments using the covalent docking module. The alkyne fragment library was refined using two different methods: high throughput non-covalent docking and size refinement. The top alkyne and azide fragments were then combined into triazoles *in silico* and the small molecules were ranked using induced-fit docking. The top 19 triazoles were successfully synthesised and were evaluated in the HTRF assay with 6 compounds (**3.47**, **3.54**, **3.56**, **3.57**, **3.58** and **3.65**) inhibiting >50% of the interaction at 100 µM after a 1 h incubation period. These are the first known small molecule inhibitors of the dyskerin-dyskerin PPI with especially compound **3.56** (HTRF IC_{50} = 19 µM) showing promise for further drug development.

Future work would focus on generating a second generation of small molecules based around **3.56**. Especially the combination of the azide fragment of **3.56** with other alkyne fragments to see whether any of those could increase the binding affinity. The reasons for the binding affinity of **3.56** could also be further explored through changing the substitution pattern on the phenyl ring, such as through moving the carboxylic acid around the ring to determine the optimal position. This could provide a better structure-activity relationship around **3.56** to understand its binding mode.

3.5 References

1 E. Cannady, K. Katyayan and N. Patel, in *Haschek and Rousseaux's Handbook of Toxicologic Pathology (Fourth Edition)*, eds. W. M. Haschek, C. G. Rousseaux, M. A. Wallig and B. Bolon, Academic Press, 2022, pp. 51–76.

2 T. D. Y. Chung, D. B. Terry and L. H. Smith, in *Assay Guidance Manual*, eds. S. Markossian, A. Grossman, K. Brimacombe, M. Arkin, D. Auld, C. Austin, J. Baell, T. D. Y. Chung, N. P. Coussens, J. L. Dahlin, V. Devanarayan, T. L. Foley, M. Glicksman, K. Gorshkov, J. V. Haas, M. D. Hall, S. Hoare, J. Inglese, P. W. Iversen, S. C. Kales, M. Lal-Nag, Z. Li, J. McGee, O. McManus, T. Riss, P. Saradjian, G. S. Sittampalam, M. Tarselli, O. J. Trask, Y. Wang, J. R. Weidner, M. J. Wildey, K. Wilson, M. Xia and X. Xu, Eli Lilly & Company and the National Center for Advancing Translational Sciences, Bethesda (MD), 2004.

3 L. Otvos and J. D. Wade, Current challenges in peptide-based drug discovery, *Front. Chem.*, 2014, **2**, 62.

4 H. Lu, Q. Zhou, J. He, Z. Jiang, C. Peng, R. Tong and J. Shi, Recent advances in the development of protein–protein interactions modulators: mechanisms and clinical trials, *Signal Transduct. Target. Ther.*, 2020, **5**, 1–23.

5 A. A. Bogan and K. S. Thorn, Anatomy of hot spots in protein interfaces, *J. Mol. Biol.*, 1998, **280**, 1–9.

6 I. S. Moreira, P. A. Fernandes and M. J. Ramos, Hot spots—A review of the protein–protein interface determinant amino-acid residues, *Proteins Struct. Funct. Bioinforma.*, 2007, **68**, 803–812.

7 T. Kortemme, D. E. Kim and D. Baker, Computational Alanine Scanning of Protein-Protein Interfaces, 2004, 2.

8 Computational Interface Alanine Scanning Server, https://robetta.bakerlab.org/alascansubmit.jsp, (accessed 24 September 2023).

9 C. W. Wood, A. A. Ibarra, G. J. Bartlett, A. J. Wilson, D. N. Woolfson and R. B. Sessions, BAlaS: fast, interactive and accessible computational alanine-scanning using BudeAlaScan, *Bioinform.*, 2020, **36**, 2917–2919.

10 F. M. Heydenreich, T. Miljuš, R. Jaussi, R. Benoit, D. Milić and D. B. Veprintsev, High-throughput mutagenesis using a two-fragment PCR approach, *Sci. Rep.*, 2017, **7**, 1–11.

11 L. Raguine, M. Ali, V. Bender, E. Diefenbach, M. R. Doddareddy, D. Hibbs and N. Manolios, Alanine Scan of an Immunosuppressive Peptide (CP): Analysis of Structure-Function Relationships, *Chem. Biol. Drug Des.*, 2013, **81**, 167–174.

12 W. Guo, J. A. Wisniewski and H. Ji, Hot spot-based design of small-molecule inhibitors for protein–protein interactions, *Bioorg. Med. Chem. Lett.*, 2014, **24**, 2546–2554.

13 I. R. Taylor, B. M. Dunyak, T. Komiyama, H. Shao, X. Ran, V. A. Assimon, C. Kalyanaraman, J. N. Rauch, M. P. Jacobson, E. R. P. Zuiderweg and J. E. Gestwicki, High-throughput screen for inhibitors of protein-protein interactions in a reconstituted heat shock protein 70 (Hsp70) complex, *J. Biol. Chem.*, 2018, **293**, 4014–4025.

14 M. M. R. Arkin and J. A. Wells, Small-molecule inhibitors of protein-protein interactions: Progressing towards the dream, *Nat. Rev. Drug Discov.*, 2004, **3**, 301–317.

15 W. H. Shin, C. W. Christoffer and D. Kihara, In silico structure-based approaches to discover protein-protein interaction-targeting drugs, *Methods*, 2017, **131**, 22–32.

16 L. Mabonga and A. P. Kappo, Peptidomimetics: A Synthetic Tool for Inhibiting Protein–Protein Interactions in Cancer, *Int. J. Pept. Res. Ther.*, 2020, **26**, 225–241.

17 D. E. Scott, M. T. Ehebauer, T. Pukala, M. Marsh, T. L. Blundell, A. R. Venkitaraman, C. Abell and M. Hyvönen, Using a Fragment-Based Approach To Target Protein-Protein Interactions, *ChemBioChem*, 2013, **14**, 332–342.

18 D. A. Erlanson, B. J. Davis and W. Jahnke, Fragment-Based Drug Discovery: Advancing Fragments in the Absence of Crystal Structures, *Cell Chem. Biol.*, 2019, **26**, 9–15.

19 A. J. Souers, J. D. Leverson, E. R. Boghaert, S. L. Ackler, N. D. Catron, J. Chen, B. D. Dayton, H. Ding, S. H. Enschede, W. J. Fairbrother, D. C. S. Huang, S. G. Hymowitz, S. Jin, S. L. Khaw, P. J. Kovar, L. T. Lam, J. Lee, H. L. Maecker, K. C. Marsh, K. D. Mason, M. J. Mitten, P. M. Nimmer, A. Oleksijew, C. H. Park, C. M. Park, D. C. Phillips, A. W. Roberts, D. Sampath, J. F. Seymour, M. L. Smith, G. M. Sullivan, S. K. Tahir, C. Tse, M. D. Wendt, Y. Xiao, J. C. Xue, H. Zhang, R. A. Humerickhouse, S. H. Rosenberg and S. W. Elmore, ABT-199, a potent and selective BCL-2 inhibitor, achieves antitumor activity while sparing platelets, *Nat. Med.*, 2013, **19**, 202–208.

20 T. Geppert, S. Bauer, J. A. Hiss, E. Conrad, M. Reutlinger, P. Schneider, M. Weisel, B. Pfeiffer, K.-H. Altmann, Z. Waibler and G. Schneider, Immunosuppressive Small Molecule Discovered by Structure-Based Virtual Screening for Inhibitors of Protein–Protein Interactions, *Angew. Chem. Int. Ed.*, 2012, **51**, 258–261.

21 A. M. Beekman, M. A. O'Connell and L. A. Howell, Peptide-Directed Binding for the Discovery of Modulators of α-Helix-Mediated Protein–Protein Interactions: Proof-of-Concept Studies with the Apoptosis Regulator Mcl-1, *Angew. Chem. - Int. Ed.*, 2017, **56**, 10446–10450.

22 D. R. Stroik, S. L. Yuen, K. A. Janicek, T. M. Schaaf, J. Li, D. K. Ceholski, R. J. Hajjar, R. L. Cornea and D. D. Thomas, Targeting protein-protein interactions for therapeutic discovery via FRET-based high-throughput screening in living cells, *Sci. Rep.*, 2018, **8**, 1–13.

23 A. M. Beekman, M. M. D. Cominetti, S. J. Walpole, S. Prabhu, M. A. O'Connell, J. Angulo and M. Searcey, Identification of selective protein-protein interaction inhibitors using efficient: In silico peptide-directed ligand design, *Chem. Sci.*, 2019, **10**, 4502–4508.

24 H. C. Kolb, M. G. Finn and K. B. Sharpless, Click Chemistry: Diverse Chemical Function from a Few Good Reactions, *Angew. Chem. Int. Ed.*, 2001, **40**, 2004–2021.

25 M. L. Blackman, M. Royzen and J. M. Fox, Tetrazine Ligation: Fast Bioconjugation Based on Inverse-Electron-Demand Diels−Alder Reactivity, *J. Am. Chem. Soc.*, 2008, **130**, 13518–13519.

26 J. Dong, L. Krasnova, M. G. Ginn and K. B. Sharpless, Sulfur(VI) Fluoride Exchange (SuFEx): Another Good Reaction for Click Chemistry - Dong - 2014 - Angewandte Chemie International Edition - Wiley Online Library, 2014, **53**, 9430–9448.

27 W. S. Brotherton, H. A. Michaels, J. T. Simmons, R. J. Clark, N. S. Dalal and L. Zhu, Apparent Copper(II)-Accelerated Azide−Alkyne Cycloaddition, *Org. Lett.*, 2009, **11**, 4954–4957.

28 V. V. Rostovtsev, L. G. Green, V. V. Fokin and K. B. Sharpless, A Stepwise Huisgen Cycloaddition Process: Copper(I)-Catalyzed Regioselective "Ligation" of Azides and Terminal Alkynes, *Angew. Chem.*, 2002, **114**, 2708–2711.

29 M. Ahlquist and V. V. Fokin, Enhanced Reactivity of Dinuclear Copper(I) Acetylides in Dipolar Cycloadditions, *Organometallics*, 2007, **26**, 4389–4391.

30 Y. Wu, F. Villa, J. Maman, Y. H. Lau, L. Dobnikar, A. C. Simon, K. Labib, D. R. Spring and L. Pellegrini, Targeting the Genome-Stability Hub Ctf4 by Stapled-Peptide Design, *Angew. Chem. - Int. Ed.*, 2017, **56**, 12866–12872.

31 X. Ran, L. Liu, C. Y. Yang, J. Lu, Y. Chen, M. Lei and S. Wang, Design of High-Affinity Stapled Peptides to Target the Repressor Activator Protein 1 (RAP1)/Telomeric Repeat-Binding Factor 2 (TRF2) Protein-Protein Interaction in the Shelterin Complex, *J. Med. Chem.*, 2016, **59**, 328–334.

32 S. A. Kawamoto, A. Coleska, X. Ran, H. Yi, C. Y. Yang and S. Wang, Design of triazole-stapled BCL9 α-helical peptides to target the β-catenin/B-cell CLL/lymphoma 9 (BCL9) protein-protein interaction, *J. Med. Chem.*, 2012, **55**, 1137–1146.

33 L. E. Hanold, K. Oruganty, N. T. Ton, A. M. Beedle, N. Kannan and E. J. Kennedy, Inhibiting EGFR dimerization using triazolyl-bridged dimerization arm mimics, *PLoS ONE*, 2015, **10**, 1–18.

34 V. Castro, H. Rodríguez and F. Albericio, CuAAC: An Efficient Click Chemistry Reaction on Solid Phase, *ACS Comb. Sci.*, 2016, **18**, 1–14.

35 G. J. Brewer, Risks of Copper and Iron Toxicity during Aging in Humans, *Chem. Res. Toxicol.*, 2010, **23**, 319–326.

36 N. J. Agard, J. A. Prescher and C. R. Bertozzi, A Strain-Promoted [3 + 2] Azide−Alkyne Cycloaddition for Covalent Modification of Biomolecules in Living Systems, *J. Am. Chem. Soc.*, 2004, **126**, 15046–15047.

37 N. J. Agard, J. M. Baskin, J. A. Prescher, A. Lo and C. R. Bertozzi, A Comparative Study of Bioorthogonal Reactions with Azides, *ACS Chem. Biol.*, 2006, **1**, 644–648.

38 R. van Geel, G. J. M. Pruijn, F. L. van Delft and W. C. Boelens, Preventing Thiol-Yne Addition Improves the Specificity of Strain-Promoted Azide–Alkyne Cycloaddition, *Bioconjugate Chem.*, 2012, **23**, 392–398.

39 J. M. Baskin, J. A. Prescher, S. T. Laughlin, N. J. Agard, P. V. Chang, I. A. Miller, A. Lo, J. A. Codelli and C. R. Bertozzi, Copper-free click chemistry for dynamic in vivo imaging, *Proc. Natl. Acad. Sci.*, 2007, **104**, 16793–16797.

40 M. F. Debets, S. S. van Berkel, S. Schoffelen, F. P. J. T. Rutjes, J. C. M. van Hest and F. L. van Delft, Aza -dibenzocyclooctynes for fast and efficient enzyme PEGylation via copper-free (3+2) cycloaddition, *Chem. Commun.*, 2010, **46**, 97–99.

41 X. Ning, J. Guo, M. A. Wolfert and G.-J. Boons, Visualizing Metabolically Labeled Glycoconjugates of Living Cells by Copper-Free and Fast Huisgen Cycloadditions, *Angew. Chem. Int. Ed.*, 2008, **47**, 2253–2255.

42 G. E. Ghanim, A. J. Fountain, A. M. M. van Roon, R. Rangan, R. Das, K. Collins and T. H. D. Nguyen, Structure of human telomerase holoenzyme with bound telomeric DNA, *Nature*, 2021, **593**, 449–453.

43 Azides for 'Click Chemistry' - Enamine, https://enamine.net/building-blocks/functional-classes/azides-for-click-chemistry, (accessed 19 September 2023).

44 J. J. Irwin, K. G. Tang, J. Young, C. Dandarchuluun, B. R. Wong, M. Khurelbaatar, Y. S. Moroz, J. Mayfield and R. A. Sayle, ZINC20—A Free Ultralarge-Scale Chemical Database for Ligand Discovery, *J. Chem. Inf. Model.*, 2020, **60**, 6065–6073.

45 J. C. Shelley, A. Cholleti, L. L. Frye, J. R. Greenwood, M. R. Timlin and M. Uchimaya, Epik: a software program for pKaprediction and protonation state generation for drug-like molecules, *J. Comput. Aided Mol. Des.*, 2007, **21**, 681–691.

46 A. Scarpino, G. G. Ferenczy and G. M. Keserű, Comparative Evaluation of Covalent Docking Tools, *J. Chem. Inf. Model.*, 2018, **58**, 1441–1458.

47 K. Zhu, K. W. Borrelli, J. R. Greenwood, T. Day, R. Abel, R. S. Farid and E. Harder, Docking Covalent Inhibitors: A Parameter Free Approach To Pose Prediction and Scoring, *J. Chem. Inf. Model.*, 2014, **54**, 1932–1940.

48 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, Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy, *J. Med. Chem.*, 2004, **47**, 1739–1749.

49 C. A. Lipinski, F. Lombardo, B. W. Dominy and P. J. Feeney, Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, *Adv. Drug Deliv. Rev.*, 2001, **46**, 3–26.

50 W. Sherman, H. S. Beard and R. Farid, Use of an Induced Fit Receptor Structure in Virtual Screening, *Chem. Biol. Drug Des.*, 2006, **67**, 83–84.

51 W. Sherman, T. Day, M. P. Jacobson, R. A. Friesner and R. Farid, Novel Procedure for Modeling Ligand/Receptor Induced Fit Effects, *J. Med. Chem.*, 2006, **49**, 534–553.

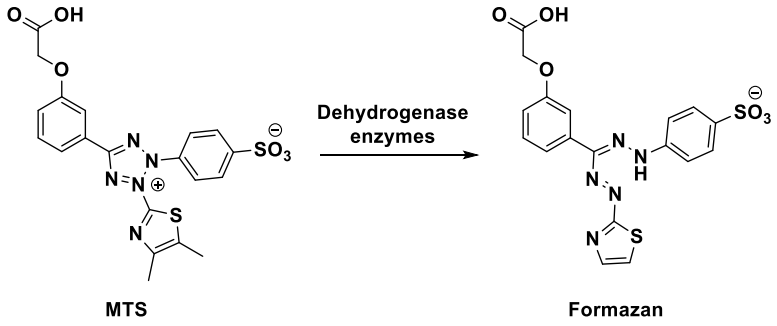
Chapter 4
The cellular evaluation of peptides and small molecules targeting the dyskerin-dyskerin PPI

4.1 Introduction

4.1.1 Biophysical vs cellular activity

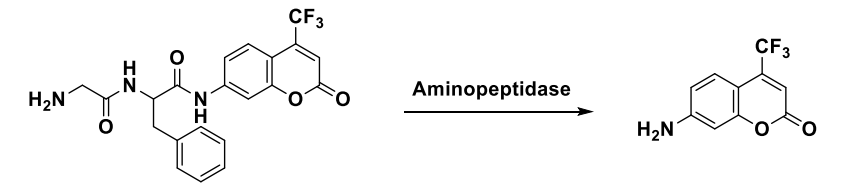
The biophysical assays explored in depth in **Chapter 2** (e.g. FP, ITC, SPR and HTRF) allow for the determination of the effect of compounds on the isolated protein. Although this provides insight into the binding and activity towards the target protein, it does not take into account other aspects of the cell, such as cell permeability and the potential for off-target protein binding which can hinder the drug discovery process.[1] As already discussed in **chapter 1** for peptides a lack of cell permeability and limited proteolytic stability is often observed, with strategies such as stapling or conjugation to a cell-penetrating peptide required for activity.[2,3]

When exploring the cellular effects of any compounds, one of the first assays that may be turned to is a cytotoxicity assay. This can provide information about the effect of the compound at different concentrations on cell viability. Different assays are available to test this, such as the colorimetric assays 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT), (3-[4,5,dimethylthiazol-2-yl]-5-[3-carboxymethoxy-phenyl]-2-[4-sulfophenyl]-2H-tetrazolium (MTS) and resazurin, the detection of proteolytic enzymes, and by ATP detection.[4–8] In the three colorimetric assays mentioned (MTT, MTS and resazurin) the viability is measured by the reduction of the compound where this reduction leads to a change in colour. For example MTS can be reduced into formazan by dehydrogenase enzymes leading to a dark blue colour of which the absorption can be measured at 490 nm (**Scheme 4.1**).[5]



Scheme 4.1: *Reduction of MTS to Formazan by hydrogenase enzymes*

The measurement of protease activity has also been noted as a marker of cell viability. This can be tested by glycylphenylalanyl-aminofluorocoumarin (GF-AFC), which is a substrate for aminopeptidases that remove the Gly and Phe residues generating aminofluorocoumarin (AFC) of which the fluorescence can be measured (**Scheme 4.2**).[7]



Glycylphenylalanyl-aminofluorocoumarin (GF-AFC)

Aminofluorocoumarin (AFC)

Scheme 4.2: *Cleavage of Gly (G) and Phe (F) of GF-AFC to give AFC*

A benefit for the GF-AFC assay over tetrazolium-based assays is that GF-AFC is non-toxic to the cells unlike the tetrazolium-based agents.[9] This would allow for a secondary assay to take place as long as there is no interference from the fluorescence.

Another method of investigating cell viability is through ATP detection which can be measured by the addition of luciferin and firefly luciferase. In the presence of luciferin and ATP, luciferases generate light which can be measured.[8] This assay is the fastest out of the cell viability assays with only 10 minutes of incubation required, making it ideal for high-throughput applications.[10]

Once a cellular effect has been established it is also important to show that the compounds have the desired mechanism of action by showing target engagement. This can be done in several different ways such as by FRET/BRET (see **chapter 2**), activity-based protein profiling (ABPP), or by cellular thermal shift assay (CETSA). In ABPP the compound of interest is functionalised with a reactive group as well as a tag, such as a fluorescent tag or biotin allowing for the detection of the compound.[11] Through affinity chromatography (for e.g. biotin) or by SDS-PAGE analysis the proteins with which the compound interacts can be identified. CETSA employs the same principle as the thermal shift assay discussed in **chapter 2** based on the thermal denaturation of proteins which can change when bound to a compound.[12] In this process cells are first treated with the compound after which the cells are heated to different temperatures. The cells are then lysed and centrifuged to remove any aggregates that have formed in the denaturation process, with only folded proteins remaining in solution. This cell lysate can be analysed by western blotting, probing for the target protein for which a melting curve can be constructed with a comparison with non-treated cells allowing the detection of any effect of the compounds on the target protein.[13]

4.1.2 Cell permeability assays

To understand the cellular activity of compounds it is important to study their cell permeability to determine whether this is interfering with their activity. The cell permeability of peptides and small molecules can be investigated with methods such as flow cytometry,

the chloroalkane penetration assay (CAPA) and artificial membranes (Caco-2 and parallel artificial membrane permeation (PAMPA)).[14–17]

For peptides the cell permeability can be investigated with a fluorescent tag such as is found on the peptides used for fluorescence polarisation.[18] The cell permeability can be explored through flow cytometry by quantifying the fluorescence seen after treating the cells with the peptide and washing away any unbound peptide.[14] This can be further complemented using confocal microscopy in which the localisation of the peptide can be observed.[19]

For small molecules without a fluorescent tag the chloroalkane penetration assay (CAPA) is a recently developed assay that uses a cell line expressing Halo-Tag GFP which reacts with a chloroalkane linker attached to a small molecule.[15] After this is washed away, a chloroalkane linked dye is added, which reacts with any leftover Halo-Tag. Using flow cytometry, the fluorescence can then be measured, which is inversely proportional to the cell permeability of the compound of interest. Artificial membranes can also be used to study the cellular uptake. Caco-2 cells are human colon epithelial cells that can be used to study absorption through mimicking the intestine.[20] PAMPA is a cell-free artificial membrane using a filter such as a polyvinylidene fluoride (PVDF), which is soaked with phospholipids, mimicking a cell membrane.[17] PAMPA measures the passive transport from one side of the membrane to the other side, but since it does not contain cells cannot consider active transport.[21]

4.1.3 Drug stability

Together with cell permeability another problem that can be encountered in *in vitro* experiments is limited proteolytic stability. If the compound is degraded before it can exert its mechanism of action, no effect may be observed in cells. Proteases are enzymes that catalyse the proteolysis of proteins by cleaving peptide bonds. This is important for protein catabolism, cell signalling and for the digestion of food.[22] However, this can pose a problem for peptides, which contain those same peptidic bonds, leaving them vulnerable to proteolytic degradation.[23]

Beyond the proteolytic stability that is important for *in vitro* experiments, the stability of compounds in plasma can also be tested to further understand whether the compound would be efficacious *in vivo*.[24] The half-life of compounds in plasma is a measure that can provide information about why certain activity may be seen for some compounds but not others and allows for the prioritisation of stable compounds in the drug development process. The plasma stability can be determined by incubation in plasmid with HPLC analysis at different time points.[24]

For orally available drugs further tests are also required to understand their stability across the gastrointestinal (GI) track. The GI track brings with it a number of challenges, such as extreme pH found in the stomach and further enzymes such as pepsin and pancreatin.[25] To mimic the GI conditions simulated gastric fluid (pH 1.2) containing pepsin and simulated intestinal fluid (pH 6.8) containing pancreatin have been used to test the stability of drug compounds.[26]

4.2 Chapter aims

In this chapter we aim to explore the effect of the linear, hydrocarbon and lactam stapled peptides synthesised in **Chapter 2** and the small molecules synthesised in **Chapter 3** on cancer cells. Specifically, we will look at their effect on cell viability and determine whether their effect in cells could be limited due to cell permeability or proteolytic stability.

4.3 Results and discussion

4.3.1 The effect of the linear and stapled peptides on cell viability

It was first investigated whether the synthesised peptides would influence the viability of cancer cells. The effect on telomerase activity might not result in a decrease in cell viability in the short-term. This was observed for BIBR1532, a telomerase inhibitor, for which no effect was noted until after 120 days of treatment.[27] However, dyskerin is also part of other H/ACA ribonucleoproteins that are involved in ribosome biogenesis. The inhibition of the catalytic pseudouridylation activity of dyskerin did show a decrease in cell viability in literature.[28] Therefore the inhibition of the dyskerin-dyskerin PPI may still result in a decrease in cell proliferation and viability from its role in this process.

The peptides were tested to see whether they showed an effect on cell viability through an MTS cytotoxicity assay.[5] For the MTS assay A549 lung adenocarcinoma cells were selected. It has previously been shown that silencing of DKC1, the gene encoding dyskerin, in A549 cells leads to significantly decreased telomerase activity and increased cell senescence and apoptosis.[29] Therefore we explored whether the inhibition of the dyskerin-dyskerin PPI could have similar effects.

All the compounds were first screened at 200 μM before IC_{50} values were determined. Work began with the screening of the linear peptides (**2.47, 2.49, 2.50**-**2.60**).

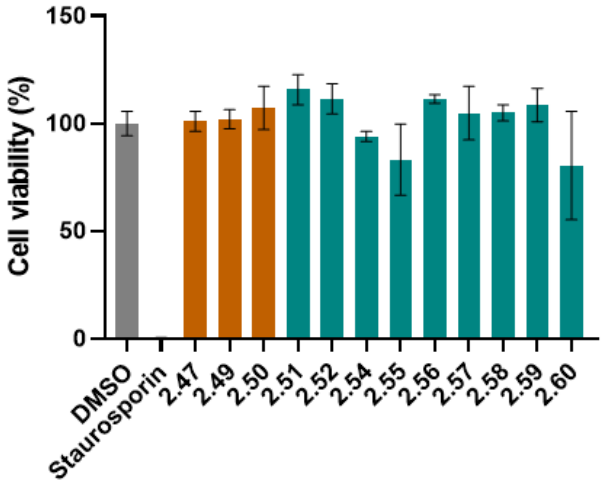


Figure 4.1: *Effect of linear peptides* ***2.47, 2.49*** *and* ***2.50*** *(orange) and alanine scan peptides* ***2.51-2.60*** *(blue) at 200 µM on A549 cell viability as determined by MTS assay*

None of the linear peptides showed any effect on the cell viability of the A549 cells at 200 µM (**Figure 4.1**). Although no effect is seen on the cell viability this could not only be due to their low affinity towards dyskerin, but it could also be due to limited cell permeability or proteolytic stability which were further explored in sections **4.3.2** and **4.3.3** respectively.

The hydrocarbon stapled peptides **2.69, 2.76**-**2.79** and the lactam stapled peptides **2.93**-**2.95** (**Figure 4.2A**) were also screened at 200 µM in the assay. Peptide **2.69** showed the greatest reduction in cell viability out of all the stapled peptides (IC_{50} = 132 µM) (**Figure 4.2B**).

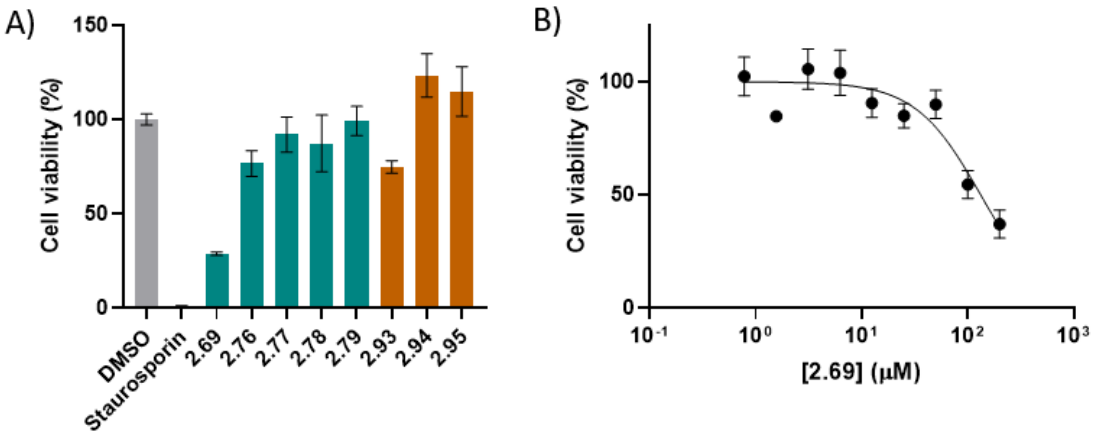


Figure 4.2: *A) Effect of hydrocarbon stapled peptides* ***2.69, 2.76-2.79*** *(blue) and lactam stapled peptides* ***2.93-2.95*** *(orange) at 200 µM on A549 cell viability as determined by MTS assay and B) Effect of peptide* ***2.69*** *on A549 cell viability as determined by MTS assay*

Peptide **2.69** had not been shown to inhibit the dyskerin-dyskerin PPI in the HTRF assay but it did show the highest binding to dyskerin in the FP assay. Therefore, it is likely this peptide is acting through a different mechanism of action other than the through the inhibition of the dyskerin-dyskerin PPI.

The lactam stapled peptide **2.93** showed a small reduction in cell viability which could be due to its ability to inhibit the dyskerin-dyskerin PPI as shown in the HTRF assay. The other two peptides in this series (**2.94** and **2.95**) did not show any effect on cell viability although they did inhibit the dyskerin-dyskerin PPI in the HTRF assay. The lower hydrophobicity of the lactam staple compared to the hydrocarbon could be influencing the cell permeability of these compounds and this was further investigated in **section 4.3.2** using flow cytometry. Alternatively, an issue faced with peptides is a lack of proteolytic stability and this could also be impacting its effect in cells, which was further explored in **section 4.3.3**.

4.3.2 The effect of peptide stapling on the cellular uptake

Important for the cellular effect of peptides is their ability to cross the cell membrane. Peptides tend to be highly hydrophilic in nature due to the high number of amide bonds as well as their side chains, depending on the residues present. As mentioned in **section 2.3.4** peptide stapling can increase cellular uptake due to the increase of hydrophobicity depending on the stapling method chosen. Alternatively, cell-penetrating peptides (CPPs) increase cell permeability by conjugating a CPP to the designed peptide. This method does not rely on the modification of the peptide and therefore if a high binding peptide has already been found but no cellular uptake is seen conjugation to a CPP can overcome that.[3]

We used flow cytometry to look at the cellular uptake of the FITC-tagged linear (**2.61**), hydrocarbon stapled (**2.68**) and lactam stapled (**2.92**) peptides. In flow cytometry the cell suspension is flowed past a laser with every cell analysed for visible light scatter as well as for fluorescence. The visible light scatter is measured in the forward direction (FSC) and at 90 °C (SSC) and a plot of these two measurements allows for the selection of live cells. The relative uptake of the peptides can be determined through the measurement of fluorescence intensity plotted against the cell count. A549 cells were treated with the three compounds for 1 h at 20 µM after which their fluorescence was quantified by flow cytometry (**Figure 4.3**).

Table 4.1: *Sequences of the fluorescent peptides,* ***2.61, 2.68*** *and* ***2.92*** *and their non-fluorescent versions,* ***2.47, 2.69*** *and* ***2.93***

Nr.	Sequence
2.61	FITC-βAla-MTTAVISTCDH-NH_2
2.47	Ac-MTTAVISTCDH-NH_2
2.68	FITC-βAla-Nle**S**$_5$TAV**S**$_5$STCDH-NH_2
2.69	Ac-Nle**S**$_5$TAV**S**$_5$STCDH-NH_2
2.92	FITC-βAla-Nle**D**TAV**K**STCDH-NH_2
2.93	Ac-Nle**D**TAV**K**STCDH-NH_2

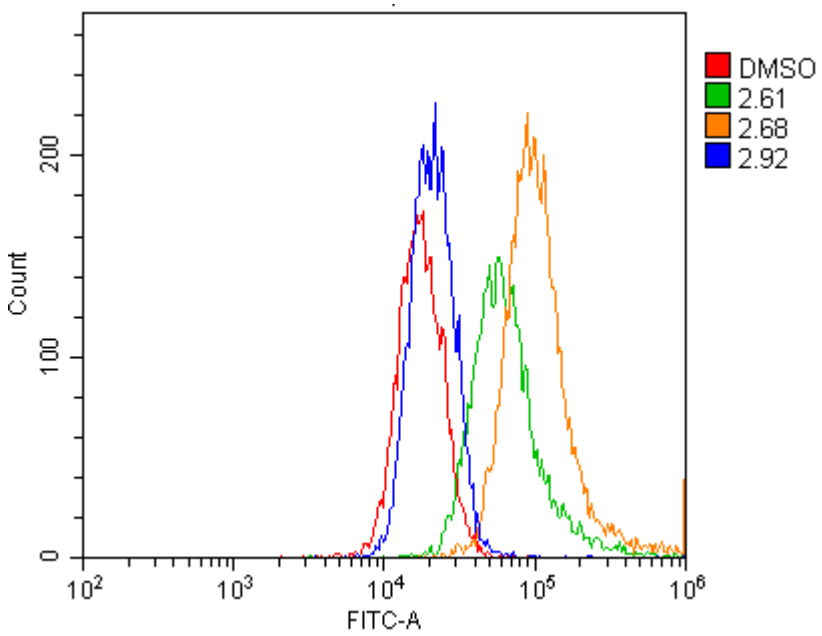


Figure 4.3: *Flow cytometry analysis of DMSO treated A549 cells compared to **2.61**, **2.68** and **2.92** treated cells*

Out of the three peptides, the highest cellular uptake was seen for **2.68** the hydrocarbon stapled version, observed as a shift of the cell population to the right, indicating increased FITC fluorescence on this cell population. This is in line with literature observations of the introduction of a hydrocarbon staple being able to increase the cell permeability of peptides.[30,31] This may also be a reason for the cytotoxic effect of **2.69** (Ac-Nle**S$_5$**TAV**S$_5$**STCDH-NH_2) compared to the linear peptide (**2.47**) and lactam stapled equivalent (**2.93**). Interestingly however the other hydrocarbon stapled peptides did not have any effect on cell viability, with especially noted peptide **2.77** which showed partial disruption of the PPI (29 ± 6% at 200 µM). Potentially this was too low to translate to a cellular effect. The lactam-stapled peptide showed no cellular uptake with only minimal difference from the DMSO treated control cells, indicating that although this staple type can enhance the affinity for dyskerin and was able to inhibit the dyskerin-dyskerin PPI, its cellular effect may be limited by its minimal cellular uptake.

4.3.3 The effect of peptide stapling on the proteolytic stability

The proteolytic stability of the peptides was tested using proteinase K, an enzyme which cleaves peptides sequences at hydrophobic residues. Using PeptideCutter it was predicted to cleave peptide **2.47** in 6 different positions, at the C-terminal side of the three Thr residues as well as the Ala, Val and Ile residues (**Figure 4.4**).[32]



Figure 4.4: *Location of peptide cleaving sites for Proteinase K (red dotted lines) as predicted by PeptideCutter[32].*

The proteolytic stability was recorded with HPLC by measuring the peak integration over time. The linear peptide **2.47** was first tested in this assay but showed immediate degradation upon exposure to proteinase K. Peptide stapling can counteract this problem as unnatural amino acids have been introduced into the sequence which cannot be recognised by the enzyme. To analyse in which position in the sequence the staple has the most effect on elongating the proteolytic stability the hydrocarbon stapled peptides **2.69**, **2.76** and **2.77** were tested in this assay (**Figure 4.5**).

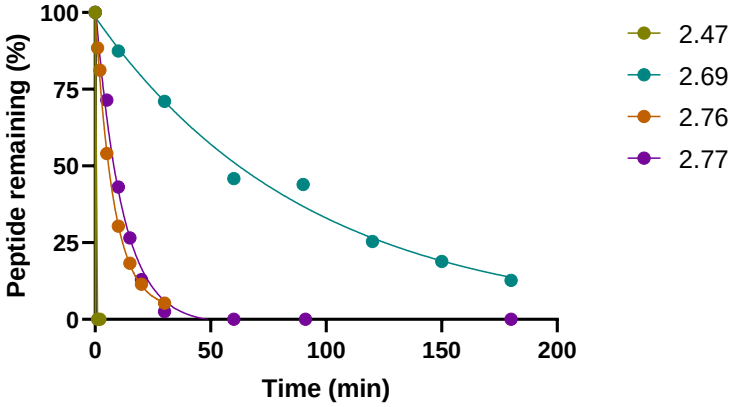


Figure 4.5: *The proteolytic degradation of hydrocarbon stapled peptides **2.47**, **2.69**, **2.76** and **2.77** by proteinase K in 10 mM phosphate buffer, pH 7.4*

This showed that hydrocarbon stapling improves the proteolytic stability of the peptides, but also that this is dependent on the staple location. The half-life of peptide **2.69** ($t_{1/2}$ = 64 min) was significantly higher than for either **2.76** ($t_{1/2}$ = 5.5 min) or **2.77** ($t_{1/2}$ = 8.0 min). Suggesting perhaps the positions which are most easily cleaved using Proteinase K are located within the stapled section in **2.69** (Ac-Nle**S**$_5$TAV**S**$_5$STCDH-NH_2) making them inaccessible to the enzyme. This could be further investigated using mass spectrometry.

Peptide **2.69** was the only compound with an effect on cell viability and together with the cell permeability data, it could be hypothesised that for a cellular effect not only an increase in cell permeability but also an increase in proteolytic stability is needed.

4.3.4 Small molecule cell viability

The small molecules were also evaluated for their effect on cell viability in the MTS assay through a screen at 200 µM to determine whether they showed any antiproliferative activity (**Figure 4.6**).

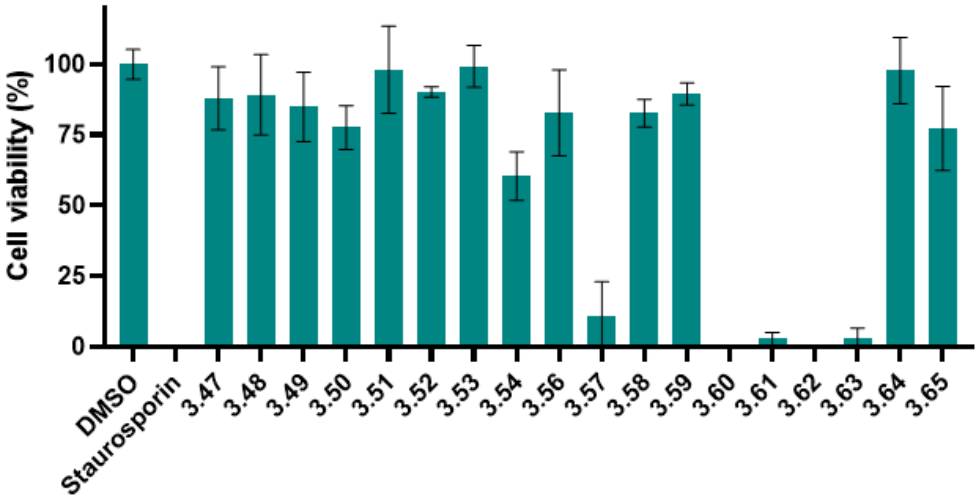
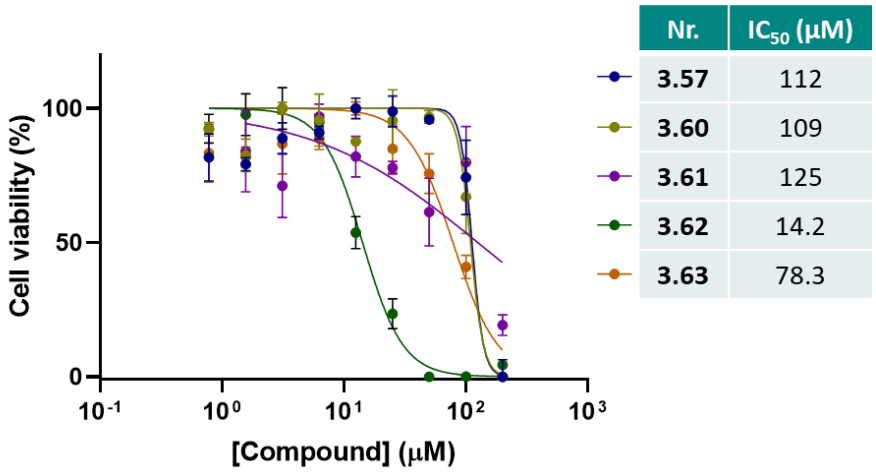


Figure 4.6: *Effect of small molecules **3.47-3.65** on A549 cell viability at 200 µM as determined by MTS assay*

Compounds **3.57**, **3.60**, **3.61**, **3.62** and **3.63** showed minimal cell viability after treatment at this concentration. Their activity was further explored through a 1 in 2 dilution series.



Nr.	IC_{50} (µM)
3.57	112
3.60	109
3.61	125
3.62	14.2
3.63	78.3

Figure 4.7: *Effect of small molecules **3.47-3.65** on A549 cell viability as determined by MTS assay*

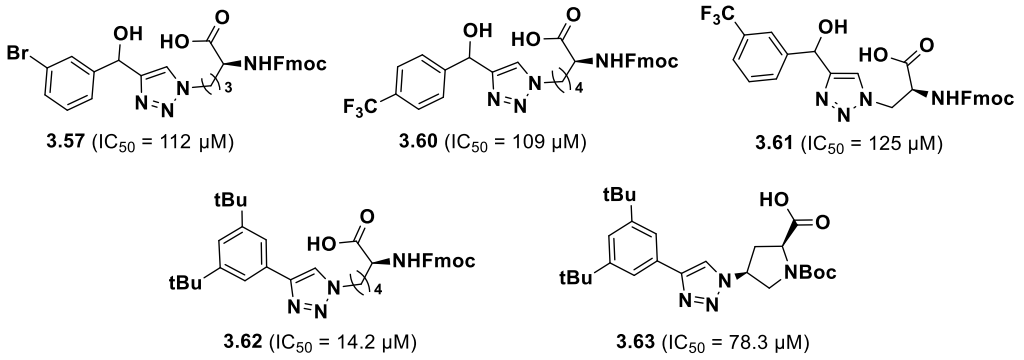


Figure 4.8: *Structures of compounds **3.57, 3.60, 3.61, 3.62** and **3.63***

From the IC_{50} it was clear that the most active compound was **3.62** (IC_{50} = 14.2 µM (95% CI = 13.0 to 15.6 µM), but this compound had shown no activity in the HTRF assay and therefore may be acting through a different mechanism from the inhibition of the dyskerin-dyskerin PPI. The five compounds which showed activity in the assay possess some structural similarity with **3.57**, **3.60** and **3.61** containing similar alkyne side of a substituted phenyl ring with an alcohol and the azide side being Fmoc-amino acid based (**Figure 4.8**). **3.62** and **3.63** both contained the di-tertbutyl substituted phenyl ring on the alkyne side, which are hydrophobic in character which may be underlying their effect in cells.

Out of this series of compounds, **3.57** was active in both the HTRF assay and in the MTS assay and could therefore be further explored as a hit compound targeting the dyskerin-dyskerin PPI. **3.54** and **3.56** that were most active in the HTRF assay did not have any activity in the MTS assay. This could be due to limited cell permeability, which would have to be further explored.

4.4 Conclusion and future work

Hydrocarbon stapled peptide **2.69** was able to inhibit A549 cell growth with an IC_{50} of 132 µM. The results from a cell permeability assay showed that hydrocarbon stapling was more effective at increasing the cell permeability than lactam stapling which could have limited the activity of the lactam stapled peptides. The optimal stapling position for increasing proteolytic stability was also explored with the hydrocarbon staple in **2.69** greatly increasing the stability against proteinase K ($t_{1/2}$ = 64 min.) compared to the linear peptide ($t_{1/2}$ < 1 min.). Since the lactam stapled series was effective at disrupting the dyskerin-dyskerin PPI but lacked cellular effects further optimisation would need to take place to increase these properties.

From the small molecules synthesised five compounds (**3.57**, **3.60**, **3.61**, **3.62** and **3.63**) were able to inhibit cell growth with the highest activity shown by **3.62** (IC_{50} = 14.2 µM). Excitingly

compound **3.57** was active in both HTRF and MTS assays providing a starting point for the further design of cell active small molecules targeting dyskerin. Future work could also be based around the understanding of the mechanism of action of both the peptides and small molecules, for example through CETSA to determine whether these compounds interact with dyskerin in cells. It would be interesting to further understand the effect of peptide **2.69**, which could be further studied through confocal microscopy to look at the localisation of peptide **2.68**. For the two most active small molecules in the HTRF assay (**3.54** and **3.56**) it would also be interesting to further study their cell permeability to understand whether their lack of activity is due to limited cell permeability. This would allow modifications to be made to these molecules to retain activity whilst increasing cell permeability.

4.5 References

1 A. Lin, C. J. Giuliano, A. Palladino, K. M. John, C. Abramowicz, M. L. Yuan, E. L. Sausville, D. A. Lukow, L. Liu, A. R. Chait, Z. C. Galluzzo, C. Tucker and J. M. Sheltzer, Off-target toxicity is a common mechanism of action of cancer drugs undergoing clinical trials, *Sci Transl Med*, 2019, **11**, 1-18.

2 P. S. Burton, R. A. Conradi, N. F. H. Ho, A. R. Hilgers and R. T. Borchardt, How structural features influence the biomembrane permeability of peptides, *Journal of Pharmaceutical Sciences*, 1996, **85**, 1336–1340.

3 H. Derakhshankhah and S. Jafari, Cell penetrating peptides: A concise review with emphasis on biomedical applications, 2018, **108**, 1090–1096.

4 T. Mosmann, Rapid colorimetric assay for cellular growth and survival: Application to proliferation and cytotoxicity assays, *Journal of Immunological Methods*, 1983, **65**, 55–63.

5 A. H. Cory, T. C. Owen, J. A. Barltrop and J. G. Cory, Use of an Aqueous Soluble Tetrazolium/Formazan Assay for Cell Growth Assays in Culture, *Cancer Commun.*, 1991, **3**, 207–212.

6 S. Ansar Ahmed, R. M. Gogal and J. E. Walsh, A new rapid and simple non-radioactive assay to monitor and determine the proliferation of lymphocytes: an alternative to [3H]thymidine incorporation assay, *Journal of Immunological Methods*, 1994, **170**, 211–224.

7 A. L. Niles, R. A. Moravec, P. Eric Hesselberth, M. A. Scurria, W. J. Daily and T. L. Riss, A homogeneous assay to measure live and dead cells in the same sample by detecting different protease markers, *Analytical Biochemistry*, 2007, **366**, 197–206.

8 S. P. M. Crouch, R. Kozlowski, K. J. Slater and J. Fletcher, The use of ATP bioluminescence as a measure of cell proliferation and cytotoxicity, *Journal of Immunological Methods*, 1993, **160**, 81–88.

9 T. L. Riss, R. A. Moravec, A. L. Niles, S. Duellman, H. A. Benink, T. J. Worzella and L. Minor, in *Assay Guidance Manual*, eds. S. Markossian, A. Grossman, K. Brimacombe, M. Arkin, D. Auld, C. Austin, J. Baell, T. D. Y. Chung, N. P. Coussens, J. L. Dahlin, V. Devanarayan, T. L. Foley, M. Glicksman, K. Gorshkov, J. V. Haas, M. D. Hall, S. Hoare, J. Inglese, P. W. Iversen, S. C. Kales, M. Lal-Nag, Z. Li, J. McGee, O. McManus, T. Riss, P. Saradjian, G. S. Sittampalam, M. Tarselli, O. J. Trask, Y. Wang, J. R. Weidner, M. J. Wildey, K. Wilson, M. Xia and X. Xu, Eli Lilly & Company and the National Center for Advancing Translational Sciences, Bethesda (MD), 2004.

10 CellTiter-Glo® Luminescent Cell Viability Assay, https://www.promega.co.uk/products/cell-health-assays/cell-viability-and-cytotoxicity-assays/celltiter_glo-luminescent-cell-viability-assay/, (accessed 21 October 2023).

11 Y. Liu, M. P. Patricelli and B. F. Cravatt, Activity-based protein profiling: The serine hydrolases, *Proceedings of the National Academy of Sciences*, 1999, **96**, 14694–14699.

12 D. M. Molina, R. Jafari, M. Ignatushchenko, T. Seki and Nordlund, Monitoring Drug Target Engagement in Cells and Tissues Using the Cellular Thermal Shift Assay, *Science*, 2013, **341**, 84–87.

13 S. Dayalan Naidu, D. Dikovskaya, T. W. Moore and A. T. Dinkova-Kostova, Detection of thermal shift in cellular Keap1 by protein-protein interaction inhibitors using immunoblot- and fluorescence microplate-based assays, *STAR Protocols*, 2022, **3**, 101265.

14 J. M. Sabatier, E. Vives, K. Mabrouk, A. Benjouad, H. Rochat, A. Duval, B. Hue and E. Bahraoui, Evidence for neurotoxic activity of tat from human immunodeficiency virus type 1, *Virol. J.*, 1991, **65**, 961–967.

15 L. Peraro, K. L. Deprey, M. K. Moser, Z. Zou, H. L. Ball, B. Levine and J. A. Kritzer, Cell Penetration Profiling Using the Chloroalkane Penetration Assay, *J. Am. Chem. Soc.*, 2018, **140**, 11360–11369.

16 C. S. Schasteen, M. G. Donovan and J. N. Cogburn, A novel in vitro screen to discover agents which increase the absorption of molecules across the intestinal epithelium, *Journal of Controlled Release*, 1992, **21**, 49–62.

17 M. Kansy, F. Senner and K. Gubernator, Physicochemical High Throughput Screening: Parallel Artificial Membrane Permeation Assay in the Description of Passive Absorption Processes, *J. Med. Chem.*, 1998, **41**, 1007–1010.

18 Y. Du, in *Protein-Protein Interactions: Methods and Applications*, eds. C. L. Meyerkord and H. Fu, Springer, New York, NY, 2015, pp. 529–544.

19 R. Enseñat-Waser, F. Martin, F. Barahona, J. Vazquez, B. Soria and J. A. Reig, Direct Visualization by Confocal Fluorescent Microscopy of the Permeation of Myristoylated Peptides Through the Cell Membrane, *IUBMB Life*, 2002, **54**, 33–36.

20 R. B. van Breemen and Y. Li, Caco-2 cell permeability assays to measure drug absorption, *Expert Opinion on Drug Metabolism & Toxicology*, 2005, **1**, 175–185.

21 P. Berben, A. Bauer-Brandl, M. Brandl, B. Faller, G. E. Flaten, A.-C. Jacobsen, J. Brouwers and P. Augustijns, Drug permeability profiling using cell-free permeation tools: Overview and applications, *European Journal of Pharmaceutical Sciences*, 2018, **119**, 219–233.

22 C. López-Otín and J. S. Bond, Proteases: Multifunctional Enzymes in Life and Disease, *Journal of Biological Chemistry*, 2008, **283**, 30433–30437.

23 R. Böttger, R. Hoffmann and D. Knappe, Differential stability of therapeutic peptides with different proteolytic cleavage sites in blood, plasma and serum, *PLOS ONE*, 2017, **12**, e0178943.

24 L. Di, E. H. Kerns, Y. Hong and H. Chen, Development and application of high throughput plasma stability assay for drug discovery, *International Journal of Pharmaceutics*, 2005, **297**, 110–119.

25 L. Di and E. H. Kerns, Stability Challenges in Drug Discovery, *Chemistry & Biodiversity*, 2009, **6**, 1875–1886.

26 E. B. Asafu-Adjaye, P. J. Faustino, M. A. Tawakkul, L. W. Anderson, L. X. Yu, H. Kwon and D. A. Volpe, Validation and application of a stability-indicating HPLC method for the in vitro determination of gastric and intestinal stability of venlafaxine, *Journal of Pharmaceutical and Biomedical Analysis*, 2007, **43**, 1854–1859.

27 K. Damm, U. Hemmann, P. Garin-Chesa, N. Hauel, I. Kauffmann, H. Priepke, C. Niestroj, C. Daiber, B. Enenkel, B. Guilliard, I. Lauritsch, E. Müller, E. Pascolo, G. Sauter, M. Pantic, U. M. Martens, C. Wenz, J. Lingner, N. Kraut, W. J. Rettig and A. Schnapp, A highly selective telomerase inhibitor limiting human cancer cell proliferation, *The EMBO Journal*, 2001, **20**, 6958–6968.

28 J. M. Barbosa, C. Onofrillo, A. D. Rio, L. Rocchi and L. Montanaro, Inhibition of Human Dyskerin as a New Approach to Target Ribosome Biogenesis, *PLoS ONE*, 2014, **9**, 1–10.

29 G. Kan, Z. Wang, C. Sheng, C. Yao, Y. Mao and S. Chen, Inhibition of DKC1 induces telomere-related senescence and apoptosis in lung adenocarcinoma, *Journal of Translational Medicine*, 2021, **19**, 161.

30 C. E. Schafmeister, J. Po and G. L. Verdine, An All-Hydrocarbon Cross-Linking System for Enhancing the Helicity and Metabolic Stability of Peptides, *J. Am. Chem. Soc.*, 2000, **122**, 5891–5892.

31 L. D. Walensky, A. L. Kung, I. Escher, T. J. Malia, S. Barbuto, R. D. Wright, G. Wagner, G. L. Verdine and S. J. Korsmeyer, Activation of apoptosis in vivo by a hydrocarbon-stapled BH3 helix, *Science*, 2004, **305**, 1466–1470.

32 E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M. R. Wilkins, R. D. Appel and A. Bairoch, in *The Proteomics Protocols Handbook*, ed. J. M. Walker, Humana Press, Totowa, NJ, 2005, pp. 571–607.

Chapter 5
Recombinant expression of dyskerin

5.1 Introduction

5.1.1 Host organisms for protein expression

During the drug discovery process the in vitro biophysical assays (e.g FP, SPR, ITC, HTRF) described in **chapter 2** are commonly used as a first test of whether a compound can bind its target. For these assays the purified target protein is required and therefore different methods have been developed to recombinantly express proteins. Protein expression can be achieved in different host organisms, including mammalian, bacterial (e.g. *E. coli*) and yeast cells (e.g. Pichia pastoris).[1]

Escherichia coli (*E. coli*) are commonly used due to their fast growth (doubling time of around 20 min) and easy transformation with exogenous DNA.[2] However, if there is a need for post-translational modifications these are not accessible. Additionally, protein expression in *E. coli.* suffers from inclusion body formation and the misfolding of proteins. Alternatively, the expression of proteins in yeast such as in *Pichia pastoris* provides access to most post-translational modifications and improved folding of proteins. Expression in yeast is however slower than in *E. coli* with a doubling time of 1-3 hours.[3] The expression of proteins using mammalian cells is much slower than using bacterial or yeast cells, but, offers the highest folding capacity of the three systems and complex post-translational modifications. Mammalian cells are commonly used for proteins which can be secreted rather than formed intracellularly with a major application found in the production of biopharmaceutical monoclonal antibodies.[4] The Chinese hamster ovary (CHO) cell line is a common mammalian cell line with expression of gram amounts of the desired protein per litre of cell culture.[5]

5.1.2 Protein expression in *E. coli.*

Although yeast and mammalian cells provide better conditions for protein expression, *E. coli.* cells provide a fast, cheap and accessible expression system for routine protein expression on milligram scale where post-translation modifications are not required.[2]

Protein expression in *E. coli* requires that the gene of interest is first cloned into an expression vector. This vector can contain antibiotic resistance genes alongside genes to express affinity tags for the purification and solubility of the target protein (**Figure 5.1**).

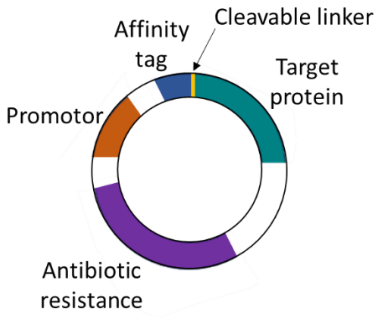


Figure 5.1: *Cartoon overview of plasmid used for protein expression*

Resistance genes ensure the growth of only the *E. coli* cells containing the desired plasmid.[2] For example, the ampicillin resistance gene encodes the production of β-lactamase, an enzyme which breaks the β-lactam ring in ampicillin deactivating it.[6] Genes encoding for affinity tags can be added to the plasmid for detection, purification and improved solubility. Common tags include small peptide tags such as a His-tag (commonly 6 x His) or a FLAG-tag (DYKDDDDK), used for affinity purification or protein detection. Larger tags such as maltose-binding protein (MBP), glutathione S-transferase (GST) and Small Ubiquitin-like Modifier (SUMO) are proteins that can enhance the solubility of the protein and enable affinity purification.[7] Often affinity tags are connected to the protein using a linker sequence which can be cleaved using a specific protease enzyme, such as Tobacco Etch Virus protease (TEV), thrombin or enterokinase. This allows for the removal of the tag from the protein if needed for subsequent analysis.[8]

Once the desired plasmid is prepared it is transformed into an *E. coli* cell line suitable for protein expression. In *E. coli* expression cell lines the *lac* operon controls the metabolism of lactose which occurs in the absence of glucose. Normally glucose is the preferred carbon source for bacteria but in the absence of glucose β-galactosidase can be expressed which can digest lactose.[9] The *lac* operon has two methods of control, using the lac repressor protein and using catabolite activator protein (CAP). In the absence of lactose, lac repressor protein binds the *lac* operon stopping the transcription of β-galactosidase (**Figure 5.2A**). In the presence of lactose the lactose binds to the lac repressor protein inducing a conformational shift that inhibits binding to the lac operon (**Figure 5.2B**). At the same time, there needs to be an absence of glucose, inducing the expression of cyclic adenosine monophosphate (cAMP), a messenger molecule, which bind CAP. This allows CAP to bind DNA, which helps RNA polymerase binding to the promotor sequence ensuring high transcription levels.[10]

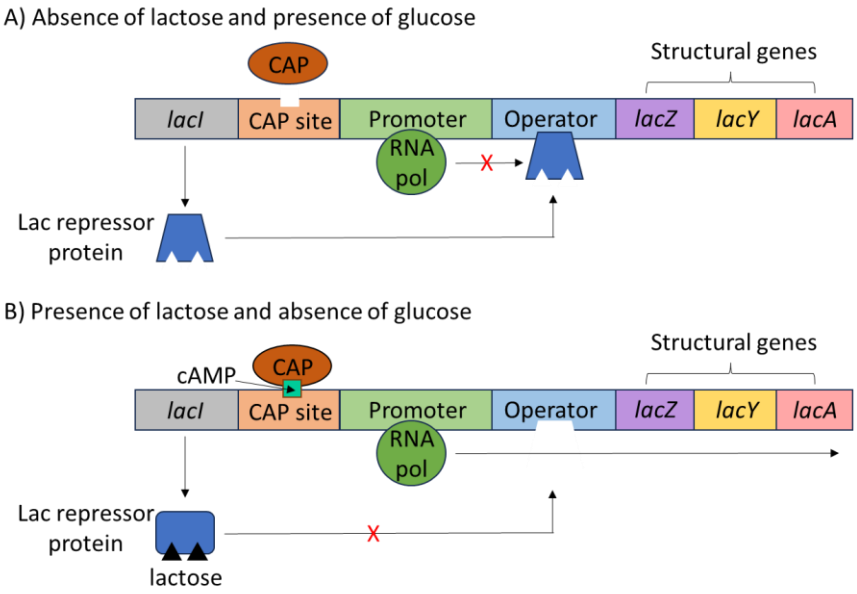


Figure 5.2: *A) lac operon in the absence of lactose and the presence of glucose, B) lac operon in the presence of lactose and the absence of glucose*

This mechanism can be exploited to control the expression of protein until an inducer is added. The presence of the lac repressor protein inhibits the transcription of the genes following the lac operon which contain the genes for the target protein. The expression can then be induced using isopropyl β-D-1-thiogalactopyranoside (IPTG), a lactose analogue, which binds the lac repressor protein, changing its conformation, stopping its binding to the lac operon which leads to the transcription of the genes encoding the protein.[2] IPTG is used over lactose since it has the same function as lactose on the lac repressor but it is not a substrate of β-galactosidase and therefore cannot be metabolised. Different amounts of IPTG, induction temperature, and induction time length can be screened to optimise the expression levels which are dependent on the target protein. After the expression process the cells can be collected using centrifugation after which they are lysed allowing the cell lysate to be purified to obtain the target protein.

5.1.3 Protein purification methods

Once the desired protein has been expressed purification of the protein is needed to separate the desired protein from all of the other proteins present in the cell lysate necessary for *E. coli* function. Affinity chromatography is a good first method of purification due to its selective nature for the protein of interest which should allow the removal of most native *E. coli.* proteins. His-tagged proteins can be purified using immobilized metal affinity chromatography (IMAC) through the chelation of the histidine residues to Ni^{2+}, commonly with agarose functionalised with a nickel nitrilotriacetic (Ni-NTA) complex.[11] The protein is

subsequently eluted from the resin using imidazole, which competes with the histidine residues for binding to the Ni-NTA resin. This may lead to pure protein, however there are common impurities found from *E. coli* proteins with high histidine content.

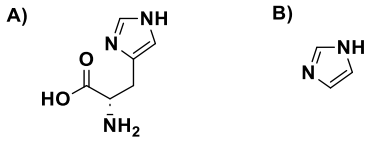


Figure 5.3: *Structures of A) L-histidine and B) imidazole*

Proteins containing a FLAG-tag can be captured through anti-FLAG antibodies immobilised on agarose resin. The FLAG-tagged protein can be removed from the resin using FLAG-tag peptide which competes with the protein for binding to the antibodies.[12] GST can selectively bind glutathione immobilised agarose and can be eluted using reduced glutathione as a competitor.[13] MBP can be purified similarly through binding to amylose immobilised agarose from which it is eluted using maltose.[14]

After affinity chromatography any remaining impurities can be removed using methods not relying on the affinity tag, such as size exclusion chromatography (SEC) or ion exchange chromatography (IEX). In SEC proteins are separated by size by a filtration through a gel resin containing pores of a specific size. The smaller the protein the more time it takes to elute due to the protein making its way through the pores leading to a longer pathlength. Larger proteins may not be able to enter these pores going instead around the resin particles and therefore having a shorter pathlength.[15]

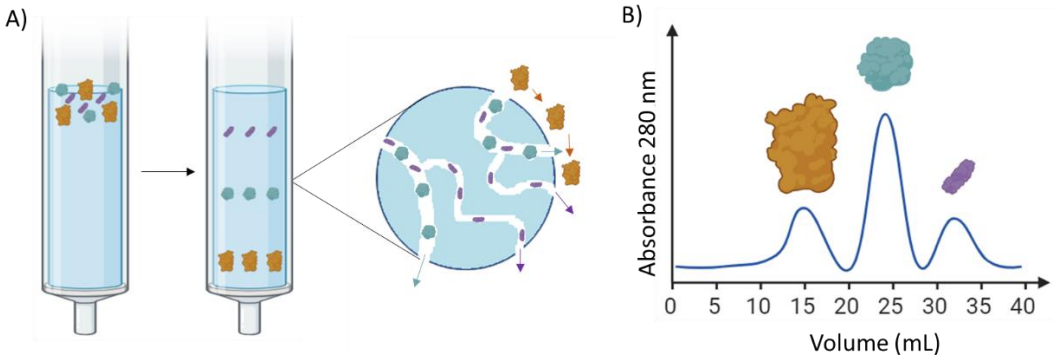


Figure 5.4: *Schematic overview of Size exclusion chromatography (SEC) with A) the separation of the proteins in the column with magnified particle and B) the elution spectrum of the three proteins over increasing volume*

In IEX the proteins are separated based on their net charge. The more positively charged the protein is the higher the affinity for the negatively charged column beads. Through increasing concentrations of sodium chloride the proteins can be eluted based on their charge.[16]

5.2 Chapter aims

In this chapter we aimed to express and purify the dyskerin protein for the use in biophysical assays needed for the drug discovery process. The expression of dyskerin will be attempted in *E. coli* with plasmids described in the literature, optimising expression cell lines and growth conditions. Subsequently, optimal isolation and purification conditions will be explored.

5.3 Results and discussion

5.3.1 His-dyskerin protein expression

To investigate whether it is possible to target the dyskerin-dyskerin protein-protein interaction with peptides or small molecules we needed to develop *in vitro* biophysical assays, which require purified dyskerin protein.

A search of the literature identified previous descriptions of the recombinant expression of a dyskerin construct without large protein tags. We thought this was the most desirable, as large tags may interfere with the dimer formation of dyskerin, or the binding site of our prepared molecules.[17] A His-tagged dyskerin plasmid was described by the Harrington group, who kindly shared a sample with us, spotted on Whatman filter paper. We began by attempting to express the His-tagged dyskerin protein (His-dyskerin) following the expression procedure reported by the Harrington group.[17] The plasmid was retrieved from the filter paper using TE buffer (10 mM tris(hydroxymethyl)aminomethane (Tris), 1 mM ethylenediaminetetraacetic acid (EDTA), pH 8.0). TE buffer solubilises the plasmid and the EDTA inactivates deoxyribonuclease (DNase) protecting the plasmid from degradation.[18]

Our work began with the transformation of *E. coli* cells with the dyskerin plasmid. We selected BL21(DE3)pLysS cells, a commonly used cell line in protein expression because of the pLysS functionality which suppresses basal expression of the target protein. This occurs through the expression of T7 lysozyme which binds T7 RNA polymerase preventing the initiation of transcription. Upon IPTG induction the expression level of T7 RNA polymerase is higher than T7 lysozyme can inhibit, allowing for protein expression to occur. The suppression of basal protein expression in beneficial because the protein may be toxic to the cells inhibiting their growth. These cell differ from the literature procedure which used BL21 Codon Plus RIL cells.[17] BL21 Codon Plus RIL cells are used for more efficient and higher levels of eukaryotic protein expression by containing extra copies of leuW, ileY and argU tRNA genes which encode tRNAs that recognize four codons.[19] These are needed for the translation of eukaryotic codons that are more rarely found in *E. coli*, limiting translation. BL21 Codon Plus RIL cells are not commonly available in the UK, but as an alternative Rosetta

2(DE3)pLysS cells could be used. These Rosetta cells contains these extra genes with the additional expression of a further three tRNAs.[20] However, long lead times were estimated for the purchase of these cells and therefore the protein expression was first explored in BL21(DE3)pLysS cells with the expectation that the protein expression may still take place but potentially at lower levels than in Rosetta 2(DE3)pLysS cells.

Initially, transformation was attempted by addition of the plasmid to the cells, followed by a heat shock at 42 °C for 30 seconds. The cells were then recovered in super optimal broth with catabolite repression (S.O.C.), a nutrient-rich medium, and incubated for one hour at 37 °C. The cells were plated on agar containing chloramphenicol (34 µg/mL) and kanamycin (50 µg/mL) for which the cells and the plasmid respectively contained resistance genes and the plates were incubated overnight. However, no colonies had formed the next day indicating an unsuccessful transformation of the plasmid into the cells.

Common troubleshooting for plasmid expression includes optimising heat shock conditions, transfection media, and cell recovery time, or the severity of the selection conditions. We screened a number of conditions, including two different heat shock lengths of 50 and 90 seconds which proved unsuccessful. The kanamycin concentration was lowered to 15 µg/mL and chloramphenicol was omitted, but again no colonies formed. Omitting the chloramphenicol is not ideal as the pLysS functionality is only operational in the presence of chloramphenicol[21], but if the cells grew in the presence of ampicillin the plasmid has been accepted, and as such this drawback was deemed acceptable as dyskerin has not been reported to be toxic to E. coli. Optimisation can continue after confirmation that protein production was possible. We added β-mercaptoethanol to the cells before incubating with the plasmid, which improves the transformation efficiency, and the cells were plated on agar containing only kanamycin (15 µg/mL).[22] Pleasingly, multiple colonies grew overnight, and single colonies were selected and grown in Luria-Bertani broth (LB) containing kanamycin (15 µg/mL) overnight at 37 °C. An overnight culture showing growth was used to continue into a larger culture (200 mL, 1:50 dilution in LB). The cells were incubated at 37 °C until an OD_{600} of 0.57 was reached. At this point IPTG was added to induce protein expression and the cells were incubated at 30 °C for four hours, before collecting the cell pellets through centrifugation. The cells were lysed using sonication in lysis buffer (50mM Tris, 10mM Imidazole, 300mM NaCl, pH 8.0) supplemented with a protease inhibitor cocktail and lysozyme (0.25 mg/mL) right before sonication. The protease inhibitor cocktail inhibits proteases, including cysteine, serine, and metalloproteases as well as calpains. These proteases present in *E. coli* would otherwise degrade the dyskerin formed. Lysozyme is a

small protein (14 kDa) derived from egg white and degrades the peptidoglycan layer of the cell wall aiding together with the sonication in the lysis of the cells.[23]

The cell lysate was clarified through centrifugation and the supernatant was purified by immobilized-metal affinity chromatography (IMAC), exploiting the polyhistidine tag of the protein and nickel nitrilotriacetic complex (Ni-NTA) functionalised agarose .[11] Analysis of the IMAC fractions was performed using sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) using NuPAGE™ MOPS SDS buffer. The gel was stained with Coomassie brilliant blue which binds with high affinity to proteins through electrostatic and hydrophobic interactions.[24] It also binds the gel itself with low affinity and therefore after the staining process the gel was destained using 10/20/70 acetic acid/methanol/water leaving only the protein bands visible.

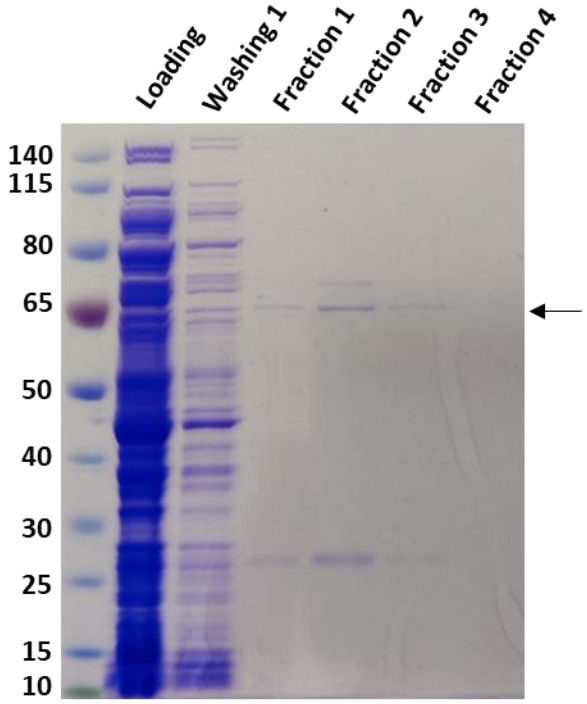


Figure 5.5: *4-12% Bis-Tris SDS-PAGE analysis of Ni-NTA resin purification with L = loading before purification, W = washing (20 mM imidazole), F1-F4 = elution fractions 1-4 (250 mM imidazole)*

The gel (**Figure 5.5**) showed a band around 65 kDa in elution fractions 1-3, which was higher than the band expected for dyskerin (57 kDa), but the closest band seen to the desired MW. The band at 65 kDa was cut out from the gel and mass spectrometry used for sequence analysis. The results showed that the protein present was glucosamine-6-phosphate synthase (GlmS). This is an *E. coli* protein known to bind to Ni-NTA resin through its surface exposed histidine residues.[25] Since no other bands were seen around the desired molecular

weight it could be assumed that His-dyskerin was not being expressed in soluble form in these BL21(DE3)pLysS cells under these conditions.

At the same time we also explored whether the lack of dyskerin expression was due to protein toxicity. Although the BL21(DE3)PLysS cells are meant for the expression of toxic proteins through the basal suppression of expression of protein without IPTG present. However, for this PLysS functionality the cells have to be grown in the presence of chloramphenicol. The transformed cells however did not grow any colonies on plates with chloramphenicol present and therefore this pLysS functionality could not be utilised. There are other methods which can suppress basal expression of the target protein, such as through the addition of glucose to the medium (see **section 5.1.2**).

The BL21(DE3)pLysS cells were grown in the presence of 0.5% glucose (w/v) in the growth of the overnight culture and during the growth stage in the large (500 mL) cultures for the protein expression. The cells were induced and lysed and the lysate was purified by IMAC purification using HisPur cobalt resin. This has the same function as Ni-NTA resin, but is cobalt based instead of nickel based and is supposed to lead to less *E. coli* protein impurities being retained on the resin. Two wash steps were added to the purification, the first with 20 mM imidazole and a second with 50 mM imidazole. Any remaining protein was then eluted with 250 mM imidazole and the fractions were analysed by SDS-PAGE (**Figure 5.6**).

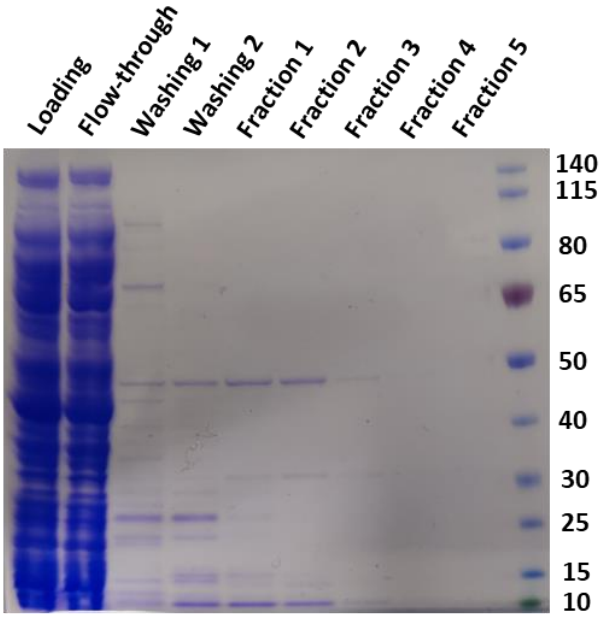


Figure 5.6: *4-12% Bis-Tris SDS-PAGE analysis of Co-resin purification with L = loading before purification, FT = flow through, W1 = washing (20 mM imidazole), W2 = washing (50 mM imidazole) and F1-F5 = elution fractions 1-5 (250 mM imidazole)*

No bands were present in any of the washing or elution fractions between 50 and 65 kDa where dyskerin would be expected. This method did show that the cobalt resin does not retain the band around 65 kDa for GlmS which was seen using Nickel resin purification.

It was further explored whether maybe the glucose would have to be present from the beginning of the cell growth such as in the agar plates. Therefore the transformation of plasmid into BL21(DE3)pLysS cells was repeated with both 0.5% glucose and chloramphenicol (34 µg/mL) as well as kanamycin (15 µg/mL) present in the agar plates. This saw the successful growth of colonies, perhaps due to the presence of glucose suppressing any expression of protein which may be toxic. The cells were grown and lysed and the cell lysate was purified with the cobalt resin, but the elution fraction showed no protein bands within the 50-65 kDa region. The flow-through was loaded onto the Ni-NTA resin to investigate if there was dyskerin protein production but that it was unable to bind the cobalt resin. However, no bands were seen in the wash or elution fractions from the Ni-NTA resin that could correspond to dyskerin.

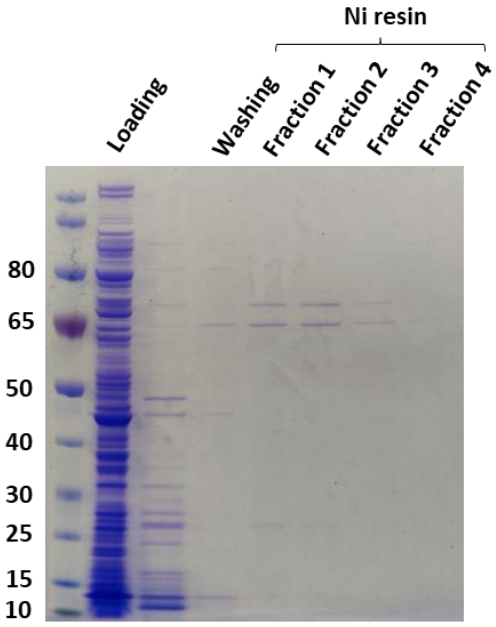


Figure 5.7: *4-12% Bis-Tris SDS-PAGE analysis of Co-resin and Ni-NTA resin purification with L = loading before purification, E = Co elution (250 mM imidazole), W = Ni washing (20 mM imidazole) and F1-F5 = Ni elution fractions 1-5 (250 mM imidazole)*

Since the suppression of basal protein expression using glucose also did not yield the desired dyskerin protein, our troubleshooting methods were becoming limited. Although a different cell line could be beneficial, the transformation of this plasmid into these other cell lines was unsuccessful, and having not yet observed any protein expression we hypothesised that perhaps this construct may not yield the amounts of dyskerin required for our desired

application in *in vitro* biophysical assays. A previous literature report for the expression of dyskerin also showed poor expression with only low levels of dyskerin seen.[26]

Therefore, we returned to the literature for an alternative strategy for dyskerin expression . utilising larger protein tags that could assist with expression, solubility, and stability.

5.3.2 MBP-dyskerin expression

The expression of dyskerin fused to maltose binding protein (MBP) has been described by Prof. Thomas Meier, whose research focusses on H/ACA RNPs and the interactions of proteins within the complex.[27] To investigate the role of dyskerin within them, the protein was expressed fused to a maltose binding protein (MBP), which can enhance the expression of soluble protein.[28] The plasmid for MBP-dyskerin was kindly shared by Prof. Thomas Meier.

The plasmid was retrieved from filter paper, and transformation into a panel of *E. coli* cell lines was performed (**Table 5.1**).

Table 5.1: *Cell lines tested for MBP-dyskerin expression with a brief description of their key properties.*

Number	Cell line	Description
1	BL21(DE3)	Routine *lac* operon controlled expression strain[29]
2	BL21(DE3)pLysS	Version of BL21(DE3) strain with basal suppression of target protein for the expression of toxic proteins[30]
3	Rosetta2(DE3)pLysS	Enhanced expression of eukaryotic proteins containing rare codons and basal suppression of target protein[20]
4	NiCo21(DE3)	Alternative to BL21(DE3) engineered to minimise *E. coli.* protein contaminations during IMAC protein purification[31]
5	NEB® 5-alpha	Used for high-quality plasmid preparation[32]

Cell lines **1**-**4** are used for protein expression and NEB® 5-alpha cells are used for cloning and plasmid extraction. Transformation into these cell lines was performed using a 45 second heat shock was performed after which the cells were plated on agar containing ampicillin (50 µg/mL – plasmid resistance) and for the pLysS strains (2 and 3) this was further supplemented with chloramphenicol (34 µg/mL – cell line resistance).

After overnight incubation at 37 °C only colony formation was seen for the NEB® 5-alpha cell line. The poor transformation of the plasmid into the cells was potentially due to impurities

present within the plasmid solution. Using a NanoDrop spectrophotometer the 260 nm/280 nm absorbance ratio was measured to be 1.29. This is lower than the expected ration of around 1.8 for pure DNA and therefore this could indicate the presence of impurities within the plasmid solution which may impact the transformation efficiency.[33] Therefore, single colonies from NEB® 5-alpha cell line were grown in LB (10 mL) overnight at 37 °C at 240 rpm. The cells were pelleted using centrifugation (6000 g, 18 min., 4 °C) and the supernatant was discarded. From this pellet the plasmid was extracted using the QIAprep spin Miniprep kit which showed a 260 nm/280 nm ration of 1.90, indicating a higher purity of DNA.

With purified plasmid now in hand, we attempted to transform this plasmid into Rosetta 2(DE3)pLysS cells which pleasingly resulted in colony formation. A single colony was picked and grown in LB overnight which was diluted the next day into larger cultures. Following literature proceedings[27], the cells were grown until an optical density at 600 nm (OD_{600}) of 0.6 was reached at which point the protein expression was induced. The cells were collected and lysed after which the cell lysate was purified using amylose resin. Amylose is a polysaccharide which specifically binds to MBP providing a way to purify the cell lysate.[14] The supernatant was loaded onto the resin using gravity after which the resin was washed with the lysis buffer to wash away any unbound protein. The desired MBP-dyskerin was then eluted using elution buffer (20 mM Tris, 200 mM NaCl, 1 mM EDTA, 1 mM DTT, 10 mM maltose pH 7.4). Maltose is a disaccharide which competes with the amylose resin for binding to MBP leading to the elution of the protein. The fractions were analysed using 4-12% SDS-PAGE (**Figure 5.8**).

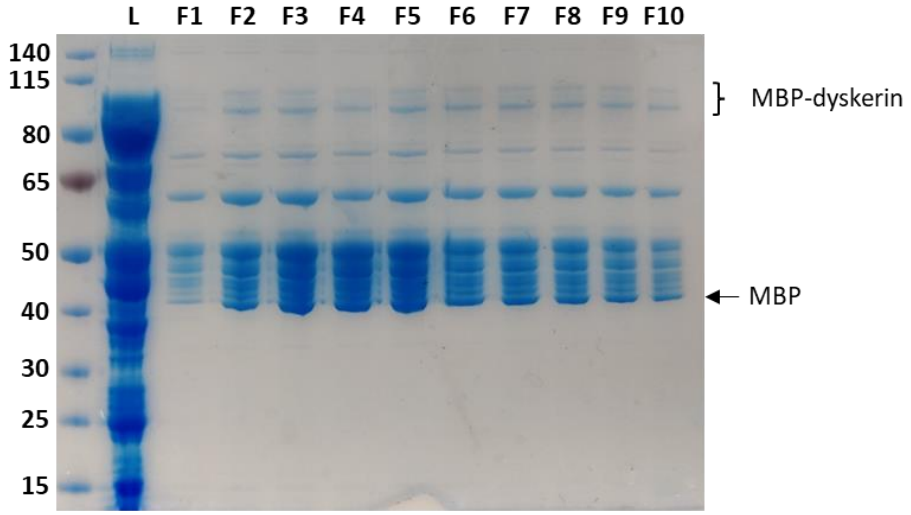


Figure 5.8: *4-12% Bis-Tris SDS-PAGE gel analysis of amylose resin purification of MBP-dyskerin, L = loading, F1-F10 = Elution fractions 1-10*

From the analysis of the fractions of the amylose resin purification the desired MBP-dyskerin product was seen, but also a number of impurities were found. The bands were all found between the molecular weight for MBP (41 kDa) and the full-length protein (108 kDa). The two top bands matched the literature expression of MBP-dyskerin in which both bands showed the same spectrum when analysed by mass spectrometry, not identifying a reason for the observation of two bands by SDS-PAGE. The lower mass bands between 40-80 kDa could be due to the incomplete formation of the MBP-dyskerin protein. The protein translation starts with the MBP-tag and the purification is selective for this tag and therefore the lowest molecular weight protein is seen for the MBP-tag itself without dyskerin. From this molecular weight until the molecular weight of the full protein several bands were found, which would correspond to the partial formation of MBP-dyskerin.

The desired MBP-dyskerin also contains a C-terminal His-tag which can be used for further purification using Ni-NTA resin. This was attempted and the fractions were analysed by SDS-PAGE (**Figure 5.9**). Although the bands thought to be MBP-dyskerin were found in the elution fractions, the same bands were also seen in the flow-through. This could potentially be mediated through purifying the protein solution using more Ni-NTA resin, longer incubation time, or a slower flow rate. However, this was not attempted as the protein that was found in the elution fractions was not completely pure with a lower band (50-60 kDa) still present.

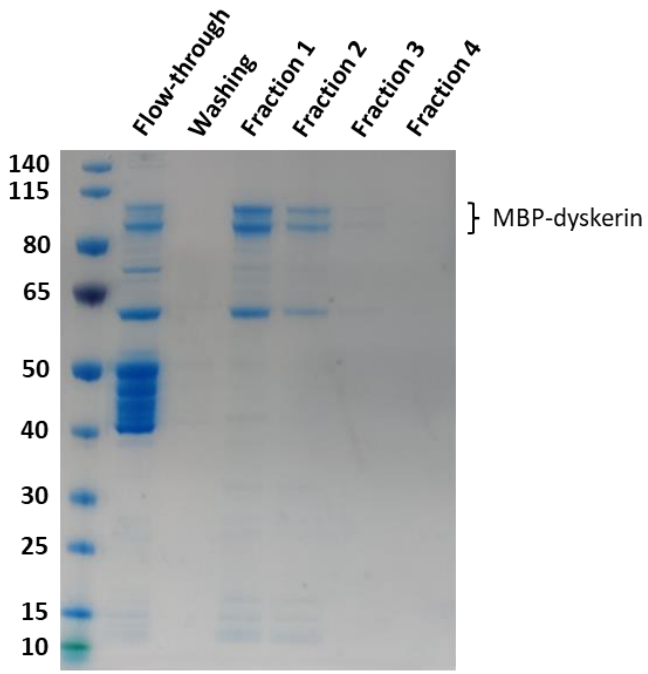


***Figure 5.9**: 4-12% Bis-Tris SDS-PAGE gel analysis of Ni-NTA resin purification of MBP-dyskerin, FT = Flow-through, W = Wash (20 mM imidazole), F1-F4 = Elution fractions 1-4 (250 mM imidazole)*

Further purification of the protein to remove the bands below 65 kDa was attempted using size exclusion chromatography (SEC). For the SEC Superdex 75 prep grade resin was used which is recommended for the separation of proteins in the 3000-70000 kDa range. Although MBP-dyskerin is heavier (108 kDa) than this range, the impurities found are within that range which therefore should allow for the separation of MBP-dyskerin from the impurities. Superdex 75 prep grade resin is a matrix composed of cross-linked dextran, needed for separation, and cross-linked agarose, providing rigidity which allows higher flow-rates to be used. The elution fractions from the Ni resin were purified by SEC and the collected fractions were analysed by SDS-PAGE (**Figure 5.10**).

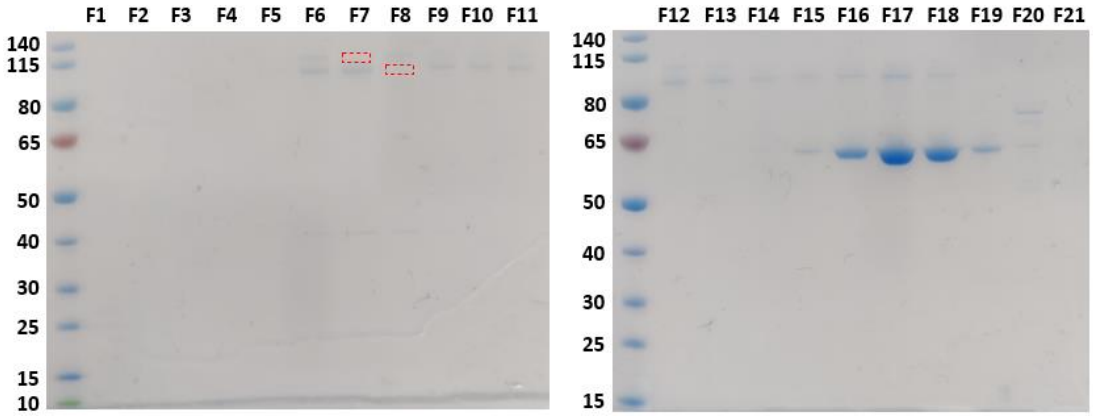


Figure 5.10: *4-12% Bis-Tris SDS-PAGE analysis of size exclusion chromatography of MBP-dyskerin, F1-F21 = Fractions 1-21. In red dotted line is indicated the spaces from which the gel was removed for peptide mass fingerprinting analysis.*

This gel analysis showed that fractions 6-14 contained the desired MBP-dyskerin without any major impurities with two bands reported as MBP-dyskerin in literature. A sample of both bands (top band from F7 and bottom band from F8) were cut from the gel and prepared for analysis by MALDI peptide mass fingerprinting (PMF) with MBP-dyskerin detected in the lower band with no protein detected in the top band, potentially due to a low concentration of protein present. However, with MBP-dyskerin found in the lower band and the two bands matching literature reports it was assumed that both bands are MBP-dyskerin.[27]

The process of purification of MBP-dyskerin from the flow-through of the Ni-NTA resin also showed that there would be a possibility to remove the Ni-NTA resin purification from the protein expression process. After the amylose-resin purification the elution fractions were concentrated to around 1-1.5 mL and immediately purified by SEC chromatography. (**Figure 5.11**).

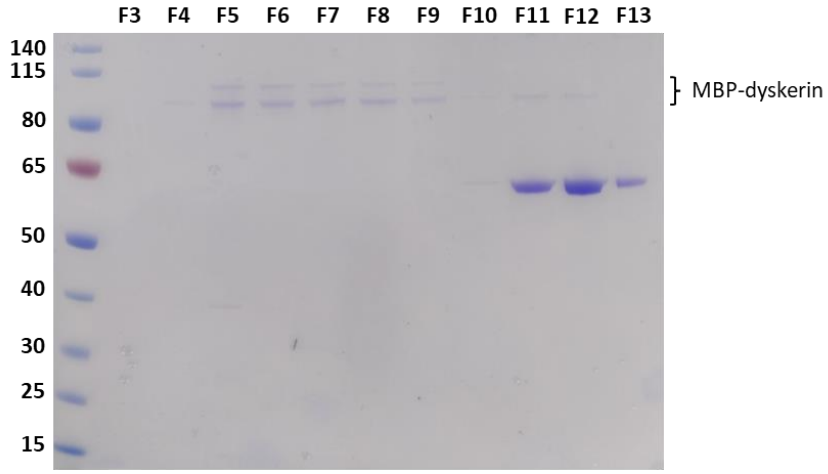


Figure 5.11: *4-12% Bis-Tris SDS-PAGE analysis of size exclusion chromatography of MBP-dyskerin*

Fractions 4-9 contained pure MBP-dyskerin and were concentrated using a 10 kDa MWCO spin column. The removal of the Ni-NTA-resin purification step from the protein preparation also saved time as there is no buffer exchange needed after the amylose resin purification. The size exclusion column is equilibrated in the desired buffer for assay applications (10 mM HEPES, 150 mM NaCl, 0.05% Tween-20, pH 7.4) and during the purification a buffer exchange also occurs leading to the isolation of approximately around 1 mg of MBP-dyskerin from 2L of cell culture.

Through the use of the MBP-dyskerin plasmid in Rosetta2(DE3)pLysS it was possible to express soluble dyskerin protein. The combination of the amylose resin purification with SEC led to the isolation of pure protein comparable to literature. Concurrently peptides were being synthesised and the production of the protein now allowed for the testing of any synthesised compounds in *in vitro* assays in **chapters 2 and 3**.

5.4 Conclusion and future work

In this chapter dyskerin was successfully expressed as an MBP-fusion protein. The expression of His-dyskerin was first attempted using BL21(DE3)pLysS cells without the presence of chloramphenicol, but no soluble expression of dyskerin was found. Exploring whether the production of the protein may be toxic to the cells, the addition of glucose to the media was explored. However, no soluble expression of His-dyskerin was seen. Further explorations of the transformation of the plasmid into the Rosetta 2(DE3)pLysS cells, which are more suitable for the expression of eukaryotic proteins was unsuccessful. Therefore instead of the expression of His-dyskerin the expression of MBP-dyskerin was explored. Although the transformation of the MBP-dyskerin plasmid into protein expression cell lines was initially unsuccessful, it was successfully transformed into NEB5α cloning cells. Using these cells pure

plasmid was extracted and successfully transformed into Rosetta 2(DE3)pLysS cells. Using a combination of purification by amylose resin followed by size exclusion chromatography the MBP-dyskerin protein was successfully purified. This protein was then used in the biophysical assays described in **Chapters 2** and **3**.

Subsequent work includes cleavage of the fusion-tag from dyskerin to assess whether it may be possible to isolate dyskerin with only a His-tag. Although different buffer compositions may need to be tried to stop the precipitation of the dyskerin protein without the MBP-tag. Discussions with Prof. Meier, who originally designed the MBP-dyskerin construct highlighted that dyskerin was unstable without the added MBP tag, although this has not been described in the literature. It be interesting to further explore the possibility for the crystallisation of the MBP-dyskerin protein by itself and with peptide and small molecules inhibitors discovered in **Chapters 2** and **3** to examine the binding mode of the different compounds.

5.5 References

1 W. H. Brondyk, in *Methods in Enzymology*, eds. R. R. Burgess and M. P. Deutscher, Academic Press, 2009, vol. 463, pp. 131–147.

2 G. L. Rosano and E. A. Ceccarelli, Recombinant protein expression in Escherichia coli: advances and challenges, *Front. Microbiol.*, 2014, **5**, 1–17.

3 M. Karbalaei, S. A. Rezaee and H. Farsiani, Pichia pastoris: A highly successful expression system for optimal synthesis of heterologous proteins, *J. Cell. Physiol.*, 2020, **235**, 5867–5881.

4 J. Zhu, Mammalian cell protein expression for biopharmaceutical production, *Biotechnol. Adv.*, 2012, **30**, 1158–1170.

5 B. Figueroa Jr., E. Ailor, D. Osborne, J. M. Hardwick, M. Reff and M. J. Betenbaugh, Enhanced cell culture performance using inducible anti-apoptotic genes E1B-19K and Aven in the production of a monoclonal antibody with Chinese hamster ovary cells, *Biotechnol. Bioeng.*, 2007, **97**, 877–892.

6 M. Babic, A. M. Hujer and R. A. Bonomo, What's new in antibiotic resistance? Focus on beta-lactamases, *Drug Resist. Updat.*, 2006, **9**, 142–156.

7 M. E. Kimple, A. L. Brill and R. L. Pasker, Overview of Affinity Tags for Protein Purification, *Curr. Protoc. Protein Sci.*, 2013, **73**, 9.9.1-9.9.23.

8 S. Costa, A. Almeida, A. Castro and L. Domingues, Fusion tags for protein solubility, purification and immunogenicity in Escherichia coli: the novel Fh8 system, *Front. Microbiol.*, 2014, **5**, 1–20.

9 D. H. Juers, B. W. Matthews and R. E. Huber, LacZ β-galactosidase: Structure and function of an enzyme of historical and molecular biological importance, *Protein Sci.*, 2012, **21**, 1792–1807.

10 W. S. Reznikoff, The lactose operon-controlling elements: a complex paradigm, *Mol. Microbiol.*, 2006, **6**, 2419–2422.

11 E. Hochuli, W. Bannwarth, H. Döbeli, R. Gentz and D. Stüber, Genetic Approach to Facilitate Purification of Recombinant Proteins with a Novel Metal Chelate Adsorbent, *Bio/Technology*, 1988, **6**, 1321–1325.

12 B. L. Brizzard, R. G. Chubet and D. L. Vizard, Immunoaffinity purification of FLAG epitope-tagged bacterial alkaline phosphatase using a novel monoclonal antibody and peptide elution, *BioTechniques*, 1994, **16**, 730–735.

13 F. Schäfer, N. Seip, B. Maertens, H. Block and J. Kubicek, Purification of GST-Tagged Proteins, *Methods Enzymol.*, 2015, **559**, 127–139.

14 K. C. Duong-Ly and S. B. Gabelli, Affinity Purification of a Recombinant Protein Expressed as a Fusion with the Maltose-Binding Protein (MBP) Tag, *Methods Enzymol.*, 2015, **559**, 17–26.

15 S. Fekete, A. Beck, J.-L. Veuthey and D. Guillarme, Theory and practice of size exclusion chromatography for the analysis of protein aggregates, *J. Pharm. Biomed. Anal.*, 2014, **101**, 161–173.

16 A. Jungbauer and R. Hahn, in *Methods in Enzymology*, eds. R. R. Burgess and M. P. Deutscher, Academic Press, 2009, vol. 463, pp. 349–371.

17 L. Gardano, L. Holland, R. Oulton, T. Le Bihan and L. Harrington, Native gel electrophoresis of human telomerase distinguishes active complexes with or without dyskerin, *Nucleic Acids Res.*, 2012, **40**, 1–12.

18 N. Yagi, K. Satonaka, M. Horio, H. Shimogaki, Y. Tokuda and S. Maeda, The Role of DNase and EDTA on DNA Degradation in Formaldehyde Fixed Tissues, *Biotech. Histochem.*, 1996, **71**, 123–129.

19 T. Kleber-Janke and W.-M. Becker, Use of Modified BL21(DE3) Escherichia coli Cells for High-Level Expression of Recombinant Peanut Allergens Affected by Poor Codon Usage, *Protein Expr. Purif.*, 2000, **19**, 419–424.

20 Rosetta 2(DE3)pLysS Singles Competent Cells - Novagen Novagen's Rosetta 2 (pLysS) host strains are BL21 derivatives designed to enhance the expression of eukaryotic proteins that contain codons rarely used in E. coli along with T7 lysozyme activity. | Sigma-Aldrich, http://www.sigmaaldrich.com/, (accessed 6 September 2023).

21 BL21(DE3) Chemically Competent Cells for protein expression | Agilent, https://www.agilent.com/en/product/protein-expression/competent-cells-for-routine-protein-expression/general-protein-expression/bl21%28de3%29-competent-cells-232943, (accessed 28 September 2023).

22 T. E. V. Aune and F. L. Aachmann, Methodologies to increase the transformation efficiencies and the range of bacteria that can be transformed, *Appl. Microbiol. Biotechnol.*, 2010, **85**, 1301–1313.

23 P. Ferraboschi, S. Ciceri and P. Grisenti, Applications of Lysozyme, an Innate Immune Defense Factor, as an Alternative Antibiotic, *Antibiotics*, 2021, **10**, 1534.

24 T. H. Steinberg, in *Methods in Enzymology*, eds. R. R. Burgess and M. P. Deutscher, Academic Press, 2009, vol. 463, pp. 541–563.

25 V. M. Bolanos-Garcia and O. R. Davies, Structural analysis and classification of native proteins from E. coli commonly co-purified by immobilised metal affinity chromatography, *Biochim. Biophys. Acta BBA - Gen. Subj.*, 2006, **1760**, 1304–1313.

26 B. Ashbridge, A. Orte, J. A. Yeoman, M. Kirwan, T. Vulliamy, I. Dokal, D. Klenerman and S. Balasubramanian, Single-Molecule Analysis of the Human Telomerase RNA·Dyskerin Interaction and the Effect of Dyskeratosis Congenita Mutations, *Biochemistry*, 2009, **48**, 10858–10865.

27 P. N. Grozdanov, S. Roy, N. Kittur and U. T. Meier, SHQ1 is required prior to NAF1 for assembly of H/ACA small nucleolar and telomerase RNPs, *RNA*, 2009, **15**, 1188–1197.

28 K. D. Pryor and B. Leiting, High-Level Expression of Soluble Protein inEscherichia coliUsing a His6-Tag and Maltose-Binding-Protein Double-Affinity Fusion System, *Protein Expr. Purif.*, 1997, **10**, 309–319.

29 F. W. Studier and B. A. Moffatt, Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes, *J. Mol. Biol.*, 1986, **189**, 113–130.

30 F. W. Studier, Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system, *J. Mol. Biol.*, 1991, **219**, 37–44.

31 C. Robichon, J. Luo, T. B. Causey, J. S. Benner and J. C. Samuelson, Engineering Escherichia coli BL21(DE3) Derivative Strains To Minimize E. coli Protein Contamination after Purification by Immobilized Metal Affinity Chromatography, *Appl. Environ. Microbiol.*, 2011, **77**, 4634–4646.

32 NEB® 5-alpha Competent E. coli (High Efficiency), https://www.neb.com/en-gb/products/c2987-neb-5-alpha-competent-e-coli-high-efficiency#, (accessed 9 October 2023).

33 S. Gallagher, Quantitation of Nucleic Acids with Absorption Spectroscopy, *Curr. Protoc. Protein Sci.*, 1998, **13**, A.4K.1-A.4K.3.

Chapter 6
Experimental

6.1 General procedures

6.1.1 Reagents and solvents

All chemicals used for small molecule synthesis were reagent grade and were purchased from Fluorochem, Sigma-Aldrich and Fisher scientific. DMF, DIPEA and piperidine used for SPPS were peptide grade and were purchased from Cambridge Reagents. Fmoc amino acids and coupling agents were purchased from AGTC Bioproducts, Novabiochem, Fluorochem and Sigma-Aldrich.

MAb anti-6HIS Tb cryptate antibody (61HISTLF) and MAb anti-6HIS d2 antibody (61HISDLF) were purchased from Revvity.

6.1.2 Instrumental techniques

^{1}H-, ^{13}C-NMR, ^{19}F- were collected using a Bruker spectrometer operating at 400 MHz (^{1}H), 100 MHz (^{13}C) or 376 MHz (^{19}F) using the specified deuterated solvent. The chemical shifts were recorded in ppm and are referenced to the residual solvent peak of $CHCl_3$ at 7.26 ppm (^{1}H), DMSO at 2.50 ppm (^{1}H), MeOH at 3.31 ppm (^{1}H) or acetone at 2.05 ppm (^{1}H). The multiplicities in the NMR spectra are described as s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet and br = broad, with coupling constants reported in Hz.

High resolution masses were recorded using QExactive mass spectrometer (Thermo) using its Vanquish UHPLC using electrospray ionisation (ESI). Low resolution mass was recorded using an Agilent 1200 Series equipped with a 6460 triple quadrupole. MALDI-MS was performed on a Kratos Analytical Axima-CFR MALDI ToF, using 2,5-dihydroxybenzoic acid as the matrix.

IR was performed using a Perkin-Elmer Spectrum BX FT-IR spectrometer and analysed using Spectrum v5.3.1 software.

Analytical RP-HPLC was performed on an Agilent 1200 using an Agilent eclipse XDB-C18 column, 4.6 x 150 mm, 5 µM and a flow rate of 1 mL/min. Solvent A = H_2O + 0.05% TFA and Solvent B = MeOH + 0.05% TFA. Gradient 5% B up to 95% B over 20 minutes. Detection wavelengths were 214 nm and 254 nm.

Preparative RP-HPLC was performed on an Agilent 1200 using an Agilent eclipse XDB-C18 column, 21.2 x 150 mm, 5 µM and a flow rate of 20 mL/min. Solvent A = 95% H_2O + 5% MeOH + 0.05% TFA and Solvent B = 95% MeOH + 5% H_2O + 0.05% TFA. Gradient 0% B up to 100% B over 20 minutes

Automated flash column chromatography was performed on a Biotage Isolera Four with MODUS columns refilled with silica gel for column chromatography (particle size 60 μm).

Automated peptide synthesis was performed using a Biotage I and Liberty Light.

Flow cytometry was performed using a Beckman Coulter CytoFLEX.

Surface plasmon resonance was performed on the Biacore S200.

Circular dichroism was performed using a Jasco J-810 spectropolarimeter.

6.2. The design, synthesis and *in vitro* biological evaluation of peptides targeting the dyskerin-dyskerin protein-protein interaction

6.2.1 General synthesis of linear peptides

Peptides **2.47, 2.49**-**2.59, 2.61** were synthesised using the Syro I automated peptide synthesiser. 100 mg Tentagel S RAM resin (0.23 mmol/g) was swollen in DMF (3 mL) for 20 min. with shaking after which the DMF was drained. Fmoc deprotection took place by the addition of 20% piperidine in DMF (2 mL) for 10 min. with shaking. The peptide column was drained after which the deprotection was repeated. The column was drained and the resin was washed with DMF (3 x 2 mL) followed by coupling of the first amino acid (4 eq.) using HOBt (4 eq.), HBTU (3.9 eq.) and DIPEA (8 eq.) in DMF for 45 min. with shaking. The column was drained and the reaction was repeated after which the resin was washed with DMF (3 x 2 mL). Subsequent deprotection and coupling reactions were carried out in the same way. After the final deprotection the resin was acetylated using acetic anhydride (4 eq.) and DIPEA (4 eq.) in DMF (2 mL) for 45 min. with shaking. The column was drained and the resin was washed with MeOH (3 x 2 mL), DCM (3 x 2 mL) and 1:1 MeOH:DCM (3 x 2 mL) to remove residual DMF. For cleavage of the peptides from the resin 92.5:2.5:2.5:2.5 TFA:DODT:H_2O:TIPS (5 mL) was added and the column was shaken for 3 h. The cleavage solution was collected and concentrated in vacuo. The peptide was precipitated using Et_2O after which the peptide was collected using centrifugation. The peptides were purified using RP-HPLC and analysed using HPLC and LCMS.

6.2.1.1 Synthesis of Ac-MTTAVISTCDH-NH_2 ***(2.47)***

2.47 was synthesised and purified following the general procedure as described in **section 6.2.1** using Fmoc-Ile-Ser(psiMe,Mepro)-OH to give a white powder (3.9 mg, 14%), 96% purity.

LRMS (ESI-QQQ): $C_{49}H_{82}N_{14}O_{18}S_2$ calculated $[M+H]^+$ = 1219.55, found = 1219.5

HRMS (ESI): $C_{49}H_{82}N_{14}O_{18}S_2$ calculated $[M+H]^+$ = 1219.5451, found = 1219.5442

HPLC: t_R = 10.2 min.

6.2.1.2 Synthesis of Ac-LMTTAVISTCDHG-NH_2 ***(2.49)***

2.49 was synthesised and purified following the general procedure as described in **section 6.2.1** using Fmoc-Ile-Ser(psiMe,Mepro)-OH to give a white powder (4.0 mg, 5%), 90% purity.

LRMS (ESI-QQQ): $C_{57}H_{96}N_{16}O_{20}S_2$ calculated $[M+2H]^{2+}$ = 695.32, found = 695.6

HRMS (ESI): $C_{57}H_{96}N_{16}O_{20}S_2$ calculated $[M+H]^+$ = 1389.6507, found = 1389.6516

HPLC: t_R = 10.9 min.

6.2.1.3 Synthesis of Ac-ALMTTAVISTCDHGI-NH_2 ***(2.50)***

2.50 was synthesised and purified following the general procedure as described in **section 6.2.1** using Fmoc-Ile-Ser(psiMe,Mepro)-OH to give a white powder (8.0 mg, 10%), 93% purity.

LRMS (ESI-QQQ): $C_{66}H_{112}N_{18}O_{22}S_2$ calculated $[M+2H]^{2+}$ = 787.38, found = 787.8

HRMS (ESI): $C_{66}H_{112}N_{18}O_{22}S_2$ calculated $[M+H]^+$ = 1573.7718, found = 1573.7716

HPLC: t_R = 11.8 min.

6.2.1.4 Synthesis of Ac-ATTAVISTCDH-NH_2 ***(2.51)***

2.51 was synthesised and purified following the general procedure as described in **section 6.2.1** using Fmoc-Ile-Ser(psiMe,Mepro)-OH to give a white powder (2.1 mg, 3%), 97% purity.

LRMS (ESI-QQQ): $C_{47}H_{78}N_{14}O_{18}S$ calculated $[M+H]^+$ = 1159.54, found = 1159.9

HRMS (ESI): $C_{47}H_{78}N_{14}O_{18}S$ calculated $[M+H]^+$ = 1159.5418, found = 1159.5419

HPLC: t_R = 9.5 min.

6.2.1.5 Synthesis of Ac-MATAVISTCDH-NH_2 ***(2.52)***

2.52 was synthesised and purified following the general procedure as described in **section 6.2.1** using Fmoc-Ile-Ser(psiMe,Mepro)-OH to give a white powder (4.2 mg, 6%), 90% purity.

LRMS (ESI-QQQ): $C_{48}H_{80}N_{14}O_{17}S_2$ calculated $[M+H]^+$ = 1189.53, found = 1189.9

HRMS (ESI): $C_{48}H_{80}N_{14}O_{17}S_2$ calculated $[M+H]^+$ = 1189.5346, found = 1189.5361

HPLC: t_R = 10.7 min.

6.2.1.6 Synthesis of Ac-MTTAAISTCDH-NH_2 ***(2.54)***

2.54 was synthesised and purified following the general procedure as described in **section 6.2.1** using Fmoc-Ile-Ser(psiMe,Mepro)-OH to give a white powder (2.0 mg, 3%), 96% purity.

LRMS (ESI-QQQ): $C_{47}H_{78}N_{14}O_{18}S_2$ calculated $[M+H]^+$ = 1191.51, found = 1191.9

HRMS (ESI): $C_{47}H_{78}N_{14}O_{18}S_2$ calculated $[M+H]^+$ = 1191.5133, found = 1191.5127

HPLC: t_R = 9.7 min.

6.2.1.7 Synthesis of Ac-MTTAVASTCDH-NH_2 ***(2.55)***

2.55 was synthesised and purified following the general procedure as described in **section 6.2.1** to give a white powder (7.1 mg, 12%), 91% purity.

LRMS (ESI-QQQ): $C_{46}H_{76}N_{14}O_{18}S_2$ calculated $[M+H]^+$ = 1177.50, found = 1178.1

HRMS (ESI): $C_{46}H_{76}N_{14}O_{18}S_2$ calculated $[M+H]^+$ = 1177.4976, found = 1177.4969

HPLC: t_R = 10.2 min.

6.2.1.8 Synthesis of Ac-MTTAVIATCDH-NH$_2$ ***(2.56)***

2.56 was synthesised and purified following the general procedure as described in **section 6.2.1** to give a white powder (2.0 mg, 3%), 96% purity.

LRMS (ESI-QQQ): $C_{49}H_{82}N_{14}O_{17}S_2$ calculated $[M+H]^+$ = 1203.55, found = 1204.1

HRMS: $C_{49}H_{82}N_{14}O_{17}S_2$ calculated $[M+H]^+$ = 1203.5497, found = 1203.5496

HPLC: t_R = 10.5 min.

6.2.1.9 Synthesis of Ac-MTTAVISACDH-NH$_2$ ***(2.57)***

2.57 was synthesised and purified following the general procedure as described in **section 6.2.1** to give a white powder (1.6 mg, 2%), 95% purity.

LRMS (ESI-QQQ): $C_{48}H_{80}N_{14}O_{17}S_2$ calculated $[M+H]^+$ = 1189.53, found = 1190.1

HRMS (ESI): $C_{48}H_{80}N_{14}O_{17}S_2$ calculated $[M+H]^+$ = 1189.5340, found = 1189.5335

HPLC: t_R = 10.6 min.

6.2.1.10 Synthesis of Ac-MTTAVISTADH-NH$_2$ ***(2.58)***

2.58 was synthesised and purified following the general procedure as described in **section 6.2.1** to give a white powder (2.1 mg, 3%), 95% purity.

LRMS (ESI-QQQ): $C_{49}H_{82}N_{14}O_{18}S$ calculated $[M+H]^+$ = 1187.57, found = 1188.2/594.6

HRMS (ESI): $C_{49}H_{82}N_{14}O_{18}S$ calculated $[M+H]^+$ = 1187.5725, found = 1187.5722

HPLC: t_R = 9.6 min.

6.2.1.11 Synthesis of Ac-MTTAVISTCAH-NH$_2$ ***(2.59)***

2.59 was synthesised and purified following the general procedure as described in **section 6.2.1** to give a white powder (1.9 mg, 3%), 94% purity.

LRMS (ESI-QQQ): $C_{48}H_{82}N_{14}O_{16}S_2$ calculated $[M+2H]^{2+}$ = 588.28, found = 588.5

HRMS (ESI): $C_{48}H_{82}N_{14}O_{16}S_2$ calculated $[M+H]^+$ = 1175.5547, found = 1175.5544

HPLC: t_R = 10.2 min.

6.2.1.12 Synthesis of FITC-βAla-MTTAVISTCDH-NH$_2$ ***(2.61)***

2.61 was synthesised and purified following the general procedure as described in **section 6.2.1**. After the deprotection of the final amino acid (Met), Fmoc-βAla-OH was coupled using the same method as for the other amino acids. After the deprotection, the fluorescent-tag was attached using FITC (1.5 eq.) and DIPEA (8 eq.) in DMF (2 mL) for 18 h with shaking. The column was drained and the resin was washed with DMF (3 x 2 mL), before cleaving and purifying the peptide as described in **section 6.2.1** to give a yellow powder (1.4 mg, 1%), 97% purity.

LRMS (ESI-QQQ): $C_{71}H_{96}N_{16}O_{23}S_3$ calculated $[M+2H]^{2+}$ = 819.31, found = 819.6

HRMS (ESI): $C_{71}H_{96}N_{16}O_{23}S_3$ calculated $[M+2H]^{2+}$ = 819.3071, found = 819.3069

HPLC: t_R = 11.4 min.

6.2.2 General synthesis of hydrocarbon stapled peptides

Hydrocarbon stapled peptides **2.68, 2.69, 2.76**-**2.83** were synthesised using the Liberty Lite automated microwave peptide synthesiser. 217 mg Tentagel S RAM (0.23 mmol/g) was swollen in DMF (3 mL) for 10 min. Fmoc deprotection took place by the addition of 20% piperidine in DMF (3 mL) for 1 min. with microwave heating at 90 °C. The reaction vessel was drained and the resin was washed with DMF (3 x 3 mL) followed by the coupling of histidine (5 eq.) using oxyma (5 eq.) and DIC (5 eq.) for 10 min. at 50 °C using microwave heating. The reaction vessel was drained and the resin was washed with DMF (3 x 3 mL). The resin was deprotected as described before and the amino acids until Fmoc-(S)-2-(4-pentenyl)Ala-OH (S_5) were coupled using AA (5 eq.), oxyma (5 eq.) and DIC (5 eq.) for 2 min. at 90 °C. S_5 was coupled using Fmoc-(S)-2-(4-pentenyl)Ala-OH (5 eq.), oxyma (5 eq.) and DIC (5 eq.) for 5 min. at 90 °C. The residues following S_5 was coupled using AA (5 eq.), oxyma (5 eq.) and DIC (5 eq.) for 5 min. at 90 °C after which the reaction vessel was drained and the coupling was repeated. Any other amino acids not directly after S_5 were coupled for 2 min. at 90 °C. After the incorporation of the final amino acid the resin was deprotected and acetylated with 10% Ac_2O (4 mL) for 30 s at 60 °C, 30 s without heating, 30 s at 60 °C and then a further 30 s without heating. The resin was washed with DMF (5 x 4 mL), DCE (3 x 4 mL) and the staple was formed using 10 mM Grubbs 1st generation catalyst in DCE (3 mL) for 30 min. at 40 °C. The reaction vessel was drained, the resin was washed with DCE (3 x 4 mL) and the stapling was repeated. The resin was washed with DMF (5 x 4 mL), MeOH (3 x 2 mL), DCM (3 x 2 mL) and 1:1 MeOH:DCM (3 x 2 mL). For cleavage of the peptides from the resin 92.5:2.5:2.5:2.5 TFA:DODT:H_2O:TIPS (5 mL) was added and the column was shaken for 3 h. The cleavage solution was collected and concentrated in vacuo. The peptide was precipitated using Et_2O after which the peptide was collected using centrifugation. The peptides were purified using RP-HPLC and analysed using HPLC and LCMS.

*6.2.2.1 Synthesis of FITC-βAla-$N_L(S_5)$TAV(S_5)STCDH-NH_2 (**2.68**)*

2.68 was synthesised and purified following the general procedure as described in **section 6.2.2** with the staple formation after the incorporation of Fmoc-βAla-OH. After the staple formation, the Fmoc group was removed and the FITC tag was attached as described in section **6.2.1.12** to give a yellow powder (3.8 mg, 5%), 92% purity.

LRMS (ESI-QQQ): $C_{76}H_{102}N_{16}O_{22}S_2$ calculated $[M+2H]^{2+}$ = 828.35, found = 828.6

HRMS (ESI): $C_{76}H_{102}N_{16}O_{22}S_2$ calculated $[M+2H]^{2+}$ = 828.3471, found = 828.3464

HPLC: t_R = 14.0 min.

*6.2.2.2 Synthesis of Ac-N_L(S_5)TAV(S_5)STCDH-NH_2 (**2.69**)*

2.69 was synthesised and purified following the general procedure as described in **section 6.2.2** to give a white powder (3.3 mg, 5%), 91% purity.

LRMS (ESI-QQQ): $C_{54}H_{88}N_{14}O_{17}S$ calculated $[M+H]^+$ = 1237.62, found = 1237.7

HRMS (ESI): $C_{54}H_{88}N_{14}O_{17}S$ calculated $[M+H]^+$ = 1237.6245, found = 1237.6230

HPLC: t_R = 13.8 min.

*6.2.2.3 Synthesis of Ac-N_LT(S_5)AVI(S_5)TCDH-NH_2 (**2.76**)*

2.76 was synthesised and purified following the general procedure as described in **section 6.2.2** to give a white powder (7.4 mg, 12%), 97% purity.

LRMS (ESI-QQQ): $C_{57}H_{94}N_{14}O_{16}S$ calculated $[M+H]^+$ = 1263.68, found = 1263.7

HRMS (ESI): $C_{57}H_{94}N_{14}O_{16}S$ calculated $[M+H]^+$ = 1263.6766, found = 1263.6758

HPLC: t_R = 12.4 min.

*6.2.2.4 Synthesis of Ac-N_LTTA(S_5)IST(S_5)DH-NH_2 (**2.77**)*

2.77 was synthesised and purified following the general procedure as described in **section 6.2.2** to give a white powder (7.8 mg, 12%), 98% purity.

LRMS (ESI-QQQ): $C_{56}H_{92}N_{14}O_{18}$ calculated $[M+H]^+$ = 1249.68, found = 1249.8

HRMS (ESI): $C_{56}H_{92}N_{14}O_{18}$ calculated $[M+H]^+$ = 1249.6787, found = 1249.6766

HPLC: t_R = 12.7 min.

*6.2.2.5 Synthesis of Ac-N_LTTAV(S_5)STC(S_5)H-NH_2 (**2.78**)*

2.78 was synthesised and purified following the general procedure as described in **section 6.2.2** to give a white powder (3.9 mg, 13%), 65% purity.

LRMS (ESI-QQQ): $C_{54}H_{90}N_{14}O_{16}S$ calculated $[M+H]^+$ = 1223.65, found = 1224.1

HRMS (ESI): $C_{54}H_{90}N_{14}O_{16}S$ calculated $[M+H]^+$ = 1223.6453, found = 1223.6413

HPLC: t_R = 13.2 min.

*6.2.2.6 Synthesis of Ac-N_LTTAVI(S_5)TCD(S_5)-NH_2 (**2.79**)*

2.79 was synthesised and purified following the general procedure as described in **section 6.2.2** to give a white powder (0.7 mg, 2%), 86% purity.

LRMS (ESI-QQQ): $C_{55}H_{94}N_{12}O_{17}S$ calculated $[M-H]^-$ = 1225.67, found = 1225.6

HRMS (ESI): $C_{55}H_{94}N_{12}O_{17}S$ calculated $[M+H]^+$ = 1227.6659, found = 1227.6616

HPLC: t_R = 13.6 min.

*6.2.2.7 Synthesis of Ac-$N_L(S_5)$WAV(S_5)STCDH-NH_2 (**2.80**)*

2.80 was synthesised and purified following the general procedure as described in **section 6.2.2** to give a white powder (2.0 mg, 6%), 90% purity.

LRMS (ESI-QQQ): $C_{61}H_{91}N_{15}O_{16}S$ calculated $[M+H]^+$ = 1322.66, found = 1322.7

HRMS (ESI): $C_{61}H_{91}N_{15}O_{16}S$ calculated $[M+H]^+$ = 1322.6562, found = 1322.6556

HPLC: t_R = 14.6 min.

*6.2.2.8 Synthesis of Ac-N_LT(S5)LVI(S5)TCDH-NH_2 (**2.81**)*

2.81 was synthesised and purified following the general procedure as described in **section 6.2.2** to give a white powder (0.6 mg, 2%), 95% purity.

LRMS (ESI-QQQ): $C_{60}H_{100}N_{14}O_{16}S$ calculated $[M+H]^+$ = 1305.72, found = 1305.3

HRMS (ESI): $C_{60}H_{100}N_{14}O_{16}S$ calculated $[M+H]^+$ = 1305.7235, found = 1305.7224

HPLC: t_R = 14.4 min.

*6.2.2.9 Synthesis of Ac-N_LT(S_5)AVI(S_5)DCDH-NH_2 (**2.82**)*

2.82 was synthesised and purified following the general procedure as described in **section 6.2.2** to give a white powder (2.4 mg, 7%), 97% purity.

LRMS (ESI-QQQ): $C_{57}H_{92}N_{14}O_{17}S$ calculated $[M+H]^+$ = 1277.66, found = 1277.8

HRMS (ESI): $C_{57}H_{92}N_{14}O_{17}S$ calculated $[M+H]^+$ = 1277.6558, found = 1277.6548

HPLC: t_R = 12.9 min.

*6.2.2.10 Synthesis of Ac-N_LT(S_5)AVI(S_5)TCDW-NH_2 (**2.83**)*

2.83 was synthesised and purified following the general procedure as described in **section 6.2.2** to give a white powder (1.3 mg, 4%), 95% purity.

LRMS (ESI-QQQ): $C_{62}H_{97}N_{13}O_{16}S$ calculated $[M+2H]^{2+}$ = 656.85, found = 657.4

HRMS (ESI): $C_{62}H_{97}N_{13}O_{16}S$ calculated $[M+H]^+$ = 1312.6975, found = 1312.6946

HPLC: t_R = 14.5 min.

6.2.3 General synthesis of lactam stapled peptides

Lactam stapled peptides **2.92**-**2.95** were synthesised using the Syro I automated peptide synthesiser. 100 mg Tentagel S RAM resin (0.23 mmol/g) was swollen in DMF (3 mL) for 20 min. with shaking after which the DMF was drained. Fmoc deprotection took place by the addition of 20% piperidine in DMF (2 mL) for 10 min. with shaking. The peptide column was drained after which the deprotection was repeated. The column was drained and the resin was washed with DMF (3 x 2 mL) followed by coupling of the first amino acid (4 eq.) using HOBt (4 eq.), HBTU (3.9 eq.) and DIPEA (8 eq.) in DMF for 45 min. with shaking. The column

was drained and the reaction was repeated after which the resin was washed with DMF (3 x 2 mL). Subsequent deprotection and coupling reactions were carried out in the same way until the incorporation of both the Lys(Mtt) and Asp(O-2-PhiPr) required for the lactam staple formation. The resin was washed with DCM (5 x 3 mL). This was followed by the side chain deprotection of the Lys(Mtt) and Asp(O-2-Phipr) using 2.25% TFA and 5% TIPS in DCM (3 mL) for 10 min., after which the column was drained and the deprotection was repeated four more times. The resin was washed with DCM (3x), 2% DIPEA in DCM (3x), DCM (3x) and DMF (3x). The lactam staple was formed using PyBOP (6 eq.) and DIPEA (10 eq.) for 1h. The resin was washed with DMF (3 x 2 mL) and the stapling was repeated. Fmoc deprotection took place by the addition of 20% piperidine in DMF (2 mL) for 10 min. with shaking. The peptide column was drained after which the deprotection was repeated. The remaining amino acids were coupled manually using AA (4 eq.), HATU (4 eq.) and DIPEA (8 eq.) for 45 min. with shaking. The column was drained and the reaction was repeated after which the resin was washed with DMF (3 x 2 mL). After the final deprotection the resin was acetylated using acetic anhydride (4 eq.) and DIPEA (4 eq.) in DMF (2 mL) for 45 min. with shaking. The column was drained and the resin was washed with MeOH (3 x 2 mL), DCM (3 x 2 mL) and 1:1 MeOH:DCM (3 x 2 mL) to remove residual DMF. For cleavage of the peptides from the resin 92.5:2.5:2.5:2.5 TFA:DODT:H_2O:TIPS (5 mL) was added and the column was shaken for 3 h. The cleavage solution was collected and concentrated in vacuo. The peptide was precipitated using Et_2O after which the peptide was collected using centrifugation. The peptides were purified using RP-HPLC and analysed using HPLC and LCMS.

*6.2.3.1 Synthesis of FITC-βAla-N_L**D**TAV**K**STCDH-NH_2 (**2.92**)*

2.92 was synthesised and purified following the general procedure as described in **section 6.2.3.** The FITC tag was attached as described in section **6.2.1.12** to give a yellow powder (2.5 mg, 7%), 87% purity.

LRMS (ESI-QQQ): $C_{72}H_{97}N_{17}O_{23}S_2$ calculated $[M+2H]^{2+}$ = 816.82, found = 816.1

HRMS (ESI): $C_{72}H_{97}N_{17}O_{23}S_2$ calculated $[M+H]^+$ = 1632.6457, found = 1632.6450

HPLC: t_R = 11.5 min.

*6.2.3.2 Synthesis of Ac-N_L**D**TAV**K**STCDH-NH_2 (**2.93**)*

2.93 was synthesised and purified following the general procedure as described in **section 6.2.3** to give a white powder (2.2 mg, 8%), 91% purity.

LRMS (ESI-QQQ): $C_{50}H_{81}N_{15}O_{18}S$ calculated $[M+H]^+$ = 1212.57, found = 1212.6

HRMS (ESI): $C_{50}H_{81}N_{15}O_{18}S$ calculated $[M+H]^+$ = 1212.5678, found = 1212.5661

HPLC: t_R = 10.9 min.

*6.2.3.3 Synthesis of Ac-N_LT**D**AVI**K**TCDH-NH_2 (**2.94**)*

2.94 was synthesised and purified following the general procedure as described in **section 6.2.3** to give a white powder (0.5 mg, 2%), 86% purity.

LRMS (ESI-QQQ): $C_{53}H_{87}N_{15}O_{17}S$ calculated $[M-H]^-$ = 1236.62, found = 1237.0

HRMS (ESI): $C_{53}H_{87}N_{15}O_{17}S$ calculated $[M+H]^+$ = 1238.6203, found = 1238.6198

HPLC: t_R = 12.0 min.

*6.2.3.4 Synthesis of Ac-N_LTTA**D**IST**K**DH-NH_2 (**2.95**)*

2.95 was synthesised and purified following the general procedure as described in **section 6.2.3** to give a white powder (3.7 mg, 13%), 99% purity.

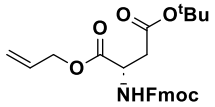
LRMS (ESI-QQQ): $C_{52}H_{85}N_{15}O_{19}$ calculated $[M+H]^+$ = 1224.62, found = 1224.7

HRMS (ESI): $C_{52}H_{85}N_{15}O_{19}$ calculated $[M+H]^+$ = 1224.6219, found = 1224.6208

HPLC: t_R = 9.8 min.

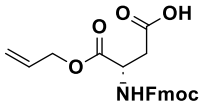
6.2.4 Synthesis of N-α-Fmoc-L-aspartic acid β-2-phenylisopropyl ester (Fmoc-L-Asp(O-2PhiPr)-OH) (2.84)

*6.2.4.1 Synthesis of L-Aspartic acid, N-[(9H-fluoren-9-ylmethoxy)carbonyl]-, 4-(1,1-dimethylethyl) 1-(2-propen-1-yl) ester (Fmoc-L-Asp(OtBu)-OAllyl) (**2.87**)*



2.87 was prepared following a literature procedure.[1] Fmoc-L-Asp(OtBu)-OH (2.00 g, 4.86 mmol, 1.0 eq.) was dissolved in DMF (10 mL) and was cooled to 0 °C. DIPEA (1.69 mL, 9.72 mmol, 2.0 eq.) was added followed by allyl bromide (0.850 mL, 9.72 mmol, 2.0 eq.) and the reaction was allowed to warm to rt. The solution was stirred at rt for 18 h after which the reaction was quenched with 0.1 M HCl (50 mL) and was extracted with EtOAc (2 x 50 mL). The combined organics were washed with 0.1 M HCl (4 x 50 mL), brine (50 mL), dried (Na_2SO_4), filtered and concentrated *in vacuo* to give crude **2.87** as a pale yellow oil, which was used in the next step without further purification.

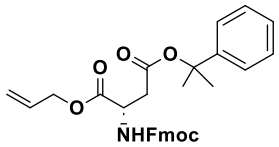
*6.2.4.2 Synthesis of L-Aspartic acid, N-[(9H-fluoren-9-ylmethoxy)carbonyl]-, 1-(2-propen-1-yl) ester (Fmoc-L-Asp(OH)-OAllyl) (**2.88**)*



2.88 was prepared following a literature procedure.[1] Crude **2.87** was dissolved in DCM (10 mL) and TFA (10 mL) was added. The solution was stirred at rt for 3 h after which it was

concentrated *in vacuo* to give crude **2.88** as a brown oil which was used in the next step without further purification.

*6.2.4.3 Synthesis of L-Aspartic acid, N-[(9H-fluoren-9-ylmethoxy)carbonyl]-3-methyl-, 4-(1-phenylethyl) 1-(2-propen-1-yl) ester (Fmoc-L-Asp(O-2PhiPr)-OAllyl) (**2.90**)*



2.90 was prepared following a literature procedure.[1] Crude **2.88** was dissolved in DCM (15 mL) and **2.89** (2.73 g, 9.72 mmol, 2.0 eq. as a 75 mg/mL solution in hexane) was added. The solution was stirred at rt for 18 h after which a further **2.89** (1.37 g, 4.86 mmol, 1.0 eq. as a 75 mg/mL solution in hexane) was added and the solution was stirred at rt for 5 h. The reaction mixture was concentrated *in vacuo* and purified by automated flash column chromatography (0-40% EtOAc in hexane) to give **2.90** (1.65 g, 66% over 3 steps) as an off-white powder.

^{1}H NMR (400 MHz; $CDCl_3$) δ 7.76 (2H, d, *J* = 7.5), 7.58 (2H, d, *J* = 7.5 Hz), 7.40 (2H, dd, *J* = 7.5, 7.5), 7.36-7.28 (7H, m), 5.83-5.78 (1H, m), 5.29 (1H, ddd, *J* = 17.1, 1.5, 1.5 Hz), 5.22 (1H, dq, *J* = 10.5, 1.5 Hz), 4.68-4.58 (3H, m), 4.42 (1H, dd, *J* = 10.5, 7.2 Hz), 4.33 (1H, dd, *J* = 10.5, 7.2 Hz), 4.23 (1H, t, *J* = 7.2 Hz), 3.08 (1H, dd, *J* = 17.1, 4.5 Hz), 2.86 (dd, *J* = 17.1, 4.5 Hz), 1.77 (6H, s)

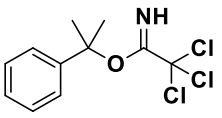
^{13}C-NMR (100 MHz; $CDCl_3$) δ 170.6, 169.8, 163.9, 156.2, 145.3, 144.0, 143.8, 141.4, 131.6, 131.4, 128.5, 127.9, 127.3, 127.2, 125.3, 124.3, 120.1, 119.2, 118.9, 91.9, 83.2, 67.5, 66.7, 66.5, 50.7, 47.2, 37.7, 36.4, 31.8, 28.8, 28.7

LRMS (ESI-QQQ): $C_{31}H_{31}NO_6$, calculated $[M+Na]^+$ = 536.22, found = 536.3

IR (cm^{-1}): 3364, 3241, 2982, 1728, 1694, 1511, 1449, 1381

Rf (20% EtOAc in hexane): 0.21

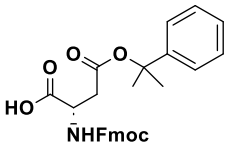
*6.2.4.4 Synthesis of 2-phenylisopropyl 2,2,2-trichloroacetimidate (**2.89**)*



2.89 was prepared according to literature procedure.[2] Sodium hydride (1.17 g, 29.37 mmol, 2.0 eq.) was suspended in anh. THF (35 mL) under a N_2 atmosphere. 2-Phenylisopropanol (2.06 mL, 14.69 mmol, 1.0 eq.) was added dropwise to the suspension and was allowed to stir at rt for 20 min. The solution was cooled to 0 °C and trichloroacetonitrile (2.94 mL, 29.37 mmol, 2.0 eq.) was added dropwise and the reaction was allowed to warm to rt over 3 h. The reaction was cooled to 0 °C and was quenched with MeOH (15 mL). The solution was

concentrated *in vacuo* and hexane (40 mL) was added to the residue. The mixture was vigorously shaken, filtered and the filtrate was concentrated *in vacuo* to give **2.89** as a brown oil which was used without further purification.

*6.2.4.5 Synthesis of L-Aspartic acid, N-[(9H-fluoren-9-ylmethoxy)carbonyl]-, 4-(1-methyl-1-phenylethyl) ester (Fmoc-L-Asp(O-2PhiPr)-OH (**2.84**)*



2.84 was prepared following a literature procedure.[1] **2.90** (0.782 g, 1.52 mmol, 1.0 eq.) was dissolved in anh. THF (10 mL) under a N_2 atmosphere. $Pd(PPh_3)_4$ (88 mg, 0.076 mmol, 0.05 eq.) was added followed by phenylsilane (376 µL, 3.05 mmol, 2.0 eq.) and the solution was stirred at rt for 15 min. The solution was concentrated *in vacuo* and purified by automated flash chromatography (10-70% EtOAc in hexane). The product containing fractions were concentrated and purified by automated reverse phase column chromatography (0-100% MeOH in H_2O). The product containing fractions were concentrated *in vacuo* to give **2.90** (160 mg, 22%) as a white powder.

^{1}H NMR (400 MHz; $CDCl_3$) δ 7.76 (2H, d, *J* = 7.5 Hz), 7.57 (2H, d, *J* = 7.5 Hz), 7.40 (2H, t, *J* = 7.5 Hz), 7.35-7.28 (7H, m), 5.78 (1H, d, *J* = 8.6 Hz), 4.68-4.62 (1H, m), 4.44 (1H, dd, *J* = 10.6, 7.3 Hz), 4.35 (1H, dd, *J* = 10.6, 7.3 Hz), 4.22 (1H, t, *J* = 7.3 Hz), 3.09 (1H, dd, *J* = 17.0, 4.7 Hz), 2.87 (1H, dd, *J* = 17.0, 4.7 Hz), 1.78 (3H, s), 1.76 (3H, s)

^{13}C-NMR (100 MHz; $CDCl_3$) δ 174.7, 170.1, 156.3, 145.2, 143.8, 141.4, 128.5, 127.9, 127.4, 127.3, 125.3, 124.1, 120.2, 83.6, 67.6, 50.4, 47.2, 37.5, 28.9, 28.6

LRMS (ESI-QQQ): $C_{28}H_{27}NO_6$, calculated $[M+Na]^+$ = 496.18, found = 496.2

IR (cm^{-1}): 3090, 2981, 1719, 1715, 1512, 1448, 1215

Rf (10% MeOH in DCM): 0.51

6.2.5 Fluorescence polarisation

Fluorescence anisotropy was carried out on a BMG Labtech clariostar microplate reader with a fluorescence anisotropy optic measuring at 482/530 nm. 384-Well black plates were used with a final volume of 40 µL and all solutions were made using MilliQ water. All FP measurements were done in 10 mM HEPES, 150 mM NaCl, 0.05% Tween-20, pH 7.4 buffer.

The anisotropy values were calculated using the following formula:

$$r = \frac{l_a - l_b}{l_a + 2l_b}$$

$$I = l_a + 2l_b$$

Where r = anisotropy, l_a = parallel emission light, l_b = perpendicular emission light and I = total intensity.

Direct FP

A 1 in 2 dilution series of MBP-dyskerin protein (11.8 µM – 0.12 µM) was added to the plate followed by the FITC-peptide (5 nM in plate). The plate was incubated at room temperature for 1 h, bubbles were removed, and the anisotropy was measured. From the collected anisotropy data the K_D was determined using GraphPad Prism Version 6.0 software, following literature methods.[3]

Competition FP

For the competition fluorescence polarisation experiments MBP-dyskerin (300 nM), FITC-peptide (5 nM) and the competitor (200 µM to 6.1 nM) were added to the plate and the plate was incubated at room temperature for 1 h. The bubbles were removed and the anisotropy was measured.

IC_{50} was determined using GraphPad Prism Version 6.0 software, using a four-parameter logistic model. Errors are the transformed greater extreme of the standard error.

$$y = r_{min} + \frac{(r_{max} - r_{min})}{1 + 10^{((logIC_{50} - x)\ x\ Hill\ Slope)}}$$

6.2.6 HTRF

For the HTRF assay 384-well low volume white plates were used with a final volume of 20 µL. Protein (5 µL, 711 nM in plate), MAb anti-6HIS Tb cryptate antibody (5 µL, Cisbio - 61HISTLF), MAb anti-6HIS d2 antibody (5 µL, Cisbio - 61HISDLF) and peptide (200 µM) were added to the plate. The antibody solutions were prepared as per the manufacturer's instructions. The plate was incubated at 4 °C for optimal signal size and kept at 4 °C for plate readings. The plate was analysed using a BMG Labtech clariostar microplate reader with standard settings for HTRF (Wavelength 1 (Tb): λ_{ex}: 340 nm, λ_{em}: 620 nm, wavelength 2 (d2): λ_{ex}: 340 nm, λ_{em}: 665 nm, lag time: 60 µs, integration time: 400 µs, number of flashes: 200, gain: 2400) The ratio of the acceptor emission signal (665 nm) over the donor emission signal (620 nm) is calculated as following:

$$Ratio = \frac{\lambda_{em}(665)}{\lambda_{em}(620)} \times 10^4$$

6.2.7 Surface plasmon resonance

MBP-dyskerin was diluted from 0.85 mg/mL in 10 mM HEPES, 150 mM NaCl, 0.05% Tween-20, pH 7.4 to 0.1 mg/mL in sodium acetate buffers pH 4.0, 4.5, 5.0 and 5.5 and analysed using the pH scouting template. MBP-dyskerin was diluted in sodium acetate pH 5.0 buffer as described for the pH scouting and was immobilised to 5000 RU density on a CM5 chip using the amine coupling kit from Cytiva. A reference flow cell was blocked with ethanolamine. Experiments were ran in the specified buffer in **Chapters 2** and **3** with a flow rate of 30 µL/min. Samples were injected for 120 s followed by a dissociation period of 240 s. The surface was regenerated using 5 mM NaOH for 30 s. A DMSO correction was applied to the results and a blank sample was subtracted from the signal after which they were analysed using steady-state affinity to determine the K_D value.

6.2.8 Circular dichroism

Peptide was dissolved in 50/50 10 mM phosphate buffer pH 7.4/trifluoroethanol to a final concentration of 0.2 mg/mL and were filtered using a 0.45 µm filtered. Experiments were performed using a 1 mm path length quartz cuvette. Samples were scanned from 260 nm to 185 nm using a scanning speed of 200 nm/min, response time of 1 s, 0.5 nm data pitch, a 2 nm bandwidth, and accumulation of 4 scans. The data collected was analysed by Dichroweb using the method CDSSTR to determine the α-helicity.[4, 5]

6.3 Design, synthesis, and in vitro biological evaluation of small molecules targeting the dyskerin-dyskerin PPI

6.3.1 Computational modelling

The Zinc compound library was used to compile a library of azide and terminal alkyne fragments which were prepared using the LigPrep application from the Schrodinger Suite 2017-2.[6,7]

The Cryo-EM structure of telomerase (PDB: 7BGB) was utilized for docking calculations with the proteins other than the two dyskerin proteins removed from the structure. From one dyskerin protein residues outside of $_{350}$MTTAVISTCDH$_{360}$ were removed and the structure was prepared using Protein Preparation.[6] For the generation of the each half-peptide the structure was modified using Maestro drawing tools. The resulting receptor was prepared using Protein Preparation.

For the non-covalent docking the binding site grid generation was performed using the Receptor Grid Generation application from the Schrodinger 2017-2 suite using the site vacated by the removed section of the peptide. The fragments were then docked using the Glide docking application in Schrodinger 2017-2.[8]

Covalent docking was performed using the CovDock tool from the Schrodinger Suite 2017-2, using the following custom reaction code to generate the desired azido-alkynyl condensation.[9]

Alkyne receptor, azide ligand:

RECEPTOR_SMARTS_PATTERN 2,C#CCCC=O

LIGAND_SMARTS_PATTERN 4,[C,c]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)))

Azide receptor, alkyne ligand:

RECEPTOR_SMARTS_PATTERN 4,CN=[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 ("<2>#C",("bond",2,(1,2)))

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

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

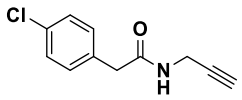
CUSTOM_CHEMISTRY ("<1>[N]",("bond",2,(1,2)))

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

Induced fit docking was performed with protein flexibility using the Induced Fit application from the Schrodinger Suite 2017-2.[10] A total of 100 possible binding conformations were generated and grouped into clusters using a 1.0 Å root-mean-square tolerance. Conformations were then ranked based on their Glide e-model scores and conformations with the lowest glide e-model score were docked again using Induced Fit to generate the lowest energy conformation.

6.3.2 Alkyne synthesis

*6.3.2.1 Synthesis of 4-chloro-N-2-propyn-1-ylbenzeneacetamide (**3.66**)*



p-Chlorophenylacetyl chloride (1.00 g, 5.29 mmol, 1.0 eq.) was dissolved in anh. DCM (30 mL) under N_2. The solution was cooled to 0 °C and propargylamine (0.370 mL, 5.82 mmol, 1.1 eq.) and triethylamine (0.890 mL, 6.35 mmol, 1.2 eq.) were added. The solution was warmed to rt and was stirred for 16 h at rt. The reaction was quenched with aq. sat. NH_4Cl (30 mL) and extracted with DCM (3 x 30 mL). The combined organic layers were washed with 1 M HCl (2 x 30 mL), brine (30 mL), dried ($MgSO_4$), filtered and concentrated *in vacuo* to give a yellow powder. The powder was purified by automated flash column chromatography (0-40% EtOAc in hexane) to give 4-chloro-*N*-2-propyn-1-ylbenzeneacetamide (0.66 g, 61%) as a pale yellow crystalline solid.

^{1}H NMR (400 MHz; $CDCl_3$) δ 7.34 (2H, dt, *J* = 8.5, 2.1 Hz), 7.21 (2H, dt, *J* = 8.5, 2.1 Hz), 5.54 (1H, br s), 4.02 (2H, dd, *J* = 5.3, 2.6 Hz), 3.55 (2H, s), 2.20 (1H, t, *J* = 2.6 Hz)

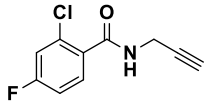
^{13}C-NMR (100 MHz; $CDCl_3$) δ 170.1, 133.7, 132.9, 130.9, 129.4, 79.3, 71.9, 42.9, 29.6

LRMS (ESI-QQQ): $C_{11}H_{10}ClNO$, calculated $[M+H]^+$ = 208.05, found = 208.1

HRMS (ESI): $C_{11}H_{10}ClNO$, calculated $[M+H]^+$ = 208.0524, found = 208.0522

Rf (20% EtOAc in hexane): 0.12

*6.3.2.2 Synthesis of 2-chloro-4-fluoro-N-2-propyn-1-ylbenzamide (**3.67**)*



2-Chloro-4-fluorobenzoyl chloride (0.630 mL, 4.66 mmol, 1.0 eq.) was dissolved in anh. DCM (25 mL) under N_2. The solution was cooled to 0 °C and propargylamine (0.300 mL, 4.66 mmol, 1.0 eq.) and triethylamine (0.780 mL, 5.59 mmol, 1.2 eq.) were added. The solution was warmed to rt and was stirred for 16 h at rt. The reaction was quenched with aq. sat. NH_4Cl (20 mL) and extracted with DCM (3 x 20 mL). The combined organic layers were washed with brine (30 mL), dried (Na_2SO_4), filtered and concentrated *in vacuo* to give a white powder. The powder was purified by automated flash column chromatography (0-40% EtOAc in hexane) to give 2-chloro-4-fluoro-*N*-2-propyn-1-ylbenzamide (0.78 g, 79%) as a white crystalline solid.

^{1}H NMR (400 MHz; $CDCl_3$) δ 7.75 (1H, dd, *J* = 8.7, 6.1 Hz), 7.15 (1H, dd, *J* = 8.4, 2.5 Hz), 7.06 (1H, ddd, *J* = 8.7, 7.7, 2.5 Hz), 6.49 (1H, br s), 4.26 (2H, dd, *J* = 5.2, 2.6 Hz), 2.29 (2H, t, *J* = 2.6 Hz)

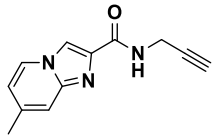
^{13}C-NMR (100 MHz; $CDCl_3$) δ 165.0, 163.5 (d, *J* = 255.3 Hz), 132.5 (d, *J* = 10.3 Hz), 132.2 (d, *J* = 9.2 Hz), 130.4 (d, *J* = 4.3 Hz), 117.9 (d, *J* = 24.5 Hz), 114.8 (d, *J* = 20.0 Hz), 79.0, 72.3, 30.0

^{19}F-NMR (376 MHz; $CDCl_3$) δ -107.08

LRMS (ESI-QQQ): $C_{10}H_7ClFNO$, calculated $[M+H]^+$ = 212.03, found = 212.1

HRMS (ESI): $C_{10}H_7ClFNO$, calculated $[M+H]^+$ = 212.0273, found = 212.0273

*6.3.2.3 Synthesis of 7-methyl-N-2-propyn-1-ylimidazo[1,2-a]pyridine-2-carboxamide (**3.68**)*



7-Methylimidazo[1,2-a]pyridine-2-carboxylic acid (250 mg, 1.42 mmol, 1.0 eq.), EDC·HCl (326 mg, 1.70 mmol, 1.2 eq.), HOBt (261 mg, 1.70 mmol, 1.2 eq.) and DIPEA (494 µL, 2.84 mmol, 2.0 eq.) were dissolved in anh. THF (10 mL) and the solution was stirred for 30 min. at rt. Propargylamine (91 µL, 1.4 mmol, 1.0 eq.) was added and the solution was stirred for 20 h at rt. The reaction was quenched with water (25 mL) and was extracted with EtOAc (2 x 30 mL). The combined organic extracts were washed with water (3 x 30 mL), brine (30 mL), dried (Na_2SO_4), filtered and concentrated in vacuo to give a pale yellow powder. The powder was purified by automated flash column chromatography (0-100% EtOAc in hexane) to give 7-methyl-*N*-2-propyn-1-ylimidazo[1,2-a]pyridine-2-carboxamide (169 mg, 56%) as a white powder.

^{1}H NMR (400 MHz; $CDCl_3$) δ 8.06 (1H, d, *J* = 0.9 Hz), 8.02 (1H, dd, *J* = 7.0, 0.9 Hz), 7.54 (1H, br s), 7.32-7.29 (1H, m), 6.69 (dd, *J* = 7.0, 1.6 Hz, 1H), 4.26 (2H, dd, *J* = 5.6, 2.5 Hz), 2.41 (3H, d, *J* = 1.1 Hz), 2.25 (t, J = 2.5 Hz, 1H)

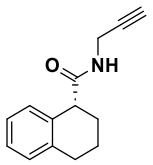
^{13}C-NMR (100 MHz; $CDCl_3$) δ 162.4, 145.0, 139.2, 137.3, 125.6, 116.3, 113.9, 79.5, 71.5, 28.8, 21.5

LRMS (ESI-QQQ): $C_{12}H_{11}N_3O$, calculated $[M+H]^+$ = 214.09, found = 214.1

HRMS (ESI): $C_{12}H_{11}N_3O$, calculated $[M+H]^+$ = 214.0975, found = 214.0970

IR (cm^{-1}): 3405, 3255, 3043, 2118, 1739, 1669, 1567, 1480

*6.3.2.4 Synthesis of 1,2,3,4-tetrahydro-N-2-propyn-1-yl-1-naphthalenecarboxamide (**3.69**)*



(R)-1,2,3,4-Tetrahydronaphthalene-1-carboxylic acid (400 mg, 2.27 mmol, 1.0 eq.), EDC·HCl (522 mg, 2.72 mmol, 1.2 eq.), HOBt (417 mg, 2.72 mmol, 1.2 eq.) and DIPEA (790 µL, 4.54 mmol, 2.0 eq.) were dissolved in anh. THF (15 mL) and the solution was stirred for 30 min. at rt. Propargylamine (145 µL, 2.27 mmol, 1.0 eq.) was added and the solution was stirred for 20 h at rt. The reaction was quenched with water (25 mL) and was extracted with EtOAc (2 x 30 mL). The combined organic extracts were washed with water (3 x 30 mL), brine (30 mL), dried (Na_2SO_4), filtered and concentrated in vacuo to give a brown powder. The powder was purified by automated flash column chromatography (0-30% EtOAc in hexane) to give 1,2,3,4-tetrahydro-*N*-2-propyn-1-yl-1-naphthalenecarboxamide (170 mg, 35%) as a white powder.

^{1}H NMR (400 MHz; $CDCl_3$) δ 7.24 – 7.07 (4H, m), 5.51 (1H, br s), 4.01 (2H, ddd, *J* = 5.3, 2.6, 2.6 Hz), 3.70 (1H, t, *J* = 5.3 Hz), 2.91 – 2.71 (2H, m), 2.36 – 2.25 (1H, m), 2.16 (1H, t, *J* = 2.6 Hz), 2.01 - 1.92 (1H, m), 1.84 – 1.72 (2H, m)

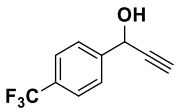
^{13}C-NMR (100 MHz; $CDCl_3$) δ 174.6, 138.0, 133.5, 130.0, 129.9, 127.5, 126.3, 79.5, 71.4, 46.9, 29.3, 29.2, 27.3, 20.3

LRMS (ESI-QQQ): $C_{14}H_{15}NO$, calculated $[M+H]^+$ = 214.13, found = 214.1

HRMS (ESI): $C_{14}H_{15}NO$ calculated $[M+H]^+$ = 214.1226, found = 214.1222

IR (cm^{-1}): 3275, 2941, 2924, 2100, 1738, 1641, 1532, 1223

*6.3.2.5 Synthesis of 1-(4-(trifluoromethyl)phenyl)prop-2-yn-1-ol (**3.70**)*



4-(Trifluoromethyl)benzaldehyde (500 mg, 2.87 mmol, 1 eq.) was dissolved in anh. THF (20 mL) under a N_2 atmosphere and the solution was cooled to 0 °C. Ethynylmagnesium bromide (12 mL of 0.5 M solution in THF, 6.00 mmol, 2.1 eq.) was added and the solution was allowed to warm to room temperature overnight. The reaction was quenched with sat. aq. NH_4Cl (20 mL) and was extracted with EtOAc (2 x 30 mL). The combined organic layers were washed with brine (30 mL), dried ($MgSO_4$), filtered and concentrated *in vacuo* to give a brown oil. The oil was purified by automated flash column chromatography (0-20% EtOAc in hexane) to give

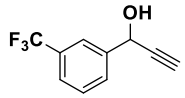
1-(4-(trifluoromethyl)phenyl)prop-2-yn-1-ol as a yellow oil (432 mg, 75%). The data matched that reported in literature.[11]

^{1}H NMR (400 MHz; $CDCl_3$) δ 7.70 – 7.63 (4H, m), 5.53 (1H, d, *J* = 2.4 Hz), 2.71 (1H, d, *J* = 2.4 Hz)

^{13}C-NMR (100 MHz; $CDCl_3$) δ 143.8 (q, *J* = 1.1 Hz), 130.8 (q, *J* = 32.2 Hz), 127.0, 125.8 (q, *J* = 3.8 Hz), 124.1 (q, *J* = 271.9 Hz), 82.9, 75.7, 63.9

IR (cm^{-1}): 3225, 2970, 2097, 1321, 1108

*6.3.2.6 Synthesis of 1-(3-(trifluoromethyl)phenyl)prop-2-yn-1-ol (**3.71**)*



3-(Trifluoromethyl)benzaldehyde (384 µL, 2.87 mmol, 1 eq.) was dissolved in anh. THF (20 mL) under a N_2 atmosphere and the solution was cooled to 0 °C. Ethynylmagnesium bromide (11.5 mL of 0.5 M solution in THF, 5.74 mmol, 2 eq.) was added and the solution was allowed to warm to room temperature overnight. The reaction was quenched with sat. aq. NH_4Cl (20 mL) and was extracted with EtOAc (2 x 30 mL). The combined organic layers were washed with brine (30 mL), dried ($MgSO_4$), filtered and concentrated *in vacuo* to give a brown oil. The oil was purified by automated flash column chromatography (0-20% EtOAc in hexane) to give 1-(3-(trifluoromethyl)phenyl)prop-2-yn-1-ol as a yellow oil (247 mg, 43%).

^{1}H NMR (400 MHz; $CDCl_3$) δ 7.84-7.82 (1H, m), 7.77-7.72 (1H, m), 7.62-7.59 (1H), 7.51 (tt, *J* = 7.7, 0.8 Hz, 1H), 5.53 (1H, d, *J* = 2.2 Hz), 2.72 (1H, d, *J* = 2.2 Hz)

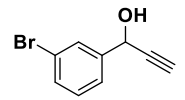
^{13}C-NMR (100 MHz; $CDCl_3$) δ 141.0, 131.2 (q, *J* = 31.6 Hz), 130.1 (q, *J* = 1.2 Hz), 129.3, 125.5 (q, *J* = 3.7 Hz), 124.1 (q, *J* = 269.9 Hz), 123.6 (q, *J* = 3.7 Hz), 82.9, 75.7, 63.9

^{19}F-NMR (376 MHz; $CDCl_3$) δ -62.9

LRMS (ESI-QQQ): $C_{10}H_7F_3O$, calculated $[M+H]^+$ = 200.04, found = 200.3

Rf (20% EtOAc in hexane): 0.45

*6.3.2.7 Synthesis of 1-(3-bromophenyl)prop-2-yn-1-ol (**3.72**)*



3-Bromobenzaldehyde (158 µL, 1.35 mmol, 1 eq.) was dissolved in anh. THF (20 mL) under a N_2 atmosphere and the solution was cooled to 0 °C. Ethynylmagnesium bromide (5.4 mL of 0.5 M solution in THF, 2.70 mmol, 2 eq.) was added and the solution was allowed to warm to room temperature overnight. The reaction was quenched with sat. aq. NH_4Cl (20 mL) and was extracted with EtOAc (3 x 20 mL). The combined organic layers were washed with brine

(20 mL), dried ($MgSO_4$), filtered and concentrated *in vacuo* to give a brown oil. The oil was purified by manual flash column chromatography (20% EtOAc in hexane) to give 1-(3-bromophenyl)prop-2-yn-1-ol as a yellow oil (171 mg, 60%). The data matched that reported in literature.[12]

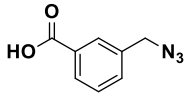
^{1}H NMR (400 MHz; $CDCl_3$) δ 7.71 (1H, t, *J* = 1.7 Hz), 7.47 (2H, ddd, *J* = 8.0, 1.7, 0.6 Hz), 7.28-7.24 (1H, m), 5.42 (1H, d, *J* = 2.3 Hz), 2.70 (1H, d, *J* = 2.3 Hz)

^{13}C-NMR (100 MHz; $CDCl_3$) δ 142.2, 131.7, 130.4, 129.8, 125.4, 122.8, 83.0, 75.5, 63.8

Rf (20% EtOAc in hexane): 0.24

6.3.3 Azide synthesis

*6.3.3.1 Synthesis of 3-(azidomethyl)benzoic acid (**3.87**)*



3.87 was prepared following a literature procedure.[13] 3-Bromomethylbenzoic acid (250 mg, 1.16 mmol, 1 eq.) and sodium azide (113 mg, 1.74 mmol, 1.5 eq.) were dissolved in DMF (10 mL) and the solution was stirred for 3h at rt. The reaction was quenched with brine (30 mL) and extracted with EtOAc (3 x 20 mL). The combined organic layers were washed with brine (5 x 20 mL), dried ($MgSO_4$), filtered and concentrated in vacuo to give 3-(azidomethyl)benzoic acid (**3.87)** as a white powder (148 mg, 72%). The data matched that reported in literature.[13]

^{1}H NMR (400 MHz; DMSO-d_6) δ 13.12 (1H, br s), 7.96-7.94 (1H, m), 7.92 (1H, ddd, *J* = 7.6, 1.5 Hz), 7.62 (1H, ddd, *J* = 7.6, 1.5 Hz), 7.53 (1H, t, *J* = 7.6 Hz), 4.56 (2H, s)

^{13}C-NMR (100 MHz; DMSO-d_6) δ 167.5, 136.7, 133.2, 131.8, 129.6, 129.5, 129.4, 53.5

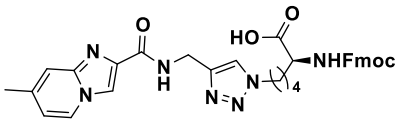
LRMS (ESI-QQQ): $C_8H_7N_3O_2$, calculated $[M-H]^-$ = 176.05, found = 176.0

IR (cm^{-1}): 2827, 2551, 2090, 1677, 1274

6.3.4 Triazole synthesis

Triazoles were prepared following a literature procedure.[14] Alkyne (1 eq.), azide (1 eq.), $CuSO_4 \cdot 5H_2O$ (0.25 eq.) and sodium ascorbate (2.5 eq.) were dissolved in DMF (0.14 mmol/mL) and stirred at rt for 18 h. The reaction mixture was purified by reverse phase flash chromatography (5-100% MeOH in H_2O + 0.05% TFA) and product-containing fractions were dried *in vacuo* to give products **3.47**-**3.65**.

*6.3.4.1 Synthesis of (S)-2-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-6-(4-((7-methylimidazo[1,2-a]pyridine-2-carboxamido)methyl)-1H-1,2,3-triazol-1-yl)hexanoic acid (**3.47**)*



3.47 was synthesised according to general protocol **6.3.4** yielding **3.47** as a yellow powder (4 mg, 10%)

^{1}H NMR (400 MHz; MeOD-d_4) δ 8.57 (1H, d, *J* = 4.9 Hz), 8.39 (1H, s), 7.98 (1H, s), 7.74 (2H, d, *J* = 7.5 Hz), 7.61 (2H, d, *J* = 6.8 Hz), 7.54 (1H, s), 7.37 (2H, dd, *J* = 7.5, 7.5 Hz), 7.31 – 7.23 (3H, m), 4.66 (2H, s), 4.42 (2H, t, *J* = 6.8 Hz), 4.29 (2H, d, *J* = 7.0 Hz), 4.16 (1H, t, *J* = 7.0 Hz), 4.11 (1H, dd, *J* = 9.3, 4.6 Hz), 2.52 (3H, s), 2.02 – 1.82 (3H, m) 1.77- 1.66 (1H, m), 1.49 – 1.33 (2H, m)

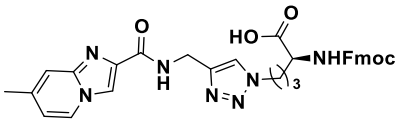
^{13}C-NMR (100 MHz; MeOD-d_4) δ 175.7, 161.8, 142.5, 128.8, 128.2, 126.2, 121.1, 120.9, 67.9, 55.1, 51.1, 35.9, 31.9, 30.7, 23.9, 21.8

LRMS (ESI-QQQ): $C_{33}H_{33}N_7O_5$, calculated $[M+H]^+$ = 608.26, found = 608.4

HRMS (ESI): $C_{33}H_{33}N_7O_5$, calculated $[M+H]^+$ = 608.2616, found = 608.2603

IR (cm^{-1}): 3250, 2947, 1663, 1180, 1130, 740

*6.3.4.2 Synthesis of (S)-2-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-5-(4-((7-methylimidazo[1,2-a]pyridine-2-carboxamido)methyl)-1H-1,2,3-triazol-1-yl)pentanoic acid (**3.48**)*



3.48 was synthesised according to general protocol **6.3.4** yielding **3.48** as a yellow powder (12 mg, 31%)

^{1}H NMR (400 MHz; DMSO-d_6) δ 8.72 (1H, t, *J* = 5.9 Hz), 8.46 (1H, d, *J* = 7.0 Hz, 1H), 8.29 (1H, s), 7.92 (1H, s), 7.86 (2H, d, *J* = 7.4 Hz), 7.70 (2H, d, *J* = 7.4 Hz), 7.54 (1H, d, *J* = 7.8 Hz), 7.39 (2H, t, *J* = 7.4 Hz), 7.34 (1H, s), 7.32-7.27 (2H, m), 6.82 (1H, dd, *J* = 7.0, 1.6 Hz), 4.51 (2H, d, *J* = 5.9 Hz), 4.34-4.17 (5H, m), 3.96-3.88 (1H, m), 2.35 (3H, s), 1.89 – 1.74 (2H, m), 1.72-1.63 (1H, m), 1.61 – 1.52 (1H, m).

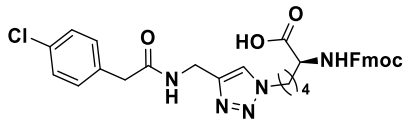
^{13}C-NMR (100 MHz; DMSO-d_6) δ 162.2, 156.0, 144.9, 143.9, 140.7, 139.2, 136.9, 127.6, 127.1, 126.8, 125.3, 122.7, 120.1, 115.7, 115.2, 114.2, 65.5, 53.6, 48.9, 46.7, 34.2, 28.0, 26.7, 20.9

LRMS (ESI-QQQ): $C_{32}H_{31}N_7O_5$, calculated $[M+H]^+$ = 594.25, found = 594.4

HRMS (ESI): $C_{32}H_{31}N_7O_5$, calculated $[M+H]^+$ = 594.2459, found = 594.2451

IR (cm^{-1}): 3316, 2980, 1681, 1650, 1553, 1446

*6.3.4.3 Synthesis of (S)-2-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-6-(4-((2-(4-chlorophenyl)acetamido)methyl)-1H-1,2,3-triazol-1-yl)hexanoic acid (**3.49**)*



3.49 was synthesised according to general protocol **6.3.4** yielding **3.49** as a pale yellow powder (9 mg, 24%).

^{1}H NMR (400 MHz; MeOD-d_4) δ 7.79 (2H, d, J = 7.3 Hz), 7.76 (1H, s), 7.66 (2H, dd, *J* = 7.3, 7.3 Hz), 7.38 (2H, *J* = 7.3, 7.3 Hz), 7.33 – 7.22 (6H, m), 4.40 (2H, s), 4.38 – 4.33 (4H, m), 4.21 (1H, t, *J* = 6.8 Hz), 4.11 (1H, dd, *J* = 9.0, 4.7 Hz), 3.50 (2H, s), 1.95 – 1.81 (3H, m), 1.76-1.65 (1H, m), 1.44 – 1.35 (2H, m).

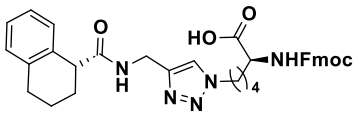
^{13}C-NMR (100 MHz; MeOD-d_4) δ 173.4, 158.6, 146.0, 145.4, 145.2, 142.6, 135.6, 133.8, 131.8, 129.6, 128.8, 128.2, 126.3, 124.2, 120.9, 67.9, 55.4, 51.1, 42.8, 35.8, 32.2, 30.7, 23.8

LRMS (ESI-QQQ): $C_{32}H_{32}ClN_5O_5$, calculated $[M+H]^+$ = 602.22, found = 602.2

HRMS (ESI): $C_{32}H_{32}ClN_5O_5$, calculated $[M+H]^+$ = 602.2165, found = 602.2153

IR (cm^{-1}): 3268, 2931, 1703, 1651, 1529, 1491, 1220

*6.3.4.4 Synthesis of (S)-2-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-6-(4-(((R)-1,2,3,4-tetrahydronaphthalene-1-carboxamido)methyl)-1H-1,2,3-triazol-1-yl)hexanoic acid (**3.50**)*



3.50 was synthesised according to general protocol **6.3.4** yielding **3.50** as a yellow powder (7 mg, 18%).

^{1}H NMR (400 MHz; DMSO-d_6) δ 8.53 (1H, t, *J* = 5.4 Hz), 7.92-7.86 (3H, m), 7.72 (2H, d, *J* = 7.5 Hz), 7.65 (1H, dd, *J* = 8.1, 2.7 Hz, 1H), 7.42 (2H, t, *J* = 7.5 Hz), 7.33 (2H, t, *J* = 7.5 Hz), 7.10-6.97 (3H, m), 4.36-4.26 (5H, m), 4.25-4.18 (1H, m), 3.97-3.87 (1H, m), 3.65 (1H, t, *J* = 6.9 Hz), 3.31 (1H, d, *J* = 6.8 Hz), 2.75-2.62 (2H, m), 2.00-1.92 (1H, m), 1.91-1.85 (2H, m), 1.84-1.70 (3H, m), 1.69-1.50 (2H, m), 1.40-1.26 (2H, m)

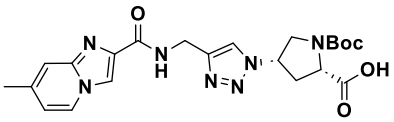
^{13}C-NMR (100 MHz; DMSO-d_6) δ 174.1, 173.8, 156.2, 143.8, 143.8, 140.7, 137.1, 135.2, 129.0, 128.3, 127.6, 127.1, 126.1, 125.5, 125.3, 120.1, 65.6, 53.7, 50.5, 49.1, 46.7, 45.1, 34.4, 29.4, 28.8, 26.9, 22.7, 20.7

LRMS (ESI-QQQ): $C_{35}H_{37}N_5O_5$, calculated $[M+H]^+$ = 608.29, found = 608.4

HRMS (ESI): $C_{35}H_{37}N_5O_5$, calculated $[M+H]^+$ = 608.2868, found = 608.2855

IR (cm^{-1}): 3275, 2931, 1705, 1650, 1524, 1450, 1217

*6.3.4.5 Synthesis of (2S,4S)-1-(tert-butoxycarbonyl)-4-(4-((7-methylimidazo[1,2-a]pyridine-2-carboxamido)methyl)-1H-1,2,3-triazol-1-yl)pyrrolidine-2-carboxylic acid (**3.51**)*



3.51 was synthesised according to general protocol **6.3.4** yielding **3.51** as a pale yellow powder (3 mg, 11%).

^{1}H NMR (400 MHz; MeOD-d_4) δ 8.79 (1H, s), 8.70 (1H, s), 8.10 (1H, s), 7.78 (1H, s), 7.37 (1H, d, *J* = 6.7 Hz), 5.36-5.25 (1H, m), 4.68 (s, 2H), 4.48-4.41 (1H, m), 4.23-4.11 (1H, m), 3.83-3.72 (1H, m), 3.43-3.33 (1H, m), 3.10-2.99 (1H, m), 2.59 (3H, s), 1.39 (9H, s)

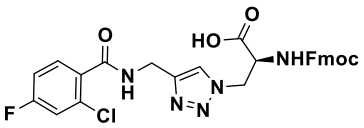
^{13}C-NMR (100 MHz; MeOD-d_4) δ 173.9, 173.2, 160.3, 159.8, 154.0, 129.4, 123.2, 120.6, 112.7, 80.5, 58.5, 58.0, 54.3, 52.4, 51.9, 35.7, 32.3, 28.4, 21.8

LRMS (ESI-QQQ): $C_{22}H_{27}N_7O_5$, calculated $[M+H]^+$ = 470.21, found = 470.3

HRMS (ESI): $C_{22}H_{27}N_7O_5$, calculated $[M+H]^+$ = 470.2146, found = 470.2137

IR (cm^{-1}): 3354, 2926, 1668, 1393, 1367, 1159, 1119

*6.3.4.6 Synthesis of (S)-2-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-3-(4-((2-chloro-4-fluorobenzamido)methyl)-1H-1,2,3-triazol-1-yl)propanoic acid (**3.52**)*



3.52 was synthesised according to general protocol **6.3.4** yielding **3.52** as a yellow powder (19 mg, 48%).

^{1}H NMR (400 MHz; DMSO-d_6) δ 8.97 (1H, t, *J* = 5.6 Hz), 7.89-7.85 (3H, m), 7.67-7.61 (2H, m), 7.52-7.43 (3H, m), 7.40 (2H, dd, *J* = 7.4, 7.4 Hz, 2H), 7.31 (2H, dd, *J* = 7.4, 7.4 Hz), 7.24-7.18 (1H, m), 6.98 (1H, d, *J* = 6.9 Hz), 4.85-4.77 (1H, m), 4.61-4.52 (1H, m), 4.46-4.40 (2H, m), 4.20-4.11 (4H, m)

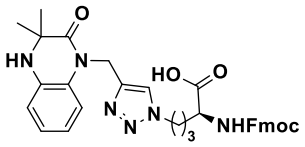
^{13}C-NMR (100 MHz; DMSO-d_6) δ 165.5, 162.0 (d, *J* = 248.4 Hz), 155.6, 144.1, 143.9, 143.8, 140.7, 133.3 (d, *J* = 3.1 Hz), 131.4 (d, *J* = 10.9 Hz), 130.7 (d, *J* = 9.5 Hz), 127.6, 127.1, 125.3, 123.3, 120.1, 116.9 (d, *J* = 23.9 Hz), 114.2 (d, *J* = 21.3 Hz), 65.7, 55.9, 50.9, 46.6, 34.8

LRMS (ESI-QQQ): $C_{28}H_{23}ClFN_5O_5$, calculated $[M+H]^+$ = 564.14, found = 564.2

HRMS (ESI): $C_{28}H_{23}ClFN_5O_5$, calculated $[M+H]^+$ = 564.1445, found = 564.1437

IR (cm^{-1}): 3296, 3035, 1715, 1601, 1535, 1450, 1260, 1216, 1055

*6.3.4.7 Synthesis of (2S)-5-{4-[(3,3-dimethyl-2-oxo-1,2,3,4-tetrahydroquinoxalin-1-yl)methyl]-1H-1,2,3-triazol-1-yl}-2-({[(9H-fluoren-9-yl)methoxy]carbonyl}amino)pentanoic acid (**3.53**)*



3.53 was synthesised according to general protocol **6.3.4** yielding **3.53** as a pale brown powder (31 mg, 79%).

^{1}H NMR (400 MHz; DMSO-d_6) δ 7.88 (2H, d, *J* = 7.2 Hz), 7.87 (1H, s), 7.71 (2H, d, *J* = 7.2 Hz), 7.67 (1H, d, *J* = 8.1 Hz), 7.40 (2H, dd, *J* = 7.2, 7.2 Hz), 7.31 (2H, dd, *J* = 7.2, 7.2 Hz), 7.11 (1H, d, *J* = 7.6 Hz), 6.84 (1H, dd, *J* = 7.6, 7.6 Hz), 6.74 (1H, d, *J* = 7.6 Hz), 6.68 (1H, dd, *J* = 7.6, 7.6 Hz), 5.09 (2H, s), 4.35-4.27 (4H, m), 4.25-4.18 (1H, m), 4.00-3.92 (1H, m), 1.87-1.77 (2H, m), 1.70-1.51 (2H, m), 1.23 (6H, s)

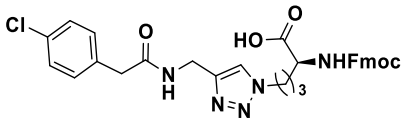
^{13}C-NMR (100 MHz; DMSO-d_6) δ 173.5, 169.5, 156.1, 143.9, 143.8, 143.2, 140.7, 135.1, 127.7, 127.3, 127.1, 125.3, 125.3, 123.3, 123.0, 120.1, 118.1, 114.6, 114.3, 65.6, 54.3, 53.3, 48.9, 46.7, 37.4, 27.7, 26.6, 24.7

LRMS (ESI-QQQ): $C_{33}H_{34}N_6O_5$, calculated $[M+H]^+$ = 595.27, found = 595.3

HRMS (ESI): $C_{33}H_{34}N_6O_5$, calculated $[M+H]^+$ = 595.2663 , found = 595.2654

IR (cm^{-1}): 3345, 2970, 1710, 1660, 1506, 1217, 740

*6.3.4.8 Synthesis of (S)-2-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-5-(4-((2-(4-chlorophenyl)acetamido)methyl)-1H-1,2,3-triazol-1-yl)pentanoic acid (**3.54**)*



3.54 was synthesised according to general protocol **6.3.4** yielding **3.54** as a pale green powder (16 mg, 41%).

^{1}H NMR (400 MHz; DMSO-d_6) δ 8.62 (1H, t, d, *J* = 5.7 Hz), 7.89 (1H, s), 7.88 (2H, d, *J* = 7.5 Hz), 7.71 (2H, d, *J* = 7.5 Hz), 7.41 (2H, t, *J* = 7.5 Hz), 7.36 – 7.24 (6H, m), 4.35 – 4.20 (7H, m), 3.98-3.88 (1H, m), 3.45 (2H, s), 1.88 – 1.76 (2H, m), 1.75-1.65 (1H, m), 1.64-1.54 (1H, m)

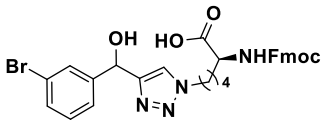
^{13}C-NMR (100 MHz; DMSO-d_6) δ 169.7, 156.1, 143.9, 143.8, 140.7, 135.3, 131.1, 130.9, 128.1, 127.6, 127.1, 125.3, 122.8, 120.1, 65.5, 49.0, 46.7, 41.3, 34.4, 28.1, 26.7

LRMS (ESI-QQQ): $C_{31}H_{30}ClN_5O_5$, calculated $[M+H]^+$ = 588.20, found = 588.3

HRMS (ESI): $C_{31}H_{30}ClN_5O_5$, calculated $[M+H]^+$ = 588.2008, found = 588.2001

IR (cm^{-1}): 3287, 2980, 1715, 1648, 1533, 1491, 1449, 1218

*6.3.4.9 Synthesis of (2S)-2-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-6-(4-((3-bromophenyl)(hydroxy)methyl)-1H-1,2,3-triazol-1-yl)hexanoic acid (**3.55**)*



3.55 was synthesised according to general protocol **6.3.4** yielding **3.55** as a pale red powder (5 mg, 13%).

^{1}H NMR (400 MHz; MeOD-d_4) δ 7.82 (1H, s), 7.79 (2H, d, *J* = 7.5 Hz), 7.67 (2H, dd, *J* = 7.2, 7.2 Hz), 7.62 (1H, s), 7.43 – 7.35 (4H, m), 7.33-7.28 (2H, m), 7.24 (1H, ddd, *J* = 7.8, 7.8, 2.9 Hz), 5.89 (1H, s), 4.42 – 4.29 (4H, m), 4.22 (1H, t, *J* = 6.9 Hz), 4.12 (1H, dd, *J* = 9.2, 4.7 Hz), 1.97 – 1.83 (3H, m), 1.75 – 1.64 (1H, m), 1.50 – 1.36 (2H, m)

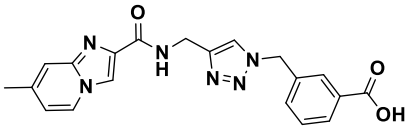
^{13}C-NMR (100 MHz; MeOD-d_4) δ 159.2, 145.7, 142.6, 131.7, 131.3, 130.5, 128.8, 128.2, 126.4, 126.3, 123.4, 120.9, 69.6, 67.5, 55.1, 51.3, 45.4, 32.0, 30.7, 24.0

LRMS (ESI-QQQ): $C_{30}H_{29}BrN_4O_5$, calculated $[M+H]^+$ = 605.14, found = 605.3

HRMS (ESI): $C_{30}H_{29}BrN_4O_5$, calculated $[M+H]^+$ = 605.1394, found = 605.1401

IR (cm^{-1}): 3327, 2948, 1705, 1699, 1521, 1450, 1218

*6.3.4.10 Synthesis of 3-((4-((7-methylimidazo[1,2-a]pyridine-2-carboxamido)methyl)-1H-1,2,3-triazol-1-yl)methyl)benzoic acid (**3.56**)*



3.56 was synthesised according to general protocol **6.3.4** yielding **3.56** as a yellow powder (9 mg, 41%).

^{1}H NMR (400 MHz; D_2O) δ 8.44 (1H, d, *J* = 6.9 Hz), 8.33 (1H, s), 8.08 (1H, s), 7.78-7.72 (2H, m), 7.57 (1H, s), 7.47 (1H, d, *J* = 7.5 Hz), 7.35 (1H, t, *J* = 7.5 Hz), 7.30 (1H, d, *J* = 6.9 Hz), 5.59 (2H, s), 4.66 (2H, s), 2.52 (3H, s)

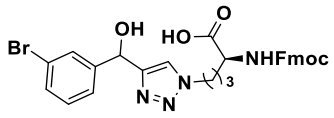
^{13}C-NMR (100 MHz; Acetone-d_6) δ 167.8, 167.8, 137.2, 133.4, 132.0, 130.2, 129.9, 119.5, 53.8, 32.1, 23.1

LRMS (ESI-QQQ): $C_{20}H_{18}N_6O_3$, calculated $[M+H]^+$ = 391.14, found = 391.2

LRMS (ESI): $C_{20}H_{18}N_6O_3$, calculated $[M+H]^+$ = 391.1513, found = 391.1503

IR (cm^{-1}): 3120, 3073, 1661, 1655, 1429, 1179, 1127

*6.3.4.11 Synthesis of (2S)-2-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-5-(4-((3-bromophenyl)(hydroxy)methyl)-1H-1,2,3-triazol-1-yl)pentanoic acid (**3.57**)*



3.57 was synthesised according to general protocol **6.3.4** yielding **3.57** as a yellow powder (6 mg, 15%).

^{1}H NMR (400 MHz; MeOD-*d*$_4$) δ 7.85 – 7.77 (3H, m), 7.71 – 7.63 (3H, m), 7.47 – 7.36 (4H, m), 7.36 – 7.22 (3H, m), 5.93 (1H, s), 4.44 - 4.38 (4H, m), 4.23 (1H, t, *J* = 7.0 Hz), 4.21 – 4.16 (1H, m), 2.01 – 1.92 (2H, m), 1.90 - 1.81 (1H, m), 1.75 – 1.63 (1H, m)

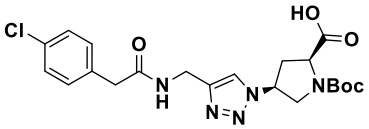
^{13}C-NMR (100 MHz; MeOD-*d*$_4$) δ 175.3, 147.0, 145.4, 145.1, 142.6, 142.6, 131.7, 131.3, 130.5, 128.8, 128.2, 126.4, 126.2, 123.5, 123.4, 120.9, 69.3, 67.8, 54.6, 50.8, 41.1, 29.6, 27.9

LRMS (ESI-QQQ): $C_{29}H_{27}BrN_4O_5$, calculated $[M+H]^+$ = 591.12, found = 591.3

HRMS (ESI): $C_{29}H_{27}BrN_4O_5$, calculated $[M+H]^+$ = 591.1238, found = 591.1240

IR (cm^{-1}): 3357, 2948, 1705, 1700, 1521, 1449, 1217

*6.3.4.12 Synthesis of (2S,4S)-1-(tert-butoxycarbonyl)-4-(4-((2-(4-chlorophenyl)acetamido)methyl)-1H-1,2,3-triazol-1-yl)pyrrolidine-2-carboxylic acid (**3.58**)*



3.58 was synthesised according to general protocol **6.3.4** yielding **3.58** as a yellow powder (7 mg, 19%).

^{1}H NMR (400 MHz; DMSO-*d*$_6$) δ 8.55 (1H, t, *J* = 5.4 Hz), 8.04 (1H, s), 7.34 (2H, d, *J* = 8.4 Hz), 7.27 (2H, d, *J* = 8.4 Hz), 5.23-5.14 (1H, m), 4.28 (2H, d, *J* = 5.4 Hz), 4.22 (1H, t, *J* = 7.9 Hz), 4.09-3.99 (1H, m), 3.55 (1H, dd, *J* = 10.7, 7.9 Hz), 2.94-2.84 (1H, m), 2.42-2.32 (1H, m), 1.36 (9H, s).

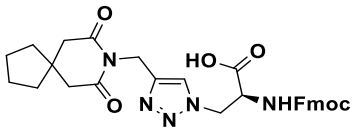
^{13}C-NMR (100 MHz; DMSO-*d*$_6$) δ 169.7, 166.6, 144.8, 135.3, 131.1, 130.9, 128.1, 122.2, 79.3, 56.6, 56.0, 41.2, 34.3, 28.0, 27.9, 24.0

LRMS (ESI-QQQ): $C_{21}H_{26}ClN_5O_5$, calculated $[M+H]^+$ = 464.17, found = 464.3

HRMS (ESI): $C_{21}H_{26}ClN_5O_5$, calculated $[M+H]^+$ = 464.1695, found = 464.1689

Rf (10% MeOH in DCM): 0.51

*6.3.4.13 Synthesis of (S)-2-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-3-(4-((7,9-dioxo-8-azaspiro[4.5]decan-8-yl)methyl)-1H-1,2,3-triazol-1-yl)propanoic acid (**3.59**)*



3.59 was synthesised according to general protocol **6.3.4** yielding **3.59** as a yellow powder (14 mg, 35%).

^{1}H NMR (400 MHz; DMSO-d_6) δ 7.88 (2H, d, *J* = 7.5 Hz), 7.74 (1H, s), 7.65 (2H, t, *J* = 6.7 Hz), 7.41 (2H, t, *J* = 7.5 Hz), 7.32 (2H, t, *J* = 7.5 Hz), 4.90 – 4.73 (3H, m), 4.52 (1H, dd, *J* = 13.7, 8.0 Hz), 4.26 – 4.12 (4H, m), 2.59 (4H, s), 1.59 – 1.54 (m, 5H), 1.44-1.33 (4H, m)

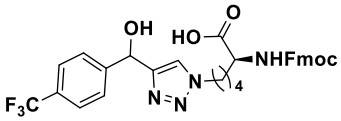
^{13}C-NMR (100 MHz; DMSO-d_6) δ 171.8, 155.6, 143.8, 142.6, 140.7, 127.6, 127.2, 125.2, 123.4, 120.1, 65.7, 55.3, 50.4, 46.6, 43.6, 36.8, 34.0, 23.6

LRMS (ESI-QQQ): $C_{30}H_{31}N_5O_6$, calculated $[M+H]^+$ = 558.23, found = 558.3

HRMS (ESI): $C_{30}H_{31}N_5O_6$, calculated $[M+H]^+$ = 558.2347, found = 558.2338

IR (cm^{-1}): 3381, 2949, 1723, 1669, 1217, 740

*6.3.4.14 Synthesis of (2S)-2-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-6-(4-(hydroxy(4-(trifluoromethyl)phenyl)methyl)-1H-1,2,3-triazol-1-yl)hexanoic acid (**3.60**)*



3.60 was synthesised according to general protocol **6.3.4** yielding **3.60** as an orange powder (22 mg, 58%).

^{1}H NMR (400 MHz; DMSO-d_6) δ 7.93 (1H, s), 7.89 (2H, d, *J* = 7.5 Hz), 7.75-7.67 (4H, m), 7.66-7.61 (3H, m), 7.41 (2H, t, *J* = 7.5 Hz), 7.33 (2H, t, *J* = 7.5 Hz), 5.91 (1H, s), 4.33-4.26 (4H, m), 4.24-4.19 (1H, m), 3.94-3.87 (1H, m), 1.86-1.68 (3H, m), 1.67-1.57 (1H, m), 1.39-1.22 (2H, m)

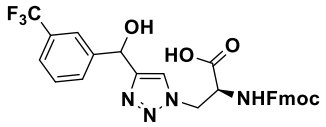
^{13}C-NMR (100 MHz; DMSO-d_4) δ 173.8, 156.2, 150.6, 148.7, 143.8, 140.7 (q, *J* = 1.1 Hz), 127.6, 127.5 (q, *J* = 46.1 Hz), 127.1, 125.3 (q, *J* = 2.6 Hz), 125.0 (q, *J* = 3.7 Hz), 124.3 (q, *J* = 271.5 Hz), 122.2, 120.1, 67.3, 65.6, 53.7, 49.1, 46.7, 30.1, 29.4, 22.7

LRMS (ESI-QQQ): $C_{31}H_{29}F_3N_4O_5$, calculated $[M+H]^+$ = 595.21, found = 595.3

HRMS (ESI): $C_{31}H_{29}F_3N_4O_5$, calculated $[M+H]^+$ = 595.2163, found = 595.2152

IR (cm^{-1}): 3378, 2948, 1714, 1710, 1521, 1450, 1324

*6.3.4.15 Synthesis of (S)-2-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-3-(4-((S)-hydroxy(3-(trifluoromethyl)phenyl)methyl)-1H-1,2,3-triazol-1-yl)propanoic acid (**3.61**)*



3.61 was synthesised according to general protocol **6.3.4** yielding **3.61** as a yellow powder (15 mg, 38%).

^{1}H NMR (400 MHz; DMSO-d_6) δ 7.93 (1H, d, *J* = 5.6 Hz), 7.90-7.86 (3H, m), 7.77 (1H, s), 7.67-7.56 (4H, m), 7.50 (1H, t, *J* = 7.8 Hz), 7.44-7.38 (2H, m), 7.34-7.28 (2H, m), 5.92 (1H, d, *J* = 5.1 Hz), 4.76 (1H, dd, *J* = 13.7, 4.3 Hz), 4.63-4.47 (2H, m), 4.25-4.14 (3H, m)

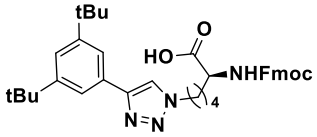
^{13}C-NMR (100 MHz; DMSO-d_6) δ 170.8, 155.9, 145.4, 143.7, 143.6, 140.7, 130.7, 129.1, 128.8 (q, *J* = 31.2 Hz), 127.7, 127.1, 125.2 (q, *J* = 3.9 Hz), 124.1 (q, *J* = 262.5 Hz), 123.8 (q, *J* = 2.8 Hz), 123.0, 122.9, 122.8, 120.1, 67.1, 65.9, 54.2, 49.5, 46.5

LRMS (ESI-QQQ): $C_{28}H_{23}F_3N_4O_5$, calculated $[M+H]^+$ = 553.16, found = 533.1

HRMS (ESI): $C_{28}H_{23}F_3N_4O_5$, calculated $[M+H]^+$ = 553.1693, found = 553.1690

IR (cm^{-1}): 3389, 2971, 1723, 1718, 1518, 1450, 1328, 1218

*6.3.4.16 Synthesis of (S)-2-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-6-(4-(3,5-di-tert-butylphenyl)-1H-1,2,3-triazol-1-yl)hexanoic acid (**3.62**)*



3.62 was synthesised according to general protocol **6.3.4** yielding **3.62** as a pale yellow powder (28 mg, 73%).

^{1}H NMR (400 MHz; MeOD-d_4) δ 8.37 (1H, s), 7.75 (2H, d, *J* = 7.5 Hz), 7.68 (2H, d, *J* = 1.9 Hz), 7.59 (1H, d, *J* = 7.5 Hz), 7.45 (1H, t, *J* = 1.9 Hz), 7.34 (2H, dd, *J* = 7.5, 7.5 Hz), 7.25 (2H, ddd, *J* = 7.5, 7.5, 1.2 Hz), 4.47 (2H, t, *J* = 6.9 Hz), 4.36 – 4.22 (2H, m), 4.19 – 4.11 (2H, m), 2.10 – 1.96 (2H, m), 1.96 – 1.85 (1H, m), 1.81 – 1.68 (1H, m), 1.51 – 1.42 (2H, m), 1.34 (18H, s)

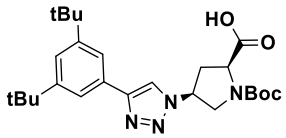
^{13}C-NMR (100 MHz; MeOD-d_4) δ 175.8, 159.0, 152.7, 149.7, 145.3, 145.1, 142.5, 130.8, 128.8, 128.1, 128.1, 126.2, 123.5, 122.2, 121.2, 120.9, 67.9, 55.0, 51.3, 35.8, 31.9, 30.6, 23.8

LRMS (ESI-QQQ): $C_{37}H_{44}N_4O_4$, calculated $[M+H]^+$ = 609.34, found = 609.4

HRMS (ESI): $C_{37}H_{44}N_4O_4$, calculated $[M+H]^+$ = 609.3435, found = 609.3424

IR (cm^{-1}): 3380, 2954, 1718, 1518, 1450, 1217

*6.3.4.17 Synthesis of (2S,4S)-1-(tert-butoxycarbonyl)-4-(4-(3,5-di-tert-butylphenyl)-1H-1,2,3-triazol-1-yl)pyrrolidine-2-carboxylic acid (**3.63**)*



3.63 was synthesised according to general protocol **6.3.4** yielding **3.63** as a white powder (17 mg, 63%).

^{1}H NMR (400 MHz; MeOD-d_4) δ 8.50 (1H, s), 7.69 (2H, d, *J* = 1.9 Hz), 7.45 (1H, t, *J* = 1.9 Hz), 5.29 (1H, p, *J* = 7.1 Hz), 4.49 – 4.41 (m, 1H), 4.18 (1H, dd, *J* = 11.2, 7.3 Hz), 3.96 (1H, dd, *J* = 11.2, 6.6 Hz), 3.10 – 2.99 (1H, m), 2.79 – 2.69 (1H, m), 1.51 – 1.44 (m, 9H), 1.37 (s, 18H)

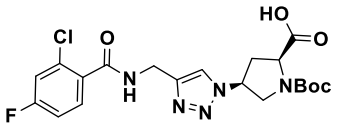
^{13}C-NMR (100 MHz; MeOD-d_4) δ 155.5, 152.7, 149.9, 130.8, 123.5, 121.2, 82.3, 59.2, 58.9, 52.0, 35.8, 31.9, 28.5

LRMS (ESI-QQQ): $C_{26}H_{38}N_4O_4$, calculated $[M+H]^+$ = 471.29, found = 471.4

HRMS (ESI): $C_{26}H_{38}N_4O_4$, calculated $[M+H]^+$ = 471.2966, found = 471.2959

IR (cm^{-1}): 2953, 1750, 1699, 1394, 1366

*6.3.4.18 Synthesis of (2S,4S)-1-(tert-butoxycarbonyl)-4-(4-((2-chloro-4-fluorobenzamido)methyl)-1H-1,2,3-triazol-1-yl)pyrrolidine-2-carboxylic acid (**3.64**)*



3.64 was synthesised according to general protocol **6.3.4** yielding **3.64** as a yellow powder (8 mg, 30%).

^{1}H NMR (400 MHz; Acetone-d_6) δ 8.06 (1H, s), 7.60 (1H, dd, *J* = 8.4, 6.1 Hz), 7.31 (1H, dd, *J* = 8.9, 2.5 Hz), 7.18 (ddd, *J* = 8.4, 8.4, 2.5 Hz, 1H), 5.39-5.24 (1H, m), 4.63 (2H, s), 4.45 (1H, t, *J* = 7.9 Hz), 4.24-4.13 (1H, m), 3.82-3.67 (1H, m), 3.14-3.02 (1H, m), 2.66-2.56 (1H, m), 1.41 (9H, s)

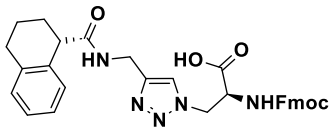
^{13}C-NMR (100 MHz; Acetone-d_6) δ 166.3, 163.7 (d, *J* = 249.0 Hz), 145.7, 134.1 (d, *J* = 4.5 Hz), 132.9 (d, *J* = 11.2 Hz), 131.9 (d, *J* = 9.1 Hz), 122.7, 117.9 (d, *J* = 25.3 Hz), 117.8, 114.9 (d, *J* = 21.3 Hz), 80.6, 58.8, 57.8, 51.9, 36.7, 36.0, 28.4, 23.3

LRMS (ESI-QQQ): $C_{20}H_{23}ClFN_5O_5$, calculated $[M+H]^+$ = 468.14, found = 468.3

HRMS (ESI): $C_{20}H_{23}ClFN_5O_5$, calculated $[M+H]^+$ = 468.1445, found = 468.1436

IR (cm^{-1}): 3310, 2934, 1737, 1697, 1366, 1217, 1158

*6.3.4.19 Synthesis of (S)-2-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-3-(4-(((R)-1,2,3,4-tetrahydronaphthalene-1-carboxamido)methyl)-1H-1,2,3-triazol-1-yl)propanoic acid (**3.65**)*



3.65 was synthesised according to general protocol **6.3.4** yielding **3.65** as a pale yellow powder (20 mg, 50%).

^{1}H NMR (400 MHz; DMSO-d_6) δ 8.55-8.50 (1H, m), 7.88 (2H, d, *J* = 7.4 Hz), 7.81 (1H, s), 7.65 (2H, t, *J* = 7.4 Hz), 7.40 (2H, t, *J* = 7.4 Hz), 7.32 (2H, t, *J* = 7.4 Hz), 7.09-6.99 (4H, m), 4.84-4.76 (1H, m), 4.62-4.50 (1H, m), 4.34-4.28 (2H, m), 4.25-4.14 (3H, m), 3.64 (1H, t, *J* = 6.9 Hz), 2.74-2.63 (2H, m), 1.99-1.82 (3H, m), 1.60-1.53 (1H, m), 1.23 (1H, s)

^{13}C-NMR (100 MHz; DMSO-d_6) δ 174.1, 144.5, 143.8, 140.7, 137.0, 135.2, 128.9, 128.4, 127.6, 127.2, 127.1, 126.0, 125.6, 125.3, 125.3, 123.1, 120.1, 65.7, 50.7, 46.6, 45.2, 34.4, 28.8, 27.0, 20.7

LRMS (ESI-QQQ): $C_{32}H_{31}N_5O_5$, calculated $[M+H]^+$ = 566.23, found = 566.3

LRMS (ESI): $C_{32}H_{31}N_5O_5$, calculated $[M+H]^+$ = 566.2411, found = 566.2394

IR (cm^{-1}): 3315, 1714, 1641, 1518, 1449, 1218

6.3.5 Surface plasmon resonance

SPR was performed as described in **Section 6.2.6.**

6.4 The cellular evaluation of peptides and small molecules targeting the dyskerin-dyskerin PPI

6.4.1 General cell culture

A549 cells were cultures in RPMI 1640 supplemented with 10% FBS, 2 mM L-glutamine, 100 U/mL penicillin and 100 µg/mL streptomycin. The cells were subcultured 1 in 10 twice a week and grown at 37 °C, 5% CO_2. During each subculture process the media was removed from the flask and the cells were washed with PBS (5 mL). TrypLE express (2 mL) was added and the cells were incubated at 37 °C, 5% CO_2 for approximately 5 minutes until the cells lifted away from the flask. To deactivate the TrypLE express, media (8 mL) was added to the flask and by pipetting up and down a homogeneous cell suspension was created. All but 1 mL of this cell suspension was removed from the flask and media (9 mL) was added before returning the cells to the incubator.

6.4.2 MTS

The effect of the compounds on cell viability was assessed by MTS assay (Promega) following the manufacturer's instructions. A549 cells were seeded in a 96 well plate (3 x 10^3 / 100 µL) and incubated overnight at 37 °C, 5% CO_2. Cells were treated in triplicate with sample using 1 in 2 dilutions (1% DMSO v/v) for 72 hours. MTS (10 µL) was added and the cells were incubated at 37 °C, 5% CO_2 for 60 min. For control growth cells were treated with DMSO (1% v/v) and for cell death control cells were treated with staurosporin (10 µM). The absorbance was measured at 492 nm using the BMG Labtech POLARstarOPTIMA microplate reader. Background absorbance was corrected for by subtracting the absorbance of the cell death control and cell viability was calculated as a percentage of DMSO treated cells. IC_{50} was determined using GraphPad Prism Version 6.0 software, using a four-parameter logistic model.

6.4.3 Flow cytometry

A549 cells were plated in 12-well plates at 1 · 10^5/well and left to adhere overnight. Cells were treated with the specified compounds for 1 h after which the cells were washed with PBS (3x). The cells were trypsinized using TrypLE (0.2 mL/well) after which PBS (0.8 mL) was added and the cells were pelleted (300 g, 5 min.). The cells were washed with PBS (2X) before being resuspended in PBS (1 mL) for the flow cytometry experiments. Live cells were gated and 10,000 events were recorded.

6.4.4 Proteolytic stability

The proteolytic stability was recorded through HPLC measurements. The peptide was dissolved to 20 mM in DMSO. From this 2 µL was diluted to 1 mL with 10 mM phosphate buffer pH 7.4. At each timepoint 100 µL of the solution was taken and added to 200 µL of 1:1 MeCN:H_2O with 1% TFA to quench the activity of proteinase K. For each timepoint 50 µL of this sample was injected into the HPLC and the area under the peak was measured to determine the percentage of peptide remaining. Before proteinase K was added a sample was taken and quenched as the rest of the timepoints to record a peak area for the t = 0 min, which was used as 100% of peptide remaining. After this sample had been taken 5 µL of 2 mg/mL proteinase K was added to start the proteolytic degradation. The data was analysed using GraphPad Prism Version 6.0 software, using One phase decay to determine the half-lives.

6.5 Recombinant expression of dyskerin

6.5.1 Recombinant expression of MBP-dyskerin

Plasmid for MBP-dyskerin-his was kindly gifted by Prof. Tom Meier and the expression procedure was adapted from literature.[15] Plasmid was transformed into Rosetta2(DE3)pLysS cells and plated on agar containing chloramphenicol (34 µg/mL) and ampicillin (50 µg/mL). A single colony was picked and cultured in LB (10 mL) containing the same antibiotics overnight at 37 °C, 200 rpm. Samples were stored following the addition of glycerol to the overnight culture to a final glycerol concentration of 25% and stored at -78 °C from which new overnight cultures were formed. Overnight cultures were diluted in LB (1 in 50) and the cells were incubated at 37 °C, 180 rpm until OD_{600} = 0.6. IPTG (1 mM) was added and the cells were incubated at 37 °C, 180 rpm for 3 h. The cells were collected by centrifugation (6000 g, 18 min., 4 °C) using minimal lysis buffer (800 mM NaCl, 20 mM Tris, 1 mM EDTA, 1 mM DTT, pH 7.4) and stored at -78 °C.

Cells were lysed using sonication in lysis buffer containing lysozyme (0.25 mg/mL) and a protease inhibitor cocktail. The solution was clarified using centrifugation (15000 g, 30 min., 4 °C) and the supernatant was loaded onto amylose resin. The resin was washed with washing buffer (200 mM NaCl, 20 mM Tris, 1 mM EDTA, 1 mM DTT, pH 7.4) and eluted with elution buffer (200 mM NaCl, 20 mM Tris, 1 mM EDTA, 1 mM DTT, 10 mM maltose pH 7.4). The eluate was concentrated using a spin column (10 kDa MWCO) and loaded onto a SEC column (Superdex 75) in 10 mM HEPES, 150 mM NaCl, 0.05% Tween-20, pH 7.4. The fractions were analysed by SDS-PAGE and the MBP-dyskerin-his containing fractions were concentrated using a spin column (10 kDa MWCO). Protein was analysed using MALDI peptide fingerprinting and shown to contain MBP-dyskerin-his.

6.5.2 SDS-PAGE

For SDS-PAGE samples were prepared by combining protein sample (10 µL), tris-glycine SDS buffer (10 µL) and NuPAGE reducing agent (2 µL). This solution was heated at 95 °C for 5 min. before loading samples onto a 4-12% SDS-polyacrylamide gel alongside Thermo Scientific PageRuler prestained protein ladder (3 µL). Electrophoresis was performed for 48 minutes using a constant voltage of 200 V in MOPS SDS running buffer.

The gel was washed with water and stained with Coomassie stain for 10 minutes with heating. The staining solution was removed and the gel was washed with water before being destained with destaining solution until the background was colourless and bands were clearly visible.

Coomassie stain was prepared by dissolving Coomassie Brilliant blue R250 (1 g) in methanol (400 mL), glacial acetic acid (100 mL) and water (500 mL) and the solution was filtered before use.

The destaining solution was prepared by mixing methanol (200 mL), glacial acetic acid (100 mL) and water (700 mL).

6.5.3 Plasmid extraction

MBP-dyskerin plasmid was extracted from 10 mL of overnight culture of NEB® 5-alpha cells. The cells were pelleted (6000 g, 18 min., 4 °C). The plasmid was extracted using the QIAprep spin Miniprep kit following manufacturer's protocol.

6.6 References

1 A. M. Giltrap, F. P. J. Haeckl, K. L. Kurita, R. G. Linington and R. J. Payne, Synthetic Studies Toward the Skyllamycins: Total Synthesis and Generation of Simplified Analogues, *J. Org. Chem.*, 2018, **83**, 7250–7270.

2 H. Martin-Gómez, F. Albericio and J. Tulla-Puche, A Lasso-Inspired Bicyclic Peptide: Synthesis, Structure and Properties, *Chemistry – A European Journal*, 2018, **24**, 19250–19257.

3 A. M. Rossi and C. W. Taylor, Analysis of protein-ligand interactions by fluorescence polarization, *Nat Protoc*, 2011, **6**, 365–387.

4 L. Whitmore and B. A. Wallace, Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases, *Biopolymers*, 2008, **89**, 392–400.

5 N. Sreerama and R. W. Woody, Estimation of protein secondary structure from circular dichroism spectra: Comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set, *Analytical Biochemistry*, 2000, **287**, 252–260.

6 G. Madhavi Sastry, M. Adzhigirey, T. Day, R. Annabhimoju and W. Sherman, Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments, *J Comput Aided Mol Des*, 2013, **27**, 221–234.

7 J. C. Shelley, A. Cholleti, L. L. Frye, J. R. Greenwood, M. R. Timlin and M. Uchimaya, Epik: a software program for pKaprediction and protonation state generation for drug-like molecules, *J Comput Aided Mol Des*, 2007, **21**, 681–691.

8 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, Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy, *J. Med. Chem.*, 2004, **47**, 1739–1749.

9 K. Zhu, K. W. Borrelli, J. R. Greenwood, T. Day, R. Abel, R. S. Farid and E. Harder, Docking Covalent Inhibitors: A Parameter Free Approach To Pose Prediction and Scoring, *J. Chem. Inf. Model.*, 2014, **54**, 1932–1940.

10 W. Sherman, H. S. Beard and R. Farid, Use of an Induced Fit Receptor Structure in Virtual Screening, *Chemical Biology & Drug Design*, 2006, **67**, 83–84.

11 H.-S. M. Siah, M. C. Hogsnes, N. Iqbal and A. Fiksdahl, Gold(I)-catalysed tandem cyclization of propargyl acetals and alkynes, *Tetrahedron*, 2016, **72**, 1058–1068.

12 A. Kolarovič and Z. Fáberová, Catalytic Decarboxylation of 2-Alkynoic Acids, *J. Org. Chem.*, 2009, **74**, 7199–7202.

13 Y. Sako, S. Ichikawa, A. Osada and A. Matsuda, Synthesis and evaluation of 5-substituted 9-hydroxypyrrolo[3,4-c]carbazole-1,3(2H,6H)-diones as check point 1 kinase inhibitors, *Bioorganic & Medicinal Chemistry*, 2010, **18**, 7878–7889.

14 A. M. Beekman, M. M. D. Cominetti, S. J. Walpole, S. Prabhu, M. A. O'Connell, J. Angulo and M. Searcey, Identification of selective protein-protein interaction inhibitors using efficient: In silico peptide-directed ligand design, *Chemical Science*, 2019, **10**, 4502–4508.

15 P. N. Grozdanov, S. Roy, N. Kittur and U. T. Meier, SHQ1 is required prior to NAF1 for assembly of H/ACA small nucleolar and telomerase RNPs, *RNA*, 2009, **15**, 1188–1197.