

The design and synthesis of peptide analogues targeting DNA four-way junctions

Eleanor Ivens

School of Chemistry, Pharmacy and Pharmacology

University of East Anglia

December 2024

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.

Eleanor Ivens

Abstract

Deoxyribonucleic acid (DNA) most commonly exists *in vivo* in its double-stranded helical form, with this being the primary target of DNA-targeting drugs. However, the substantial size of mega base DNA, and its high level of flexibility, means that it can also form other three-dimensional shapes, known as higher-order DNA. Many of these structures have been linked to diseases and are increasingly being explored as novel therapeutic targets.

The DNA four-way junction (4WJ) is a higher order DNA structure that is formed by the crossing over of DNA strands belonging to two separate dsDNA segments. In low salt environments, the 4WJ exists in the open conformation, whereas in high salt environments, the X-stacked conformation is adopted. The 4WJ is an important intermediate in DNA pressing and repair, such as during homologous recombination and site-specific recombination. These pathways are often relied upon by both bacterial and cancer cells.

Peptides have previously been discovered that bind to the 4WJ in its open conformation and have antibacterial effects. In this thesis, a series of novel biological assays were developed, to further access the 4WJ activity, double-stranded DNA activity and cancer cell cytotoxicity of these peptides. A combination of solid-phase peptide synthesis and multi-step organic synthesis were then used to produce novel peptide analogues, including cyclic peptides and linear peptides containing non-canonical amino acids. These molecules were predominantly shown to have similar biological results to the original peptides, including their ability to maintain 4WJ activity.

The rational design of 4WJ-targeting peptides was also investigated, based on protein/4WJ crystal structures. Although the two peptides designed, synthesised and tested were shown to have no 4WJ activity, this approach could be worth a further look, as more structural information on the DNA 4WJ is obtained.

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.

Acknowledgements

I would like to thank my primary supervisor Prof Mark Searcey for all his help during the last three years. It has been amazing to learn from his wealth of knowledge, whilst getting the opportunity to have great discussions about the direction of my research.

I am grateful to my secondary supervisor Dr Andrew Beekman for helping me stay on track and having great advice whenever I needed it. I would like to thank him for reminding me to have a life during my PhD, especially with the sacred Friday pub session.

Alongside my supervisors, I'd like to thank Dr Zoë Goddard and Dr Marco Cominetti for their incredible support with lab work and anything else I needed during my PhD. It has been wonderful to be a part of the Searcey/Beekman lab and work alongside friends; I could not have done it without them.

During my PhD climbing has always been there to help keep me sane, even if I did break my finger at one point, forcing me to switch to XL gloves. I have made great friends through climbing who have stuck by me, even through the endless stress rants.

Finally, I would like to give a big thank you to my family for their unwavering support. In particular, my dad for always being on the end of the phone to talk sense into me and guide me through this experience. Also, my mum for the coffees, dog walks and dinners that I have enjoyed so much. I love you both.

Contents

DECLARATION	II
ABSTRACT	III
ACKNOWLEDGEMENTS.....	IV
CONTENTS.....	V
ABBREVIATIONS.....	VIII
CHAPTER 1 – INTRODUCTION.....	1
1.1 DNA structure	2
1.2 Higher-order DNA.....	4
1.3 The DNA four-way junction.....	8
1.3.1 Targeting 4WJs in cancer	10
1.3.2 Targeting 4WJs in bacterial infections	12
1.3.3 4WJ conformations	14
1.4 4WJ-binding ligands.....	17
1.4.1 Peptides.....	17
1.4.2 Metal complexes.....	24
1.4.3 Small molecules	27
1.4.4 Crystal structures of 4WJ/ligand complexes.....	36
1.4.5 Summary	38
1.5 Thesis aims.....	38
1.6 Thesis outline	39
CHAPTER 2 - THE SYNTHESIS AND BIOLOGICAL TESTING OF WRWYRGGRYWRW AND RELATED PEPTIDES.....	41
2.1 Introduction	42
2.1.1 Solid-phase peptide synthesis.....	42
2.1.2 Fluorescence assays	48
2.1.3 Assays for identifying 4WJ conformations.....	54
2.2 Chapter aims.....	59
2.3 Results and discussion	60
2.3.1 Synthesis of peptides using SPPS	60
2.3.2 Synthesis of acridines.....	63

2.3.3 PAGE assay	65
2.3.4 FRET	74
2.3.5 Fluorescence polarisation	83
2.3.6 Ethidium bromide displacement assay	87
2.3.7 MTS assay	89
2.4 Conclusion and future work	92
 CHAPTER 3 - THE SYNTHESIS AND BIOLOGICAL TESTING OF PEPTIDE ANALOGUES BASED ON WRWYRGGRYWRW	 94
3.1 Introduction	95
3.1.1 Peptidomimetics	95
3.1.2 Cyclic peptides	99
3.2 Chapter aims	102
3.3 Results and discussion	103
3.3.1 The design of peptide analogues	103
3.3.2 Choice of linkers	105
3.3.3 Cyclic peptide analogues	107
3.3.4 Solution phase synthesis and SPPS of peptide analogues	108
3.3.5 Synthesis of cyclic peptides	116
3.3.6 PAGE assay	118
3.3.7 Fluorescence polarisation	120
3.3.8 Ethidium bromide displacement assay	121
3.3.9 FRET	123
3.3.10 MTS assay	124
3.4 Conclusion and future work	128
 CHAPTER 4 – DEVELOPING 4WJ-BINDING PEPTIDES FROM PROTEIN-4WJ CRYSTAL STRUCTURES	 130
4.1 Introduction	131
4.1.1 Identifying therapeutically active peptides	131
4.1.2 4WJ resolvases	135
4.1.3 The MOC1-4WJ complex	135
4.2 Chapter aims	139
4.3 Results and discussion	139
4.3.1 Design of crystal structure-based 4WJ peptides	139

4.3.2 Peptide synthesis	141
4.3.3 PAGE assay	142
4.4 Conclusions and future work.....	144
CHAPTER 5 – GENERAL CONCLUSIONS AND FUTURE PERSPECTIVES	146
5.1 General conclusions.....	147
5.2 Future perspectives	149
CHAPTER 6 – EXPERIMENTAL	151
6.1 Materials and analysis techniques	152
6.1.1 Solvents and reagents	152
6.1.2 NMR and spectroscopic techniques.....	152
6.1.3 Chromatography techniques	153
6.1.4 MALDI sample preparation	153
6.2 Peptide synthesis.....	154
6.2.1 General peptide synthesis procedure.....	154
6.2.2 General procedure for Kaiser test.....	155
6.2.3 Peptide data	155
6.2.4 Peptide analogue data	158
6.3 Chemical synthesis and characterisation.....	164
6.3.1 Synthesis of acridines.....	164
6.3.2 Synthesis of triazole-linked peptide analogue (20).....	170
6.3.3 Synthesis of alkene-linked peptide analogue (17)	171
6.3.4 Synthesis of dibenzene-linked peptide analogue (19)	173
6.4 Biological Assays	175
6.4.1 PAGE assay	175
6.4.2 FRET	176
6.4.3 Fluorescence polarisation assay	178
6.4.4 Ethidium bromide displacement assay	178
6.4.5 MTS assay.....	179
REFERENCES	180

Abbreviations

2AP	2-aminopurine
3WJ	Three-way junction
4WJ	Four-way junction
Acm	Acetamidomethyl
ALT	Alternative lengthening of telomeres
Boc	<i>tert</i> -Butoxy-carbonyl
bp	Base pair
BRM	Base-recognition motif
C₅₀	Half maximal displacement concentration
CI	Confidence interval
CuAAC	Copper(I)-catalysed azide alkyne cycloaddition
DIC	Diisopropylcarbodiimide
DIPEA	<i>N,N</i> -Diisopropylethylamine
Dmab	4-{ <i>N</i> -[1-(4,4-dimethyl-2,6-dioxocyclohexylidene)-3-methylbutyl]amino}benzyl ester
DNA	Deoxyribonucleic acid
dsDNA	Double-stranded DNA
<i>E. coli</i>	<i>Escherichia coli</i>
EC₅₀	Half maximal effective concentration
EDTA	Ethylenediaminetetraacetic acid
Equiv	Equivalents
EtBr	Ethidium bromide
FAM	6-carboxyfluorescein
Fmoc	Fluoren-9-ylmethyloxycarbonyl
FP	Fluorescence polarisation
FRET	Förster resonance energy transfer
G4	Guanine quadruplex
HBTU	<i>N</i> -[(1 <i>H</i> -benzotriazol-1-yl)(dimethylamino)-methylene]- <i>N</i> -methylmethanaminium hexafluorophosphate <i>N</i> -oxide
HL-60	Human promyelocytic leukaemia cells
HOBt	1-Hydroxybenzotriazole
HPLC	High-performance liquid chromatography

HRMS	High-resolution mass spectrometry
IC₅₀	Half-maximal inhibitory concentration
IR	Infrared spectroscopy
J3	Junction 3
K_D	Dissociation constant
LCMS	Liquid chromatography mass spectrometry
MALDI	Matrix-assisted laser desorption/ionisation
MOC1	Monokaryotic chloroplast 1 protein
MTS	(3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl) 2-(4-sulfophenyl)-2H-tetrazolium)
NMR	Nuclear magnetic resonance
Oxyma pure	Ethyl 2-cyano-2-(hydroxyimino)acetate
PAGE	Polyacrylamide gel electrophoresis
Pbf	2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl
PDB	Protein data bank
PEG	Polyethylene glycol
RP	Reverse phase
RT	Room temperature
SPPS	Solid-phase peptide synthesis
ssDNA	Single-stranded DNA
TAMRA	6-carboxytetramethylrhodamine
TBE	Tris/Borate/EDTA
^tBu	<i>Tert</i> -Butyl
TFA	Trifluoroacetic acid
TFO	Triplex-forming oligonucleotide
TIPS	Triisopropylsilane
Tris	Tris(hydroxymethyl)aminomethane
Trt	Trityl
UV-VIS	Ultraviolet–visible spectroscopy

Chapter 1 – Introduction

1.1 DNA structure

Deoxyribonucleic acid (DNA) is a macromolecule that carries the genetic information necessary for the growth, reproduction and functioning of all living organisms and many viruses. DNA consists of a polymer made up of repeating units of nucleotides. These subunits each contain a 2'-deoxyribose sugar, a phosphate group, and a nitrogenous base. Alternating sugar and phosphate groups form the backbone of a polynucleotide chain, with the 3' and 5' positions of the sugar each linked to separate phosphate groups. The base is connected to the first position of the sugar and can be one of four different types; adenine (A) or guanine (G), that are purine derivatives, or thymine (T) and cytosine (C), that are pyrimidine derivatives (**Figure 1-1**).^{1,2}

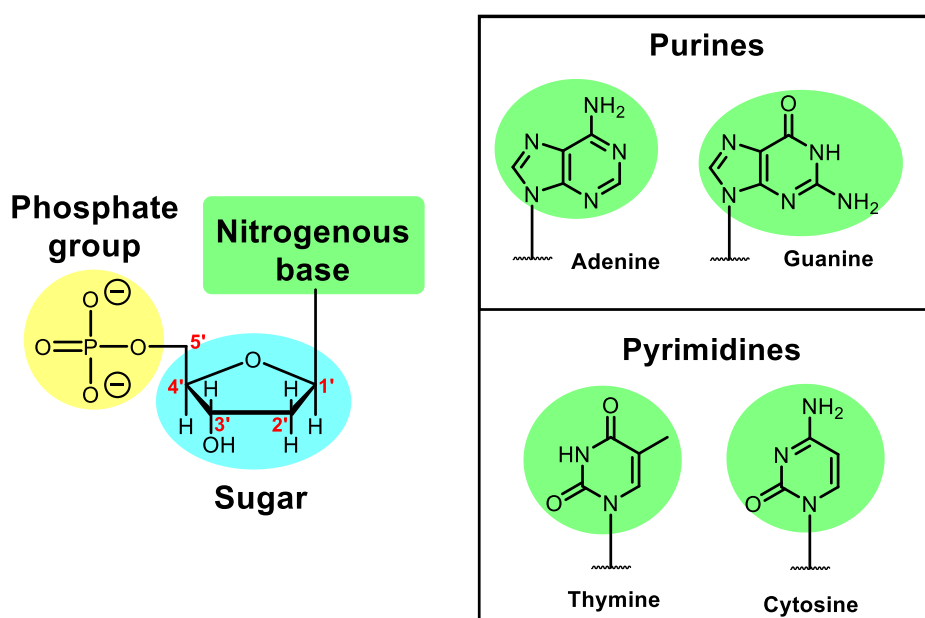


Figure 1-1: Chemical structure of a nucleotide and the four different nitrogenous bases that make up DNA.

DNA usually exists as a double helix, formed by the coiling of two polynucleotide chains around each other (**A, Figure 1-2**). Due to twisting, the double helix contains a major groove, that is wide and shallow, and a narrow and deep minor groove. The structure of the double helix was first solved in 1953 by Francis Crick and James Watson, using X-ray diffraction data obtained by Rosalind Franklin.³ The most common conformation is known as B-form DNA and has right-handed helicity. The two chains are antiparallel, with one chain running from 5'-3' and the other 3'-5' and contain a 2-fold symmetry, perpendicular to the sequences of nucleotides (**B, Figure**

1-2). The helical DNA phosphate-sugar backbones are intertwined, with the bases projecting inwards between the strands. When two separate polynucleotides bind together, base pairs are formed between A-T and C-G. The double helix is stabilised *via* the hydrogen bonding (H-bonding) and base-stacking interactions of the aromatic bases (**B, Figure 1-2**). Each base pair is made up of a purine and a pyrimidine, with A-T forming two H-bonds and C-G forming three. The base-pairs are stacked upon each other 3.4 Å apart along the helix axis, regardless of the identity of the bases. This mean any sequence of A-T and C-G base pairs can be accommodated into the double helix structure.³

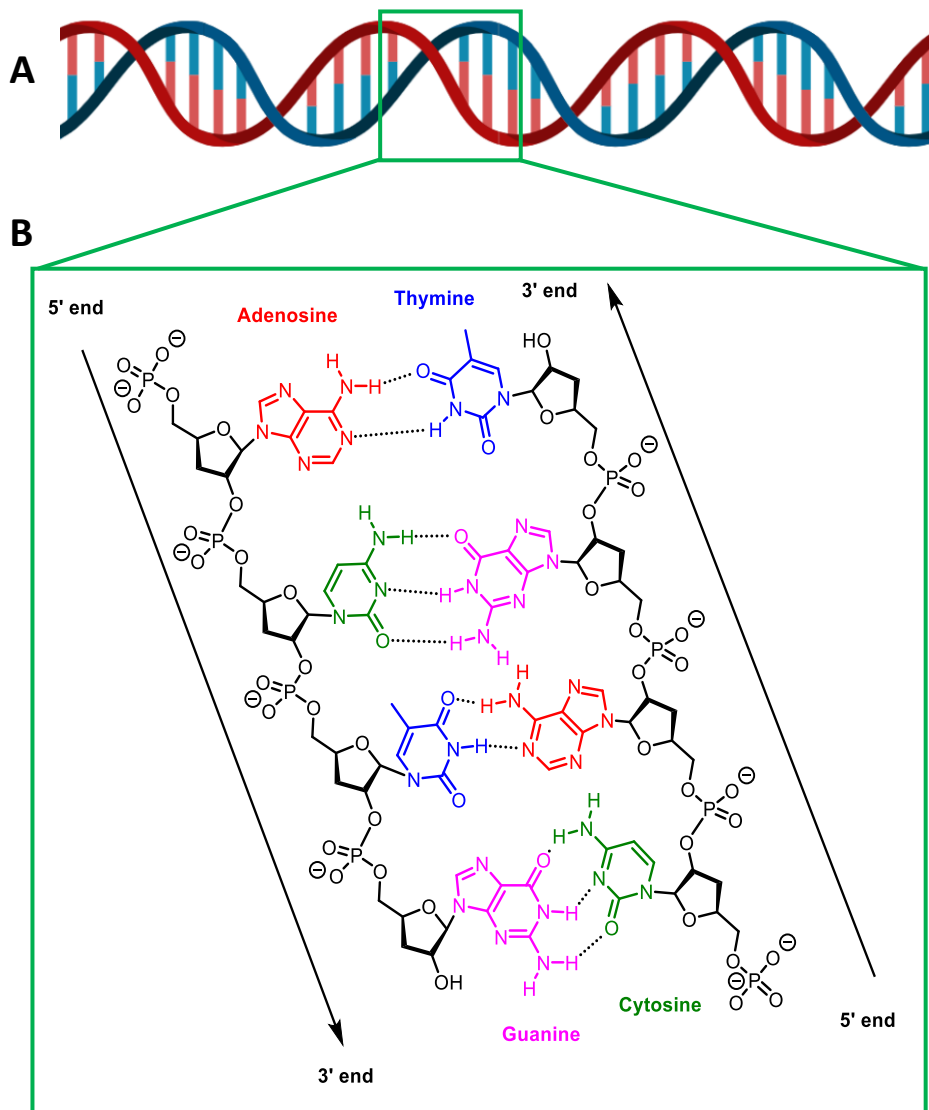


Figure 1-2: The top image shows the overall structure of a DNA double helix, and the bottom image shows a zoomed in section of the DNA double helix, demonstrating base pairing through H-bonding.⁴

The canonical B-form of DNA occurs in humid conditions and in the presence of counterions, such as sodium ions (Na^+), to neutralise the negative phosphate

backbone. However, double-stranded DNA can adopt alternative conformations (**Figure 1-3**).⁵ When relative humidity is reduced to 75%, duplex DNA can change to the A-form conformation.⁶ This remains right-handed but has a wider and flatter conformation, causing the overall length of the DNA to be shorter.⁷ This is because the increase in relative salt concentration causes the negative phosphate backbone to be further shielded, meaning it can become more compact. Specific DNA sequences can form the Z-form conformation of duplex DNA, such as long sections of G-C repeat units. Z-form DNA is left-handed, causing the minor groove to be deep and the major groove to be wide and shallow, the opposite way to B-form DNA.⁸

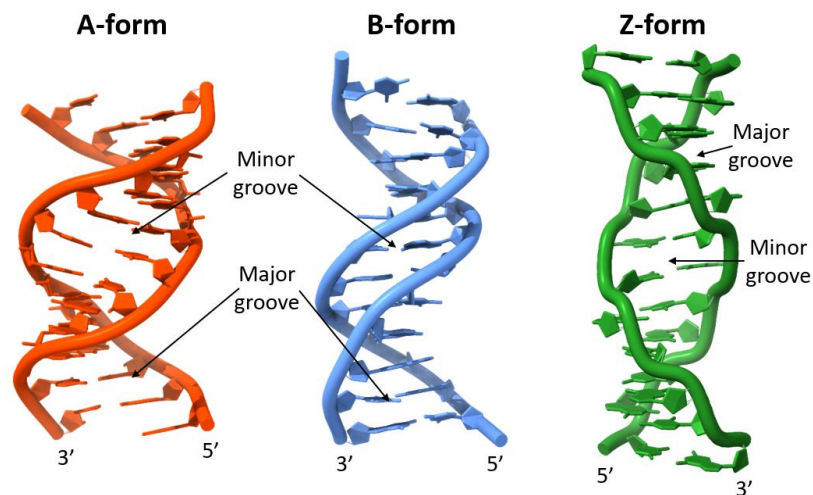


Figure 1-3: Crystal structures of A-form DNA (red, PDB code: 413D⁷), B-form DNA (cyan, PDB code: 6CQ3) and Z-form DNA (green, PDB code: 4OCB).⁵

1.2 Higher-order DNA

DNA primarily exists as duplex DNA, because of the stability of this structure. However, the huge size of mega base DNA, and its high level of flexibility, means that it can also form other three-dimensional shapes, known as higher-order DNA. These include triplex DNA, guanine-rich G-quadruplexes, cytosine-rich i-motifs and DNA four-way junctions (**Figure 1-4**). This structural polymorphism, beyond the canonical B-form DNA, means that these DNA shapes disobey the Watson-Crick rules of hybridisation, that depend on A-T and C-G binding. Non-canonical DNA structures are often formed when DNA is taking part in cellular processes, such as transcription, replication, repair, and recombination. This is because in eukaryotic cells, DNA is supercoiled and stabilised by various proteins, until it is required by the cell. At this point the DNA is unwound and can change into non-canonical structures.⁹

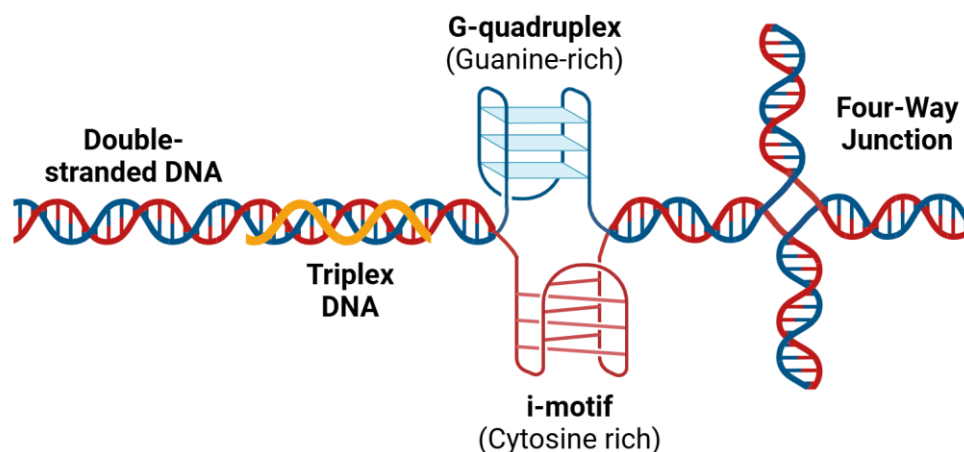


Figure 1-4: Examples of higher-order DNA structures.

The formation of DNA polymorphs depends upon several factors, including specific sequences, lengths, orientations, strand numbers and solution conditions. It is also affected by the presence of proteins and ligands. Due to this, sequence specificity may in fact relate to the formation of DNA structures and not just the actual sequence.⁹ Research into the design of ligands, with the ability to recognise and induce the formation of these structures, has taken off in recent years.^{9,10} Three examples of higher-order DNA structures, their biological roles and the therapeutic potential of ligands that bind to them will be discussed.

Triplex DNA

Triplex DNA is a simple structure that forms when a third strand joins a DNA duplex, binding to the major groove to produce three-strands. The strand adding to the duplex DNA is known as a triplex-forming oligonucleotide (TFO). Although the duplex DNA maintain classical Watson-Crick bonds, the TFO is attached *via* Hoogsteen or reverse Hoogsteen bonds (**Figure 1-5**). The formation of triplex DNA can inhibit transcription and induce site-specific DNA cleavage, with high affinity and sequence specificity.¹¹ These properties have been investigated for anti-gene therapies against cancer, to reduce oncogene expression.¹² However, their development has been hindered by the poor bioavailability,¹¹ although the use of artificial nucleotides could overcome this issue.¹¹ Alternatively, TFOs could offer promise for drug delivery when conjugated to therapeutic compounds, to target specific DNA sequences.¹³

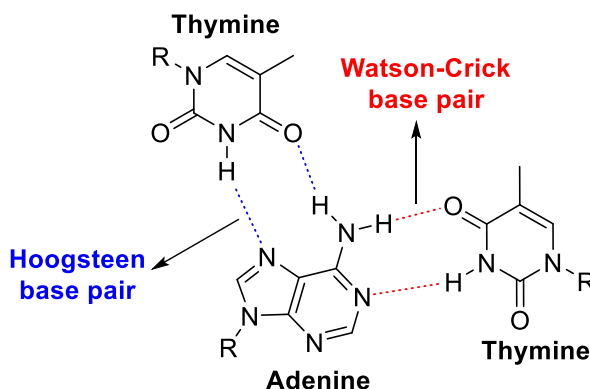


Figure 1-5: Example of a base triplet, formed from a classical Watson-Crick base pair and a Hoogsteen base pair. R indicates the sugar moiety of DNA.¹¹

Guanine quadruplex

The guanine quadruplex (G4) is formed in guanine rich areas of DNA. These cyclic structures contain groups of four guanines, assembled into G-tetrads using Hoogsteen bonds and stacked on top of each other (**A, Figure 1-6**). G4s come in many different polymorphisms, based on intra-/inter-molecular and parallel/antiparallel arrangements. This DNA structure is commonly found in the telomeric region of chromosomes. Telomeres are highly repetitive nucleotide sequences, found at the end of chromosomes, and associated with specific telomerase proteins. The repair and elongation of telomeres has been linked to cancer onset and increased tumour progression. In 1997, Neidle and Hurley developed small molecules capable of stabilising G4s and inhibiting telomerase.¹⁴ This led to an explosion of interest into G4 binding compounds, with several progressing into the clinic for anti-cancer therapies. So far, no compounds have continued past phase II clinical trials, with several compounds withdrawn due to lack of efficacy or issues with toxicity.¹⁵ However, two compounds, CX-5461¹⁶ and QN-302¹⁷ (**B, Figure 1-6**) are showing initial promise in early-stage clinical evaluations.

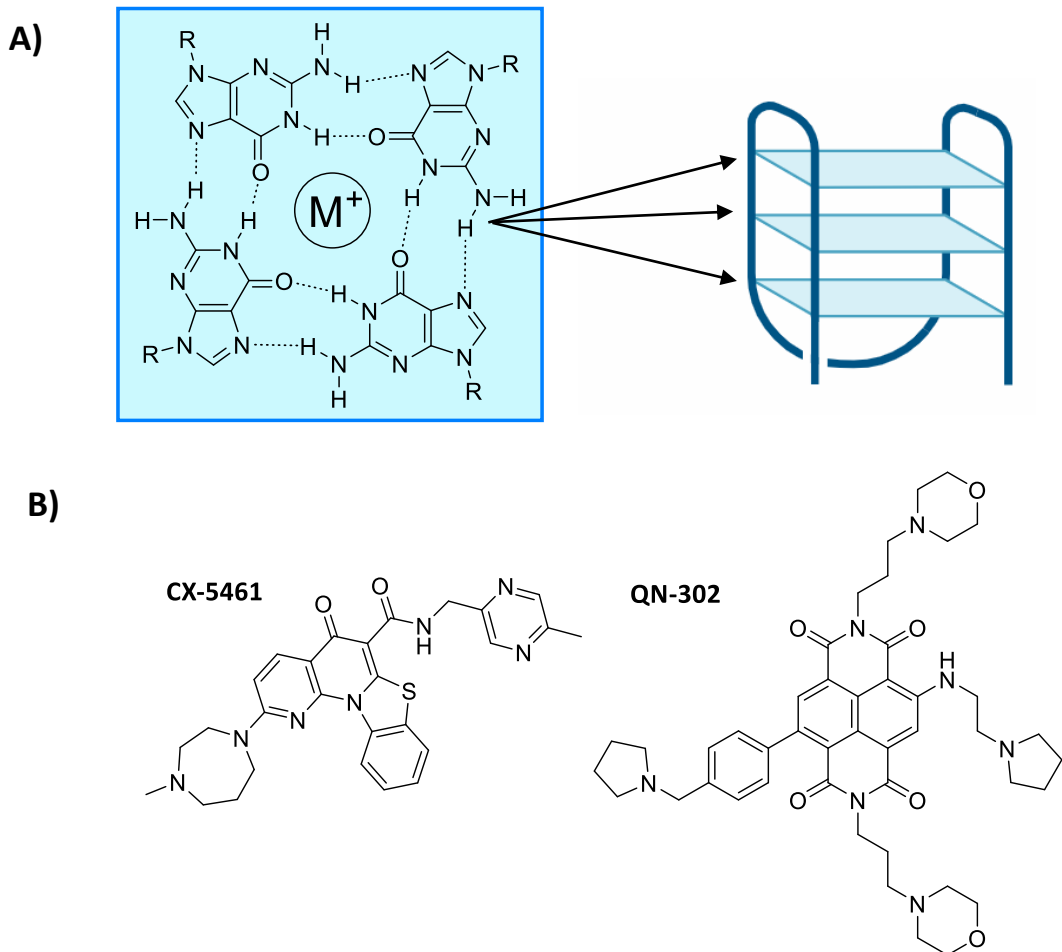


Figure 1-6: (A) The G-tetrad chemical structure, shown stacking in a G4 complex. R indicates the sugar moiety of DNA.¹⁸ **(B)** The chemical structures of G4 targeting ligands.^{16,17}

i-Motif

The DNA i-motif is a four-stranded secondary structure formed in cytosine-rich regions of the genome, complementary to the guanine-rich sequences that make up G4s. The i-motif is composed of two parallel duplexes that are intercalated in an antiparallel orientation *via* hemiprotonated cytosine–cytosine (C–C⁺) base pairs.¹⁹ These tetrameric structures stack into a compact, folded configuration (**Figure 1-7**), generally under acidic conditions.²⁰ However, stable i-motifs have been identified at neutral pH, that may play a regulatory role in transcription and oncogene expression. This includes in telomeric regions of DNA, alongside their complementary G4 sequences.²¹ Ligands have been developed that target i-motifs, for the study of their structure *in vivo* and biological functions. Although no ligands specifically targeting i-motifs have progressed to the clinic thus far, it has been found that the promising

preclinical compound QN-302 (**B, Figure 1-6**), stabilises G4s whilst destabilising i-motifs at the same time, providing synergistic anti-cancer effects.²²

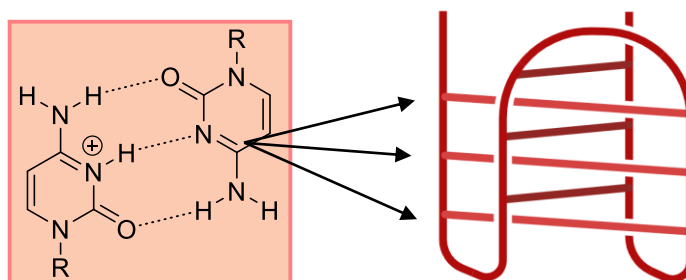


Figure 1-7: The hemiprotonated C-C⁺ base pair, shown stacking in an i-motif. R indicates the sugar moiety of DNA.¹⁹

The study of non-canonical DNA structures, such as triplex DNA, G4s, and i-motifs, is crucial for understanding their role in cellular processes and their potential as therapeutic targets. Ongoing research holds promise for the future of anti-cancer therapies that exploit these unique DNA forms.

1.3 The DNA four-way junction

The focus of this project will be on the higher-order structure, the DNA four-way junction (4WJ). 4WJs are formed by the mutual exchange of DNA strands belonging to four helices. This can allow DNA to be rearranged, making 4WJs important intermediates in DNA processing and repair.²³ 4WJs are an integral part of homologous recombination, a process that allows the cell to access homologous DNA, with shared sequence identity. This is used for chromosomal segregation during meiosis to create genetic diversity, for the processing of stalled replication forks, and the repair of DNA double-strand breaks.²³ Homologous recombination is used universally by eukaryotes, bacteria, and viruses. Many different pathways are known, with the DNA double-stranded break model given as an example (**Figure 1-8**). Firstly, the gap caused by the break (**a, Figure 1-8**) is extended by exonucleases (**b, Figure 1-8**) before the broken end of the DNA is paired with homologous DNA (**c, Figure 1-8**). This results in the formation of a 4WJ intermediate where the DNA crosses over. Next, branch migration causes the 4WJ to be moved along the DNA as repair occurs. When this is complete, a second 4WJ is formed. Lastly, resolution of both 4WJs takes place *via* cleavage across the centre of the junctions (**d, Figure 1-8**), by 4WJ resolvase enzymes, to reform two separate double-stranded DNA molecules (**e, Figure 1-8**). Based on the direction of junction cleavage, different

double-stranded DNA (dsDNA) molecules can form, allowing for increased genetic diversity.²⁴

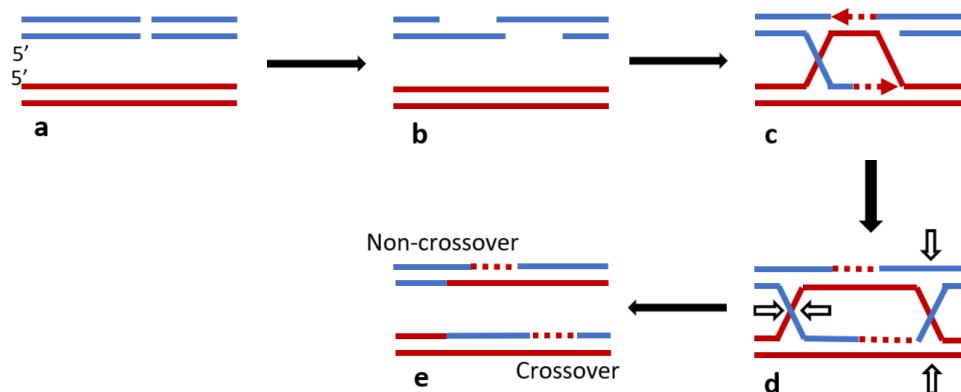


Figure 1-8: DNA double-strand break repair model: **(a)** A cut is made across a dsDNA molecule **(b)** The gap is extended by exonucleases **(c)** One of the broken strands invades a homologous dsDNA molecule, forming a 4WJ, and DNA repair takes place **(d)** Once the DNA repair is complete a second 4WJ is formed **(e)** The 4WJs are resolved to give either the crossover or non-crossover configuration.²⁵

Site-specific recombination is a process that plays a crucial role the life cycle of yeast, bacteria, and bacteriophage.²⁶ 4WJs are involved in site-specific recombination, when two DNA segments with target sequences are rearranged by specific recombinase enzymes. Unlike homologous recombination, site-specific recombination requires minimal homology between the DNA molecules. An example of the site-specific recombination mechanism, using tyrosine site-specific recombinases, is shown (**Figure 1-9**). This begins with the interaction of four recombinase molecules with the DNA target sequences (**a, Figure 1-9**), followed by DNA cleavage from the attack of nucleophilic tyrosine residues on the scissile phosphates (**b, Figure 1-9**). Strand exchange then occurs when the phosphotyrosine bond is attacked by the free OH from the other DNA molecule, to form a 4WJ intermediate (**c, Figure 1-9**). The 4WJ subsequently isomerises (**d, Figure 1-9**), to allow the nucleophilic tyrosine residues to attack adjacent DNA strands and complete the final strand exchange (**e, Figure 1-9**). Recombination is finished when the phosphotyrosine bond is attacked by the free OH on the same DNA (**f, Figure 1-9**). Each step of this reaction mechanism is reversible, meaning that the same duplex DNA segment can be exchanged back to its original sequence without recombination occurring. However, host-cofactors are often used to shift the equilibrium of the reaction and drive recombination.²⁷

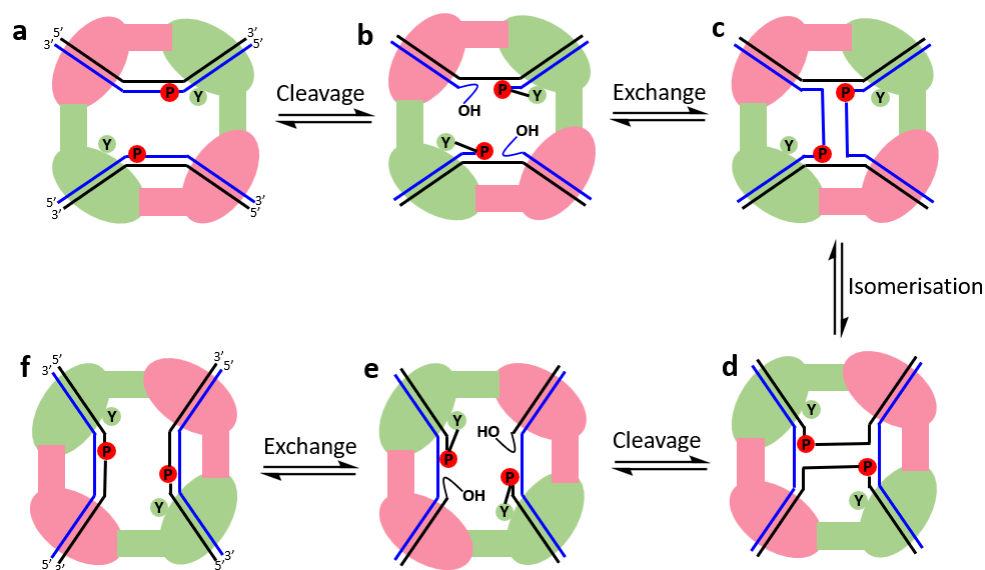


Figure 1-9: Site-specific recombination mechanism using tyrosine site-specific recombinases. Showing duplex DNA strands (black and blue lines), active conformer recombinase molecules (green) and inactive conformer recombinase molecules (pink). **(a)** The interaction of four recombinase molecules with two duplex DNA molecules. **(b)** DNA cleavage from the attack of nucleophilic tyrosine residues (Y) on scissile phosphates (P). **(c)** Strand exchange when the phosphotyrosine bond is attacked by the free OH from the other DNA molecule, to form a 4WJ intermediate. **(d)** Isomerisation of the 4WJ. **(e)** DNA cleavage from the attack of nucleophilic tyrosine residues on the adjacent DNA strands. **(f)** Recombination is complete when the phosphotyrosine bond is attacked by the free OH on the same DNA molecule.²⁷

1.3.1 Targeting 4WJs in cancer

Cancer is used to describe a collection of diseases characterised by the uncontrolled growth and spread of abnormal cells in the body.²⁸ This is caused by alterations in signalling and metabolism within the cancerous cells.^{29,30} Cancer is a complex disease that can be caused by several factors, including genetic mutations, exposure to environmental carcinogens, lifestyle choices, infections, and hormone expression.²⁸ This is one of the most destructive diseases in humans, resulting in millions of deaths across the world every year.³¹ There are >100 genetically distinct conditions that cause cancer to occur in many different areas of the body, meaning a variety of treatments are necessary.²⁸

Duplex DNA has been targeted extensively by anticancer drugs, such as with nitrogen mustards, DNA-reactive agents, and antimetabolites.³² However, these are general cytotoxic agents, with minimal discrimination for cancerous cells over healthy dividing cells. This results in serious side effects and highlights the importance of developing DNA-targeting anticancer drugs that are more selective, such as compounds that bind to specific higher-order DNA structures including the 4WJ.³²

A feature of cancer cells is their increased frequency in DNA double-strand breaks that can cause genomic instability. This is caused by the higher proliferation rates of cancer cells, their increased metabolism, and mutations affecting the efficiency of DNA repair, leading to an accumulation of damaged DNA.³³ The homologous recombination pathway is the most used for the repair of DNA double-strand breaks (**Figure 1-8**) and requires the appropriate resolution of the DNA 4WJ intermediate, for accurate strand exchange and the continued proliferation of tumour cells.³⁴ Many current anticancer treatments increase rates of DNA double-strand breaks, including radiotherapy and DNA-targeting anticancer drugs, making the homologous recombination pathway a good target for combination therapies.³⁵ Two examples showing the possible anticancer effects of targeting 4WJs in different homologous recombination pathways will now be discussed.

Telomeres were previously mentioned in **section 1.2** as promising anticancer targets. The alternative lengthening of telomeres (ALT) is a homologous recombination pathway that plays a critical role in telomere repair in human glioblastomas and osteosarcomas cancer cells. 4WJs have been shown to be present in telomeric DNA *in vitro* and could be targeted to inhibit the ALT pathway. This could offer a new anticancer approach for the treatment of ALT-positive cancers.³⁶

In eukaryotic cells the tumour suppressor protein p53 acts as a cell cycle checkpoint, resolving unwanted DNA structures, and repairing DNA damage, before the cell proceeds to the next part of the cell cycle.³⁷ Higher expression levels of p53 have been correlated with increased levels of homologous recombination *in vivo*. This may show that p53 limits the amount of recombination taking place in a cell and monitors any errors caused by the process. p53 binds strongly to the 4WJ, with high specificity, and enables junction resolution. This prevents 4WJs from proceeding to the next stage of the cell cycle and causing further DNA damage. Targeting these 4WJs could prevent their resolution by p53 and result in a build-up of DNA damage in cancer cells.³⁸

The development of 4WJ-targeting drugs holds promise as a new anticancer therapeutic approach.^{36,39,40} By focusing on 4WJs, particularly in pathways like ALT and homologous recombination involving p53, it may be possible to develop more selective treatments for cancer cells than to healthy ones, when compared to duplex DNA targeting drugs.

1.3.2 Targeting 4WJs in bacterial infections

The introduction of the first antibiotics in the 20th century led to a belief that pathogenic bacteria could be defeated.⁴¹ However, it was quickly discovered that bacteria can build-up resistance to these drugs through a variety of mechanisms, including drug inactivation by chemical modification and active efflux of the drug out of the cell.⁴² Antibiotic resistance is a major course of global morbidity and mortality and is becoming worse over time.⁴³ Current antibiotics inhibit a variety of processes within the bacteria, including protein synthesis, cell wall synthesis, and metabolic pathways.⁴⁴ Duplex DNA has also been targeted by fluoroquinolones, that cause DNA breaks and stop DNA from unwinding by inhibiting bacterial DNA gyrase and topoisomerase IV.⁴¹ Resistance to these drugs highlights the need to develop new types of antibiotics with different biological targets.

In bacterial cells, crucial DNA repair *via* homologous recombination is carried out by nucleases of bacteria, such as the RuvABC resolvosome. This complex is present in almost all bacterial species and resolves 4WJs formed during single-/double-stranded DNA breaks.⁴⁵ RuvABC is made up of three proteins that carry out distinct roles (**Figure 1-10**). RuvA is a DNA-binding protein that forms a tetramer and has higher affinity for the 4WJ. RuvB is an ATPase that forms two hexamers either side of the junction centre and moves the RuvAB complex along the duplex DNA for branch migration and DNA unwinding (**A, Figure 1-10**). Lastly, RuvC is a resolvase that forms a dimer and cleaves the 4WJ for effective recombination (**B, Figure 1-10**).⁴⁶ By targeting the DNA 4WJ, resolution of the junction by the RuvABC could be prevented and lead to an increase in DNA damage. This could be a promising new antibacterial approach if 4WJ binders can be developed that specifically target the 4WJ-RuvABC complex over 4WJ-nuclease complexes present in eukaryotic cells.

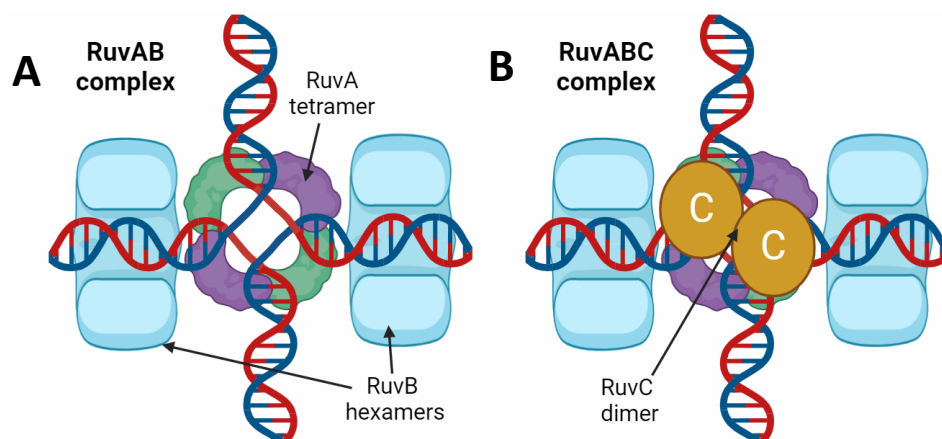


Figure 1-10: Model of RuvAB and RuvABC complexes bound to 4WJ DNA (red and dark blue). **(A)** The RuvAB complex, with a RuvA tetramer (green and purple) bound to the centre of the 4WJ and two RuvB hexamers (pale blue) bound to the duplex DNA. **(B)** The RuvABC complex, with the RuvC resolvase dimer (yellow) bound to the centre of the 4WJ.⁴⁶

As mentioned in **section 1.3**, tyrosine site-specific recombinases (T-SSR) are a family of enzymes that carry out site-specific recombination (**Figure 1-9**). T-SSRs are commonly found in bacteria and very rarely in higher eukaryotes, making them ideal targets for antibiotic development. T-SSRs have various functions, including controlling gene expression, increasing genetic diversity, and making sure chromosomal segregation is carried out correctly. These proteins typically bind to palindromic DNA recognition sites, such as in the well-studied Cre recombinase, found in bacteriophage P1, that binds to the *loxP* DNA recognition sequence. The *loxP* site is 34 bp long, containing two 13 bp palindromic sequences at each end and a central 8 bp spacer that decides the direction of recombination (**Figure 1-11**).²⁷ By targeting the 4WJ intermediate in the T-SSR pathway, this would prevent the Cre recombinase from being able to carry out its functions and lead to a build-up of unresolved 4WJs. The specificity of DNA sequences involved in site-specific recombination mean that the 4WJ intermediate contains certain base pairs. It could be interesting to explore if compounds could be developed that have both sequence and structural selectivity for these unique 4WJs.

5' – ATAAC**TCGTATAATGTATGCT**ATACGAAGTTAT – 3'

3' – TATTGAAGCATAT**TACATACGATATGCTTCAATA** – 3'

Figure 1-11: *LoxP* DNA sequences required for Cre-*loxP* recombination. Palindromic sequences (blue) and spacer sequences (red).²⁷

Targeting 4WJs offers a novel and promising strategy for developing new antibiotics, particularly in the fight against antibiotic-resistant bacteria. The RuvABC

resolvasome and T-SSRs are critical in bacterial DNA repair and recombination processes, making them attractive targets. By inhibiting the resolution of 4WJs, these essential pathways, specific to bacterial cells, could potentially be disrupted.

1.3.3 4WJ conformations

The structure of the DNA 4WJ was first proposed by R. Holliday in 1964, providing its former name of the Holliday Junction.⁴⁷ Subsequent studies of the 4WJ were described by several groups during the 1980's, using asymmetrical DNA sequences to create immobile 4WJs. This prevented the migration and resolution of the junction, for in-depth study of their structures.^{48–53}

The 4WJ was found to adopt two main conformations - an open conformation, where a square-planar shape is formed in the middle of the junction, and a more compact X-stacked form (**Figure 1-12**). The open conformation occurs at low salt concentrations, caused by the repulsion between negatively charged phosphate groups on the DNA backbone, and has a 4-fold symmetry. The X-stacked conformation occurs in solutions with high salt concentrations, where two pairs of DNA strands stack coaxially at a 60° angle to each other and have a 2-fold symmetry. Two different X-stacked isoforms are possible, depending on which pairs of DNA strands stack on top of each other (**Figure 1-12**). The duplex strands come together in an antiparallel orientation and one DNA strand from each pair forms a U-turn at the junction. In this X-stacked conformation the negative phosphates are shielded by the positive salt cations in solution, allowing for more compact shape.⁵⁴

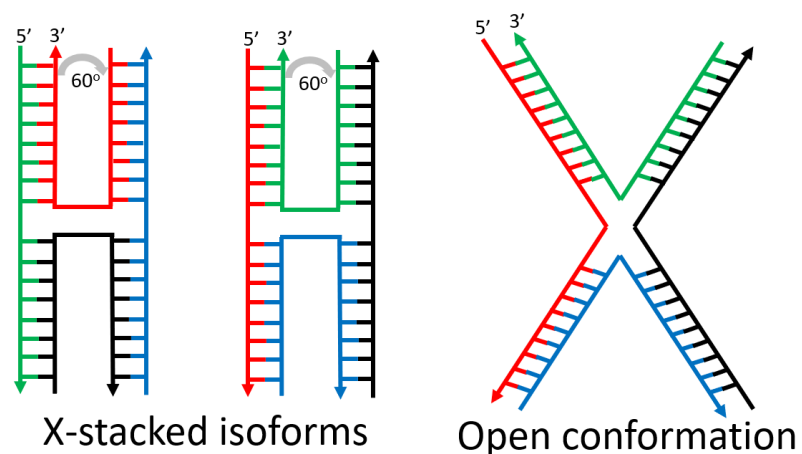


Figure 1-12: 4WJ X-stacked and open conformations, with both X-stacked isoforms shown.

Generally, one of the X-stacked isoforms is preferred. For example, using single-molecule Förster resonance energy transfer (FRET) experiments, an immobile 4WJ that has been studied extensively, known as Junction 3, showed a 77.4% preference for isoform II (**Figure 1-13**).⁵⁵ The difference in prevalence for each isoform reflects the energetic differences in the central stacking interactions of each structure, that are also affected by nucleotides 2 – 3 bases from the junction centre. The favoured isoform can affect the binding of junction-resolving enzymes, determining where cleavage occurs and the final recombination products. During the exchange of two isoforms, an open 4WJ conformation intermediate is transiently formed. In the presence of 50 mM Mg^{2+} ions the rate of conformer exchange is significantly reduced, due to destabilisation of the open transition state in the high salt environment.²³

Figure 1-13: The structure and sequences of the two Junction 3 4WJ isoforms.⁵⁵

mostly found to contain NCC or ANC trinucleotide motifs at the junctions' centre, showing the formation of stable X-stacked 4WJs is sequence dependent.⁵⁴

Proteins that bind to 4WJs, during processes like homologous recombination and site-specific recombination, can be classified by whether they bind to the X-stacked or open conformations (**Table 1-1**). Generally, enzymes bind to the 2-fold symmetrical X-stacked 4WJ as dimers, and to the 4-fold symmetrical open 4WJ as tetramers. However, there is also a group of homodimers that recognise the X-stacked 4WJ but induce the DNA into the open 4WJ upon binding. Most proteins that bind or induce the open 4WJ have some level of sequence specificity, whilst those binding to the X-stacked 4WJ are relatively non-sequence specific. This may reflect the sequence dependence of the X-stacked conformation to form in the first place, allowing for indirect selectivity for X-stacked binding proteins.⁵⁴

Table 1-1: 4WJ binding enzymes (N represents any nucleic acid base).⁵⁴

Enzyme	Organism	Oligomeric state	Sequence specificity
Open 4WJ substrate			
Cre	Bacteriophage P1	Tetramer	<i>LoxP</i> sequence
RuvAB	<i>E. Coli</i>	Tetramer	Damaged DNA
λ Integrase	Bacteriophage λ	Tetramer	TNNNTTNNNTNNNANNAANNNG
Induced open 4WJ substrate			
RuvC	<i>E. Coli</i>	Dimer	(A/T)TT(G/C)
CCE1	<i>S. Cerevisiae</i>	Dimer	ACTA
Ydc2	<i>S. Pombe</i>	Dimer	CT and/or TT
X-stacked 4WJ substrate			
T4 nuclease VII	Bacteriophage T4	Dimer	None
Hjc	<i>P. furiosu</i>	Dimer	None
Hje	<i>S. solfataricus</i>	Dimer	None

Overall, the conformational state of the 4WJ, influenced by environmental factors like salt concentration and DNA sequence, plays a crucial role in the binding and function of recombination and repair proteins. Information known about X-stacked and open 4WJs can be used to guide the design of ligands that specifically interact with one of these conformations.

1.4 4WJ-binding ligands

1.4.1 Peptides

1.4.1.1 WKHYN Y

Segall's group pioneered the area of peptides targeting 4WJs, with an evolving set of peptides (**Figure 1-14**) and mechanistic investigations into their binding and biology. These peptides were synthesised with neutral C-terminal amide groups, to avoid the negative charge from a carboxylic acid group, which could repel against the negative charge of the DNA phosphate backbone. Additionally, the majority of naturally occurring bioactive peptides contain the post-translational modification of C-terminal amidation.⁶²

After conducting a large-scale library screen of 300 million hexapeptides, they found that the peptide WKHYN Y was able to stabilise the 4WJ intermediate involved in Phage λ integrase (Int)-mediated recombination. This occurred when the junction was in its open conformation, preventing proteins from binding to, and resolving, the 4WJ. This compound was used as a biological tool when studying bacterial pathways. Although the peptide was unable to inhibit the overall growth of bacteria, they did find WKHYN Y inhibited the bent-L pathway resulting in the build-up of 4WJ intermediates.⁶³

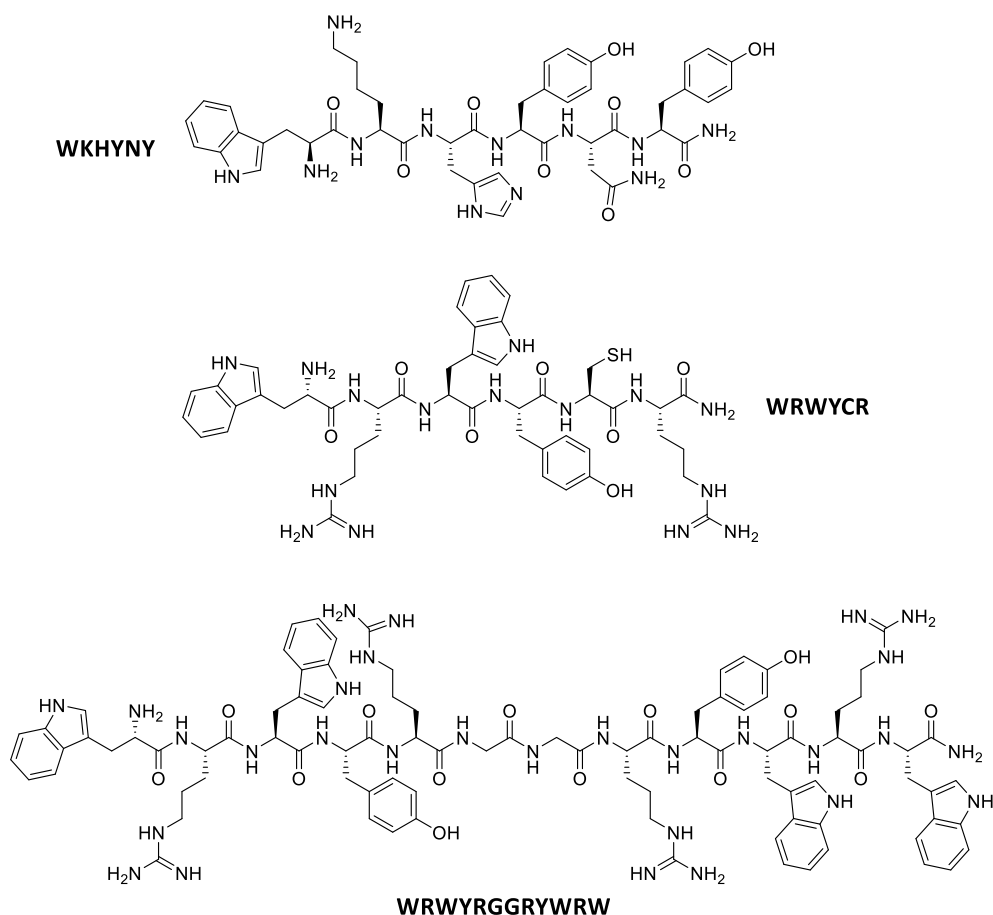


Figure 1-14: Chemical structures of peptides developed by Segall's research group, synthesised with C-terminal amidation.

To further understand the binding of WKHYNY to the 4WJ, a crystal structure was partially resolved, although the exact binding of the ligand remained unclear. This showed that the ligand bound to the centre of the square planar 4WJ pocket. This was further demonstrated by replacement of adenine bases at the centre of the 4WJ with 2-aminopurine (2AP) residues, a fluorescent analogue. Peptide binding was shown to quench the fluorescence of the solvent exposed residues at the centre of the 4WJ, due to the stacking of aromatic rings.⁶⁴

1.4.1.2 WRWYCR (1)

An additional screening, focused on the inhibition of the excision pathway for site-specific recombination in Int, led to the discovery of peptide **1**, that had nanomolar potency against both free and protein bound 4WJ.⁶⁵ This peptide was strongly inhibited by the presence of Mg^{2+} , showing the peptide binds to the open conformation. **1** was found to have more than 50 times the potency in the excision pathway when compared to previously made hexapeptides.⁶⁶

Further analysis of **1** binding showed that the formation of a dimer was required for the peptide to be active. This was confirmed when adding the reducing agent dithiothreitol drastically reduced the potency of the peptide, by preventing disulfide bridge formation between the cysteine residues. The replacement of the cysteine residues with lysine also drastically reduced the activity of the peptide. To confirm the formation of a dimer, two monomers of peptide **1** were irreversibly cross-linked using bismaleimido-hexane to bind to both cysteine sulfide groups (**Figure 1-15**). Although the amount of 4WJ trapping reduced slightly, the addition of dithiothreitol did not affect activity.⁶⁵

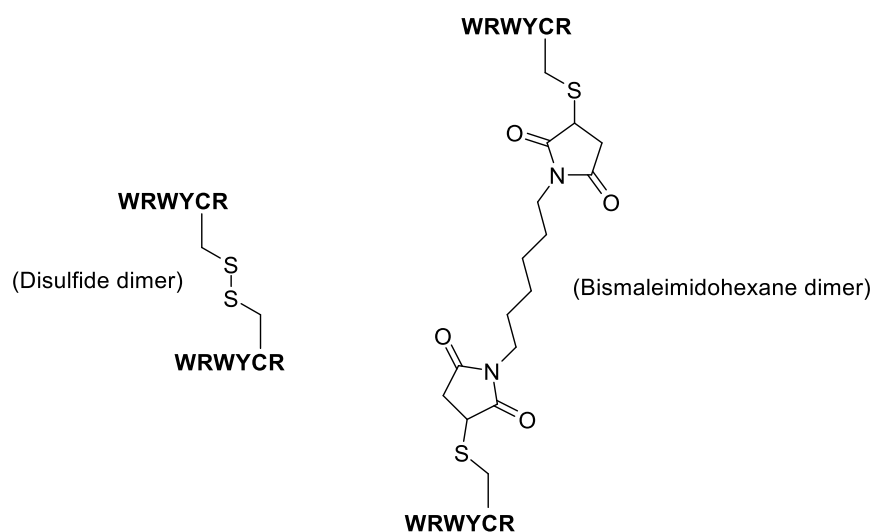


Figure 1-15: Structure of peptide **1** dimers with different cross-linking groups.

Modelling of the active dimer to a Cre-Lox 4WJ complex predicted that the dimer forms both intramolecular and intermolecular stacking interactions with DNA bases, and H-bonds with both deoxyribose and base residues (**Figure 1-16**).⁶⁶

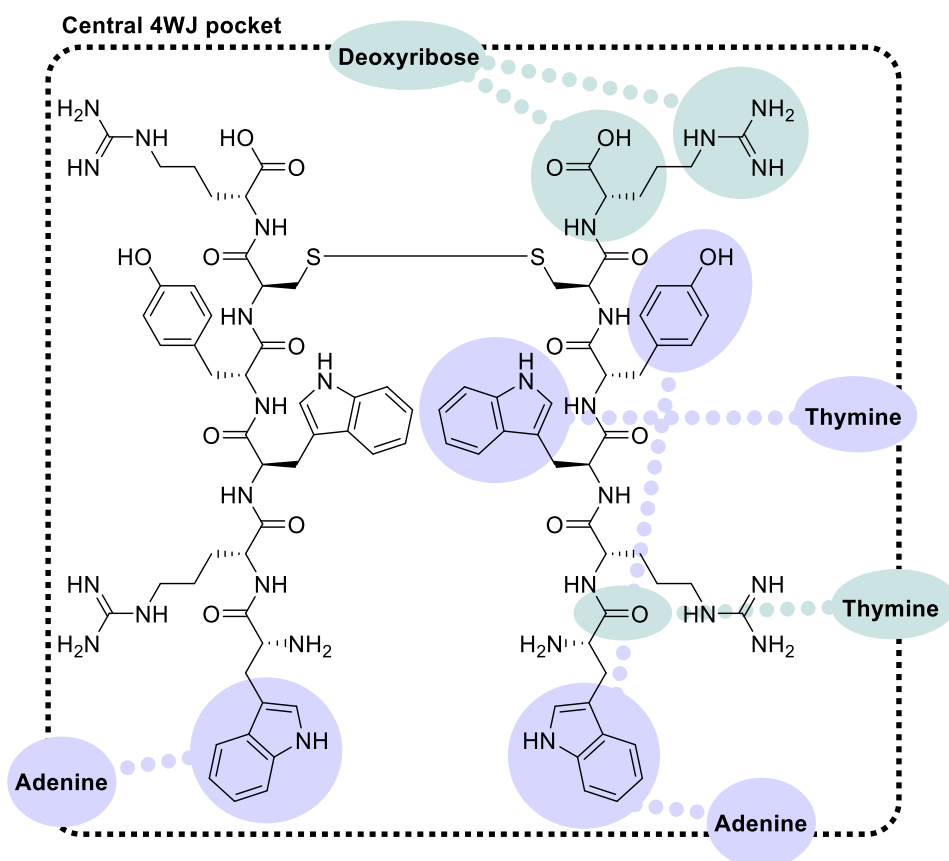


Figure 1-16: Predicted binding interactions from molecular modelling of the Cre-Lox 4WJ and the peptide **1** dimer, showing π - π interactions (purple) and H-bond interactions (green).⁶⁶

Extensive studies have been carried out into the biological effects of **1**. The peptide was tested against the other three Int recombination pathways and found to have low nM IC₅₀ values against the bent-L and integration pathways and a higher nM value against the straight-L pathway.⁶⁶ When testing **1** against the Cre-Lox recombination complex, found in phage P1, the peptide was found to have a slightly reduced potency, when compared to the excision pathway of Int. Other peptides screened in the library of hexapeptides were found to have similar potency against both the Int pathways and Cre-Lox pathway.⁶⁶

Peptide **1** was able to inhibit the binding of several proteins, including the RecG helicase and RuvABC complex of *E. coli*. This peptide acted as a competitive inhibitor of RecG-4WJ binding in competition assays using gel mobility shift.⁶⁶

The selectivity of **1** for 4WJs over three-way junction (3WJ) replication fork substrates was demonstrated to be 4- to 9-fold higher. This was confirmed by the higher concentration of **1** required to quench the fluorescence of 2-aminopurine (2AP) bases at the centre of the 3WJ. In addition, **1** has a 240-fold preference for 4WJs over non-specific double-stranded DNA. Throughout these selectivity assays it

was shown that **1** had no apparent sequence specificity but was highly structurally selective for 4WJs.⁶⁷

Peptide **1** was found to have antimicrobial activity against both gram-positive and gram-negative bacteria, with increased activity against gram-positive bacteria.⁶⁸ The D-isoform of peptide **1** (**1-D**), that contains right-handed amino acids, rather than the left-handed L-amino acids in the original peptide, was also tested. It was shown that both peptides **1** and **1-D** were able to enter the bacterial cell and are capable of *in vivo* inhibition of site-specific recombination.⁶⁸ Both forms were shown to inhibit the growth of *Salmonella* cells in mammalian cells, reducing the number of bacterial cells recovered after 24 h by 100 – 1000 times.⁶⁸ These peptides were shown to be non-toxic to murine cells at 50 µM, with no increase in cell death observed.⁶⁹

When comparing the two isoforms, peptide **1-D** was shown to have higher overall antibacterial activity, with the growth of the bacteria *Bacillus subtilis* fully inhibited after 8 – 10 h by 5 µM of **1-D** compared to 8 µM of **1**.⁷⁰ This higher activity of the D-isoform could result from increased resistance to cellular peptidase degradation.⁷⁰ The ability for these peptides to exhibit antimicrobial effects make it a promising starting point for developing novel antibiotics targeting DNA repair pathways.

1.4.1.3 WRWYRGGRYWRW (2)

In 2013, the Segall group investigated making a single chain analogue of peptide **1**.⁴⁵ The idea was to develop a compound that does not have to form a dimer in the 4WJ, with the disulfide bond being susceptible to reduction in the cellular environment. 10 – 12 amino acid analogues were designed, resulting in the development of peptide **2**. This peptide is of a similar molecule weight to the dimer, has the same number of positive charges and avoids the need for dimerisation by the removal of cysteine groups.

Peptide **2** was able to bind to the open conformation of the 4WJ with a similar affinity to **1**, had comparable potencies in excision pathways and remained active in the presence of dithiothreitol. Overall, it was found that the D-isoform of peptide **2** (**2-D**) had more effective 4WJ binding *in vivo*,⁴⁵ in a similar manner to **1-D** compared to **1**. When testing antimicrobial activity, **2-D** had an equal or greater potency against both gram positive and gram-negative bacteria when compared to **1-D**. However, **2-D** had increased non-specific DNA binding. This may be caused by the increased number of positive charges and aromatic rings resulting in **2-D** interacting more with

other forms of DNA. The non-specific binding is reflected by the higher level of toxicity against eukaryotic cells. Despite this, when testing the peptide against eukaryotic cells the concentration necessary for bacterial growth inhibition remained sufficiently low to be non-toxic.⁴⁵

The peptides **1-D** and its single-chain analogue **2-D** have also been tested for their cytotoxicity against cancer cells. In MTT cell viability assays against several cancer cell lines, **1-D** was found to be active at 50 – 200 μM and **2-D** from 10 – 50 μM . Although the dodeca-peptide was active at lower concentrations, its higher level of nonspecific binding makes it hard to confirm whether this results from 4WJ binding. Co-treatment of the cells with known chemotherapeutic agents that cause DNA damage was also tested. Synergistic activity was found when combining peptide **1-D** with drugs that act on the S phase of the cell cycle, such as etoposide, doxorubicin, or hydroxyurea.⁷¹ This signifies that the peptide similarly acts on the S phase, affecting the synthesis of new copies of DNA required for replication. This was further shown when **1-D** was combined with docetaxel, a drug that inhibits the M phase of the cell cycle, where an additive effect was seen. This is presumably due to docetaxel blocking the cell cycle in the M phase, for any cells that have not already been stalled by the peptide in the S phase. Further studies are required to understand the full mechanism of cancer cell death after peptide treatment; however, accumulation of double-stranded breaks *in vitro* indicate 4WJ stabilisation may be involved and may lead to cell cycle arrest.⁷¹

Cancer cells have been shown to have higher membrane potentials and increased concentrations of anionic residues on phospholipid membrane proteins, when compared to healthy cells.⁷² Both peptides **1-D** and **2-D** contain hydrophobic and basic residues that may make it easier for them to pass through cancer cell membranes more efficiently than healthy cells. This has previously been shown to be the case with cell penetrating peptides as a strategy to help with the drug delivery and shows how the 4WJ-targeting peptides are promising starting points for drug development.⁷¹

1.4.1.4 Cyclic peptides

Following the discovery of 4WJ-targeting peptides, an investigation to produce cyclic peptides with increased specificity for 4WJs compared to other DNA isoforms was disclosed. In two separate studies, in 2003⁷³ and 2004,⁷⁴ small cyclic peptide libraries

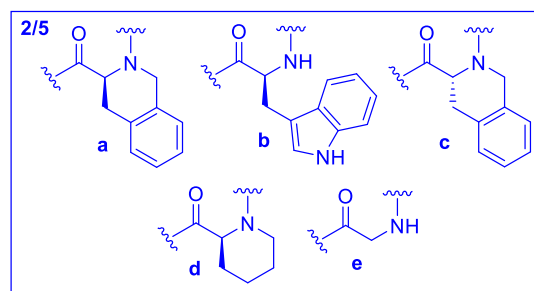
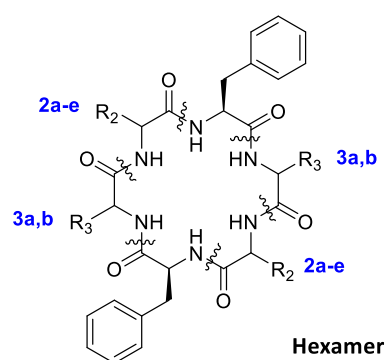
containing hexa- and octamers were produced. The design of the cyclic peptides was focused around a 2-fold symmetry to mimic the symmetry of the 4WJ.

The initial screening of eight cyclic peptides showed that three of the compounds were able to trap 4WJs in the open conformation. These compounds contained the aromatic amino acids **2a-3a**, **2a-3b** and **2c-3b** (**Figure 1-17**). The ability to trap the 4WJs was assessed by the blocking of RuvC endonuclease, in *E. coli*, from eliminating 4WJs, when at a 1 μ M concentration of cyclic peptide, as confirmed by gel electrophoresis.⁷⁵

For the second screen, an effort was made to introduce hydrophilic, polar residues, to increase solubility and permeability. Eighteen compounds were synthesised with eleven tested against RuvC. It was found that all compounds containing tyrosine residues were able to trap 4WJs, showing the importance of this residue for 4WJ binding. It was also shown that both hexa- and octa-peptides could be accommodated, demonstrated by the more effective 4WJ trapping of the octamer with amino acids **4a-5a-6a** when compared to the similar hexamer **4a-5d-6a** (**Figure 1-17**).

When the cyclic peptides were tested against both gram negative and positive bacteria, they were found to have no effect on cell growth. This could result from the compounds remaining too hydrophobic to cross bacterial cell walls, with the introduction of further hydrophilic residues required for anti-bacterial activity.⁷⁵

First generation cyclic peptides



Second generation cyclic peptides

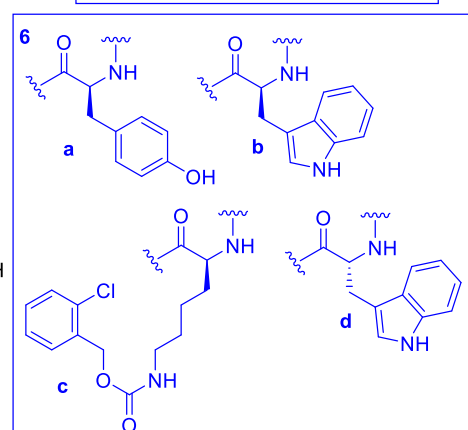
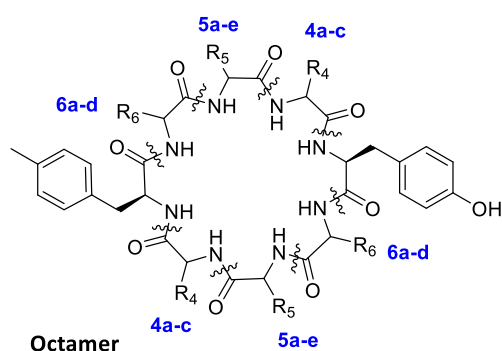
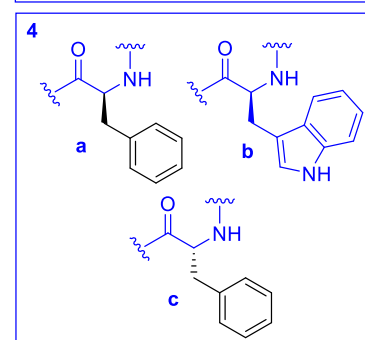
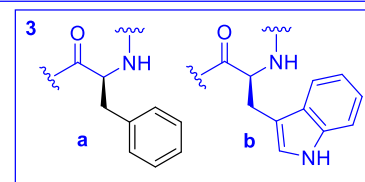
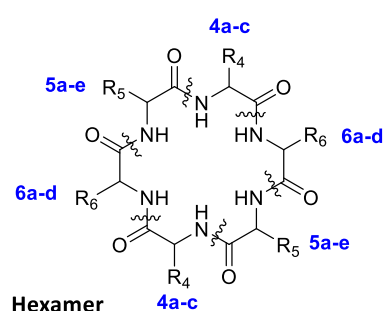


Figure 1-17: Chemical structure of 4WJ-binding cyclic peptides.⁷⁵

1.4.2 Metal complexes

Early studies into the binding of metal complexes to an immobile 4WJ were carried out by Kallenbach *et al.*⁷⁶ They found that methidiumpropyl-EDTA-Fe(II) (MPE-Fe^{II}) (Figure 1-18) was able to cleave a 4WJ within two residues of the central pocket, with intercalation occurring next to the branch point.

Well-known DNA binding dyes, that appeared to bind in a different fashion in presence of 4WJs, were characterised using a combination of competitive binding, DNA cleavage and footprinting assays. It was suggested that the Stains-All dye and

propidium iodide (**Figure 1-18**) bind to the centre of the 4WJ.⁷⁷ On the other hand, tetrapyrridyl porphyrins with varying metal groups and/or axial ligands, were suggested to bind near the branch point, although the binding mode has not yet been fully elucidated.⁷⁸

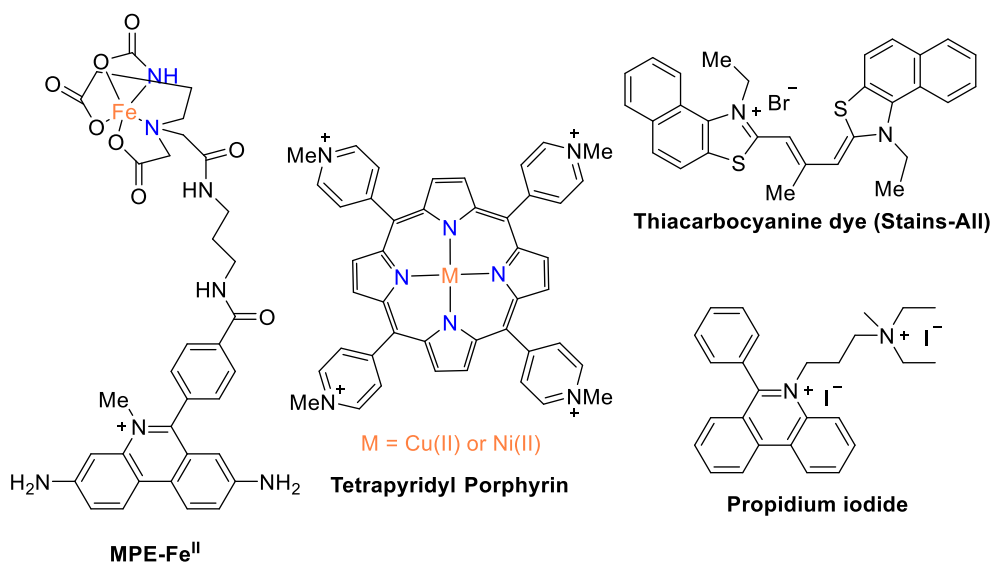


Figure 1-18: Chemical structures of metal complexes and DNA-binding dyes, investigated by KallenBach *et al.* for 4WJ binding.^{75,77 79}

1.4.2.1 Platinum (II) complex

Extensive research has been carried out into the ability of trans-Pt complexes to bind to various DNA motifs. This started with the development of cisplatin, a commonly used chemotherapeutic that targets duplex DNA. Following this Hannon and co-workers developed binuclear Fe/Ni/Ru complexes capable of forming a 3WJ with DNA 6-mer oligomers of 5'-d(CGTACG)-3'.^{80,81} In 2019, Bonnet *et al.* discovered a mononuclear platinum (II) complex, [Pt(H₂bapbpy)](PF₆)₂ (**A**, **Figure 1-19**) that was able to cause four oligomers of this same DNA sequence, 5'-d(CGTACG)-3', to crystallise into a 4WJ-like motif (**B+C**, **Figure 1-19**).⁸² They found that two molecules of the Pt complex bound to form an X-stacked 4WJ, sitting at the centre of the junction. These complexes adopted a distorted square planar structure when binding that were relatively flat, and most likely adopt a co-planar arrangement in relation to each other. The Pt complexes were found to form π - π interactions with the DNA bases and are stabilised by coulombic interactions between the positively charged Pt atom and negatively charged DNA backbone. The two NH groups may be involved

in H-bonding, although the longer length of the NH...O contacts (2.6 – 4.1 Å) indicate a weaker interaction.

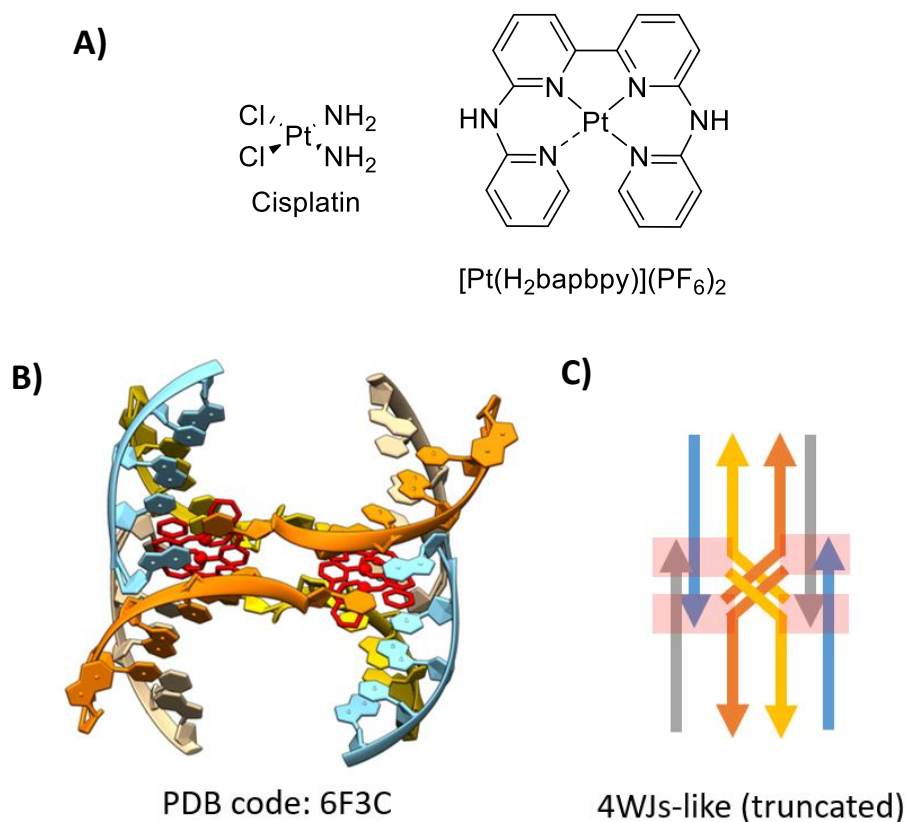


Figure 1-19: **(A)** Chemical structures of DNA binding Pt complexes. **(B)** Crystal structure of two $[\text{Pt}(\text{H}_2\text{bapbpy})](\text{PF}_6)_2$ molecules (red) bound to a 4WJ, formed by 5'-d(CGTACG)-3' DNA sequences. **(C)** Schematic of crystal structure.⁸²

When comparing the antiproliferative activity of this Pt complex to cisplatin in cancer cell lines, it was found that the new Pt complex was more potent than cisplatin and retained low μM potency against cisplatin resistant cells. However, this could result from interactions with dsDNA, with further analysis required to look at the selectivity of the compound.⁸²

1.4.2.2 Organometallic pillarplexes

In addition to their work with 3WJs, in 2023 Hannon's group discovered the first synthetic molecules that bind to the open 4WJ conformation.⁸³ These are Au and Ag tetracationic organometallic pillarplexes that have aryl faces around a cylindrical core (**Figure 1-20**). The Au pillarplex was used for further studies, with the Ag pillarplex showing decreased solution stability. Molecular dynamics simulation demonstrated that the Au pillarplex sits in the centre of the 4WJ, with the aryl groups making a π -surface to interact with the DNA. Experimental studies showed the Au

pillarplex preferentially bound to the 4WJ over other DNA structures and in a low salt concentration, when the 4WJ was present in the open conformation. The Au pillarplex did also show a high level of binding to Y-shaped forks, although optimisation of the compound, to fully fill the 4WJ centre, could improve selectivity. When testing the Au pillarplex in cancer cell lines it was shown to have comparable cellular uptake and antiproliferic activity to cisplatin. This could be a promising starting point for developing selective 4WJ binding molecules, for both therapeutic use and as tool compounds.⁸³

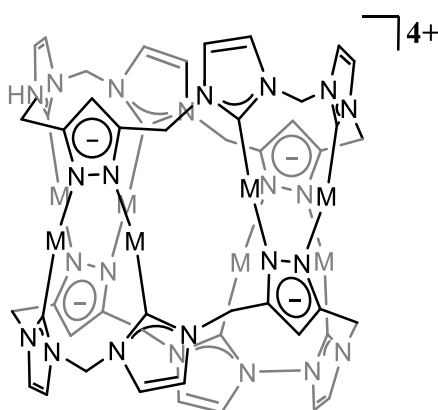


Figure 1-20: Chemical structure of an organometallic pillarplex.⁸³

1.4.3 Small molecules

1.4.3.1 Acridines

Acridines possess aromatic systems and basic/cationic groups and have been studied extensively for their DNA intercalating abilities. These compound that have been known for their therapeutic properties since the 19th century.⁸⁴ Interest in the use of acridine molecules in relation to 4WJs began in 1996, when Cook *et al.* investigated their use as probes for the DNA structure, alongside the acridine isomer phenanthridine. This involved using various rigid linkers to bind one or two intercalating acridine or phenanthridine groups (**Figure 1-21**). All compounds, except for the acridine dimer **APEPA**, showed non-specific binding between the 4WJs centre and duplex arms. In contrast, **APEPA** increased the sensitivity of the 4WJ DNA to cleavage by deoxyribonuclease 9 bp's from the centre of the junction. This may indicate that the compound was selectively binding nearby and changing the conformation of the 4WJ.⁸⁵

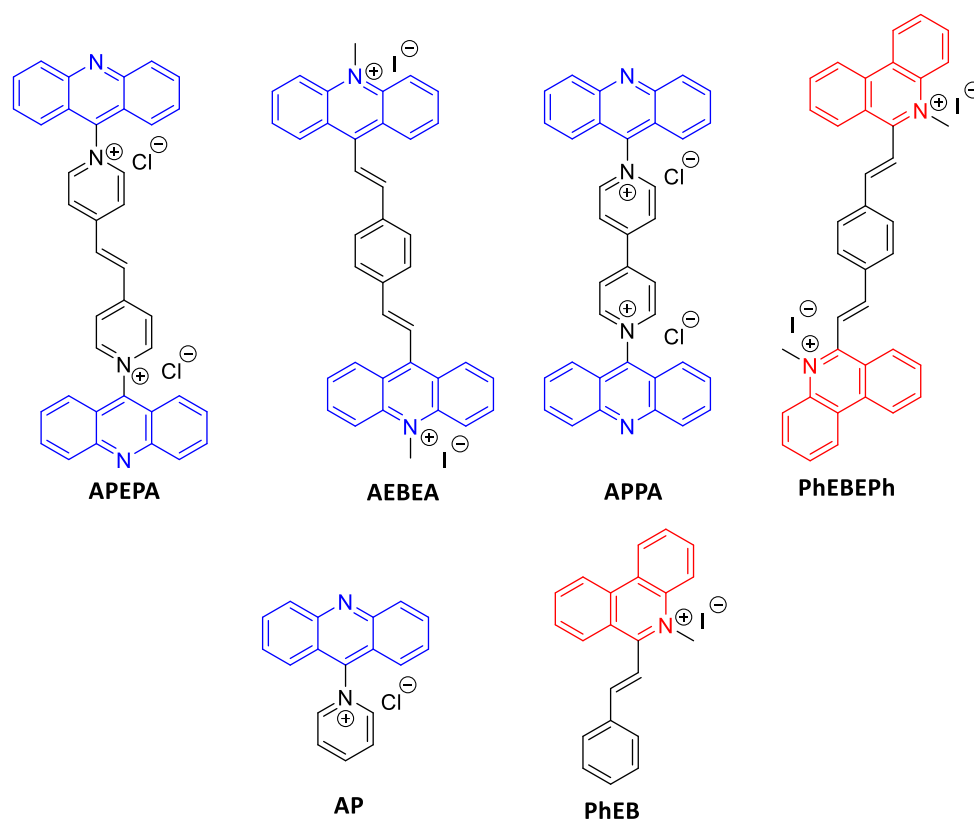


Figure 1-21: Chemical structures of acridine (blue) and phenanthridine (red) monomers and dimers, containing rigid linkers (black).⁸⁵ The compounds are named after the chemical groups they contain; A = acridine, B = benzene, E = ethylene, P = pyridinium and Ph = phenanthridinium.

Dimeric acridines were also investigated by Searcey *et al.*⁸⁶ Compared to the previous structures, the linker is more flexible and the acridines carry an additional *N,N*-dimethylaminoethyl carboxamide in position 4. This type of substituted acridines are well-known cytotoxic compounds that act as topoisomerase inhibitors and have been shown to be active against tumour cells.⁸⁷ C6 acridine **3 (A, Figure 1-22)**, with a six-carbon flexible linker, was shown to bind to a 4WJ in the central region. A crystal structure of the compound bound to a 4WJ (**B, Figure 1-22**) clarified this non-covalent binding mode across the centre of the 4WJ when in the X-stacked conformation. A detailed study of the crystal structure showed that the two acridine chromophores bound with a 2-fold symmetry through the centre of the 4WJ. The adenine bases were flipped out to the major groove and replaced by the acridine groups. Extensive π - π interactions occur between the acridines and DNA bases. Additionally, H-bonding was observed between the amide hydrogen of the acridine sidechain and the oxygen of a cytosine base. This results in the cytosine being misaligned with its base-paired guanine, causing H-bonding between the base pairs to become longer than usual.⁸⁸

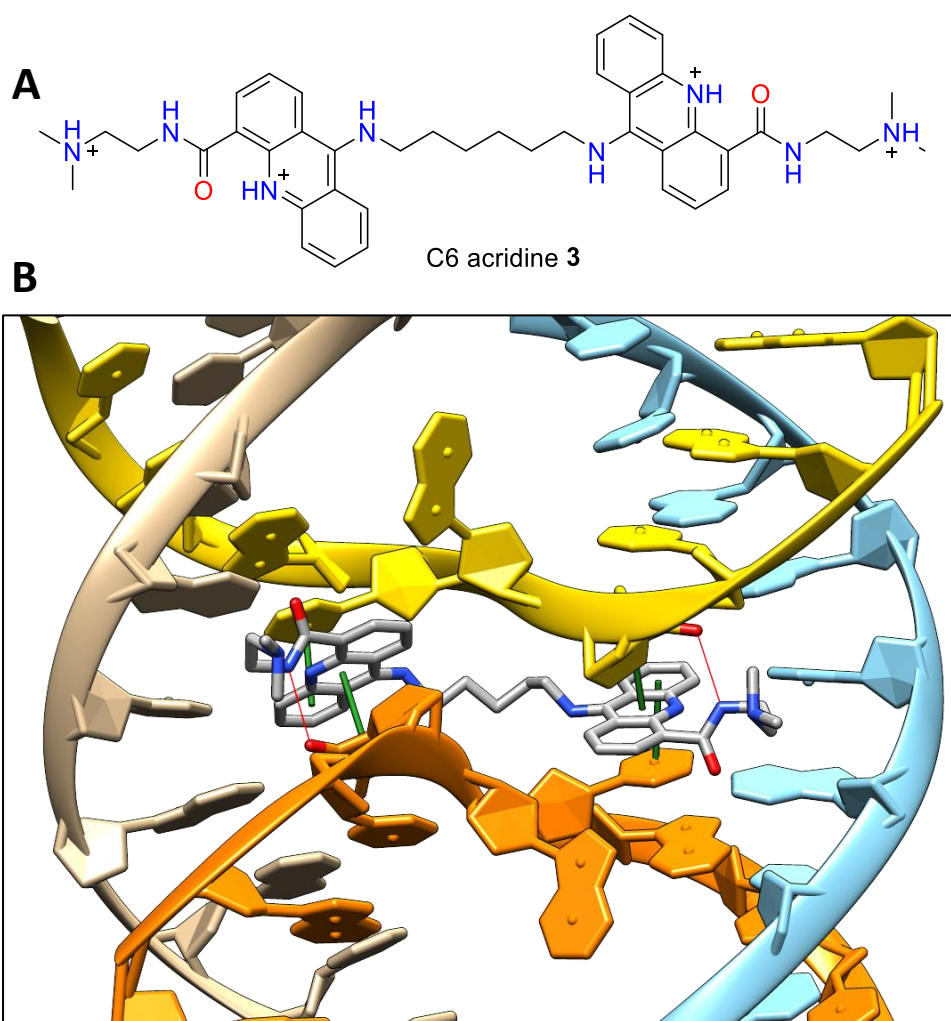


Figure 1-22: (A) Chemical structure of C6 acridine **3**. **(B)** Crystal structure of **3** (grey) bound to 4WJ DNA (yellow, orange, cyan and beige). H-bonds (red) and π - π interactions (green) are shown (PDB code: 2GWA).⁸⁶

A different set of acridine dimers (**Figure 1-23**), with a short flexible chain and two aromatic groups between the monomers, was obtained *via* click chemistry. Of these, **BA3** and **BA4** displayed 4WJ binding. Interestingly, compound **BA3**, is the only compound in the group which does not display antiproliferative activity, despite its ability to bind to both 4WJs and dsDNA.⁸⁹

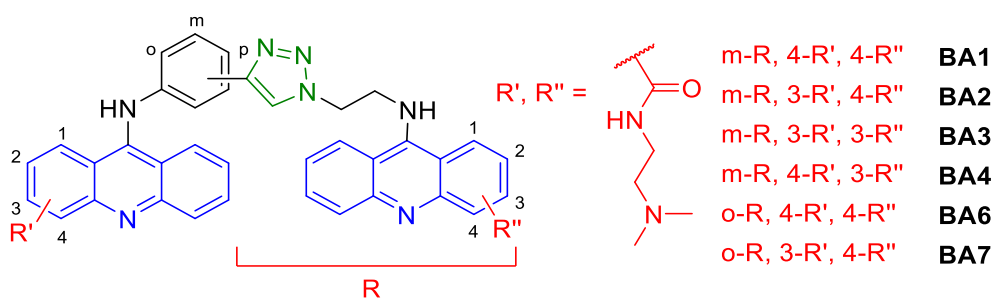


Figure 1-23: Chemical structures of triazole-containing acridine dimers.⁸⁹

A key advance in 4WJ binding was the discovery of compound **RTA** (**Figure 1-24**). Normally, folding of an open 4WJ into its X-stacked form requires high temperature annealing in the presence of Mg^{2+} . **RTA**, unique among compounds discovered to date, not only targets 4WJs, but induces folding of the X-stacked form at RT. This effect was confirmed *via* gel electrophoresis and circular dichroism. The compound was also found to be active against tumour cells, although this may result from the low level of selectivity of the compound for 4WJ and duplex DNA.⁹⁰

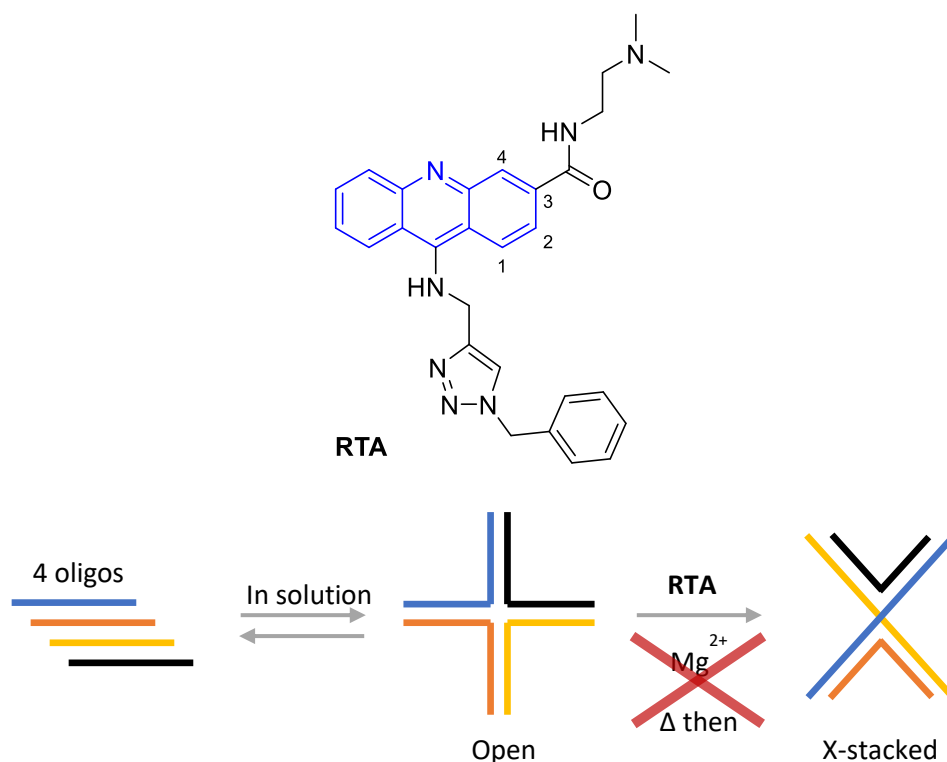


Figure 1-24: Chemical structure of **RTA** and folding of an X-stacked 4WJ through annealing in the presence of Mg^{2+} or in the presence of **RTA**.⁹⁰

Following attempts to synthesise compounds with the same ability to fold 4WJ, three new structures with similar features, **A14**, **A15** and **A21**, were identified (**Figure 1-25**). Similarly, to **RTA**, they all have the dimethylamino-side chain in position 3, and the distance between the aromatic groups is six atoms – 2 atoms shorter than the original C6 acridine **3** (6 carbons + 2 nitrogen atoms). It is interesting to note that compounds **A14** and **A15** displayed antiproliferative activity, differently from **A21**, which is inactive.⁸⁹

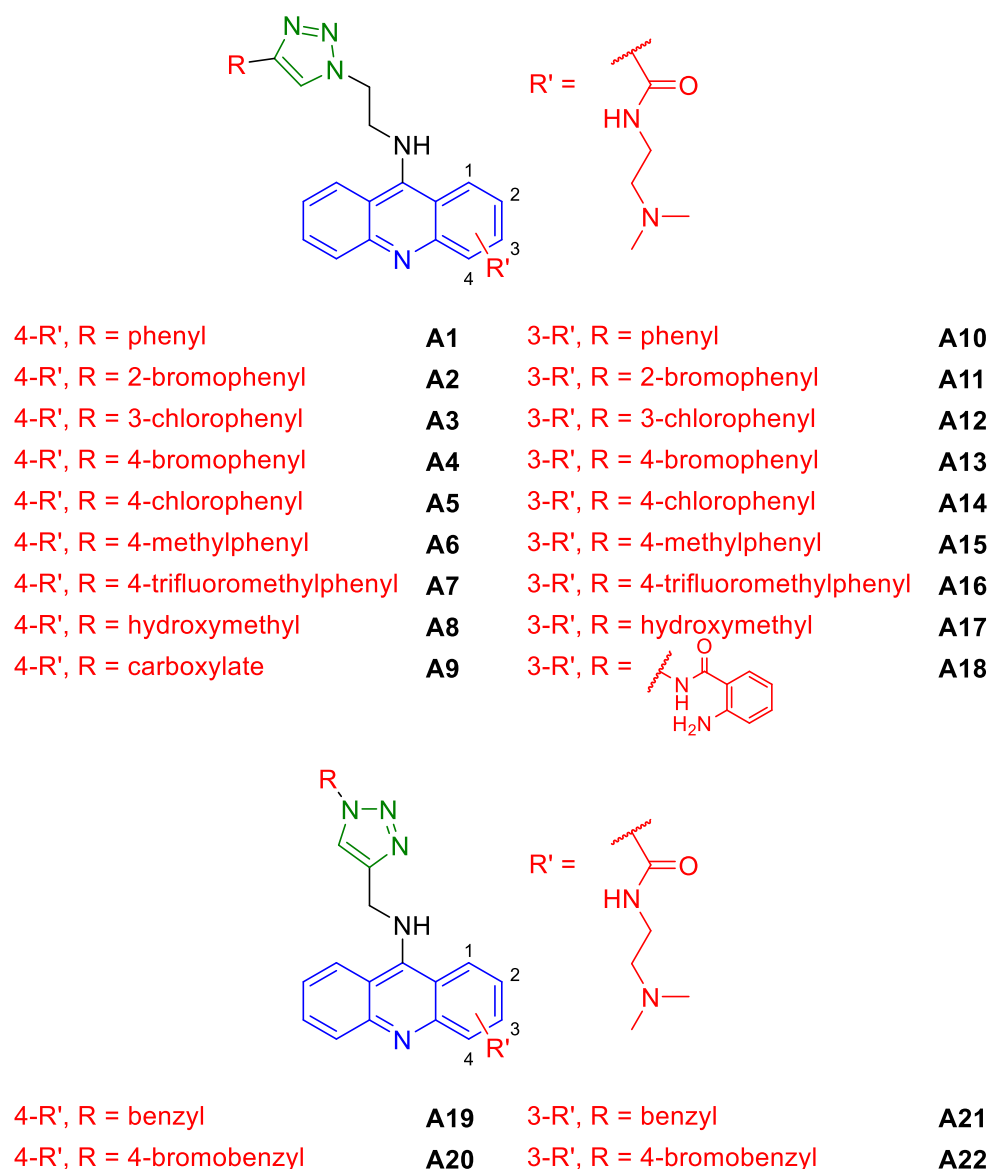


Figure 1-25: Chemical structures of clicked acridine monomers.⁸⁹

Modulating 4WJ folding at RT is a fundamental achievement and could allow regulation and folding in cells and material applications. The major limitation of acridines is their affinity for dsDNA, and advancement in their selectivity or discovery of new classes of compounds with the same folding ability would be a welcome addition in the field of 4WJ targeting.

1.4.3.2 Triaminotriazine-acridine

Recently, the study of acridines in relation to 4WJ was continued by Chien *et al.* when looking into a known DNA-binder; the triaminotriazine-acridine conjugate (**A**, **Figure 1-26**). This compound targets CTG trinucleotide repeats containing T:T mis-matches, a target that has been found to cause various neurological diseases. With the

mechanism of binding remaining unknown, the team crystallised the compound with three CTG trinucleotides containing T:T mis-matches and found that a noncanonical 4WJ was formed (**B, Figure 1-26**). Instead of having the standard X-stacked conformation, they found that both duplex strands performed a unique U-turn when facing head-to-head with the other duplex (**C, Figure 1-26**). This is formed by the binding of two acridine-conjugate molecules sit at the centre of two overlapping -GCTGC- 5-mers. The conformation adopted by the ligands shows the acridine and triaminotriazine stack over each other, with the acridine further stacking with a guanine base. The insertion of the triaminotriazine moiety between DNA bases causes the flipping out of thymine residues, resulting in the DNA backbone being able to bend more than 90° to form a 4WJ structure.⁸⁴

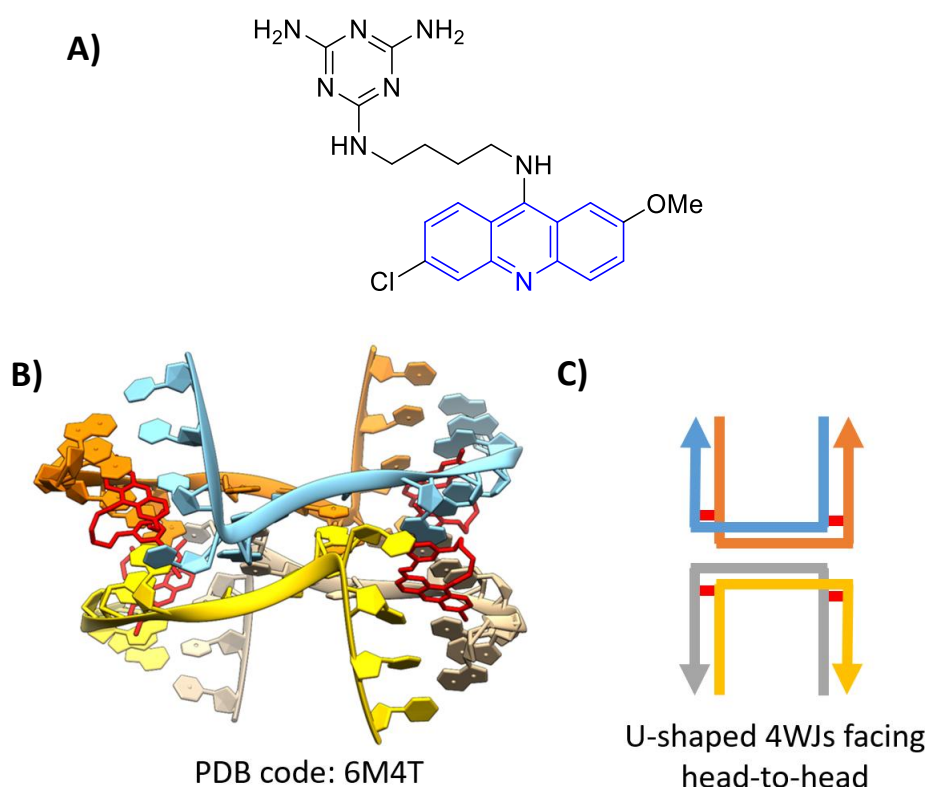


Figure 1-26: **(A)** Chemical structure of triaminotriazine-acridine conjugate, with the acridine moiety highlighted in blue. **(B)** Crystal structure of two triaminotriazine-acridine conjugate molecules (red) bound to a 4WJ, formed by CTG repeats. **(C)** Schematic of crystal structure.⁸⁴

1.4.3.3 VE-822

In late 2021, Wang *et al.* established that the known drug **VE-822** (**Figure 1-27**), an ATR inhibitor in clinical trials for cancer with the commercial name Berzosertib, also promotes the formation of 4WJs.³⁶ It was discovered in a screening of 160

compounds from a focussed library of DNA damage and repair compounds. From carrying out a CD assay and docking studies they confirmed that the compound converts a 4WJ to its X-stacked form. VE-822 has single digit μM potency for 4WJ formation and low levels of binding to double and single-stranded DNA. The group showed that **VE-822** had significant anti-proliferation effects in siATR cells, which could indicate that stabilisation of 4WJs is contributing to the anti-cancer effects. The compound was shown to reduce levels of homologous recombination, triggering an increase in DNA damage, leading to potential apoptosis.³⁶

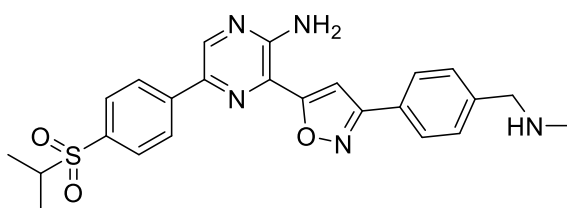


Figure 1-27: Chemical structure of **VE-822**.

It has also been shown that **VE-822** can bind to telomere-4WJs, albeit with 3-fold less potency when compared to 4WJs only. This could be of interest due to the alternative lengthening of telomeres (ALT) pathway containing 4WJ intermediates, as discussed in **section 1.3.1**. VE-822 has been shown to cause telomeric DNA damage by assembling telomere-4WJs. When using **VE-822** in combination with the anticancer agent doxorubicin against osteosarcoma cells, known to rely on the ALT pathway, they found that VE-822 was able to sensitise the cells to doxorubicin treatment.³⁶ This could lead to a new approach for targeting telomeres.

1.4.3.4 Cyclic thioureas and guanidines

In addition to their work on peptides, Segall's group investigated the development of small molecules for 4WJ inhibition. Initially the group carried out the screening of 9 million compounds with low molecular weights. They identified a compound containing an *N*-methyl aminocyclic thiourea scaffold (**T1**, **Figure 1-28**) able to trap 4WJs during site-specific recombination. However, due to low permeability, the compound did not exhibit antibiotic activity.⁹¹

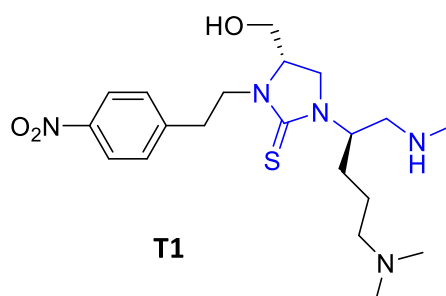


Figure 1-28: Chemical structure of **T1**, with the *N*-methyl aminocyclic thiourea scaffold highlighted in blue.⁹¹

Following this initial study, a more focused library was screened, involving compounds closer to the larger size and 2-fold symmetry of the hexapeptide **1** dimer (**Figure 1-15**). From this, a series of compounds containing a pyrrolidine bis-cyclic guanidine scaffold were identified that affect DNA restriction (**G1 – 4, Figure 1-29**). Among these, only **G3** bound to protein-free 4WJ with nanomolar affinity, showing a K_D of 300 nM versus 12.5 nM for **1**. It also exhibited weaker inhibition of RecG-mediated DNA repair, with an IC_{50} of 0.94 μ g/mL compared to 0.16 μ g/mL for **1**.⁶⁸ This may be caused by the reduced stability of the small molecules when binding to 4WJs, caused by the absence of strong H-bonds. This is reflected by the reduction in flexible, basic groups when compared to **1**.⁶⁸

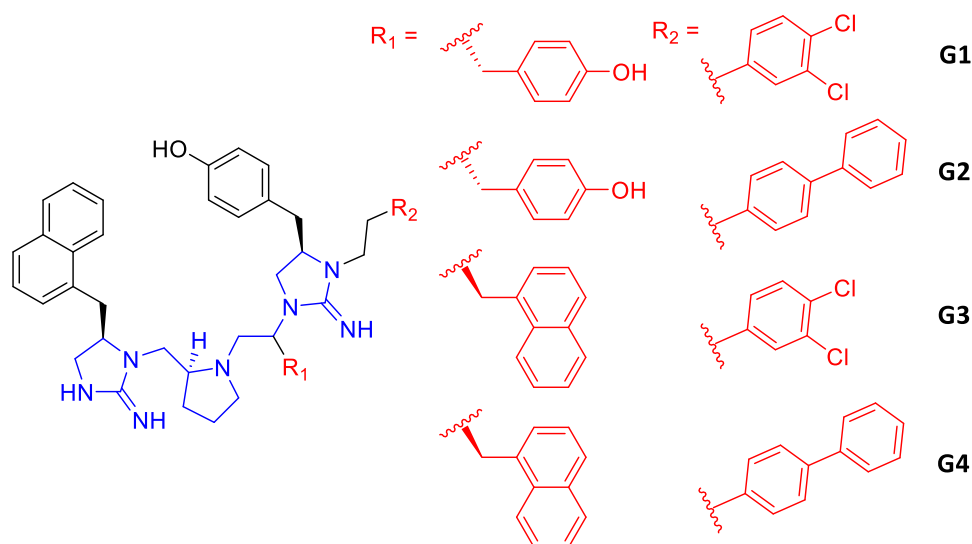


Figure 1-29: Chemical structure of **G1 – 4**, with the pyrrolidine bis-cyclic guanidine scaffold highlighted in blue and changing *R*-groups in red.⁶⁸

Despite their lower activity, the compounds were found to be more potent inhibitors of both gram-negative and gram-positive bacteria than the peptide hits. This could result from the increased permeability of the small molecules. However, this series of compounds were found to inhibit eukaryotic cells, showing possible issues with

toxicity.⁶⁸ These promising compounds could be developed further with the addition of more basic functionalities to overcome difficulties with selectivity and stability.

1.4.3.5 Psoralens

In 2001, Eichman *et al.* found that 4'-hydroxymethyl-4,5',8-trimethylpsoralen (**HMT**) was able to induce the formation of a 4WJ.⁹² **HMT** is a type of psoralen (**A, Figure 1-30**), a class of established photo-chemotherapeutics for the treatment of skin related diseases. To study the effects of cross-links formed by psoralens on DNA conformation, **HMT** was covalently cross-linked between two thymine bases in various DNA sequences and crystallised. It was found that a 4WJ could be formed, with two **HMT** compounds bound (**B and C, Figure 1-30**).⁹² The 4WJ contained four d(CCGGTACCGG) sequences, with the ACC trinucleotides forming the stable centre of the junction. The binding of **HMT** to this 4WJ had little effect on the structure when compared to the unbound form. In contrast, the compound also induced a noncanonical 4WJ from four d(CCGGTACCGG) DNA sequences. This is caused by the cross-linking of the **HMT**, which leads to the destabilisation of the dsDNA, allowing the exchange of strands and formation of the 4WJ. It could be conjectured that part of the therapeutic effects of psoralens may result from their ability to interfere with homologous recombination, potentially initiating cross-link repair in damaged cells.⁹²

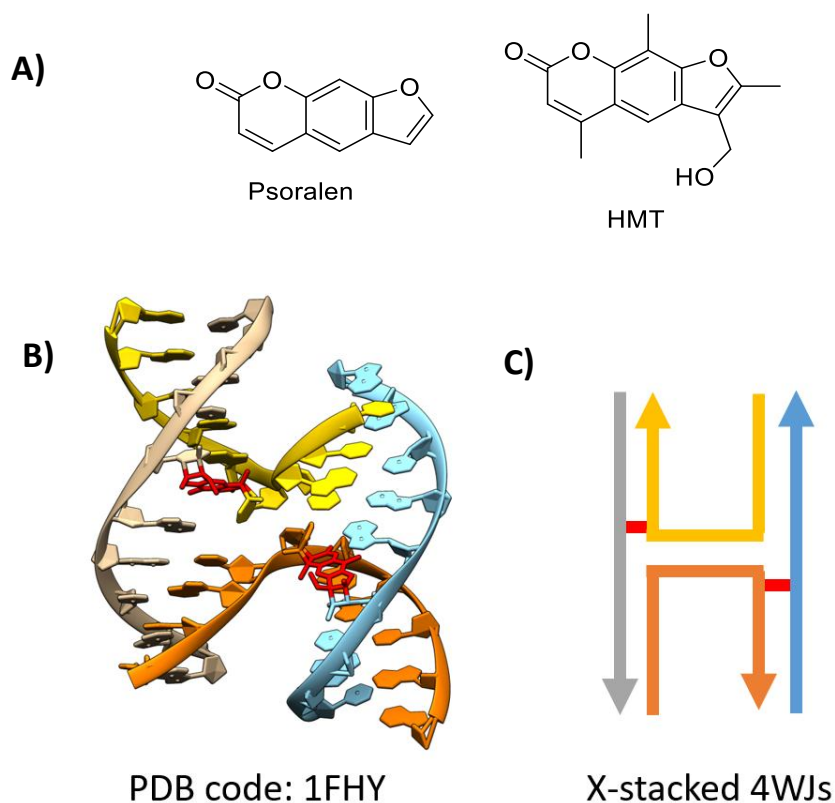


Figure 1-30: **(A)** Chemical structures of psoralen and HMT. **(B)** Crystal structure of two HMT molecules (red) bound to a 4WJ, formed by d(CCGGTACCGG) DNA sequences. **(C)** Schematic of crystal structure.⁹²

1.4.4 Crystal structures of 4WJ/ligand complexes

With the emergence of multiple crystal structures containing X-stacked 4WJs bound to various ligands, Chien's team set out to compare the 4WJ/ligand complexes (**Figure 1-31**).⁸⁴ They found that the binding of each ligand results in a varying DNA topology. In addition, [Pt(Hbapbpy)]PF₆,⁸² the triaminotriazine-acridine conjugate,⁸⁴ and psoralen-based HMT compound⁹² were able to form noncanonical 4WJs. The ability to induce the formation of specific higher-order DNA structures could have applications in the field of nanotechnology, allowing for the development of novel DNA sensors and superstructures.

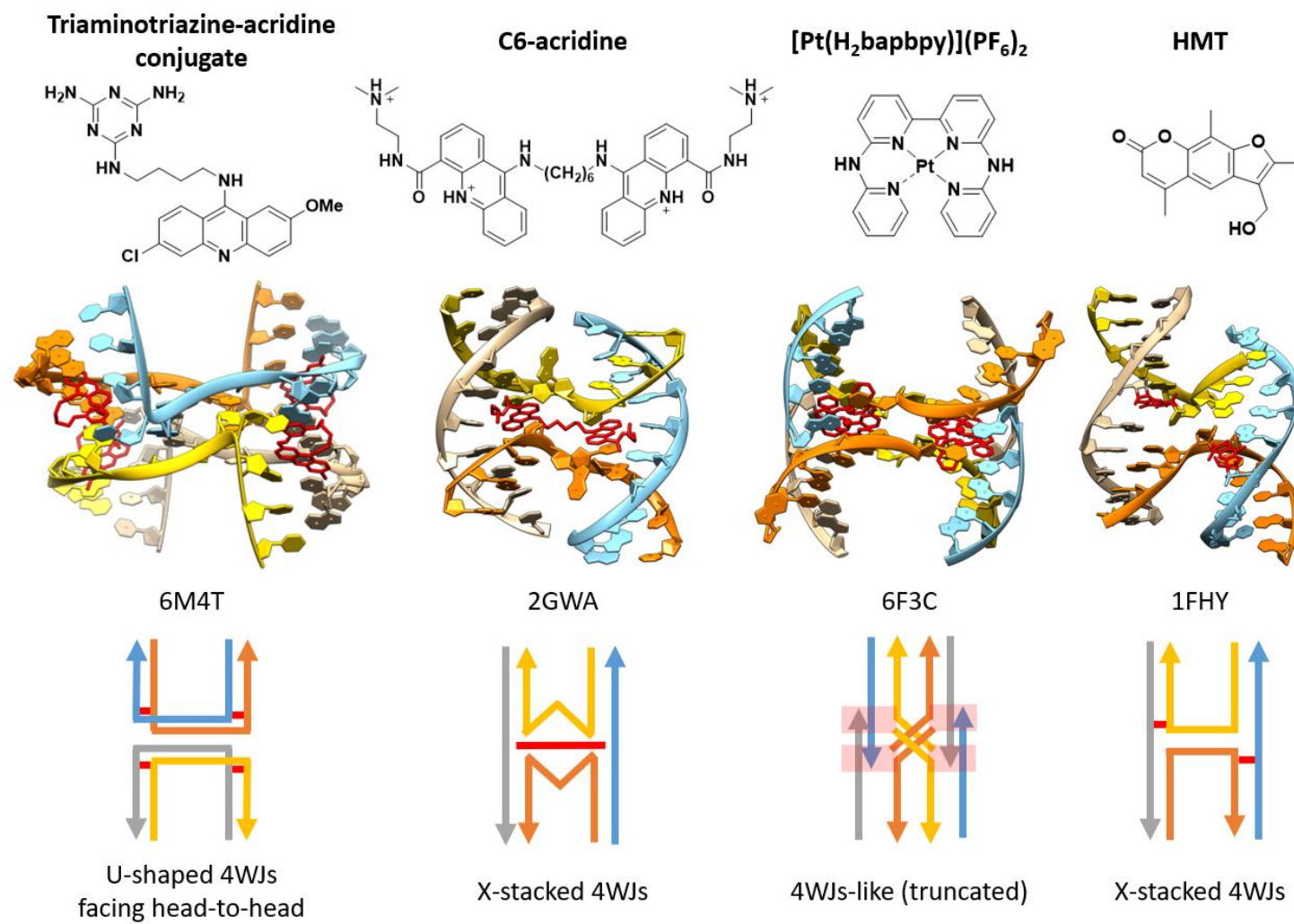


Figure 1-31: Crystal structures of various ligands bound to 4WJs.⁸⁴

1.4.5 Summary

Analysis of the compounds developed so far for targeting 4WJ show overall similarities in the pharmacophoric features required for binding. These include the presence of cationic groups and aromaticity to form stacking interactions with DNA. However, these features are also found with duplex DNA binding, showing the difficulties with selectively targeting 4WJs over other forms of DNA. Another feature often shared is the symmetry of the inhibitors, or the ability of the inhibitors to bind two molecules symmetrically within the 4WJ pocket. This reflects the 2-fold symmetry of the 4WJ and could be a useful starting point for the development of future inhibitors, although it is not always required.

Many compounds bind to just one of the 4WJ conformations, including X-stacked binders illustrated in **Figure 1-31** and open 4WJ binders, such as Segall's peptides^{45,65} and Hannon's organometallic pillarplexes.⁸³ The ability of each compound to target just one form shows the marked difference between the two conformations. This feature could be used to create tuneable 4WJs that can be altered between the two conformations in a controlled manner, with ligands acting like a molecular switch.

Biological results obtained so far have shown how 4WJ inhibitors could potentially be developed into therapeutic agents targeting areas such as cancer and antibiotics. However, further development of selective compounds is required, specifically over other DNA structures, to confirm the therapeutic effects result from 4WJ binding.

1.5 Thesis aims

The overall aims for this thesis are to:

- Synthesise previously discovered 4WJ-binding peptides, developing novel techniques to test their biological effects on DNA 4WJs, using polyacrylamide gel electrophoresis (PAGE), Förster resonance energy transfer (FRET) and fluorescence polarisation (FP) assays.
- Test the selectivity of these known 4WJ-binding peptides for 4WJ DNA, compared to dsDNA, using an ethidium bromide (EtBr) displacement assay.
- Examine the activity of these known 4WJ-binding peptides against human leukaemia cancer cell line HL-60.
- Produce novel peptides based on the structure of WRWYRGGRYWRW (**2**), a peptide previously shown to bind to the 4WJ and have antibacterial and

anticancer effects (**Figure 1-32**).⁴⁵ A series of peptidomimetic analogues have been synthesised, containing a variety of linker groups that replace the central Gly-Gly residues of dodeca-peptide **2**. In addition, two cyclic peptides have been prepared, where the ends on peptide **2** are linked together. The biological activities of these compounds have been evaluated against 4WJ DNA, dsDNA and HL-60 cancer cells using the assays developed for testing the known 4WJ-binding peptides.

- Explore an alternative approach to designing novel 4WJ-binding peptides, using information from previously published protein/4WJ X-ray crystallography structures.

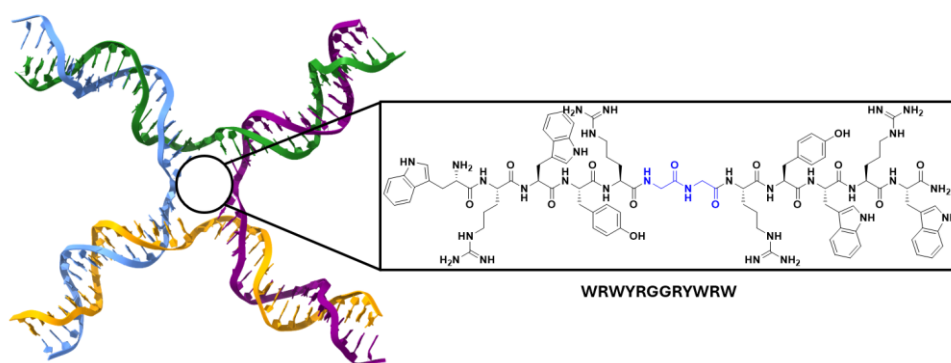


Figure 1-32: Image representing WRWYRGGRYWRW (**2**) targeting the centre of the open 4WJ (PDB code: 3CRX). The glycine amino acids that will be replaced with a variety of linker groups are shown in blue.

1.6 Thesis outline

In **chapter 2**, peptide **1**, **2** and related peptides have been synthesised using solid-phase peptide synthesis (SPPS), and their subsequent biological testing discussed. The activity of these peptides against the 4WJ are compared to two acridine compounds. The ability of peptide **2** to bind to the open conformation of the 4WJ is also examined using FRET analysis.

Chapter 3 looks at the design of peptide analogues based on the structure of peptide **2**. This includes linear peptides containing unnatural amino acids and cyclic peptides. The synthesis of these peptide analogues, using a combination of organic synthesis and solid-phase peptide synthesis (SPPS), is discussed. These compounds are then tested in the same suite of assays as those used in **chapter 2**, to allow for comparison to the original peptide **2**.

Chapter 4 examines an alternative approach to designing 4WJ-targeting peptides, using information from protein/4WJ crystal structures. By identifying protein fragments that interact with the 4WJ, two peptides are designed, synthesised, and tested against 4WJ DNA.

Finally, **chapter 5** provides an account of the experimental procedures, methods, instruments, and materials used throughout this project.

**Chapter 2 - The
synthesis and biological
testing of
WRWYRGGRYWRW and
related peptides**

2.1 Introduction

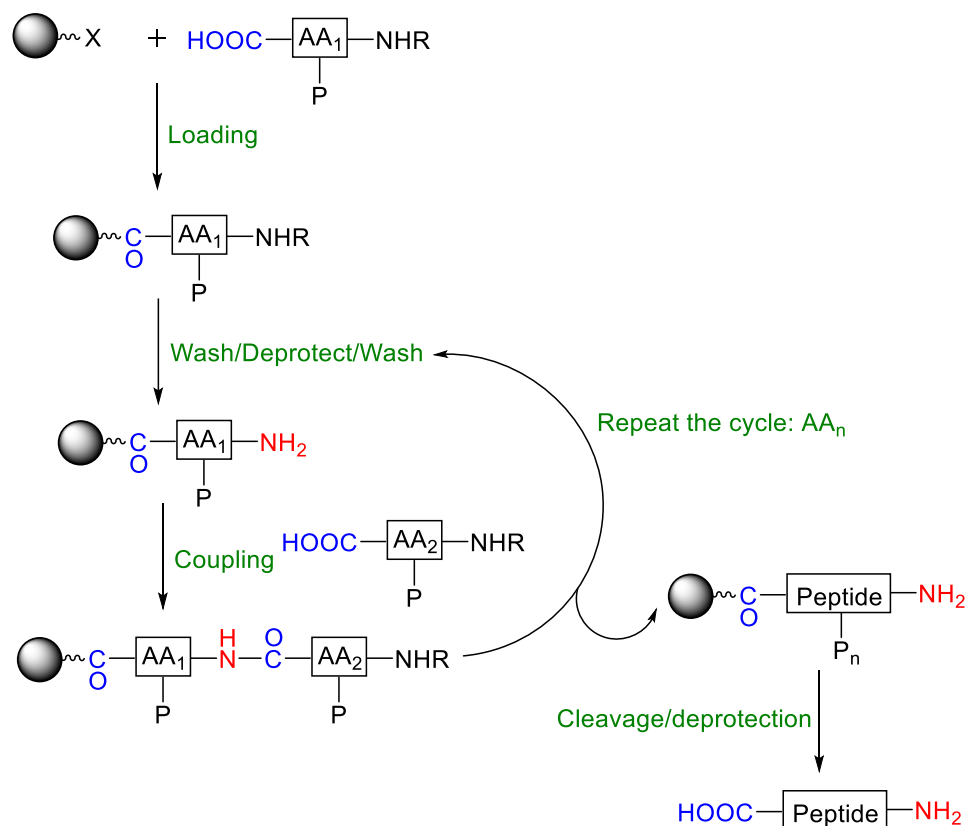
2.1.1 Solid-phase peptide synthesis

Peptides consist of short amino acid chains, generally with molecular weights between 500-5000 Da. These biomolecules are of growing interest as pharmaceutical agents, with >400 peptides having reached clinical trials. Therapeutic peptides sit between the size of small molecules and large biologics. They can achieve higher affinity and specificity for targets with large surface areas, such as protein-protein interactions, when compared to small molecules, whilst showing lower immunogenicity and production costs when compared to biologics. However, peptides often have problems with poor membrane permeability and low *in vivo* stability. These issues can be overcome using a variety of approaches, including the introduction of unnatural residues and peptide cyclisation.⁹³

Peptides are synthesised *via* repetitive amidation reactions between the amino group of one residue and the carboxylic acid group of another. This was originally carried out in solution, with all reagents fully dissolved, and purification carried out after each amino acid addition. However, as peptides increased in length, successive purifications became more time consuming and solubility issues increased.⁹⁴ To overcome these issues, in 1963 Merrifield established solid-phase peptide synthesis (SPPS).⁹⁵ This involved synthesising peptides on a solid bead, also known as a resin, that is insoluble in the solvents used during peptide synthesis. This allowed reagents to be added sequentially, with excess reagents simply washed from the resin after reaction completion. This means that purification only occurred once all amino acids have been added, increasing the speed and ease of peptide synthesis.

The general steps of SPPS are outlined in **Scheme 2-1**. Synthesis starts with the attachment of the first amino acid (AA), *via* its C-terminal (carboxylic acid), to a reactive group (X) attached to the resin. Subsequent amino acids are added to the N-terminal (α -amine), with each amino acid containing temporary protecting groups (P) on reactive side chains and at the α -amino group (R). After coupling, the *N*- α -protecting group is removed before the next coupling takes place, until the sequence is complete. At the end of the synthesis, all protecting groups are removed and the peptide is cleaved from the resin.⁹⁶ Peptides are typically synthesised from the C-to-

N terminus, to avoid higher levels of epimerisation experienced during N-to-C peptide syntheses.⁹⁷



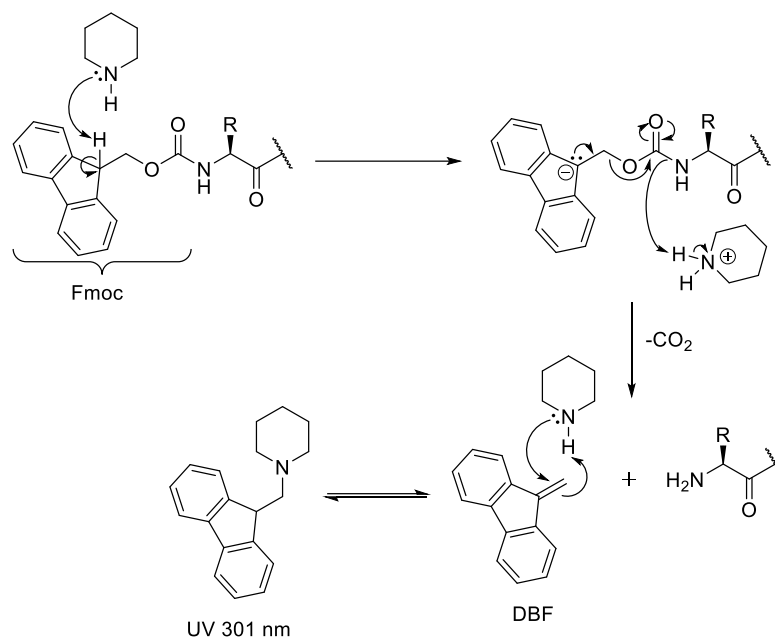
Scheme 2-1: General scheme of SPPS, containing resin (black ball) with a reactive group attached (X), amino acids (AA) and protecting groups (P).⁹⁶

SPPS chemistry is commonly defined by the identity of the temporary *N*- α -protecting group. The two most commonly used are the acid sensitive *tert*-butoxy-carbonyl (Boc) group and the base sensitive fluoren-9-ylmethoxycarbonyl (Fmoc) group. The Fmoc group is generally preferred because it can be removed under milder conditions, using weak bases. Resin linkers and amino acid side chains can then be orthogonally deprotected using acids, with trifluoroacetic acid (TFA) commonly used.⁹⁸

Fmoc Deprotection

Deprotection of the *N*- α -Fmoc group during SPPS is often carried out with piperidine (**Scheme 2-2**).⁹⁹ The mechanism starts with the piperidine nitrogen lone pair removing the acidic proton at position-9 of the Fmoc fluorene ring. The resultant carbanion then causes a cascade reaction to produce the desired free amine, whilst liberating CO₂ and forming dibenzofulvene (DBF). This side product subsequently

reacts with piperidine to form a UV active adduct, that absorbs at 301 nm, that can be monitored to check deprotection is complete or simply washed away.⁹⁶



Scheme 2-2: Fmoc deprotection mechanism using piperidine, where R indicates the amino acid side chain.⁹⁶

Amino acid coupling

Before the coupling of amino acids *via* amidation can be carried out, the α -carboxylic acid group must be activated. Initially Merrifield used carbodiimides as activating groups, finding diisopropylcarbodiimide (DIC) (**Figure 2-1**) to be useful for SPPS due to the solubility of its urea byproduct.¹⁰⁰ However, the high reactivity of the *O*-acylisourea intermediate group resulted in partial racemisation of the amino acid side chain, through the transient formation of an oxazolone ring.¹⁰¹ Subsequently different coupling group salts have been developed that reduce epimerisation, including *N*-[(1*H*-benzotriazol-1-yl)(dimethylamino)-methylene]-*N*-methylmethanaminium hexafluorophosphate *N*-oxide (HBTU) and *N*-[(Dimethylamino)-1*H*-1,2,3-triazolo-[4,5-*b*]pyridin-1-ylmethylene]-*N*-methylmethanaminium hexafluorophosphate *N*-oxide (HATU) (**Figure 2-1**). These coupling reagents are used with 1-hydroxybenzotriazole (HOBt) (**Figure 2-1**), that reacts with the initial intermediate to form a less reactive benzotriazolyl ester intermediate and reduce epimerisation. The addition of a base, such as *N,N*-diisopropylethylamine (DIPEA), is also required to deprotonate the carboxylic acid before reacting with the coupling reagents.⁹⁶ More recently coupling reagents based

on the oxyma scaffold, such as Ethyl 2-cyano-2-(hydroxyimino)acetate (oxyma pure) (**Figure 2-1**), have been combined with DIC for utilisation in SPPS. Oxyma reagents can inhibit racemisation and replace HOBt, a reagent that has potentially explosive properties. This is especially useful for microwave assisted SPPS carried out at elevated temperatures.¹⁰²

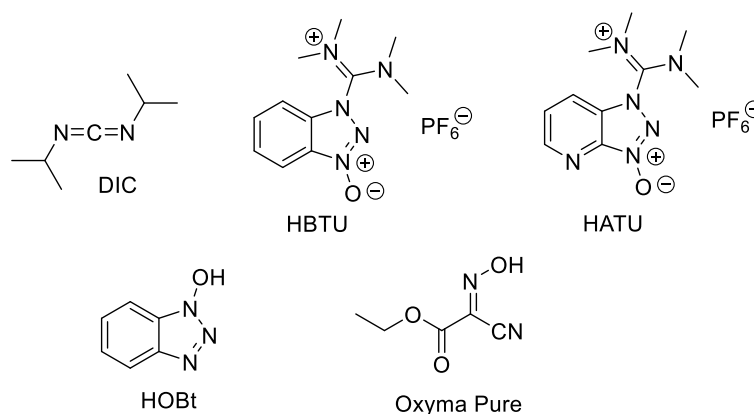
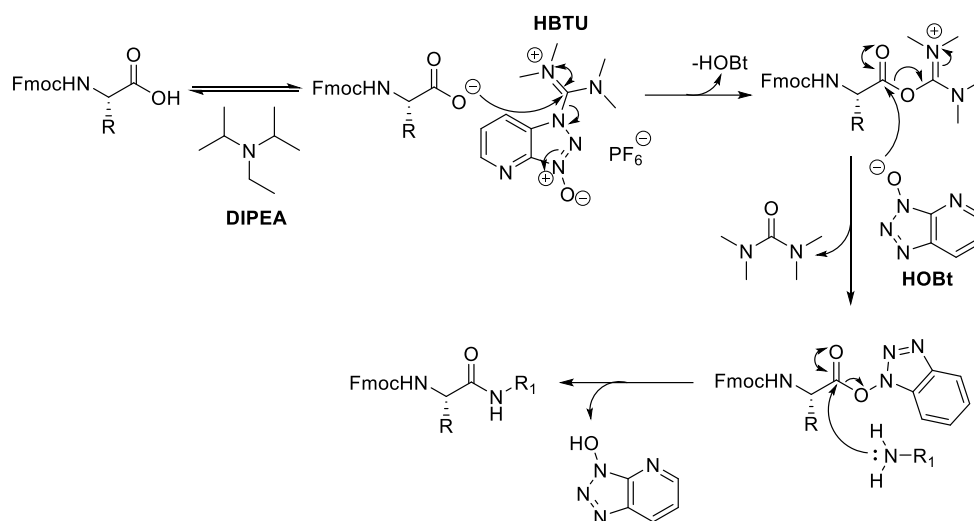


Figure 2-1: Coupling reagents used during SPPS, for carboxylic acid activation before amino acid coupling.

A typical reaction mechanism for carboxylic acid activation followed by amino acid coupling is shown in **Scheme 2-3**, using HBTU, HOBt and DIPEA as coupling reagents. Firstly, the α -carboxylic acid is deprotonated using DIPEA. The subsequent carboxylate attacks the imine group of HBTU, displacing the HOBt group. The carbonyl group of the *O*-acylisourea intermediate is attacked by HOBt, resulting in the formation of the diisopropylurea side product and benzotriazolyl ester

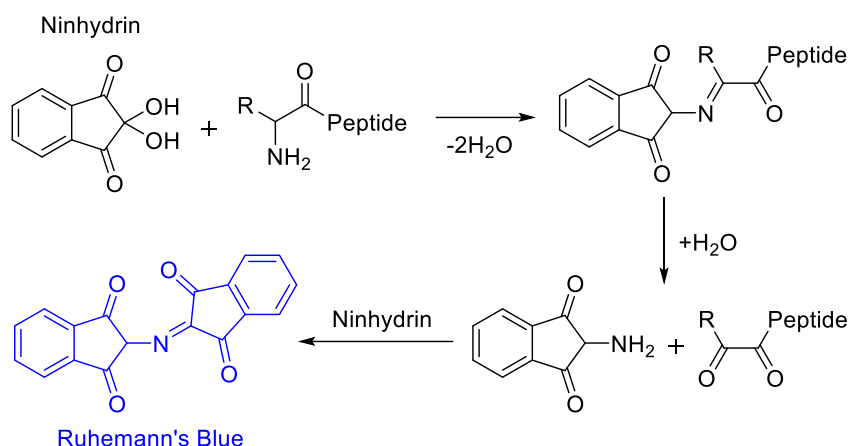
intermediate. The activated carboxylic acid is then attacked by the amine group of another amino acid, to form HOBT and the desired amide bond.



Scheme 2-3: Reaction mechanism of carboxylic acid activation using DIPEA/HBTU/HOBt, where *R* indicates the amino acid side chain.⁹⁶

Monitoring amino acid couplings

During SPPS it is important to monitor difficult coupling reactions, to ensure all sites on the resin have the desired amino acid attached before moving onto the next reaction. This prevents the formation of unwanted products, containing differing peptide sequences that can be difficult to remove during purification. The Kaiser test is often used to qualitatively determine whether a solution contains free primary amino groups, that should be absent if a coupling reaction is complete.^{103,104} This test is based on the colour change that occurs when ninhydrin reacts with a primary amino group, forming a dark blue solution (**Scheme 2-4**). The Kaiser test needs minimal peptide for the colour change to be observed and can be carried out directly on peptides attached to resin, without the need for cleavage.



Scheme 2-4: Reaction of peptides and ninhydrin for the Kaiser test, where R indicates the side chain of the final amino acid.⁹⁶

Amino acid protecting groups

When using Fmoc SPPS, amino acid side chains must be resistant to treatment with base during Fmoc removal and acid liable for deprotection at the end of peptide synthesis. Commonly used protecting groups include *tert*-butoxy-carbonyl (Boc), *tert*-Butyl (*t*Bu), trityl (Trt) and 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl (Pbf) (**Figure 2-2**). These groups can be removed by treatment with TFA solutions, containing additional scavenger groups, such as triisopropylsilane (TIPS) and water, to prevent modification or reattachment to the unprotected side chains.¹⁰⁵

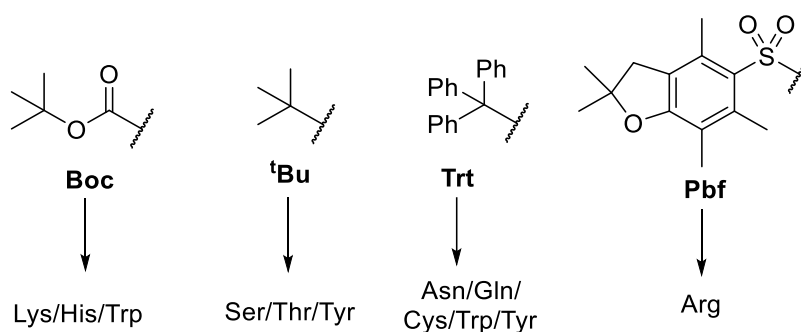
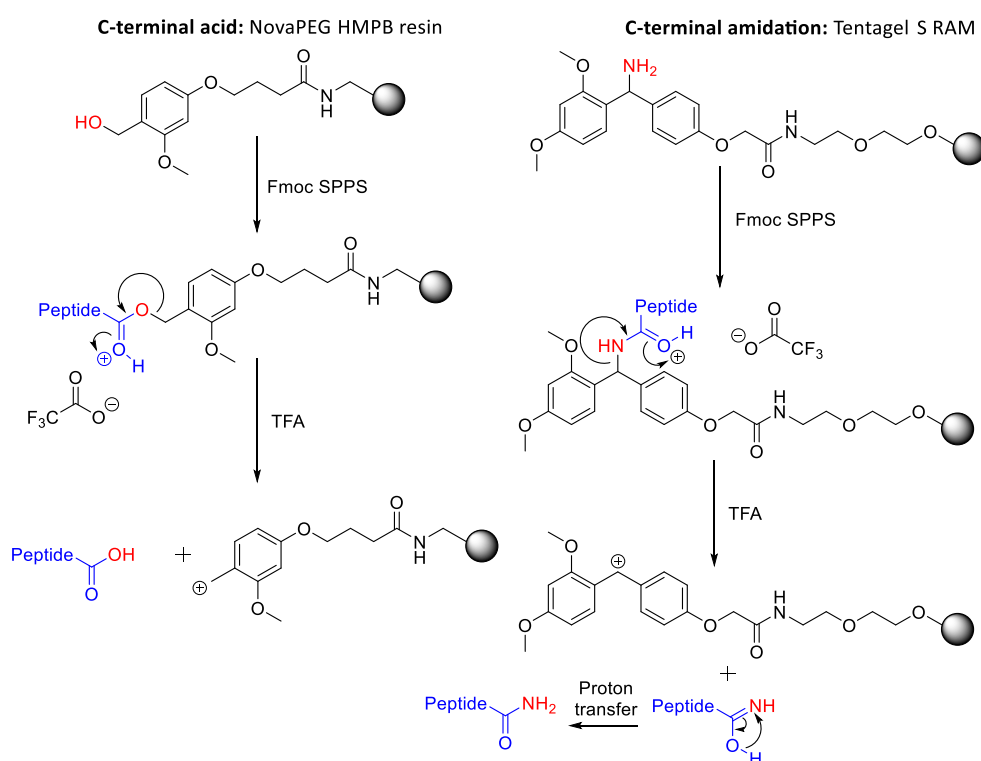


Figure 2-2: Common amino acid protecting groups.¹⁰⁵

Resins and linkers

The core resin is often made from polystyrenes, polyamines or polyethylene glycols (PEGs) containing 1-2% divinylbenzene to act as a cross-linking agent. The small spherical beads are typically either 100 – 200 mesh (75 – 150 μm) or 200-400 mesh

(35 – 75 μm) in size. Resins have different loading capacities, indicating the number of reactive sites the resin has per gram (mmol g^{-1}). Higher loading capacities can increase synthetic efficiency; however, this must be balanced with an increased risk of aggregation. A variety of linker groups can be used between the first amino acid and resin, including linkers that result in either acid or amide C-terminus functionalisation after cleavage. Once synthesis is complete, cleavage from the resin can be achieved using the same TFA cocktails used for amino acid side chain deprotection (**Scheme 2-5**).⁹⁶



Scheme 2-5: Examples of linkers resulting in acid/amide C-terminal functionalisation, showing resins (black balls), linker handles (red) and peptides (blue). After Fmoc-SPPS the peptide is cleaved from the resin using TFA to form the desired-terminal functionalisation and a positively charged linker/resin group.

2.1.2 Fluorescence assays

After the synthesis of peptides using SPPS, their effect on biological targets of interest must be evaluated. Fluorescence assays are routinely used to probe the interactions of peptides with other biomolecules. Fluorescence and phosphorescence are both types of luminescence, that occur when light is emitted from a substance in an electronically excited state. Fluorescence happens when an electron is in an excited singlet state, meaning it has opposite spin to the electron in

the ground state. This allows for the rapid emission of a photon when the electron returns to ground state, generally at a rate of $\sim 10^8 \text{ s}^{-1}$. Alternatively, phosphorescence takes place when the excited electron has the same spin as the ground state electron, known as the triplet excited state. Emission is then slowed to rates of $10^3 - 10^0 \text{ s}^{-1}$ since the transition to the ground state is forbidden.¹⁰⁶

These processes can be represented by a Jablonski diagram (**Figure 2-3**).¹⁰⁷ This shows three separate singlet energy states (S_0 , S_1 and S_2), each containing three vibrational energy states (0 – 2). After light absorption ($h\nu_A$), a molecule is excited from the ground state (S_0) to a higher energy state (S_1 or S_2). Internal conversion then rapidly occurs, in $\leq 10^{12} \text{ s}^{-1}$, as molecules relax into the lowest S_1 vibrational level. The molecule can emit fluorescence ($h\nu_F$) and return to a vibrational level of S_0 or undergo intersystem crossing, where the spin of the electron is reversed into the first triplet state (T_1). Subsequently, phosphorescence of the electron takes place at a slower rate ($h\nu_P$).

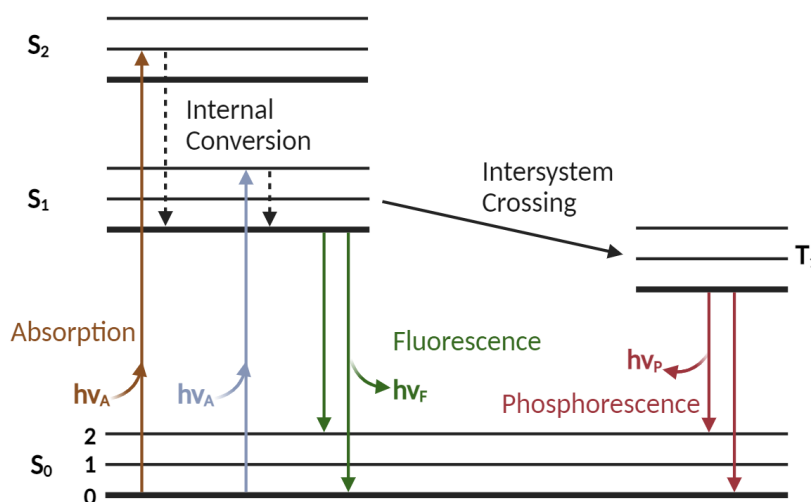


Figure 2-3: A Jablonski diagram illustrating the processes involved in luminescence, showing singlet energy states (S_x), the first triplet energy state (T_1) and vibrational energy states (0 – 2).¹⁰⁷

Fluorescent molecules, known as fluorophores, are generally aromatic compounds that contain several conjugated π -bonds. Common examples of fluorophores, used in biological assays include ethidium bromide (EtBr), 6-carboxyfluorescein (FAM) and 6-carboxytetramethylrhodamine (TAMRA) (**A, Figure 2-4**). These photoreactive compounds absorb light within specific wavelength ranges, before emission at longer wavelengths. The difference between the maximum excitation and emission wavelengths is known as their Stokes shift (**B, Figure 2-4**).¹⁰⁸ A larger value can help to reduce the reabsorption of emitted photons, to give a higher contrast for

fluorescence imaging.¹⁰⁹ The fluorescence process is cyclic, allowing a single fluorophore molecule to be repeatedly excited causing the emission of thousands of photons. This means that fluorophores can be detected at very low concentrations, making fluorescence spectroscopy highly sensitive.¹⁰⁶ The efficiency of a fluorophore to convert an absorbed photon into fluorescence emission in a specific environment, provides its quantum yield. This is a ratiometric number, with values closest to 1 showing the brightest fluorescence.¹¹⁰

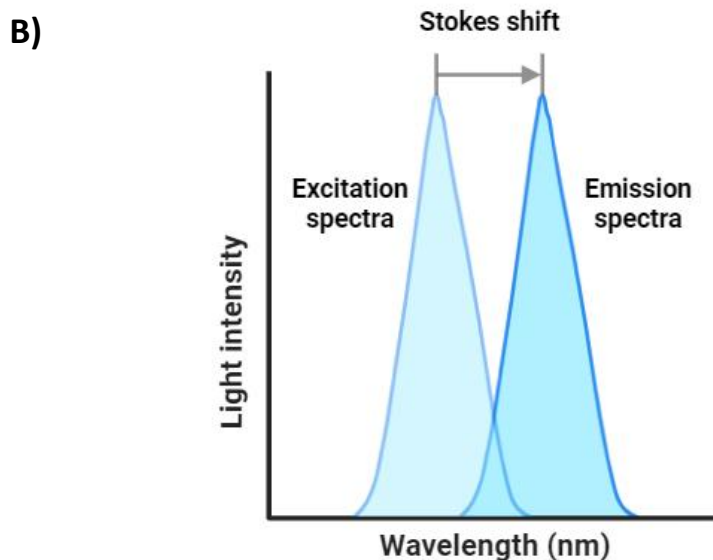
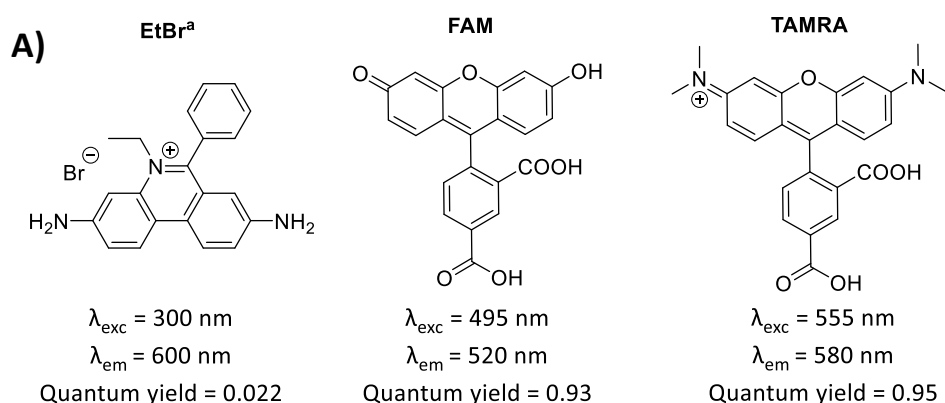


Figure 2-4: (A) Examples of fluorophores used in biological assays, with characteristic excitation/emission maxima and quantum yields provided.¹⁰⁶ (B) The excitation and emission of a fluorophore, with the Stokes shift indicated.¹⁰⁸
^aValues measured in water.¹¹¹

The amino acids tryptophan and tyrosine can act as intrinsic fluorophores in proteins and peptides, although their quantum yields in aqueous solutions of 0.13 ± 0.01 and 0.14 ± 0.01 respectively are much lower than the 0.95 ± 0.03 quantum yield of

fluorescein.¹⁰⁶ The intrinsic fluorescence of DNA is even weaker, with quantum yields observed between 10^{-4} – 10^{-5} .¹¹² Due to their low levels of fluorescence, biomolecules are often labelled with fluorophores such as FAM, to enhance their fluorescent signals.

Two issues are often encountered when using fluorophores, photobleaching and quenching. Photobleaching is caused by the structural instability of a fluorophore in its excited state, causing vulnerability to degradation. After repeated exposure to light, a fluorophore can undergo a structural change that permanently prevents fluorescence. This process is observed as the dimming in a fluorescent image or a reduction in signal intensity.¹⁰⁶ Quenching is another process that results in a decrease in fluorescence intensity *via* a variety of different mechanisms. Collisional quenching is caused by contact between the excited fluorophore and another compound in solution, for example solvent or buffer components. These quencher molecules can enable transition of the electron to the ground state *via* a non-radiative transition. Alternatively static quenching can occur when the fluorophore forms a non-fluorescent complex with a quencher before the excited state can be reached. For example, this can happen when a fluorophore and quencher aggregate *via* π -stacking interactions.¹¹³ Therefore, it is important to carefully design fluorophores or fluorescently labelled complexes to minimise these issues.

Several techniques can be utilised for fluorescence-based assays. These include assays that monitor changes in fluorescence intensity, such as FRET, and those that detect changes in molecular rotation, such as FP. The following section will describe these different techniques.

Ethidium bromide displacement assay

EtBr (**Figure 2-4**) is a non-selective DNA intercalator, that has rapid equilibration rates and low binding affinity.¹¹⁴ This fluorophore shows an increase in intrinsic fluorescence when bound to DNA, potentially because EtBr is then shielded from water, that acts as a quencher.¹¹⁵ When a non-fluorescent DNA binding compound is added to a solution containing EtBr prebound to DNA, a decrease in fluorescence intensity is observed, as the EtBr is displaced and quenched (**A and B, Figure 2-5**). This can be quantified by titrating the competitor compound into solution, to find its apparent binding affinity.¹¹⁴

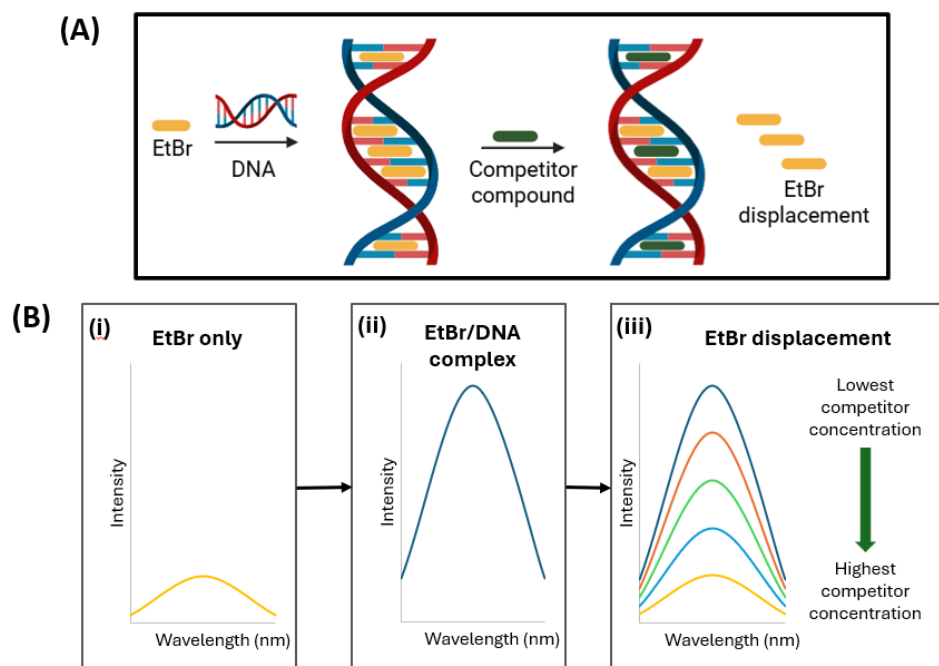


Figure 2-5: (A) Schematic representation of EtBr displacement assay. **(B)** Effect on fluorescence intensity throughout the EtBr displacement assay: (i) Initially the EtBr is unbound and has low fluorescence. (ii) EtBr binds to DNA via intercalation and fluorescence is increased. (iii) A competitor compound is added into the solution that intercalates with DNA, displacing EtBr and reducing fluorescence.¹¹⁶

Fluorescence polarisation (FP)

FP detects binding events in solution by exciting a fluorescently tagged molecule (tracer) with polarised light. When the tracer is unbound, it tumbles quickly in solution, emitting mostly unpolarised light. Upon binding to another molecule, the subsequent complex tumbles more slowly, causing the emitted light to have increased polarisation (**Figure 2-6**). To quantify the level of polarisation (P), the following calculation is used:

$$P = \frac{(F_{\parallel} - F_{\perp})}{(F_{\parallel} + F_{\perp})}$$

Equation 2-1: Polarisation (P) calculated from parallel ($F_{\parallel}$) and perpendicular ($F_{\perp}$) fluorescence emission intensities.¹¹³

This describes the difference between parallel ($F_{\parallel}$) and perpendicular ($F_{\perp}$) fluorescence emission intensities, in relation to the excitation light plane, divided by the total fluorescence emission intensity (**Equation 2-1**).¹⁰⁶ The value of P is dimensionless because it is a ratio of light intensities and is commonly expressed in milli-polarisation (mP), where 1 P = 1000 mP. Fluorescence anisotropy (FA) is a term used interchangeably with FP. Although anisotropy has a slightly different equation, the information obtained from both FP and anisotropy is similar.¹¹³

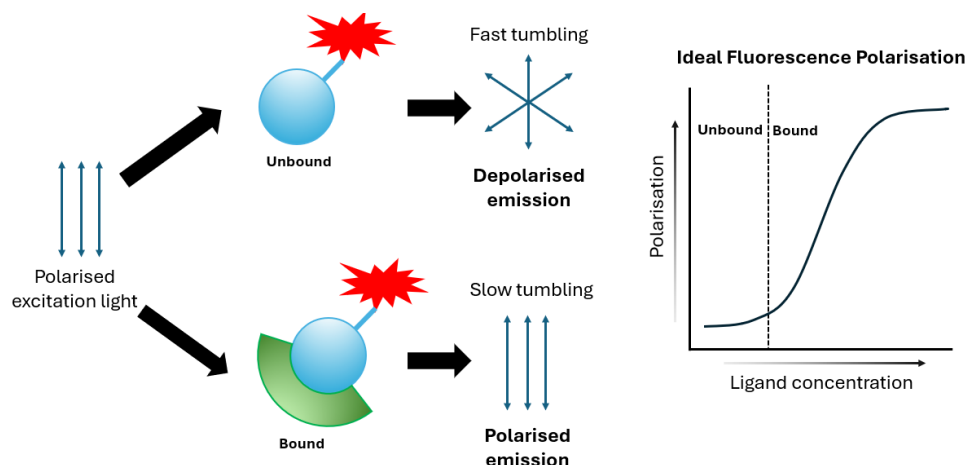


Figure 2-6: On the left is a FP schematic, showing an unbound tracer molecule (blue), tagged with a fluorophore (red star), tumbling quickly in solution with low polarisation. When a second molecule (green) binds to the tracer, tumbling is slowed, and polarisation increases. On the right is an idealised FP graph, with polarisation increasing with ligand concentration.¹¹⁷

Förster resonance energy transfer (FRET)

FRET measures the energy transfer between two suitable fluorophores when in close proximity (<10 nm). FRET occurs when a donor fluorophore is excited, by absorbing light at one wavelength, followed by a non-radiative energy transfer to an acceptor fluorophore, that subsequently emits a photon at a longer wavelength (**Figure 2-7**). For FRET to take place, considerable spectral overlap of >30% between the donor's emission wavelengths and acceptor's absorption wavelengths is required.¹¹⁸ This process causes a loss donor fluorophore emission and an increase in acceptor fluorophore emission and can be used to calculate FRET efficiency (E). FRET can be used to measure the interaction between fluorophores in different molecules, such as when looking at the binding of ligands to a target biomolecule, or changes in conformation within a single molecule.¹¹⁹

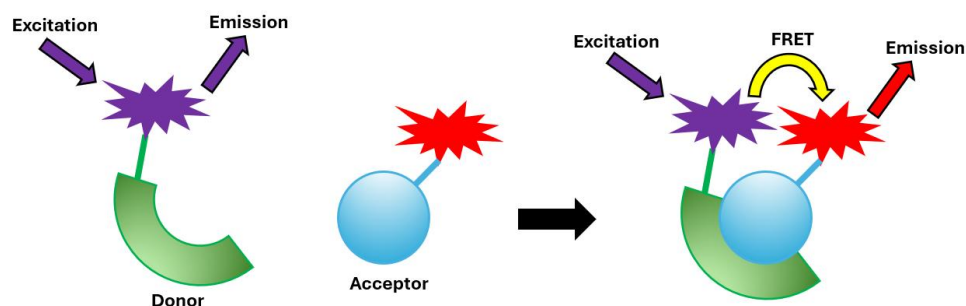


Figure 2-7: FRET schematic showing a molecule (green) tagged with a donor fluorophore (purple) that on its own emits fluorescence after excitation. Upon addition of a second molecule (blue) tagged with an acceptor fluorophore (purple), when the donor fluorophore is excited, FRET occurs, and emission of the acceptor fluorophore is observed.¹¹⁹

2.1.3 Assays for identifying 4WJ conformations

It is important to identify whether small molecules/peptides target the open or X-stacked 4WJ conformation. This can help predict their biological effect, with various proteins being structurally selective for only one 4WJ conformation (**Table 1, section 1.3.3**). Both gel electrophoresis and FRET assays have been utilised to probe 4WJ conformation and will be discussed further.

Gel electrophoresis works by pushing an electric charge through a gel containing small pores. This causes molecules, such as DNA, RNA, and proteins, to be separated based on size and charge, with smaller and less charged molecules moving further. The gel used is generally a cross-linked polymer, with a specific composition selected based on the size and type of biomolecules being separated. Polyacrylamide gels are typically used for short DNA sequences (<1000 bp), making them ideal for analysing 4WJ made up of short oligonucleotides. PAGE is a native technique that allows complexes, such as the 4WJ, to be maintained during the assay. The net charge of DNA is negative therefore it will travel towards the positively charged anode during PAGE (**Figure 2-8**).¹²⁰

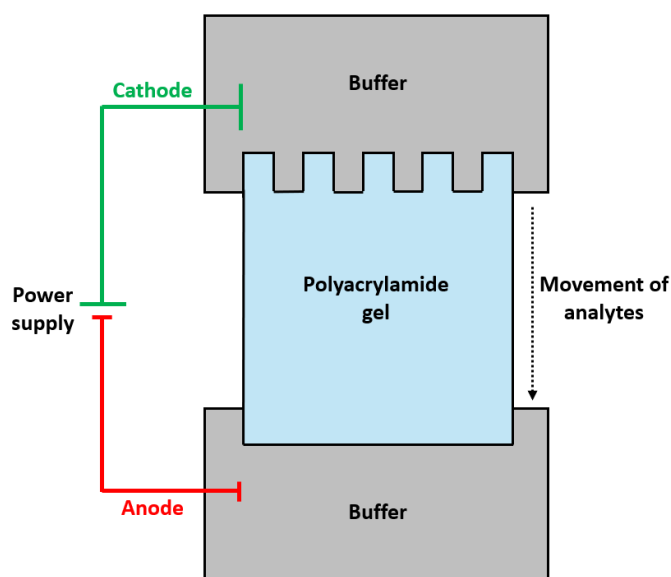


Figure 2-8: Representation of a PAGE assay set-up.¹²⁰

When DNA is bent, such as at the centre of the 4WJ, electrophoretic mobility is reduced.⁵⁰ The retardation of bent DNA is attributed to its larger hydrodynamic radius relative to linear DNA, leading to greater frictional resistance with the gel matrix. Angles closer to 90° have the largest hydrodynamic radius and are therefore the slowest to pass through gels.¹²¹ Comparing the retardation of different DNA species during PAGE can be used to determine their shape and symmetry. This method was used to confirm the structure of the 4WJ conformations a decade before X-ray crystal structures were solved.¹²²

To study the 4WJ structure using PAGE, an immobile 4WJ was designed by Lilley *et al.*,⁵³ with sequences chosen that could form only one 4WJ at the centre of each of the 80 bp oligonucleotides. As expected, migration of the full junction during gel electrophoresis was very retarded, running to the same point as the 510 bp marker when using a standard DNA ladder. On each arm of the junction, a unique restriction site was added, around 12 bp from the junction's centre. Restriction sites are short DNA sequences of 4-8 bp that are recognised by particular restriction enzymes, with each of the 4WJ arms named after its specific enzyme (**Figure 2-9**). This allowed for the selective cleavage of each arm by ~25 bp and could be used to study the effects of this on electrophoretic mobility.⁵³

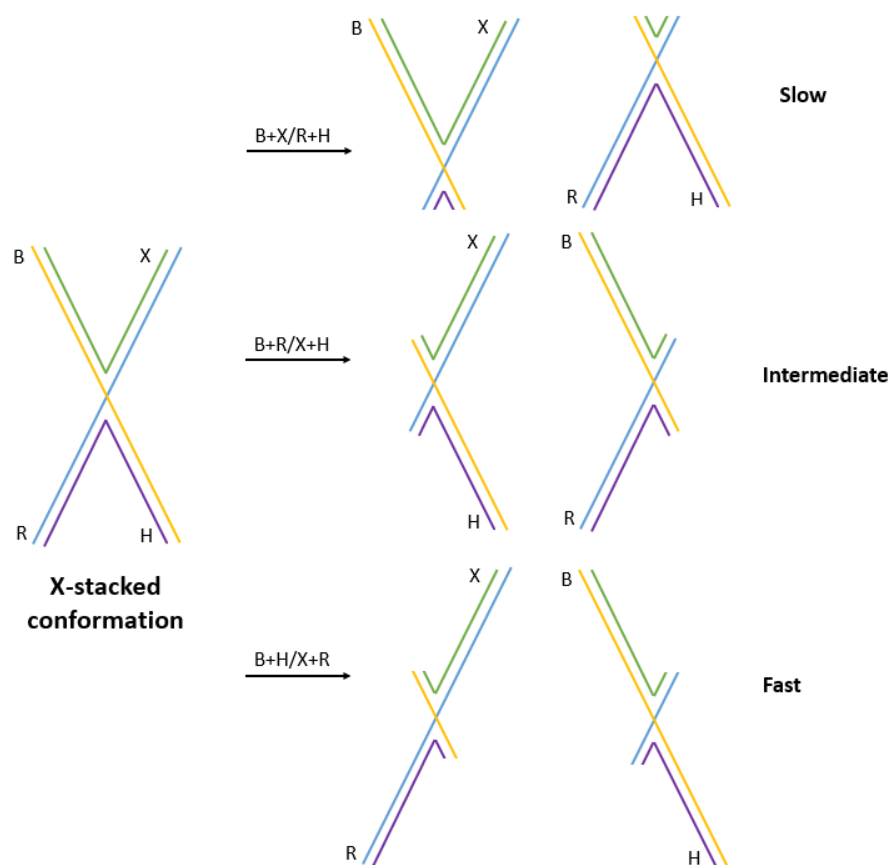


Figure 2-10: Analysis of different gel electrophoresis species, formed by enzyme cleavage of the immobile junction when in a high salt environment.⁵³

The square-planar structure of the open conformation was also established using the same method (**Figure 2-11**). Two bands were formed, one containing 4 species and the other 2 species, indicating a 4-fold symmetry. The band containing 4 species ($B+R/X+H/B+X/R+H$) travelled slowly through the gel due to the 90° angle between the duplex arms. The other band containing 2 species ($B+H/R+X$) contained the co-linear arms and travelled quickly through the gel.⁵³

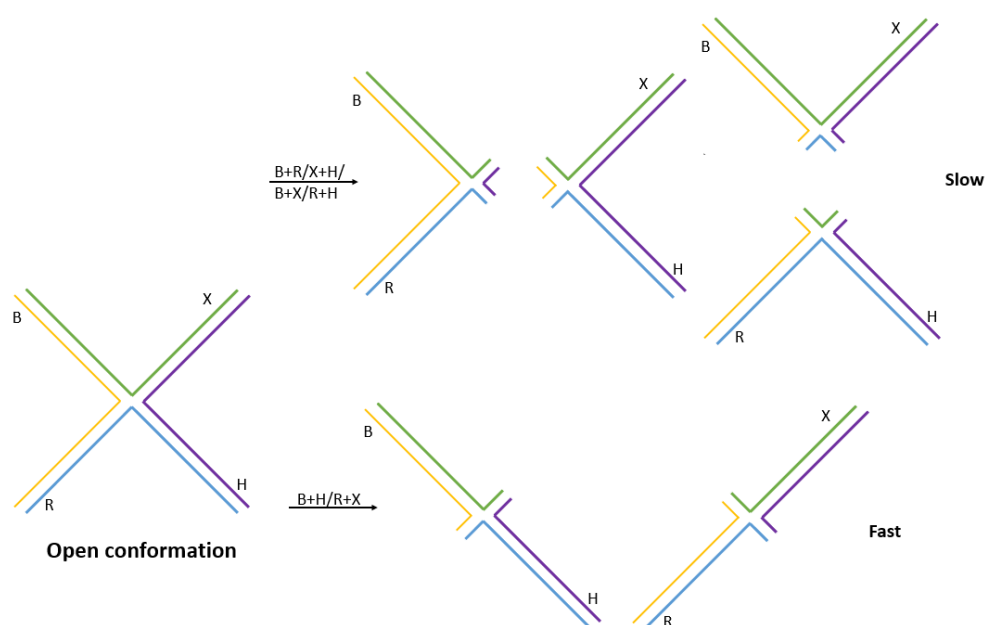


Figure 2-11: Analysis of different gel electrophoresis species, formed by enzyme cleavage of the immobile junction when in a low salt environment.⁵³

Another technique developed for studying 4WJ conformations is FRET. This assay was first used to observe changes in 4WJ conformations upon the addition of different types of cations.¹²³ Junction 3 (J3) is a well-studied 4WJ, that was shown to have a 77.4% preference for isoform II when in the X-stacked conformation, using single-molecule fluorescence techniques.^{55,124} A donor fluorophore (FAM) and acceptor fluorophore (TAMRA), were carefully placed on the 5' terminus of two separate oligonucleotides in J3, so that when isoform II was formed, the co-axial stacking of the duplex arms brought the fluorophores close enough together for FRET to occur. In contrast, when J3 was either in isoform I or the open conformation, the fluorophores were too far apart for FRET to occur (**Figure 2-12**). This was demonstrated when solutions of NaCl, CaCl₂ or MgCl₂ were titrated into a low salt J3 solution, with FRET shown to increase as the concentration of cations increased, reflecting the change from the open to X-stacked conformation.¹²³ FRET using J3 was subsequently used to confirm that the *E. coli* histone-like protein HU preferred to bind to the X-stacked 4WJ conformation.¹²⁵

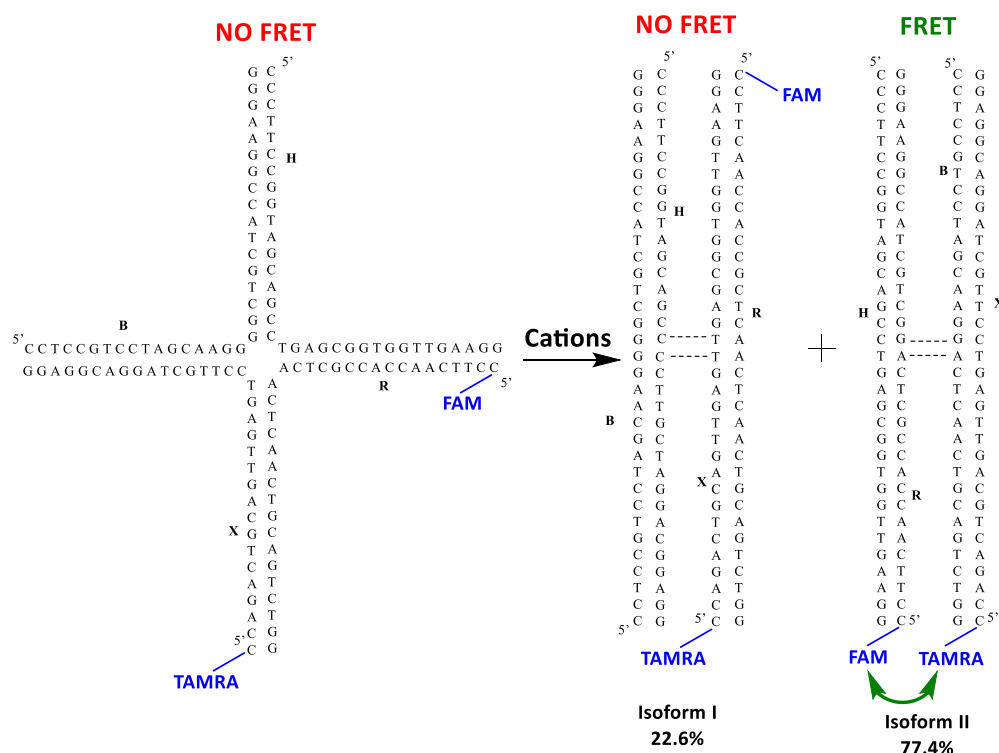


Figure 2-12: J3 conformations, showing low FRET in the open conformation or X-stacked isoform I and high FRET in the X-stacked isoform II.¹²³

2.2 Chapter aims

The purpose of this chapter is to synthesise previously identified 4WJ-binding ligands, before testing them in a series of novel biological assays. Firstly, SPPS will be used to synthesise WRWYCR (**1**) and WRWYRGGRYWRW (**2**), peptides that were discussed in **section 1.4.1**. Secondly, two examples of acridine small molecules, mentioned in **section 1.4.3.1**, will be synthesised using multi-step organic synthesis, to allow for their activities to be directly compared to those of the peptides. A series of robust assays will be developed, to test the ligands for 4WJ activity, dsDNA activity and cancer cell cytotoxicity. To ensure the validity of these assays, two novel peptides will be designed and synthesised to act as negative controls.

2.3 Results and discussion

2.3.1 Synthesis of peptides using SPPS

Peptides were made using solid-phase peptide synthesis (SPPS) on either a MultisynTech Syro I or a CEM Liberty automated peptide synthesiser. The general procedure involved first swelling the specified resin in DMF, before Fmoc deprotection, using 20% piperidine in DMF. The required amino acids were then coupled twice, to ensure all sites on the resin were successfully coupled. Between each reaction on the resin, DMF was used to wash away any excess reagents. Fmoc deprotection was repeated before the addition of the next amino acid. Once the sequence was complete, the resin was washed with DMF, MeOH and CH₂Cl₂ before the peptide was cleaved from the resin, and side chains deprotected, using a cleavage cocktail containing TFA and scavenger compounds. When the peptide contained cysteine residues, TIPS, water and 2,2'-(ethylenedioxy)diethanethiol (DODT) were used as scavenger compounds, with the DODT removed for peptides not containing cysteine. The peptides were purified by prep-HPLC or RP-column chromatography. Successful peptide synthesis was confirmed by analysis using MALDI or LCMS, HPLC, and HRMS. The results from the synthesis of peptides are summarised in **Table 2-1**.

Table 2-1: Literature peptides (black) and novel peptides (red) synthesised using SPPS. Observed mass taken from HRMS analysis.

No.	Sequence	Exact mass (g/mol)	Observed mass [M+H] ⁺	Purity (%)	Yield (%)
1 ⁶⁵	H-WRWYCR-NH ₂	967.4599	968.4677	91	50
2 ⁴⁵	H-WRWYRGGRYWRW-NH ₂	1825.9178	1826.9235	95	26
4 ⁶⁵	H-WRWYCR-OH	968.4439	969.4432	95	51
5	H-WRWYC(Acm)R-NH ₂	1038.4970	1039.5049	93	31
6	H-WRWYR-NH ₂	864.4507	865.4597	100	9

The WRWYCR sequence was initially synthesised with a carboxylic acid on the C-terminus, using the pre-loaded arginine resin; H-Arg(Pbf)-HMPB NovaPEG. After SPPS and subsequent purification, a 51% yield was obtained for the synthesis of peptide **4** (**Table 2-1**). This can be deemed a high yield, considering the number of reactions involved in SPPS, and reflects the high crude purity of the peptide of ~70%

(A, Figure 2-13). This allowed for the straightforward purification of **4**, using RP-column chromatography, to obtain the desired product with 95% purity (B, Figure 2-13).

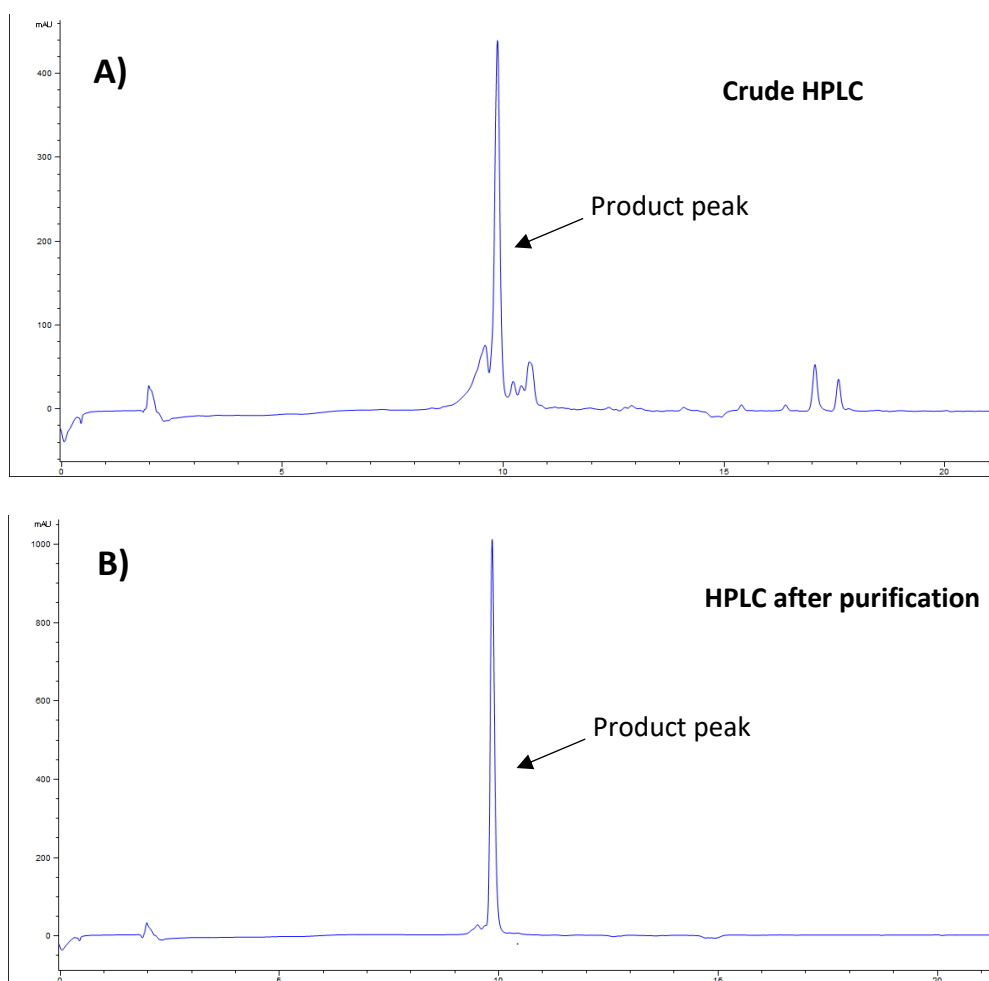


Figure 2-13: HPLC traces of peptide **4** (214 nm, 5-95% MeCN in H₂O + 0.05% TFA over 15 min, then 5 min at 95%). **A)** The crude peptide with ~70% purity. **B)** The peptide after purification with 95% purity and a retention time of 9.84 min.

Following the synthesis of **4**, focus moved to synthesising the peptide with C-terminus amidation, to follow the work of Segall's group.⁴⁵ This was to allow for the direct comparison of biological results to those previously obtained and neutralise the C-terminus in an attempt to improve DNA binding. SPPS using the H-Rink amide ChemMatrix[®] resin resulted in peptide **1** being obtained, with similar yield and purity to the acid functionalised peptide (Table 2-1).

Peptide **2** was mentioned in section 1.4.1.3 as a single-chain analogue of the peptide **1** dimer (Figure 1-15), that has previously been shown to have similar biochemical and antibacterial activity to **1**.⁴⁵ Peptide **2** was successfully synthesised with C-terminal amidation using NovaPEG Rink amide resin (Table 2-1). This peptide had

~2-fold lower yield when compared to **1**. This may be caused by the increased length of **2**, requiring a higher number of steps during SPPS that can introduce impurities commonly observed during the process. These can include the deletion of amino acids in the sequence, caused by unreacted sites on the resin beads, or insertion of additional amino acids, resulting from by the 4 – 5 equiv excess of amino acid reagents used. Additionally, diastereomeric impurities can occasionally be observed after unwanted racemisation of amino acids.¹²⁶

As mentioned in **section 1.4.1.2**, for peptide **1** to bind to the 4WJ the peptide must form a dimer, *via* a disulfide bridge between two cysteine residues. Using this information, two peptides were designed and synthesised that are unable to form this crucial disulfide bridge. Therefore, these peptides should be unable to bind to the 4WJ and can function as negative controls in assays. The first peptide designed was **5**, that contains an acetamidomethyl (Acm) protecting group on the cysteine, blocking the residue from potential disulfide bridge formation (**Figure 2-14**). The Acm group cannot be cleaved by treatment with piperidine and TFA, meaning it will remain attached to the peptide after SPPS. The second peptide was **6**, as this sequence matches peptide **1** with just the cysteine residue removed.

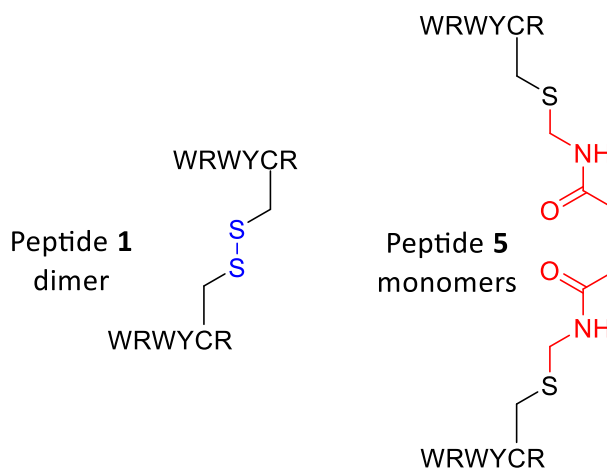


Figure 2-14: The chemical structures of the peptide **1** dimer with a disulfide bridge (blue), and two peptide **5** monomers containing Acm groups (red), that block disulfide bridge formation.

Both peptides were successfully synthesised using SPPS (**Table 2-1**). The low yield of **6** resulted from the co-elution of the product with impurities during RP-column chromatography, remaining after reducing the gradient from 0 – 100% to 0 – 40% MeCN in water with 0.05% TFA. To ensure the peptide was obtained with high purity, only one fraction in the centre of the product peak was collected, with impure fractions discarded, reducing overall peptide yield (**Figure 2-15**).

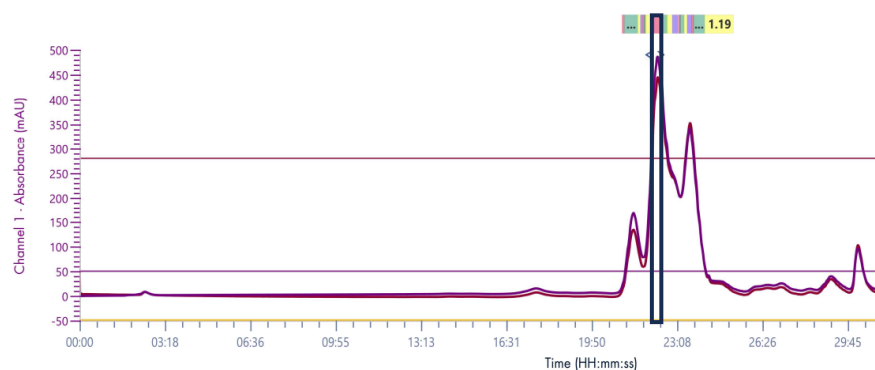
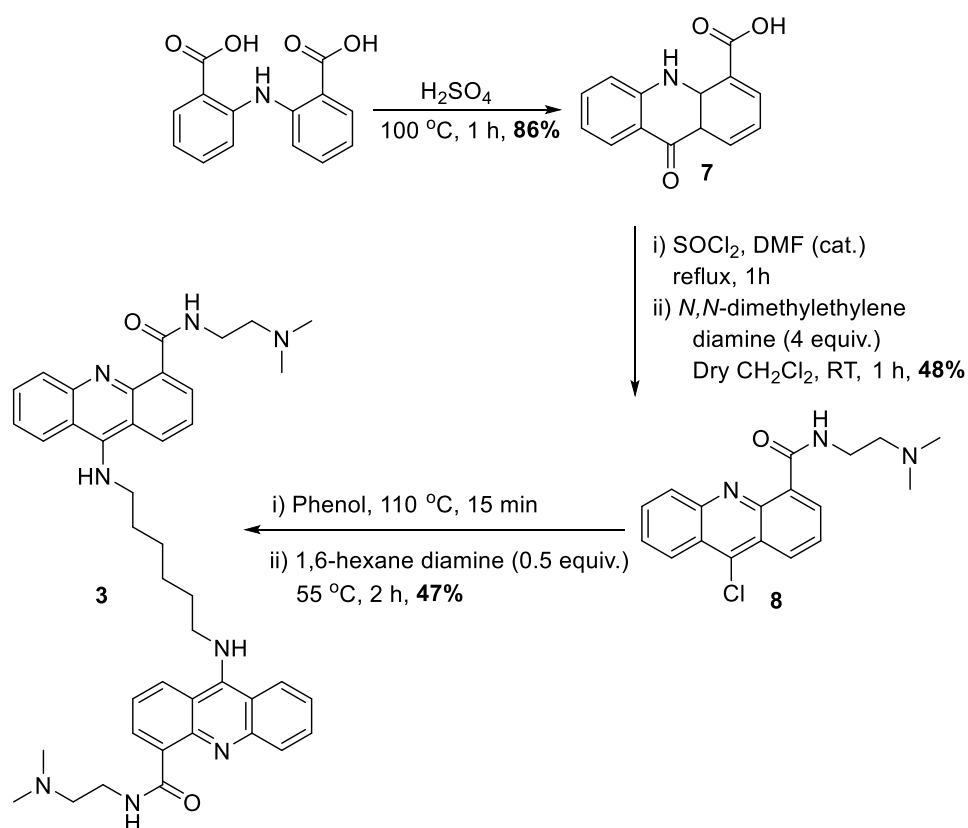


Figure 2-15: UV-VIS trace from RP-column chromatography of peptide **6**, showing 214 and 254 nm wavelengths, with a black box indicating the pure product fraction collected.

2.3.2 Synthesis of acridines

The dimeric C6 acridine **3**, mentioned in **section 1.4.3.1**, was developed by Searcey *et al.* and found to bind to the X-stacked conformation of the 4WJ through X-ray crystallography studies.⁸⁶ This compound was synthesised to allow for its 4WJ activity to be directly compared to the peptides.

The synthetic route shown (**Scheme 2-6**) for **3**, was taken from literature.¹²⁷ The synthesis began with the formation of an acridone from 2,2'-iminodibenzoic acid, by refluxing in sulfuric acid to give **7**. The intermediate 9-chloroacridine-4-carbonyl chloride was prepared by refluxing in thionyl chloride, before amidation with *N,N*-dimethylethylene diamine to afford **8**. Treatment with phenol resulted in the formation of a phenoxy intermediate, to act as a leaving group for the subsequent nucleophilic aromatic substitution (S_NAr) reaction with 1,6-hexane diamine, for the formation of the C6 acridine **3** as a bright orange solid (25 mg, 95% purity, 47% yield).

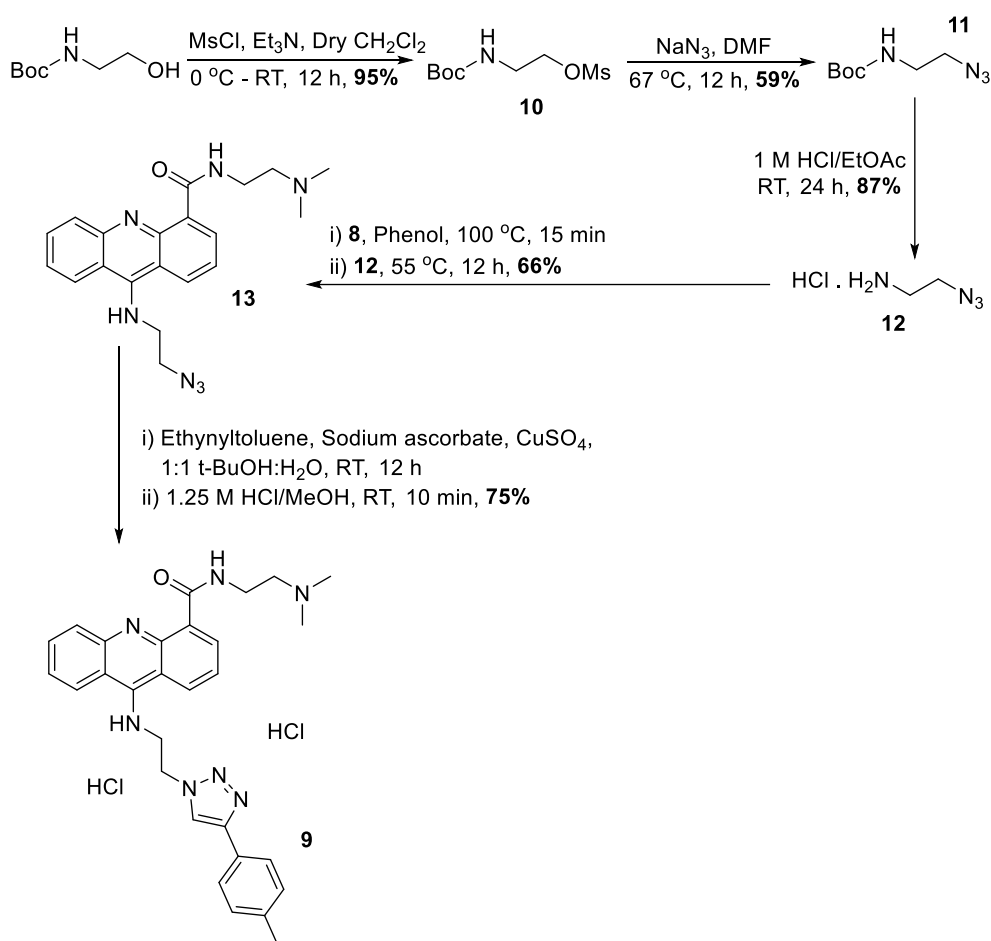


Scheme 2-6: Synthesis of C6 acridine **3**.¹²⁷

From previous screening of triazole-containing unsymmetrical acridine analogues, **9** was found to destabilise the 4WJ structure during a gel electrophoresis assay.⁸⁹ **9** was synthesised, to compare its activity against dsDNA and cancer cells to the 4WJ-binding peptides. This was to see if any of peptides showed similar results to **9**, that could indicate activity unrelated to 4WJ-binding. A five step synthesis was used to produce the dihydrochloride salt of **9**, taken from literature (**Scheme 2-7**).¹²⁷ The synthesis started with the formation of a mesylate, by adding methanesulfonylchloride (MsCl) drop-wise into a $0\text{ }^\circ\text{C}$ solution containing Boc-ethanolamine, before stirring at RT for 12 h, to give **10**. An $\text{S}_\text{N}2$ reaction then displaced the mesylate leaving group with an azide group, by heating sodium azide in DMF, to give **11**. The Boc group was deprotected under acidic conditions, in 1 M HCl/EtOAc, to afford azide **12** as a HCl salt. Following the final step from the synthesis of the C6 acridine dimer **3**, an $\text{S}_\text{N}\text{Ar}$ reaction was carried out by treating acridine **8** with phenol, before adding azide **12** to afford the azide acridine **13**. The final step in the synthesis was a click reaction with commercially available 4-ethynyltoluene, in the presence of sodium ascorbate and copper sulfate. Following treatment with

1.25 M HCl/MeOH, the di-HCl salt of triazole acridine **9** was obtained as a dark yellow solid (66 mg, 95% purity, 75% yield).

The successful synthesis of **9** was confirmed by the presence of a singlet in the ^1H NMR spectrum at 8.62 ppm, characteristic of the triazole CH group. Also, the azide signal at 2083 cm^{-1} for **13**, disappeared in the IR spectrum of **9**. Finally HRMS analysis (ES+) calculated for $\text{C}_{29}\text{H}_{32}\text{N}_7\text{O}$ ($\text{M}+\text{H}^+$) 494.2663 was found at 494.2668, further proving **9** was isolated.



Scheme 2-7: Synthesis of triazole acridine **9**.¹²⁷

2.3.3 PAGE assay

After the synthesis of different peptides and acridines, the first assay used to evaluate their ability to bind to the 4WJ was polyacrylamide gel electrophoresis (PAGE). As mentioned in **section 2.1.3**, gel electrophoresis was previously used to analyse the structures of open and X-stacked 4WJ conformations.⁵³ Subsequently, this technique was adapted to evaluate the binding affinity of ligands for 4WJ DNA.

For this, four 22 bp oligonucleotides were combined, based on sequences shown to arrange into a stable 4WJ (**Figure 2-16**).⁵³ Prior analysis showed these oligonucleotides formed a mixture of ssDNA and 4WJ DNA when annealed in a low salt buffer.¹²⁷ Due to the 4-fold size difference between the ssDNA and 4WJ DNA, these bands could easily be separated by gel electrophoresis, with the 4WJ moving more slowly in the gel. By looking at the shift in equilibrium, from a mixture of ssDNA/4WJ to 4WJ only, the concentration of salt/compound required to fully stabilise the DNA as the 4WJ could be detected.¹²⁷

To visualise the DNA bands on the gels, one of the DNA strands was tagged with two fluorophores. Strand X was labelled with FAM (6-carboxyfluorescein) on the 5' terminus and TAMRA (tetramethyl-6-carboxyrhodamine) at 3' terminus (**Figure 2-16**). This sequence was originally developed for Förster resonance energy transfer (FRET) analysis, that required a FAM and TAMRA label. For the PAGE assay only the FAM fluorophore was used for visualisation.¹²⁷ Owing to the fact that both the X-stacked and open conformation 4WJs have the same mass, their bands should travel a similar distance through PAGE gels. However, due to its more compact shape, compared to the extended planar shape of the open conformation, the X-stacked conformation can be predicted to have slightly increased electrophoretic mobility.¹²⁸

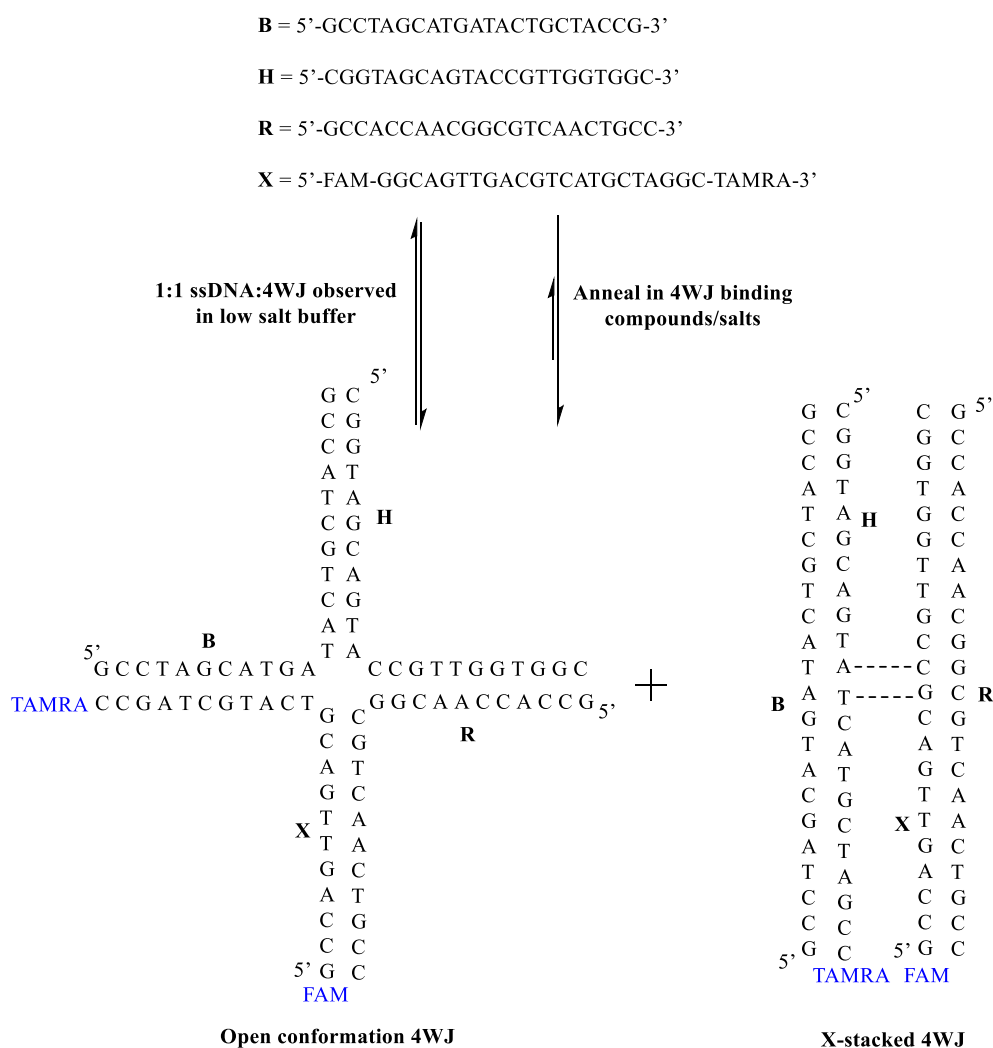


Figure 2-16: DNA sequences of oligonucleotides used in PAGE, showing how the addition of 4WJ binding compounds and salts causes the oligonucleotides to form more 4WJ, in either the open or X-stacked conformation.

The PAGE assay was initially carried out with C6 acridine **3**, using literature procedures.¹²⁷ To a series of Eppendorfs, containing equal amounts of the four oligonucleotides B, H, R and X (100 μ L, 1.25 μ M), were added different concentrations of **3** dissolved in DMSO (2 μ L), at 50-fold the desired final concentration (100 nM – 5 mM). The solutions were annealed in 1 x Tris-Borate-EDTA (TBE) buffer at 90 °C for 5 min followed by gradual cooling to 20 °C, over the course of 12 h. After this, gel electrophoresis was carried out at 50 V for 2 h, using 10% TBE gels and 1 x TBE as the running buffer. A gel loading dye was used to observe how far the bands had travelled in the gel. After the gel electrophoresis had finished, the bands were visualised by detecting FAM (460 nm), using a biomolecular imager. This showed that at 40 molar equiv all ssDNA formed into the 4WJ (**Figure 2-17**) and matched results previously obtained.¹²⁷ The 4WJ bands containing 40 – 80 molar

equiv of **3** were slightly higher than the previous 4WJ bands and may result from the acridine cross-linking 4WJs at higher concentrations. This could also account for the additional band at 16 molar equiv, that may be caused by the crosslinking of partially formed 4WJs.¹²⁷

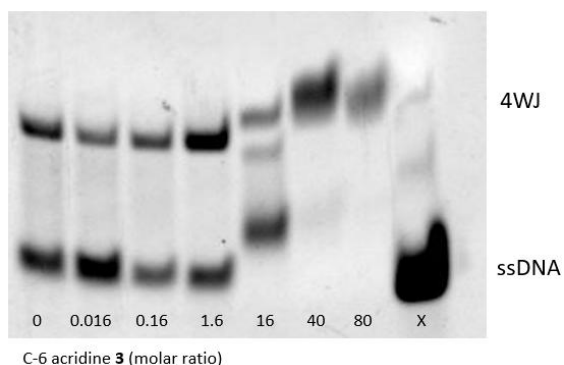


Figure 2-17: PAGE image showing the effect on 4WJ formation from increasing concentrations of C6 acridine **3**, using a DNA concentration of 1.25 μ M. A solution containing only the labelled X-strand was used as a control for ssDNA. FAM-labelled DNA was used and visualised at 460 nm.

To find the concentration of hexapeptide **1** and dodeca-peptide **2** required to fully trap the ssDNA as 4WJ DNA, a range of molar ratios were initially tested (0 – 320-fold final concentrations). This showed that **1** was able to fully trap the 4WJ at 40 molar equiv and **2** at 16 molar equiv. However, as the peptide concentrations increased, the DNA-peptide complex became trapped at the top of the gel wells (**Figure 2-18** – red boxes), causing the bands to become weaker. At the highest concentrations nothing was observed, indicating that the DNA remained in the Eppendorf after annealing. This could be caused by precipitation at higher peptide concentrations, or it may show aggregation of the complex.

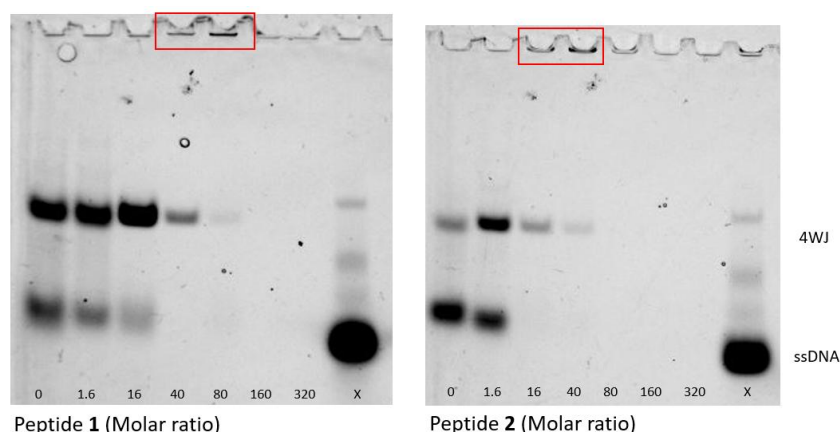


Figure 2-18: PAGE image showing the effect on 4WJ formation from increasing concentrations of peptide **1** (left) and peptide **2** (right), using a DNA concentration of 1.25 μ M. A solution containing only the labelled X-strand was used as a control for ssDNA. FAM-labelled DNA was used and visualised at 460 nm.

To avoid the disappearance of bands and get a more accurate concentration of peptide required for the 4WJ to fully form, a lower range of peptide concentrations was tested (0 – 16 molar equiv). From this it was determined that **1** fully trapped the 4WJ at 16 molar equiv and **2** at 9.6 molar equiv (**Figure 2-19**). This shows that **2** has higher activity against the 4WJ than **1** but lower solubility, as demonstrated by the slight fading of the band at 16 molar equiv in the peptide **2** gel.

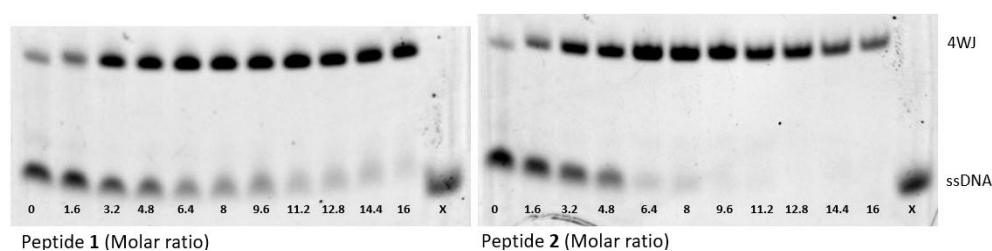


Figure 2-19: PAGE image showing the effect on 4WJ formation from lower concentrations of peptide **1** (left) and peptide **2** (right), using a DNA concentration of 1.25 μ M. A solution containing only the labelled X-strand was used as a control for ssDNA. FAM-labelled DNA was used and visualised at 460 nm.

The PAGE assay was repeated under the same conditions for peptides **5** and **6** (**Figure 2-20**). Up to 16 molar equiv, AcM-containing peptide **5** had minor effects on the intensities of the ssDNA and 4WJ bands, with the ssDNA band still clearly visible at the highest peptide concentration. This validates the hypothesis that **5** will be unable to bind to the 4WJ, because it cannot form the required peptide dimer, due to the AcM group on the cysteine blocking the formation of a disulfide bridge. Pentapeptide **6** did show some activity, with a slight increase in 4WJ band intensity from ~8 molar equiv. However, a faint ssDNA band was remaining at the highest concentration of peptide, showing **6** was unable to fully trap DNA as the 4WJ <16 molar equiv. This reduced activity suggests **6** could not maintain full 4WJ-binding interactions, when compared to peptide **1**.

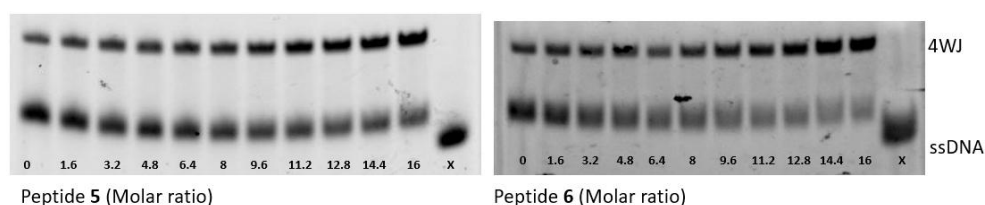


Figure 2-20: PAGE image showing the effect on 4WJ formation from increasing concentrations of peptide **5** (left) and peptide **6** (right), using a DNA concentration of 1.25 μ M. A solution containing only the labelled X-strand was used as a control for ssDNA. FAM-labelled DNA was used and visualised at 460 nm.

Both peptides **1** and **2** have an amide group on their C-terminus, following literature synthesis.⁴⁵ Replacing the amide group with a carboxylic acid at the C-terminus

caused a decrease in 4WJ activity in the PAGE assay, with the lower ssDNA band appearing faintly at 16 molar equiv (**Figure 2-21**). This shows that an amide C-terminus functionalisation is preferred by the WRWYCR sequence for 4WJ binding. This could be because the carboxylic acid group was negatively charged and repelled the DNA, whereas the amide group is neutral. Consequently, in later assays only the amide functionalised peptide **1** was tested.

Close examination of the peptide 4 PAGE image (**Figure 1-21**) revealed a slight retardation of the 4WJ band at a molar ratio of 3.2. This could be caused by the weight increase of the peptide-4WJ complex, compared to the 4WJ DNA on its own. Alternatively, this could signify a change in 4WJ conformation, resulting from peptide binding. This is unlikely to indicate a full change of the conformation, from open to X-stacked, which would result in the band travelling further through the gel.¹²⁸ Instead, a smaller localised structural rearrangement, such as a change in 4WJ branch angle causing changes in hydrodynamic radii,¹²⁹ could explain the band retardation.

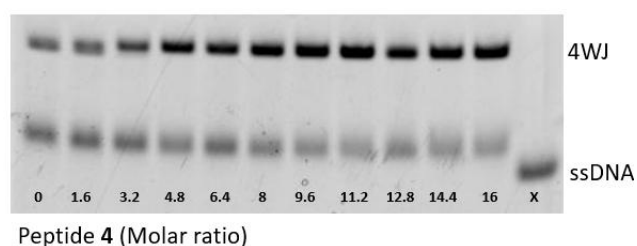


Figure 2-21: PAGE image showing the effect on 4WJ formation from increasing concentrations of peptide **4**, using a DNA concentration of 1.25 μ M. A solution containing only the labelled X-strand was used as a control for ssDNA. FAM-labelled DNA was used and visualised at 460 nm.

To determine whether having fluorescent dyes attached to the DNA effected peptide binding, PAGE was carried out using unlabelled DNA and dodeca-peptide **2**, followed by staining with SYBR[™] gold and visualisation at 460 nm. Due to the high sensitivity of the SYBR[™] gold stain for DNA detection, the final concentration of 4WJ DNA was reduced from 1.25 μ M to 250 nM. PAGE using the unlabelled DNA showed that most of the DNA was trapped as the 4WJ at ~9.6 molar equiv (**Figure 2-22**), the same concentration observed when using fluorescently labelled DNA (**Figure 2-19**). Faint bands were still visible for non-4WJ DNA > 9.6 molar equiv however this is likely to be trace amounts, with the SYBR[™] gold stain able to detect as little as 1 pM of DNA.¹³⁰ This establishes that the presence of the fluorescent tags on the DNA had minimal effect of peptide binding.

The SYBR™ gold-stained gel (**Figure 2-22**) clearly has two lower bands, one that is consistent with ssDNA and the other with forked-DNA, a DNA construct formed by two strands that share 50% complementarity. In the absence of any peptide a mixture of forked DNA/4WJ is mostly present. This is to be expected due to the increased stability of DNA when in a duplex.

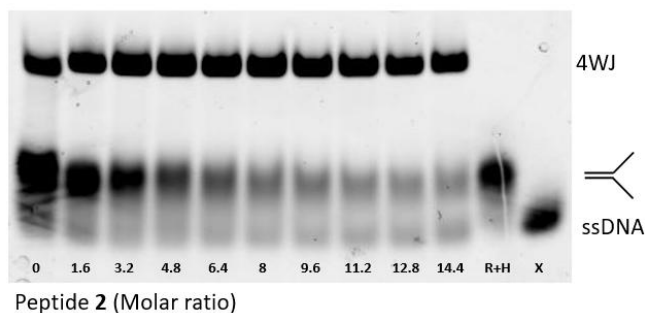


Figure 2-22: PAGE image using unlabelled DNA, staining with SYBR™ gold before visualisation at 460 nm. This shows the effect on 4WJ formation from increasing concentrations peptide **2**, using a DNA concentration of 250 nM. A solution containing only the R-strand was used as a control for ssDNA and another containing annealed R+H strands for forked-DNA.

As mentioned in **section 1.4.3.1**, RTA is an acridine with the ability to induce the formation of the 4WJ without the need for annealing.⁹⁰ To investigate the activity of **2** on 4WJ formation at reduced temperature, different concentrations of the peptide were incubated with the 4 unlabelled oligonucleotides (B, H, R and X), for 1 h at RT (**Figure 2-23**). PAGE analysis, using SYBR™ gold staining, showed that all DNA was trapped as the 4WJ at 24 molar equiv, a 2.5-fold reduction in activity compared to when the experiment was carried out with annealing. This may reflect a reduced binding rate due to the lower availability of thermal energy at reduced temperatures, which results in slower molecular motion and hinders the conformational flexibility required for effective DNA–peptide binding.

In **Figure 2-23**, two images of PAGE gels are shown, for the binding of **2** to 4WJ DNA without annealing. Notably, when no peptide is present in solution, image **A** has only ssDNA present and image **B** has ssDNA and 4WJ DNA present. Difference in the ratio of ssDNA and 4WJ DNA was often observed when no other compound was present, throughout the PAGE analysis. This could show that the ratio of the different species is highly sensitive to minor experimental differences in buffer composition, oligonucleotide ratios or changes in room temperatures. However, in **Figure 2-23** both images show that all DNA is trapped as the 4WJ at 24 molar equiv of **2**, showing that results in binding activities can be repeated. Despite this, it should be mentioned

that in most cases the PAGE assay was carried out once for each compound, meaning that reproducibility cannot be discussed.

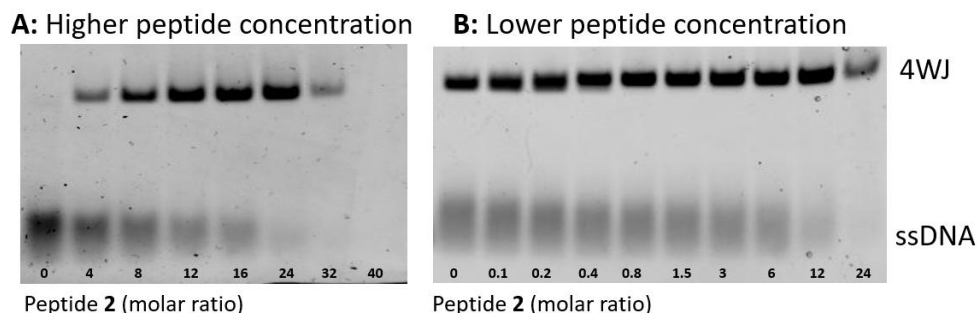


Figure 2-23: PAGE images obtained without annealing. This showed the effect on 4WJ formation from increasing concentrations peptide **2** after incubating in the presence of DNA (250 nM) for 1 h at RT, with image **A** having a higher range of peptide concentrations and image **B** a lower range. Unlabelled DNA was used and stained with SYBR™ gold, before visualisation at 460 nm.

The focus then moved to using PAGE to examine the selectivity of **2** for 4WJ DNA compared to dsDNA. A competition assay was set up, where the peptide first binds to the 4WJ DNA, before dsDNA is titrated into the solution. Successful displacement of the peptide from the 4WJ, caused by dsDNA binding, would result in a reduction in the 4WJ band and increase in the lower DNA bands. For this experiment, 12 molar equiv of peptide **2** were used, to allow for the majority of the DNA to be trapped as the 4WJ, whilst avoiding solubility issues. After incubating the peptide with the 4 unlabelled oligonucleotides (B, H, R and X) for 1 h at RT, varying concentrations of dsDNA was added. The dsDNA used was 25 bp in length and contained no sequences capable of hybridising with the 4WJ DNA strands, avoiding the formation of unwanted DNA species. The resultant solution was incubated at RT for a further 1 h, before analysis by PAGE followed by SYBR™ gold staining.

Results from the 4WJ/dsDNA PAGE competition assay, using **2**, showed there was no change in the strength of lower DNA bands with increasing concentrations of dsDNA, up to 25 molar equiv. Also, the 4WJ band showed no reduction in strength, in fact increased with dsDNA concentrations >3.13 molar equivalents, indicating potential issues with DNA cross-linking (**Figure 2-24**). These outcomes show that the dsDNA was unable to compete for **2**, demonstrating the peptides selectivity for 4WJ binding. However, it should also be noted that the DNA sequences used for the 4WJ, and dsDNA are very different meaning that sequence selectivity cannot be ruled out.

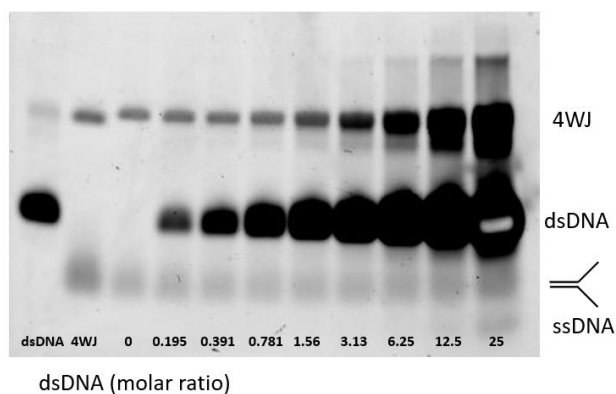


Figure 2-24: 4WJ/dsDNA PAGE competition assay. This showed the inability of dsDNA to compete for peptide binding, in the presence of 12 molar equiv peptide **2** and 250 nM 4WJ-forming DNA, after incubation at RT. Unlabelled DNA was used and stained with SYBR™ gold, before visualisation at 460 nm.

Attempts were made to use the well-studied J3 4WJ (**Figure 2-12**) for the PAGE assay. This 4WJ has 17 bp arms, compared to the 11 bp arms of the 4WJ used so far. When the four J3-forming oligonucleotides were annealed in 1 x TBE, in the absence of any salt or 4WJ-binding compounds, the DNA was found to fully form the 4WJ. This may reflect the increased stability of this 4WJ due to its longer duplex arms and means that J3 cannot be used to monitor changes in 4WJ stability in the presence of compounds of interest using the PAGE technique.

To summarise, the ability of the C6 acridine **3** and peptides to bind and trap the DNA 4WJ was evaluated using PAGE. From testing the peptides, it was found that **2** had the highest activity, followed by **1**. In contrast, the other 3 peptides tested showed incomplete 4WJ trapping <20 μ M. When comparing peptide **2** and the C6 acridine **3**, the peptide showed ~4-fold higher activity, exhibiting its higher affinity for 4WJ binding (**Table 2-2**). Additionally, the versatility of PAGE was demonstrated, with the assay being used to monitor the effects of fluorescent labelling, annealing and the addition of dsDNA, on the ability of peptide **2** to bind to 4WJ DNA.

Table 2-2: Results from PAGE assay, using 1.25 μM FAM labelled DNA.

No.	Name/sequence	Gel trapping (Molar equiv)	Gel trapping (μM)
3	C6 acridine	40	50
1	H-WRWYCR-NH ₂	16	20
2	H-WRWYRGGRYWRW-NH ₂	9.6	12
4	H-WRWYCR-OH	>16	>20
5	H-WRWYC(Acm)R-NH ₂	>16	>20
6	H-WRWYR-NH ₂	>16	>20

2.3.4 FRET

To find out whether these peptides bound to the open or X-stacked conformation of the 4WJ, a FRET assay was carried out, using similar conditions to previous studies.^{123,125} As mentioned in **sections 1.4.1.2 – 1.4.1.3**, both peptides **1** and **2** have been shown to bind to the open conformation of the 4WJ. Therefore, when these peptides bind to the 4WJ in a high salt environment a decrease in FRET can be predicted, caused by the change in conformation from the X-stacked to open form. To follow other FRET studies on 4WJs, J3 was used, with FAM as the donor fluorophore and TAMRA as the acceptor fluorophore.¹²⁵ These dyes are commonly used for FRET due to the good spectral overlap of FAM emission and TAMRA excitation. By exciting FAM at 490 nm, energy can be transferred to TAMRA, resulting in a loss of FAM emission at 520 nm and an increase in TAMRA emission at 580 nm (**Figure 2-25**).

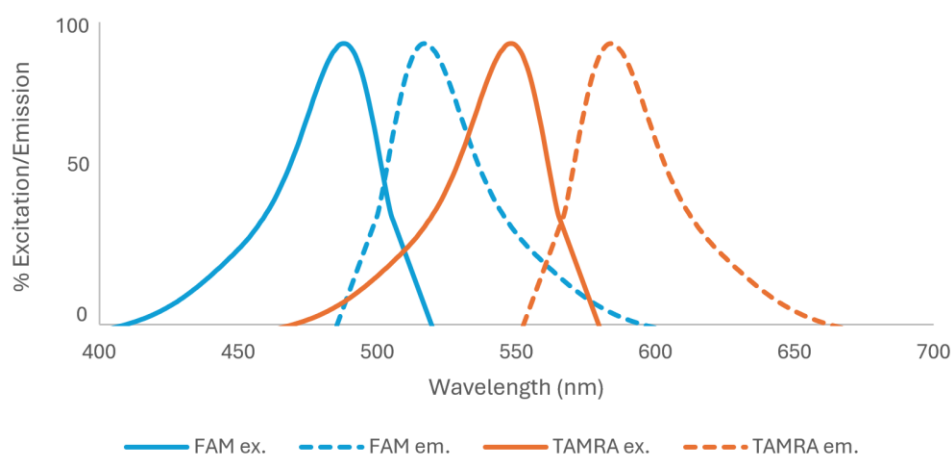


Figure 2-25: Excitation and emission spectra of FAM and TAMRA fluorophores.

The J3 oligonucleotide sequences and 4WJ combinations are as follows:

Sequences:

B = 5'-CCTCCGTCCTAGCAAGGGGCTGCTACCGGAAGGG-3'

H = 5'-CCCTTCCGGTAGCAGCCTGAGCGGTGGTTGAAGG-3'

R = 5'-CCTTCAACCACCGCTCAACTCAACTGCAGTCTGG-3'

X = 5'-CCAGACTGCAGTTGAGTCCTTGCTAGGACGGAGG-3'

X(TAMRA) = 5'-TAMRA-CCAGACTGCAGTTGAGTCCTTGCTAGGACGGAGG-3'

R(FAM) = 5'-FAM-CCTTCAACCACCGCTCAACTCAACTGCAGTCTGG-3'

4WJs:

(F+T) FAM and TAMRA = X(TAMRA)/R(FAM)/B/H

(F) FAM-only = X/R(FAM)/B/H

(T) TAMRA-only = X(TAMRA)/R/B/H

To begin with, the titration of NaCl, into a low salt solution of J3 DNA was repeated, following procedures previously described.¹²³ Na⁺ was chosen as the titrant because it forms the tightest junctions when compared to Mg²⁺ and K⁺, allowing for the highest levels of FRET.¹²⁵ Firstly three 4WJs, based on J3 sequences and containing differing fluorescent labels, were constructed *via* annealing at 90 °C for 5 min followed by gradual cooling to 20 °C over 12 h in 10 mM Tris, 1 mM EDTA and 300 mM NaCl (pH 7.6). The high level of NaCl added was to ensure all DNA formed the 4WJ whilst annealing. Each 4WJ was then buffer exchanged 3 times into a low salt buffer containing 10 mM Tris-HCl and 0.1 mM EDTA (pH 7.6).

Experiments were performed using a 3.5 mL cuvette with a 1 cm path length in a fluorimeter, with an excitation bandwidth of 4 nm and an emission bandwidth of 8 nm. Measurements were taken at 20 °C, with the polarisers set to 0° for excitation and 55° for emission. The “magic angle” was used for the emission polarisation to avoid artifacts caused by anisotropy effects.¹³¹ The FAM and TAMRA (F+T) and FAM-only (F) labelled junctions were excited at 490 nm and the TAMRA-only (T) junction was excited at 565 nm. All emission spectra were recorded from 500 – 650 nm at 0.2 nm/s. For the NaCl titration, to each 4WJ (50 nM, 2 mL) was added increasing volumes of NaCl (5 M). The titration was repeated in triplicate and control emission scans, with NaCl titrated into buffer, were subtracted from the DNA emission scans.

FRET efficiency (E) provides a ratio of the acceptor emission after donor excitation or acceptor excitation. This can be calculated using the following equation:

$$E_{580} = F_{DA} \div F_A$$

Equation 2-2: FRET efficiency (E) calculation at 580 nm.

Where F_{DA} is the emission of the acceptor (TAMRA) in the presence of the donor (FAM), after excitation at 490 nm, and F_A is the emission of the acceptor after excitation at 565 nm (**Equation 2-2**).

The contribution of FAM emission to the F_{DA} value for each NaCl addition was removed by first calculating the conversion factor required to match the spectra of **F+T** to **F** between 500 – 530 nm, where only FAM emission is observed (**A, Figure 2-26**). This conversion factor was then used to multiply each wavelength of **F+T** to give the “calculated fit” (**B, Figure 2-26**).

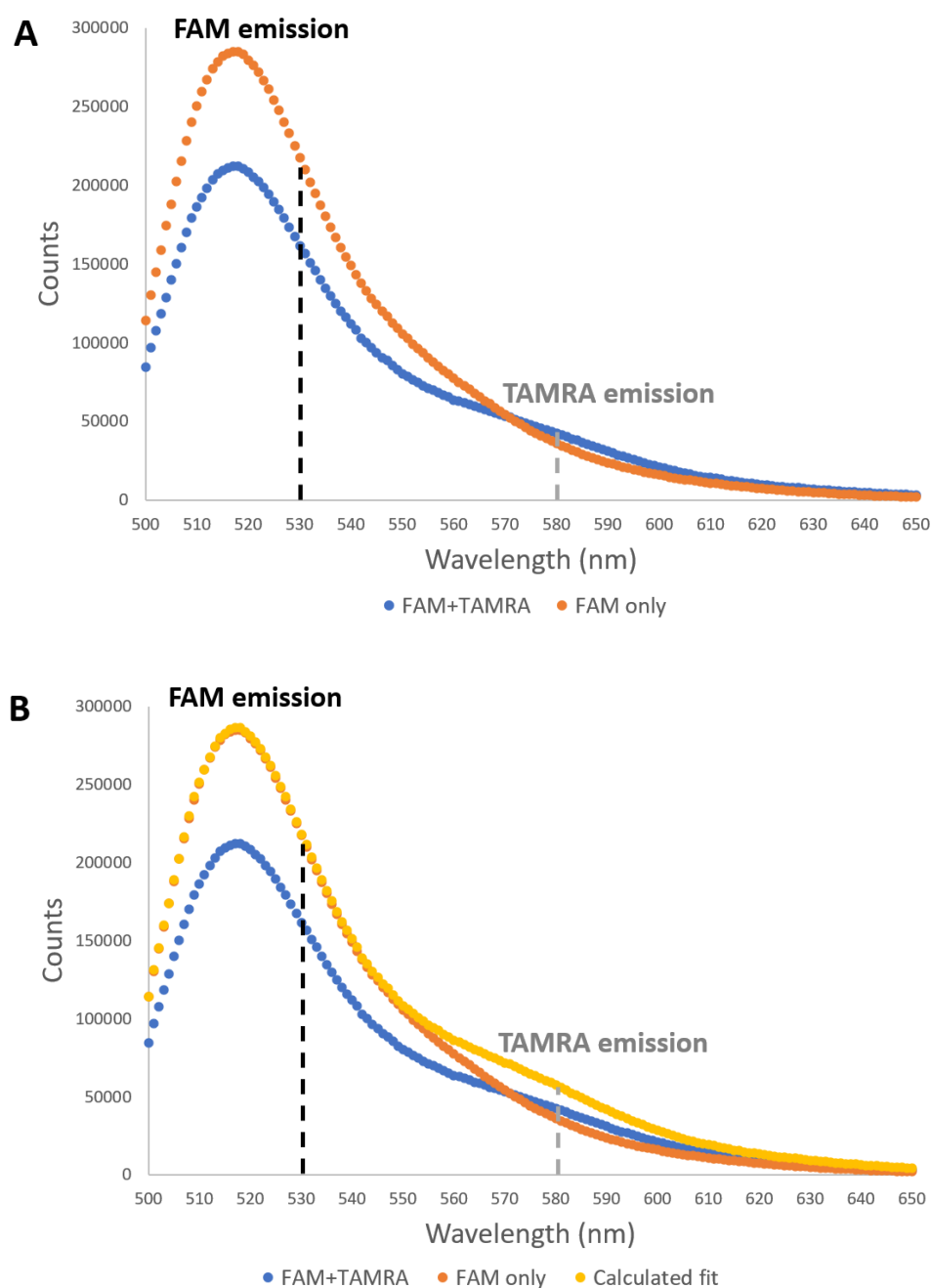


Figure 2-26: Example of the deconvolution of FAM and TAMRA signals during FRET analysis, using J3 after the addition of 200 mM NaCl. **(A)** Emission spectrum of **F+T** (blue) and **F** (orange). **(B)** Emission spectrum with the “calculated fit” (yellow) added, after finding the conversion factor (1.34) to match the **F+T** and **F** spectra from 500 – 530 nm, then multiplying every **F+T** value by this.

The values at each wavelength of **F** were then subtracted from the **F+T** “calculated fit” spectrum to provide the “extracted acceptor emission” (**A**, **Figure 2-27**).¹²³ The “extracted” value at 580 nm was then divided by the **T** emission value at 580 nm, after excitation at 565 nm to provide the FRET efficiency (**B**, **Figure 2-27**). The

average values of triplicates from **F** and **T** were used for the calculations and each **F+T** value was analysed separately.

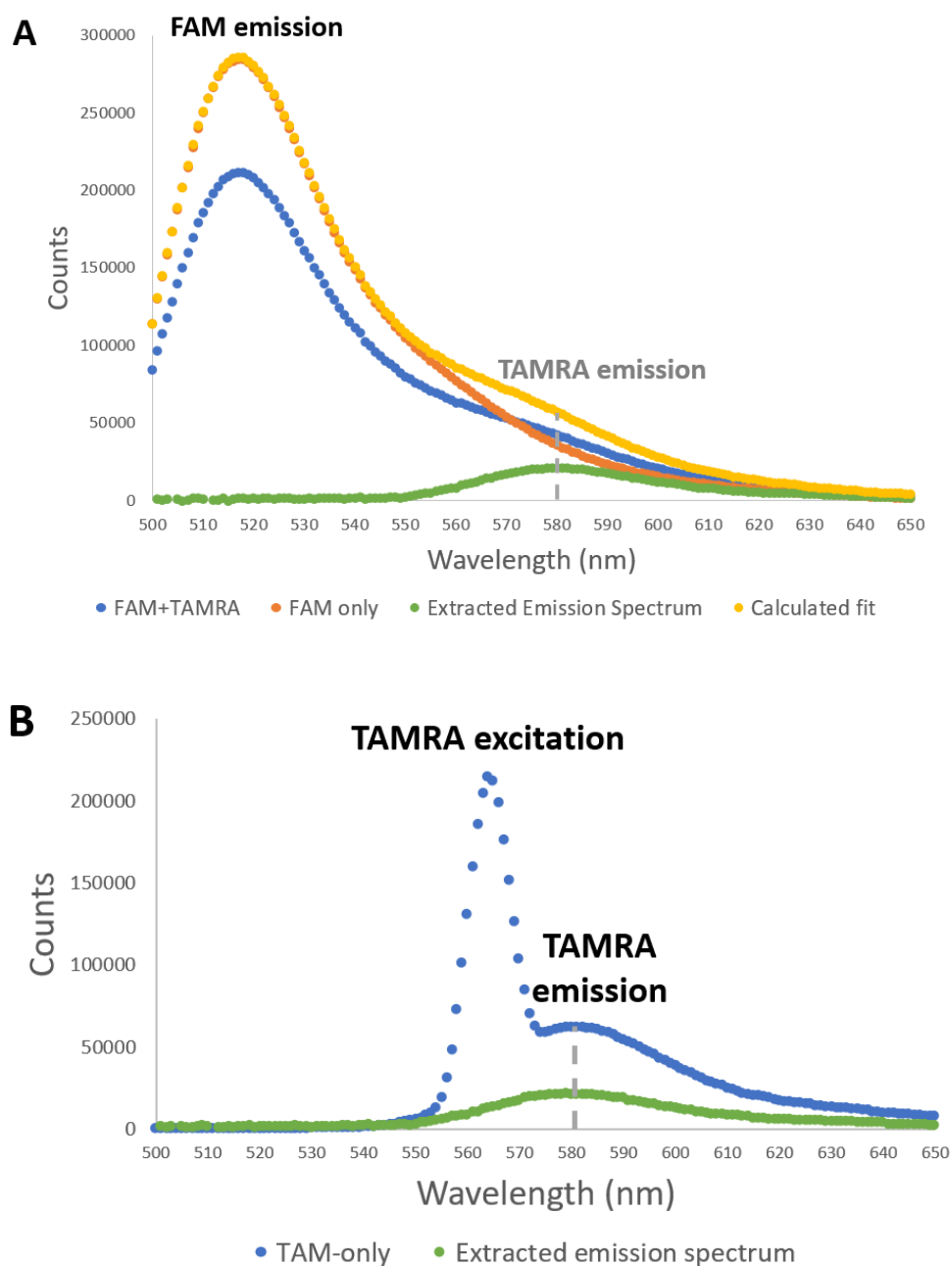


Figure 2-27: Example of FRET data analysis, using J3 after the addition of 200 mM NaCl. **(A)** Emission spectrum with the “extracted acceptor emission” (green) added, after taking away the emission of **F** from each wavelength of the “calculated fit”. **(B)** Emission spectrum of **T** compared to the “extracted acceptor spectrum”.

An automated method for carrying out these calculations was developed in collaboration with Marco Cominetti, using Python coding. This allowed for the negative controls, of NaCl titrated into buffer, to automatically be subtracted for emission values after either 490 nm or 565 nm excitation. The conversion factor was

calculated by multiplying the 500 – 530 nm values of the **F+T** 4WJ by every number between 0.01 – 100, until an R-squared value closest to 1 was obtained, when compared to the **F** 4WJ values. The value for the conversion factor was calculated for each **F+T** scan individually, using the averaged **F** scans from triplicate readings.

The results from the titration of NaCl into the J3 showed that FRET increased as the concentration of NaCl increased, with a dissociation constant of 24.4 ± 2.47 mM (**Figure 2-28**). This shows how the 4WJ is changing from the open conformation to the X-stacked conformation as the concentration of salt increases. Due to the high volume of NaCl (up to 500 μ L) required to reach a plateau, DNA blanks were also carried out, where buffer was added directly to each of the 3 different 4WJs. This showed that the dilution effect of the NaCl solution had no effect on the FRET values obtained. This is because the FRET value was obtained from the ratio of emission at 580 nm from the doubly labelled FAM and TAMRA 4WJ DNA and the singly labelled TAMRA only junction, that would have been equally affected by dilution.

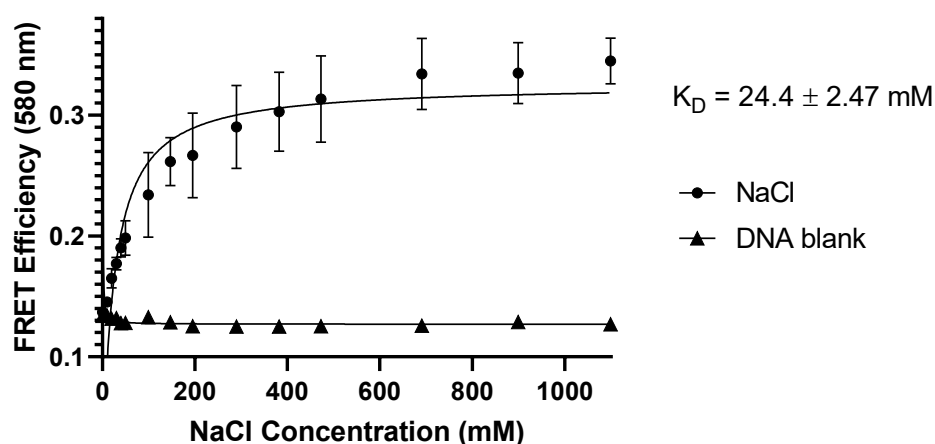


Figure 2-28: FRET efficiency of J3 (50 nM) upon the addition of 5 M NaCl aq. (circles) and buffer only (triangles). Error bars and K_D error represent the mean $\pm$ standard deviation from triplicate readings.

With the curve of the line becoming relatively flat at ~ 600 mM, it was decided that 300 mM ($\sim 50\%$ of $B_{\max}$) would be put into the DNA solution when testing the peptides. This was so the majority of 4WJ would be in the X-stacked conformation, meaning the peptide binding would be observed by a decrease in FRET, as the 4WJ is opened. A value below 100% $B_{\max}$ was chosen to reduce the amount of peptide required to compete against NaCl, as well as the effects of non-specific binding of NaCl to the duplex arms of the 4WJ.

Dodeca-peptide **2** has previously been shown to bind to the open conformation of the 4WJ,⁴⁵ and was the first peptide tested in FRET assay. All conditions were kept the same as for the NaCl titration, except that the J3 DNA was dissolved in a high salt buffer, containing 300 mM NaCl, 10 mM Tris-HCl and 0.1 mM EDTA (pH 7.6). The peptide was dissolved in this same buffer without NaCl, due to reduced solubility in the presence of the salt. To the 4WJ DNA (50 nM, 2 mL) was added aliquots of peptide **2** (1 μ L) at varying stock concentrations (10 – 1000 μ M). A total of 12 peptide additions were made making the change in volume <1%, reducing the effect of changing volume on NaCl or DNA concentrations.

The results from the titration of **2** into J3 showed that FRET decreased with increasing peptide concentration (**Figure 2-29**). A K_D value of 96.4 ± 22.8 nM was obtained, showing the high affinity of the peptide for the 4WJ DNA. This could indicate that the 4WJ was being converted from the X-stacked to the open conformation, in the presence of this peptide. This also shows that **2** can recognise the 4WJ in the X-stacked conformation and outcompete NaCl for binding.

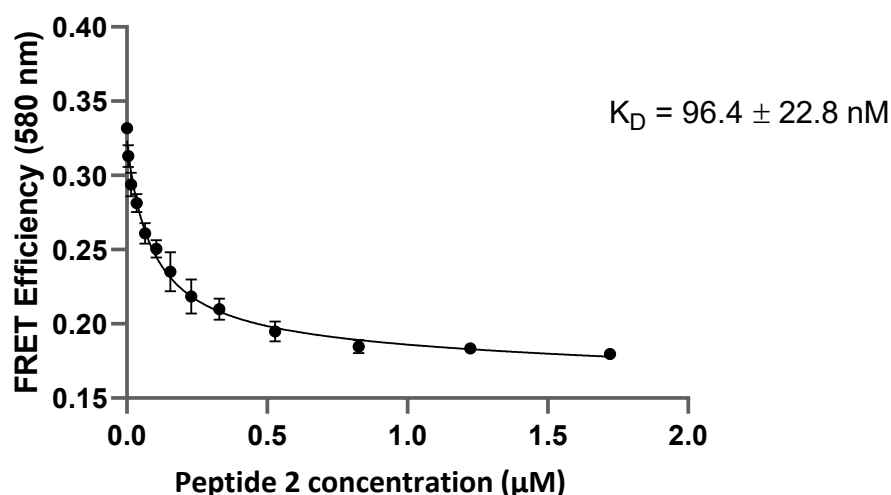


Figure 2-29: Change in the FRET efficiency of J3 (50 nM), dissolved in 300 mM NaCl, 10 mM Tris-HCl and 0.1 mM EDTA (pH 7.6), with the titration of peptide **2**. Error bars and K_D error represent the mean $\pm$ standard deviation from triplicate readings.

To confirm that the results from FRET analysis of **2** were caused by 4WJ-binding, AcM-containing peptide **5** was tested next. As mentioned in **section 2.3.1**, this peptide is predicted to be unable to bind to the 4WJ because it cannot form the cysteine disulfide bridge required for activity. **5** was shown to have minimal effect on the FRET efficiency of J3 (**Figure 2-30**), up to a concentration of 1.7 μ M. This confirms that **5** was unable to convert the 4WJ from the X-stacked to open

conformation. However further analysis is required to show this peptide does not bind to the X-stacked 4WJ.

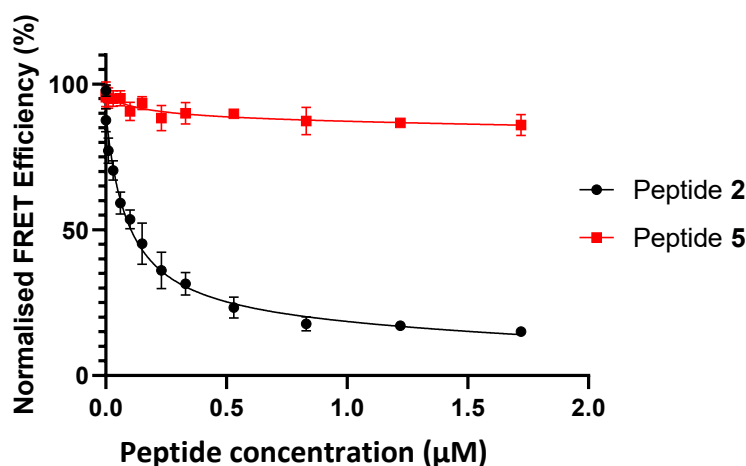


Figure 2-30: Change in the FRET efficiency of J3 (50 nM), dissolved in 300 mM NaCl, 10 mM Tris-HCl and 0.1 mM EDTA (pH 7.6), with the titration of peptide **2** (black circles) or peptide **5** (red squares). Error bars represent the mean $\pm$ standard deviation from triplicate readings.

After confirming the validity of the FRET assay, next hexapeptide **1** was tested, as this peptide has also been shown to bind to the open conformation 4WJ.⁴⁵ When testing **1**, a small decrease in FRET efficiency was observed, with 69% still observed at 1.72 µM, compared to 0% for peptide **2** (**Figure 2-31**). A linear drop in FRET was observed for **1**, compared to the exponential decrease of **2**. This could be caused by the time required for the equilibrium shift of **1**, from the inactive monomer to the 4WJ-binding dimer. Each emission scan was taken ~10 sec after the peptide addition and may not have been long enough to form the active dimer. For each point of the peptide titration, 3 emission scans were required with the different DNA constructs and 2 emission scans of peptide blanks, in the absence of DNA. With each emission scan taking 30 sec, and triplicates needed, it would be highly inefficient to increase the length of time between additions. Therefore, other assays may be more appropriate for testing **1**, such as gel electrophoresis and FP, that both involve incubation of the compound for longer periods of time.

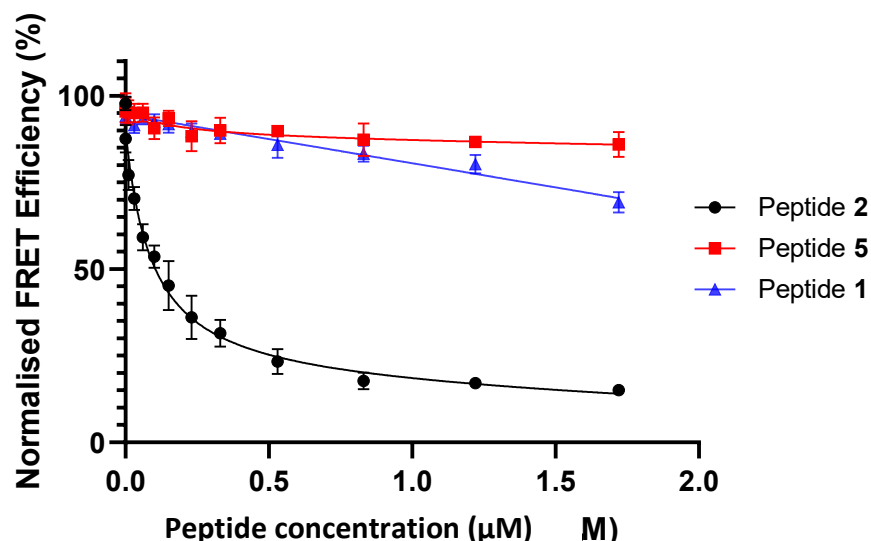


Figure 2-31: Change in the FRET efficiency of J3 (50 nM), dissolved in 300 mM NaCl, 10 mM Tris-HCl and 0.1 mM EDTA (pH 7.6), with the titration of peptide **2** (black circles), peptide **5** (red squares) or peptide **1** (blue triangles). Error bars represent the mean $\pm$ standard deviation from triplicate readings.

Unfortunately, acridines could not be tested under the conditions used for the FRET assay, due to their high levels of intrinsic fluorescence, with acridines often used as fluorescent dyes for the detection of DNA.¹⁰⁶ With a low concentration of 50 nM fluorescently labelled DNA used for FRET, interference from acridine fluorescence occurred at low concentrations, making this assay non-viable. This was not an issue in the PAGE assay, as any unbound acridine was washed out of the gel because of its small size (<1 kDa).

To summarise, FRET was used to test the effect of peptides on the 4WJ conformation. From this, **2** was found to have a K_D of 90 nM for converting the 4WJ from X-stacked to open in a high salt environment. On the other hand, **1** and **5** both showed minimal effect <1.7 μM.

Although the results from this experiment were presumed to result from FRET, based on the approach originally described by Lilley *et al.*,¹²³ a potential limitation is that TAMRA emission following excitation at 490 nm was not explicitly accounted for. To conclusively demonstrate that FRET is occurring between FAM and TAMRA, further experiments are required. These should include measuring emission at 580 nm after excitation at 490 nm for both the **F+T** and **T** 4WJs, to assess any direct excitation of TAMRA and confirm that the observed signal is due to energy transfer rather than from directly emission from TAMRA. Until such controls are performed, the

interpretation of these results as definitive evidence of FRET must be treated with caution.

2.3.5 Fluorescence polarisation

FP was previously carried out to find the binding affinity of proteins for the 4WJ, using J3 4WJ DNA as the fluorescent tracer molecule. Lilley's group showed the binding of the J3 4WJ to two different 4WJ resolvase enzymes; *E. Coli* RuvC and *Sulfolobus Solfataricus* Hjc.¹³² During this study, J3 was labelled with either Cyanine3 or FAM, terminally attached to one of its helical arms. When the proteins were titrated into separate solutions containing 25 nM J3, increases in FP were observed. This is indicative of binding and matches the FRET and gel electrophoresis results obtained during this study.¹³²

Mukerji's group showed the ability of Integration Host Factor (IHF) to bind to the DNA 4WJ using FP.¹³³ IHF is an architectural protein found in *E. Coli*, that has sequence specific duplex DNA-binding. Upon binding, IHF induces DNA bending, with an angle of ~160° observed in IHF-DNA co-crystal structures. When carrying out FP using 2 nM FAM-labelled J3, it was shown that IHF had a K_D of 2 nM. This study proves that IHF binds to the DNA 4WJ with high affinity and no sequence specificity, showing this protein recognises the overall 4WJ structure.¹³³

The proteins RuvC, Hjc and IHF bind to the 4WJ as dimers, with overall molecular weights of 19 kDa, 30 kDa, and 21 kDa respectively, making them lighter than the 40 kDa J3 4WJ. This is unusual for FP studies, where traditionally the smaller binding partner acts as the tracer, with the binding of a larger molecule causing a more substantial increase in polarisation.¹³⁴ However, these studies from both the groups of Lilley¹³² and Mukerji,¹³³ show it is possible to observe the binding of smaller proteins to larger fluorescently-labelled 4WJ DNA, with these results backed-up by orthogonal assays, such as FRET, gel electrophoresis and single-molecule techniques.

With the availability of J3 DNA after FRET analysis, an assay was set up using FAM-labelled J3 as a fluorescent tracer and monitoring changes in FP upon the addition of different 4WJ-binding peptides (**Figure 2-32**). The DNA sequences matched those used in FRET for the FAM-only junction, containing X/R(FAM)/B/H oligonucleotides. The FAM-J3 tracer was used at a final concentration of 10 nM, to

allow for the use of minimal reagents whilst getting a strong enough signal for high quality results. As with the FRET assay, FP could not test the binding of acridine molecule due to their intrinsic fluorescence and the low concentration of fluorescently labelled DNA used.

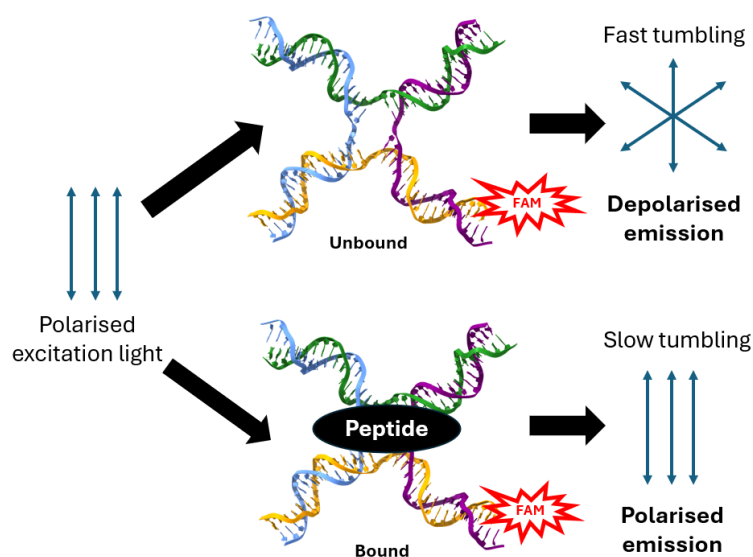


Figure 2-32: FP schematic, showing unbound 4WJ DNA, tagged with FAM (red), tumbling quickly in solution with low polarisation. When a 4WJ-binding peptide molecule (black) binds to the 4WJ, tumbling is slowed, and polarisation increases.

For the FP assay, a plate reader was used, enabling rapid analysis and the generation of full binding curves for multiple peptides simultaneously. This contrasts with the FRET assay, where a fluorimeter was used to analyse samples in a cuvette. Although this was a slower and more labour-intensive technique, using a cuvette gave more sensitive and precise results.¹⁰⁶ This was required during FRET because the signal was weaker than direct fluorescence intensity signals, with the acceptor and donor fluorophores needing to be within ~10 nm of each other for energy transfer to occur. Due to this weak signal, background noise from other sources of fluorescence had to be reduced. By using a cuvette, emission measurements could be made perpendicular to the excitation beam, minimising detection of excitation light. The cuvettes also had a fixed path length of 1 cm, whereas variation caused by meniscus effects or slight volume changes between wells in microplates can reduce the accuracy of fluorescence readings. Finally, fluorimeters can analyse exact wavelengths with narrow bandwidths, unlike the plate reader available during this PhD, that contained a limited set of filters with wider bandwidths. By using a fluorimeter, wavelengths were chosen that maximised spectral overlap between the donor and acceptor fluorophores whilst reducing interference through direct

excitation.¹⁰⁶ For the FP assay, strong fluorescence intensity signals meant that the satisfactory results could be obtained quickly using a plate reader, with minimal interference from background noise.

Initial FP experiments used a buffer containing only 1 mM phosphate-buffered saline (PBS) and 10 mM NaCl (pH 6.5) and gave highly variable results with large error bars. However, after the addition of 0.05% of the detergent Tween 20, improvements were observed in the reproducibility of the assay, presumably due to the ability of Tween 20 to reduce the adsorption of biomolecules to the well surfaces.¹¹³ The FP assay was prepared by pipetting 40 μ L of 20 nM FAM-J3 into a black 96-well plate, before adding 40 μ L of peptide (25 nM – 3 μ M final concentrations). Both DNA and peptide solutions were dissolved in the same buffer, containing 1 mM PBS, 10 mM NaCl and 0.05% Tween 20 (pH 6.5). After incubation for 20 min, the samples were excited at 482 ± 16 nm and emission recorded at 530 ± 40 nm. FP was calculated using **Equation 2-1**, to provide half maximal effective concentration (EC_{50}) values, the concentration required to get a 50% response in polarisation, in triplicate readings (**Table 2-3**).

Table 2-3: Results from FP using 10 nM FAM-J3 4WJ. Error in the EC_{50} and polarisation change values are from the mean $\pm$ standard deviation from triplicate readings.

No.	Sequence	EC_{50} (nM)	Polarisation change (mP)
2	H-WRWYRGGRYWRW-NH ₂	310 ± 2.54	249 ± 0.939
1	H-WRWYCR-NH ₂	553 ± 30.1	202 ± 12.8
5	H-WRWYC(Acm)R-NH ₂	>3000	N/A
6	H-WRWYR-NH ₂	>3000	N/A

Results from the FP assay (**Figure 2-33**) showed that peptide **2** had the highest activity and caused the largest change in polarisation, with an increase of 249 ± 0.939 mP observed. This was surprising, given that **2** weighs ~20 times less than the J3 DNA, at 2 kDa compared with 40 kDa. This large increase in polarisation could be explained by the peptide potentially causing the 4WJ to adopt a more rigid conformation or could signify that multiple peptides have bound to the 4WJ at once. Compared to **2**, peptide **1** had ~2-fold lower activity and showed a smaller increase in polarisation of 202 ± 12.8 mP. This hexapeptide is known to bind to the 4WJ as a

dimer, therefore the reduction in activity may reflect the need for two molecules to bind to a single 4WJ, unlike a single monomer of **2**. Additionally, **1** is ~2-fold smaller than the dodeca-peptide and ~40-fold smaller than the J3 4WJ. This small size may have caused the reduced change in polarisation.

As expected, both **5** and **6** showed minimal activity in the FP assay up to 3 μM concentrations (**Figure 2-33**). This matches the results from the PAGE assay, confirming that these peptides have little activity against DNA 4WJs. This also validates the FP assay, suggesting that the effects of the changes in polarisation come from 4WJ binding and not an assay artifact.

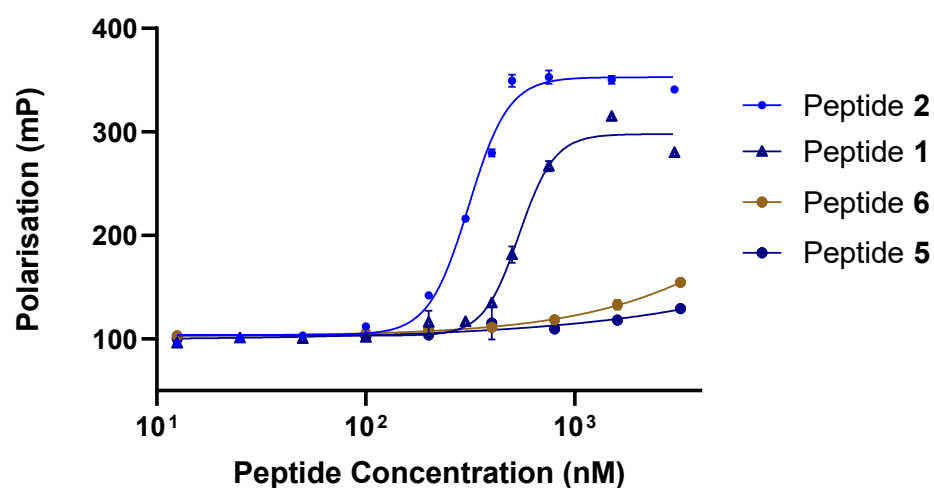


Figure 2-33: Dose-response curves from FP, using 10 nM FAM-J3 4WJ and titrating in peptides. Polarisation was measured in millipolarisation (mP). Error bars represent the mean $\pm$ standard deviation from triplicate readings.

2.3.6 Ethidium bromide displacement assay

This project focusses on the ability of different ligands to bind to 4WJ DNA. However, when carrying out assays using immobile 4WJ models, the arms of the junctions are in fact duplex DNA. It is important to understand the selectivity of these acridines and peptides in comparison to biomolecules, such as 4WJ-resolving enzymes that generally have a 1000-fold higher affinity for the centre of 4WJ DNA over duplex DNA.¹³⁵ Also, the vast majority of DNA in the cell is in the stable duplex form, as discussed in **section 1.1**. Therefore, a high level of selectivity would be required to target the low abundance of 4WJs *in vivo*.

For this purpose, the ethidium bromide (EtBr) displacement assay was used to determine the duplex DNA binding affinities of both the acridines and peptides. Firstly, EtBr (4.4 μM in 0.1 M Tris, 0.1 M NaCl, pH 7) was prebound onto calf thymus DNA (ctDNA, 62.3 μM (bp)final concentration), a natural source of duplex DNA of varying lengths. Next was added the DNA-binding compound dissolved in DMSO (2 μM), at 50-fold the desired final concentration (10 nM – 400 μM). After incubation at RT for 2 h, the sample was excited at 555 ± 20 nm and emission recorded at 610 ± 40 nm. Previous UV-Vis readings showed that the acridines do not absorb light at 555 ± 20 nm and should not interfere with the fluorescence of EtBr.¹²⁷ Readings for the half maximal displacement concentration (C_{50}) value were carried out in triplicate, providing the concentration required to get a 50% reduction in EtBr fluorescence (**Table 2-4**).

Table 2-4: C_{50} values from the EtBr displacement assay. Error in the C_{50} values is from the mean $\pm$ standard deviation from triplicate readings.

No.	Name/sequence	C_{50} (μM)
3	C6 acridine	2.03 ± 0.109
9	Triazole acridine	>400
2	H-WRWYRGGRYWRW-NH ₂	1.48 ± 0.0454
1	H-WRWYCR-NH ₂	17.0 ± 0.649
5	H-WRWYC(Acm)R-NH ₂	115 ± 18.7
6	H-WRWYR-NH ₂	92.3 ± 12.0

When testing the acridines in the EtBr displacement assay, **3** was found to have a C_{50} value of 2.03 ± 0.109 μM , whereas the C_{50} of **9** could not be determined below 400 μM , due to an incomplete binding curve (**Figure 2-34**). This reflects a similar

result obtained by Searcey *et al.* that showed **9** had a large reduction in activity compared to **3**.⁸⁹

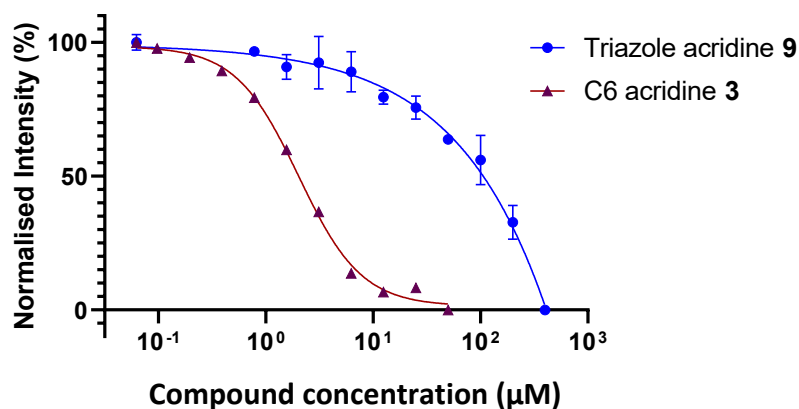


Figure 2-34: EtBr displacement assay dose-response curves for the triazole acridine **9** and C6 acridine **3**.

The peptides were then tested in the EtBr displacement assay (**Figure 2-35**). The results showed that both **5** and **6** had low binding affinity to duplex DNA, with C_{50} values $>90\ \mu\text{M}$, whereas hexapeptide **1** had a higher affinity of $17.0 \pm 0.649\ \mu\text{M}$. This may indicate that the formation of a dimer, *via* a disulfide bridge, is important for both 4WJ and duplex DNA activity. The highest activity observed in this assay was for dodeca-peptide **2**, with a C_{50} value of $1.48 \pm 0.0454\ \mu\text{M}$. This shows the ability of this peptide to bind to dsDNA and may indicate that some of its activity in previous assays comes from binding to the duplex arms of the 4WJ rather than its centre.

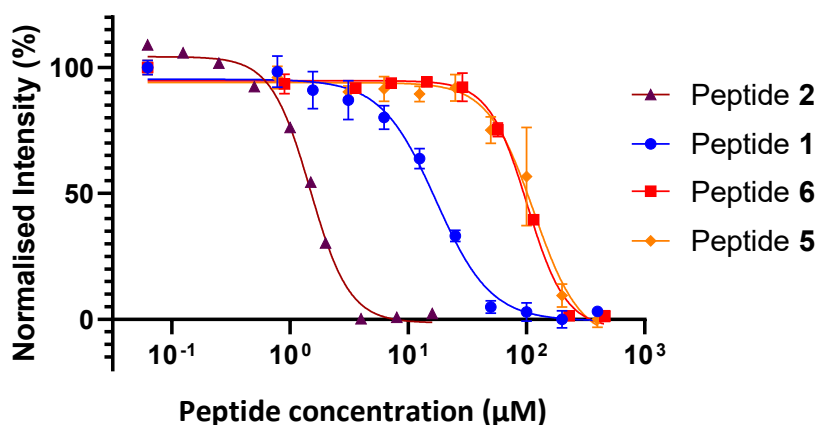


Figure 2-35: EtBr displacement assay dose-response curves for peptides **1**, **2**, **5** and **6**.

A major drawback of the EtBr displacement assay is that it cannot be used to directly compare the binding of ligands to duplex DNA and 4WJs. This is because EtBr

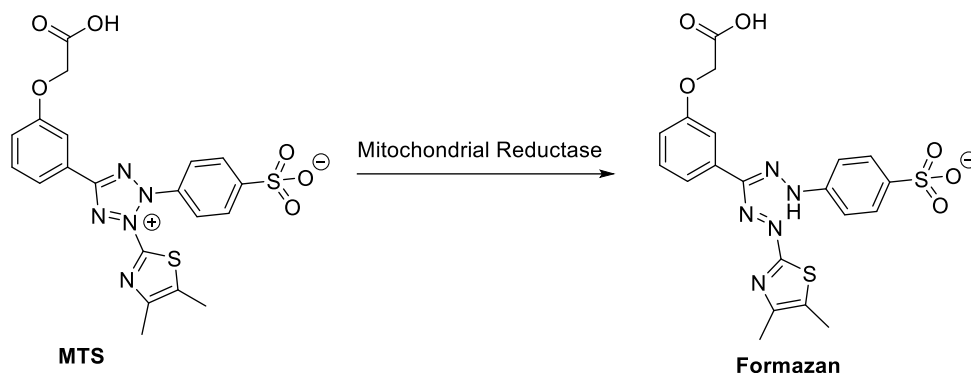
intercalates into helical DNA, making it unlikely to bind to the centre of the open conformation 4WJ, that contains unpaired bases. However, looking at the C6 acridine **3** compared to **2**, the peptide has ~4-fold higher activity in the 4WJ PAGE assay compared to 1.3-fold higher activity in the duplex DNA EtBr displacement assay. This may indirectly show that the peptide has overall higher selectivity for 4WJ DNA than acridine **3**. The ctDNA could also contain inverted repeats that are capable of forming 4WJs. Therefore, the activity of the ligands may be a consequence of binding to these 4WJs, rather than duplex DNA.

The binding activities of the peptides in the FRET, FP and EtBr assays, provide a consistent trend, with **2** being the most active peptide and **5** being the least. Although the activity of **1** was found somewhere between those of **2** and **5**, each assay gave a slightly different result. The FRET assay showed a large reduction in activity of **1** compared to **2**, with no K_D value obtained for **1** within the assay concentration range of <1.7 μ M. However, this could result from the short equilibration times in this assay being unfavourable for the formation of the active dimer of **1**. Also as mentioned in **section 2.3.5**, the FRET results should be interpreted with caution as further controls are required to confirm the readings are irrefutably caused by FRET occurring. FP analysis showed a 2-fold fall in activity for **1** in relation to **2**, compared to an ~10-fold fall in relative activities in the EtBr assay. These findings imply that although **1** has lower 4WJ binding activity in the FP assay, this peptide may have higher selectivity for 4WJ DNA over dsDNA.

2.3.7 MTS assay

Cell viability is defined as the quantity of healthy cells in a specimen and can be used to determine the response of cells to different compounds. This can be used to determine the degree of cytotoxicity a compound has against a specific cell type. A common measure for the cell viability of cancer cell lines, to test for cell survival and proliferation, is the MTS colorimetric assay. (3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl) 2-(4-sulfophenyl)-2H-tetrazolium) (MTS) is a tetrazolium salt that is reduced to formazan in viable cells, *via* mitochondrial reductase (**Scheme 2-8**). Dead cells are unable to carry out the reduction of MTS, required to produce formazan. This change in cell viability can be visualised by a colour change from yellow, when only MTS is present, to purple when formazan is formed and can be quantified by measuring absorbance of formazan at 492 nm. The MTS assay was

developed from the original 3-[4,5-dimethylthiazol-2-yl]-2,5 diphenyl tetrazolium bromide (MTT) assay, that produced an insoluble formazan compound upon reduction, unlike the soluble formazan compound produced in the MTS assay. This means the MTT assay requires an additional step, of adding a solubilising agent, before analysis.¹³⁶



Scheme 2-8: Reduction of MTS by mitochondrial reductase to form a formazan.¹³⁶

As discussed in **section 1.4.1.3**, the MTT assay was previously used to test the D-isomer peptides, **1-D** and **2-D**, against cancer cell lines PC3, Du145, LnCAP, PPC-1, DuPro-1, HeLa and A549. This showed that at 25 μ M, **1-D** was able to reduce cell viability by 0 – 10%, whilst **2-D** showed reductions between 50 – 80%.⁷¹ The MTS assay has also been used to test acridines, with the C6 acridine **3**, showing an IC₅₀ value >1 μ M and a ~10-fold drop-in activity observed for the triazole acridine **9**, when tested against the HL-60 human leukaemia cell line.¹²⁷ The idea in this project was to carry out an MTS assay to directly compare the activities of both peptide and acridines against the human leukaemia cell line HL-60.

To prepare the assay, the HL-60 cells were seeded (100 μ L, 3 x 10⁵ cells/mL) on a 96-well plate and incubated overnight at 37 °C, 5% CO₂. After this, 1 μ L of ligand in DMSO (25 μ M final concentration) was added to the cells, followed by incubation for another 3 days. MTS (10 μ L) was then added and after incubation for 3 h, absorbance was measured at 492 nm. Cell viability (%) values were obtained from triplicate readings, looking at the relative fluorescence of the treated cells, compared to untreated cells containing 1 μ L DMSO (100%) and wells containing MilliQ water with MTS added (0%).

The screening of compounds at 25 μ M (**Figure 2-36**) showed that treatment with the C6 acridine **3** resulted in only 8.00% $\pm$ 0.686% cell viability remaining. The triazole

acridine **9** caused a drop to $28.1 \pm 6.78\%$ cell viability, reflecting the loss in activity previously observed when compared to **3**.¹²⁷ Both peptides **1** and **2** had minimal effect on cell viability. When comparing the cell viability results of the compounds to the DMSO negative control, results from a two-tailed unpaired t-test confirmed that both of the acridine compounds caused a significant difference in cell viability ($p\text{-value} = <0.05$), whereas the peptide did not bring about a significant change ($p\text{-value} = >0.05$) (**Table 2-5**). This may show that the L-isomers of both peptides have lower activity than the D-isomers, potentially caused by the higher propensity of L-isomers to degradation by cellular peptidase.⁷⁰ However, to confirm this, the D-isomer peptides would need to be tested in the HL-60 cell line for direct comparison.

Table 2-5: Results from the MTS assay against HL-60 cancer cells. Errors in the cell viability (%) values are from the mean $\pm$ standard deviation from triplicate readings. P-/T-values were obtained from a two-tailed unpaired t-test, comparing samples to cells treated with DMSO.

No.	Name/ sequence	Cell viability (%)	p-value	T-value
3	C6 acridine	8.00 ± 0.686	0.000614	40.3
9	Triazole acridine	28.1 ± 6.78	0.000878	13.5
Peptide 1	H-WRWYCR-NH ₂	99.6 ± 2.75	0.958	-0.0557
Peptide 2	H-WRWYRGGRYWRW-NH ₂	101 ± 3.73	0.768	-0.315

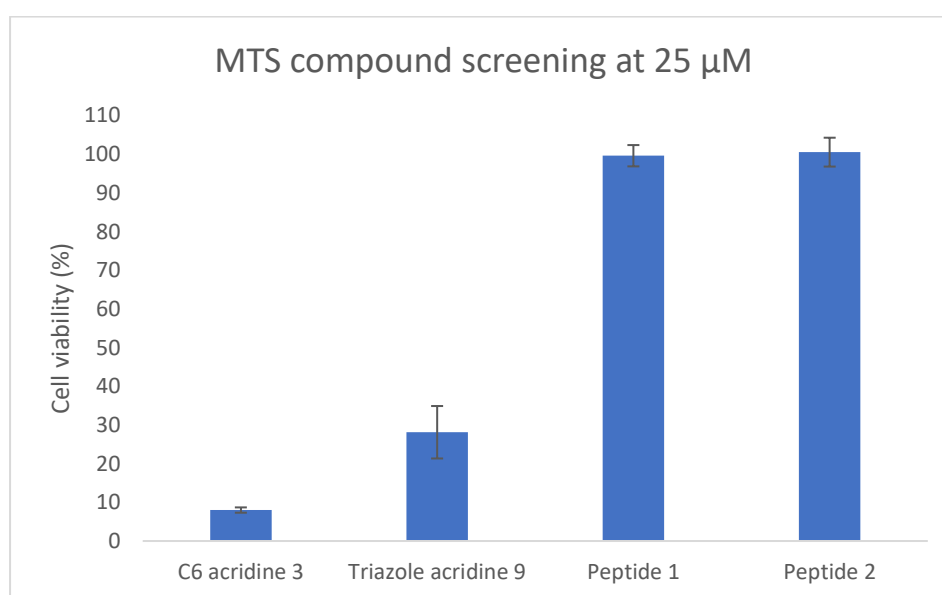


Figure 2-36: MTS assay using different compounds (25 μM) against HL-60 cells (3×10^5 cells/mL). Error bars represent the mean $\pm$ standard deviation from triplicate readings.

2.4 Conclusion and future work

In conclusion, a series of 5 peptides were successfully synthesised using SPPS and two acridines using multi-step organic synthesis. These ligands were then tested in a series of biological assays, to determine their ability to bind to 4WJ/dsDNA and their cytotoxicity against HL-60 cancer cells. From this, peptide **2** was shown to have the highest activity against the 4WJ in the PAGE, FRET and FP assays, making it a molecule of interest for further studies.

Peptide **2** also had the highest activity against dsDNA in the EtBr displacement assay, indicating it may be unselective for 4WJ binding. However, when carrying out a PAGE competition assay, **2** showed only 4WJ binding, even with a large excess of dsDNA added. Mega base pair dsDNA was used for the EtBr displacement assay, containing a high variation in DNA sequences, compared to the 25 bp dsDNA used in the PAGE competition assay, that contained no sequence similarity to the DNA used to form the 4WJ. This may indicate that **2** has some DNA sequence selectivity. Repeating the PAGE assay using a range of different DNA sequences, for both the dsDNA and 4WJ DNA, could be helpful to check whether the peptide has DNA sequence or structure specificity. This could be further tested by analysing peptide binding against with a range of 4WJ sequences in the FP assay.

As mentioned in **section 2.3.4**, the FRET assay carried out for testing the peptides synthesised against the J3 4WJ DNA, requires further testing before results can be validated. This is to ensure that energy transfer from FAM to TAMRA is occurring, instead of direct emission from TAMRA, after excitation at 490 nm. In addition to this control experiment, other FRET-based techniques could be employed to further investigate the 4WJ conformations induced by peptide binding. Both time-resolved FRET^{123,133} and single-molecule FRET⁵⁵ have previously been used in research into the conformational changes of 4WJ DNA. Time-resolved FRET measures changes in donor fluorescence over time, allowing accurate quantification of energy transfer independent of fluorophore concentrations or background fluorescence. This technique could help to overcome the relatively weak signal observed in the steady-state FRET, in relation to TAMRA emission (**B, Figure 2-27**). Single-molecule FRET offers the ability to observe individual 4WJ molecules, capturing their dynamic conformational changes in real time as peptides bind. This could give insights into

whether the peptides are able to trap the 4WJ DNA into a single conformation upon binding.

Despite the activity of **2** against 4WJ and dsDNA, the peptide showed no cytotoxicity when tested against HL-60 cancer cells at a concentration of 25 μ M. This could result from several different processes, such as degradation by cellular peptidase or from the peptide being unable to penetrate the cells. The next chapter will focus on exploring the effects of changing the structure of **2** on 4WJ/dsDNA activity and HL-60 cancer cell cytotoxicity.

**Chapter 3 - The
synthesis and biological
testing of peptide
analogues based on
WRWYRGGRYWRW**

3.1 Introduction

3.1.1 Peptidomimetics

As noted in **section 2.1.1**, peptides are an important and growing class of therapeutic molecules. However, the use of peptides in the clinic has been limited by their instability *in vivo*, in the presence of endogenous proteases that cause amide bond hydrolysis.¹³⁷ This results in values of <2% bioavailability¹³⁸ and short half-lives often being obtained for peptides, meaning nasal or parenteral administration is required.¹³⁹ Additionally, the intrinsic flexibility of the rotatable amino bonds N-C α and C α -CO can cause issues with poor selectivity, increasing unwanted side effects.¹⁴⁰ Therefore, efforts must be made to transform peptides into more drug-like molecules and stabilise their metabolic vulnerabilities. These include modifying their C-/N-termini, peptide cyclisation and adding D-amino acids or unnatural amino acids into the sequence.¹⁴¹

This section will focus on the replacement of the standard 20 amino acids, with non-canonical chemical groups, using peptidomimetics. This is where the natural peptide/protein is structurally mimicked, using different chemical groups that match the original pharmacophore, whilst retaining the same biological activity.¹⁴⁰ One method used in peptidomimetics is the replacement of natural L-amino acids for their D-isomers. This has been shown to reduce proteolysis, by shifting the conformation of the peptide away from those recognised by proteases.¹⁴² Using D-amino acids can improve the antimicrobial activity of peptides, such as for peptide **1-D** and **2-D** (see **section 1.4.1.3**).

In proteins, linker/spacer groups are short sequences of amino acids that separate different domains within a single protein. They are generally Gly-rich, to allow for the formation of flexible loop-type structures and do not perform any additional functions for the protein.¹⁴³ This means that these irrelevant amino acids can be replaced with isosteric spacers, that do not contain the peptide backbone. This can help improve bioavailability and reduce the peptides susceptibility to proteolysis, by reducing the number of amide bonds present. This is known as backbone

prosthesis.¹⁴⁴ Several different chemical groups have been used as linkers and will now be discussed.

A commonly used backbone spacer is the 6-aminohexanoic acid (Ahx) group, a hydrophobic and highly flexible amino acid.¹⁴⁴ Ahx was used to replace two adjacent glycine residues in Met-enkephalin, an opioid neuropeptide. This resulted in the discovery that a single hydrogen bond is required within the spacer for activity to be maintained, leading to an improved understanding of the binding of the peptide to δ -/ μ -opioid receptors. This study also explored the use of different isosteric spacers of Gly-Gly, including 4 – 7 carbon alkane linkers and a PEG-type linker (**Figure 3-1**).¹⁴⁵ Additionally, Ahx has been used as a spacer between Arg amino acids, to act as a hydrophobic linker between the two charged guanine groups. In the development of cell-penetrating peptides, this has been found to increase their ability to enter the cell and metabolic stability.¹⁴⁴ Subsequent attachment of the oligoarginine (R-Ahx-R)₄ to synthetic oligonucleotides increased their cell penetration and ability to produce functional fully spliced mRNA, for the treatment of diseases including β -thalassemia and various cancers.¹⁴⁶

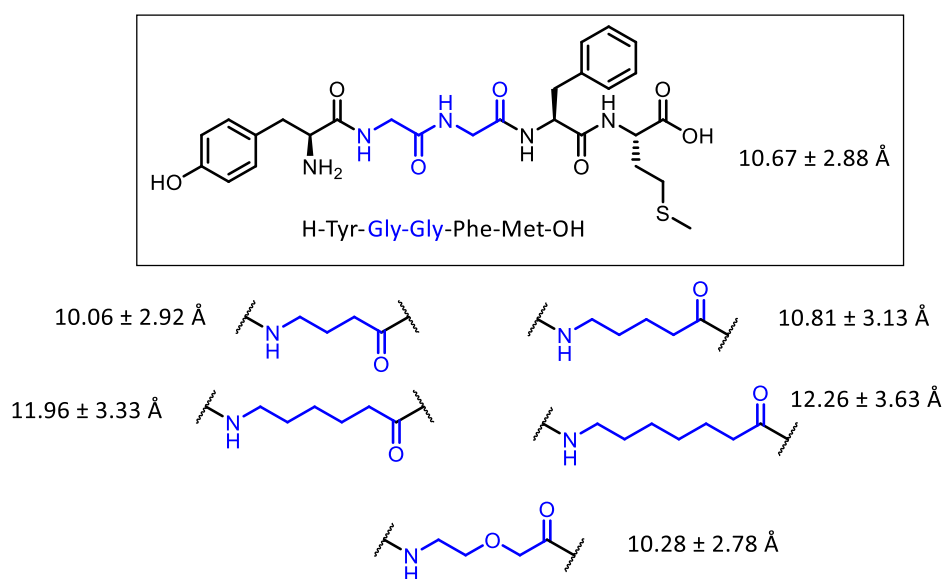


Figure 3-1: The replacement of the Gly-Gly residues of Met-enkephalin with different backbone spacers. Distances were calculated using molecular dynamics simulation and were measured from the centre of the Tyr and Phe aromatic rings.¹⁴⁵

Triazoles are of interest in peptidomimetics because of their ability to act as amide bioisosteres. When disubstituted at the 1,4 positions, triazoles can be compared to *trans* (Z) amide bonds, as demonstrated by their similar features. Both contain a H-bond acceptor and a H-bond donor, an electrophilic carbon and relatively similar

distances between R-groups (**Figure 3-2**).¹⁴⁷ Triazoles are useful in drug development because of their stability during acid/base hydrolysis and in reductive/oxidative conditions, making them metabolically resistant. This heterocycle can also take part in H-bonding, dipole-dipole interactions and π - π stacking.¹⁴¹

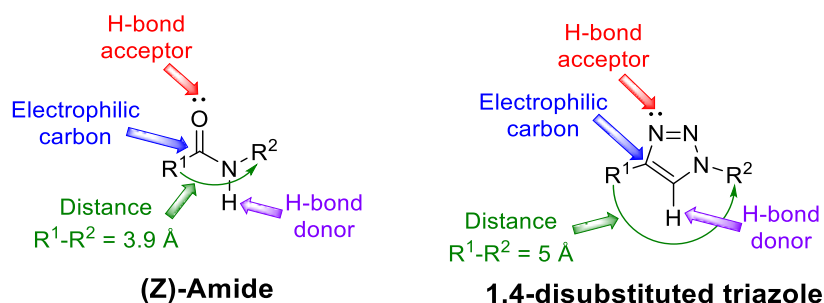
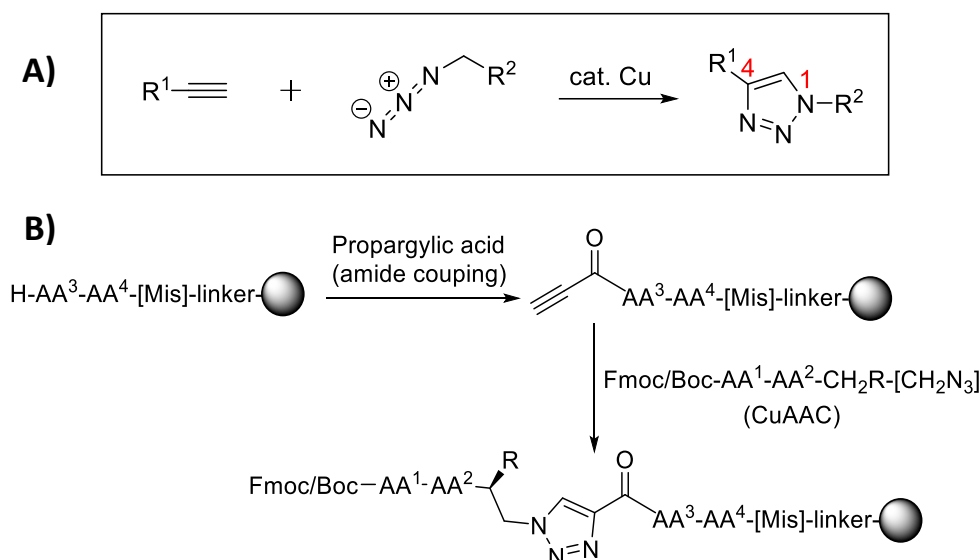


Figure 3-2: Comparison of electronic and topological features of (Z)-amides and the 1,4-disubstituted triazole bioisostere.¹⁴⁷

The introduction of click chemistry by Sharpless *et al.* provided an approach to organic synthesis that utilises reactions that are quick and reliable.¹⁴⁸ One of the most commonly used click reactions is the copper catalysed 1,3-dipolar cycloaddition between an azide and an alkyne, that allows for the selective synthesis of 1,4-disubstituted triazoles (**A**, **Scheme 3-1**). Copper-catalysed azide-alkyne cycloaddition (CuAAC) has been used extensively during peptide synthesis, for producing unnatural amino acids, the ligation of peptide fragments and peptide cyclisation.¹⁴⁹ For example, Meldal's group produced a library of >400,000 compounds, using CuAAC to form a variety of peptidotriazoles with possible biological activity. For this, peptides were immobilised on solid support, before being reacted with propargylic acid, to form a terminal alkyne. Next a series of Fmoc or Boc protected azides were clicked onto the peptide, using CuAAC, to provide library compounds in high yields (**B**, **Scheme 3-1**).¹⁵⁰



Scheme 3-1: **(A)** The synthesis of a 1,4-disubstituted triazole using CuAAC.¹⁵¹ **(B)** The synthesis of peptidotriazoles for a compound library using CuAAC on solid support (black ball).¹⁵⁰ The compounds included a mass/ioniser spacer (Mis) and amino acid (AA).

As mentioned in **section 1.4.3.1**, rigid linkers were used to connect acridine or phenanthridinium groups during the development of 4WJ bis-intercalators. This was to investigate the effect of forcing the DNA-binding groups to face opposing directions.⁸⁵ Rigid linkers have also been used for the tethering of one or two α -melanocyte stimulating hormone (α -MSH) neuropeptide ligands.¹⁵² α -MSH is a 13 amino acid peptide that is produced after exposure to ultraviolet light, before taking part in melanin synthesis required for skin pigmentation.¹⁵³ This peptide has been found to have anti-inflammatory and anti-microbial effects and is still under investigation for its role in malignant melanoma.¹⁵⁴ By using different rigid linkers, containing carboxylic acid and amine terminal functionalities (**Figure 3-3**), α -MSH peptides could be linked together using SPPS. Overall, the linkers were well tolerated when binding was evaluated against human melanocortin 4 receptor, with minimal change to IC₅₀ values observed.¹⁵²

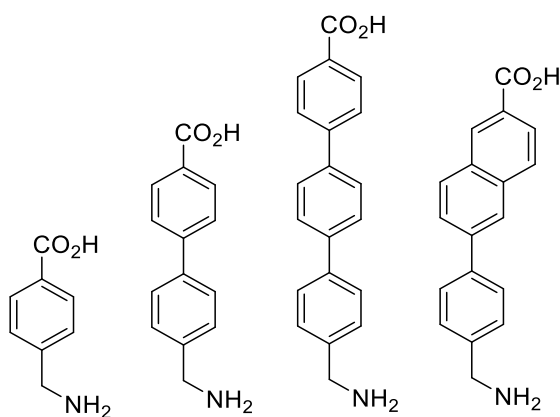


Figure 3-3: Different rigid linker groups for the tethering of α -MSH peptides.¹⁵²

3.1.2 Cyclic peptides

Cyclic compounds are highly notable as therapeutic molecules, with ~20% of known natural products having cyclic structures.¹⁵⁵ The same can be said for cyclic peptides, with their biological function in nature often associated with their ability to cyclise.¹⁵⁶ Cyclic peptides have had success as drugs, with nine approved in the clinic between 2006 – 2015. For example, telavancin used as an antibacterial drug and romidepsin as an anti-cancer drug (**Figure 3-4**).¹⁵⁷ Linear peptides often suffer from low bioavailability and high metabolic instability. By cyclising these peptides, thermal stability, membrane permeability and proteolytic resistance can all be increased.¹⁵⁶ In addition, cyclic peptides can be preorganised into the active conformation, reducing entropic penalties and enhancing binding.¹⁵⁵

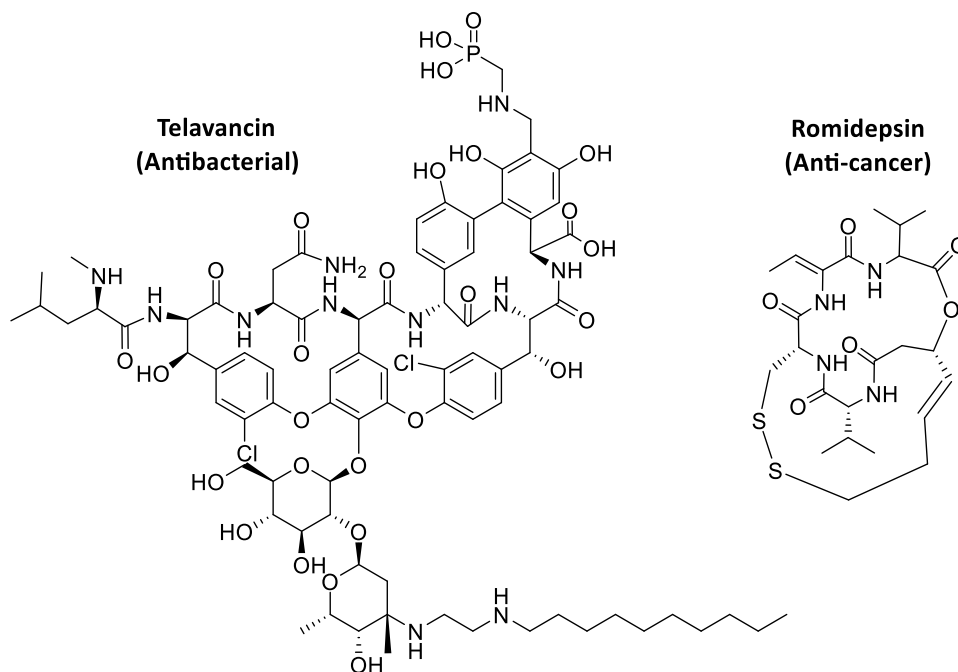


Figure 3-4: Chemical structures of two cyclic peptide drugs.¹⁵⁷

A variety of methods are known for peptide cyclisation and can be categorised into whether the N-terminus (head), C-terminus (tail) or side chains are involved in the cyclisation (**Figure 3-5**). A common example of side chain-to-side chain cyclisation is the formation of disulphide bridges between two thiols from cysteine residues, as observed with insulin.¹⁵⁶ Side chain-to-terminus cyclisation can take place between an appropriate side chain and terminus, for example between the C-terminus and lysine side chain of the antibiotic bacitracin.¹⁵⁸ Finally, head-to-tail cyclisation is formed by the amide coupling of the peptides' backbone termini and is widespread in normal flora found in the human gut.¹⁵⁹ A benefit specific to head-to-tail cyclisation is the higher resistance these cyclic peptides have to exopeptidase hydrolysis, without the presence of reactive amine and carboxylic acid termini groups.¹⁵⁶ This makes head-to-tail cyclisation an attractive approach for the design of synthetic cyclic peptides.

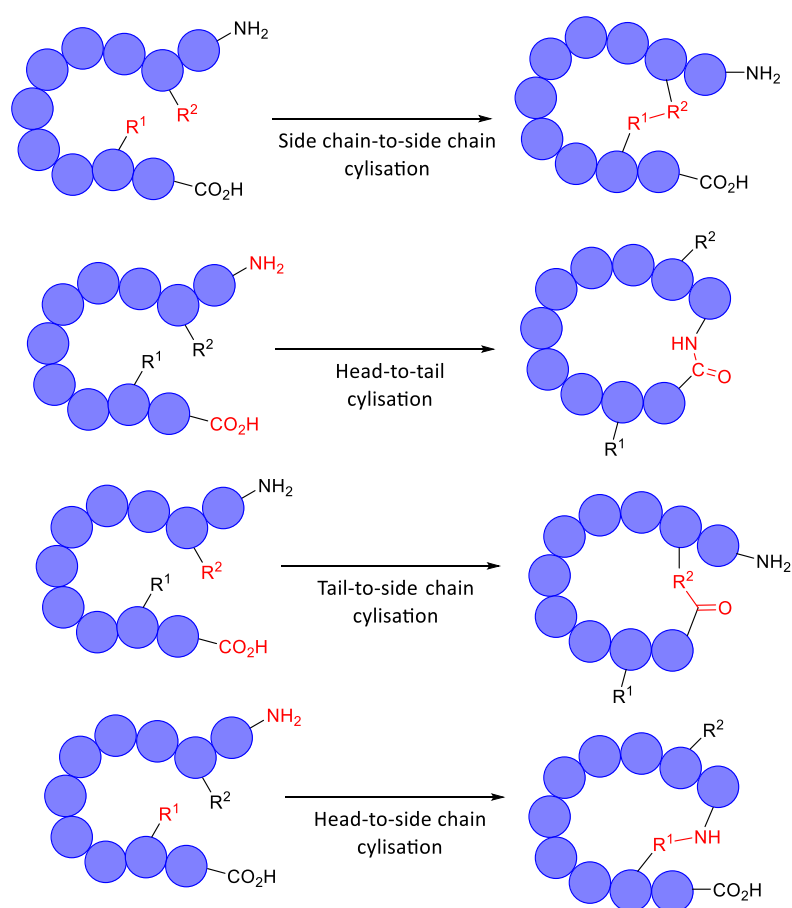
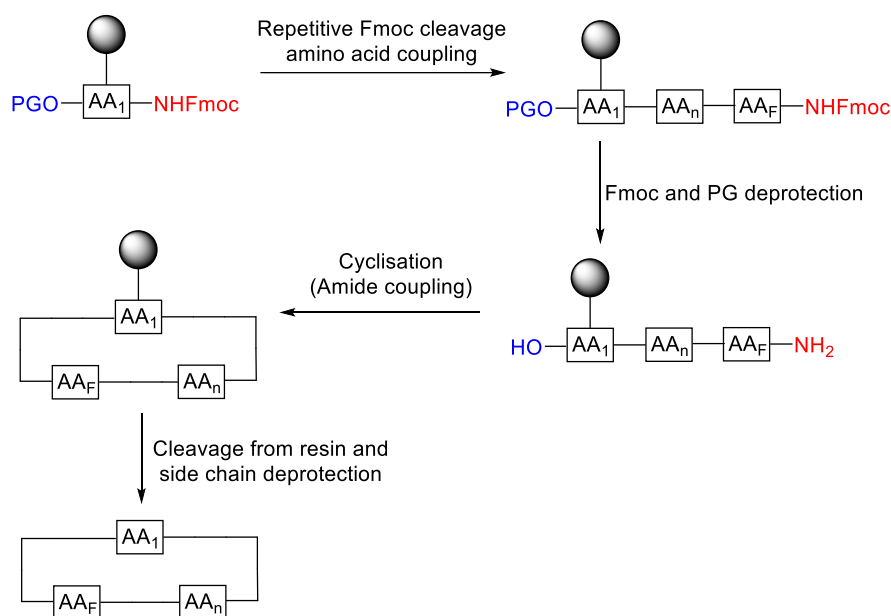


Figure 3-5: Methods for peptide cyclisation.¹⁵⁶

When synthesising cyclic peptides, unwanted dimerisation or cyclodimerisation often occurs, resulting from intramolecular ring-closing reactions generally being

slower than intermolecular reactions.¹⁶⁰ To suppress intermolecular reactions, cyclisation reactions are generally performed in highly diluted solutions, to reduce the chance of separate molecules coming together to react. An alternative approach is to keep the peptide attached to a low loading resin during cyclisation, to get the pseudo-dilution effect, caused by the large size of the resin preventing molecules approaching each other.¹⁶¹ Another advantage of using on-resin cyclisation is that excess coupling reagents and by-products can be used and easily washed away, as with the use of SPPS for linear peptides.¹⁶⁰ To obtain a head-to-tail cyclised peptide, the initial amino acid is immobilised on the resin *via* a side chain and its carboxylic acid protected with an orthogonal protecting group, that can be selectively removed before cyclisation.¹⁶² After the cyclisation is complete the peptide can simply be cleaved from the resin and purified to give the desired product (**Scheme 3-2**).



Scheme 3-2: On-resin cyclisation of a cyclic peptide synthesised using SPPS and side chain immobilisation. The first amino acid is AA_1 , the central chain of variable length is AA_n and the final amino acid is AA_F . PG represents an orthogonal protecting group that can be selectively removed.¹⁶⁰

To facilitate the cyclisation of peptides, the presence of D-amino acids that induce turns are often required. Without this, cyclisation may not be possible or become a much slower process.¹⁶³ For example, D-amino acids are found in two cyclic peptides known to inhibit the p53-MDM2 protein-protein interaction. p53 is a tumour suppressor protein that is mutated in >50% of tumours¹⁶⁴ and MDM2 is a regulatory protein that binds to p53, causing downstream ubiquitination and degradation.¹⁶⁵ The overexpression of MDM2 in cancer cells causes p53 activity to be suppressed. By blocking this protein-protein interaction, p53-mediated apoptosis can be restored in

tumours.¹⁶⁶ Chlorofusin is a naturally derived cyclic peptide that contains a D-Asn-L-Asn dipeptide, two D-Leu and a D-2-aminodecanoic acid moiety (**A**, **Figure 3-6**). This compound is derived from the fungus *Microdochium caespitosum* and inhibits the p53/MDM2 protein-protein interaction with an IC_{50} of 4.7 μ M.¹⁶⁷ A peptidomimetic compound, based on the p53 sequence in the MDM2 binding domain, has also been developed and binds to MDM2 with a K_D of 40 nM (**B**, **Figure 3-6**). This cyclic peptide contains a D-Pro-L-Pro dipeptide, that acts as a β -hairpin mimetic.¹⁶⁸

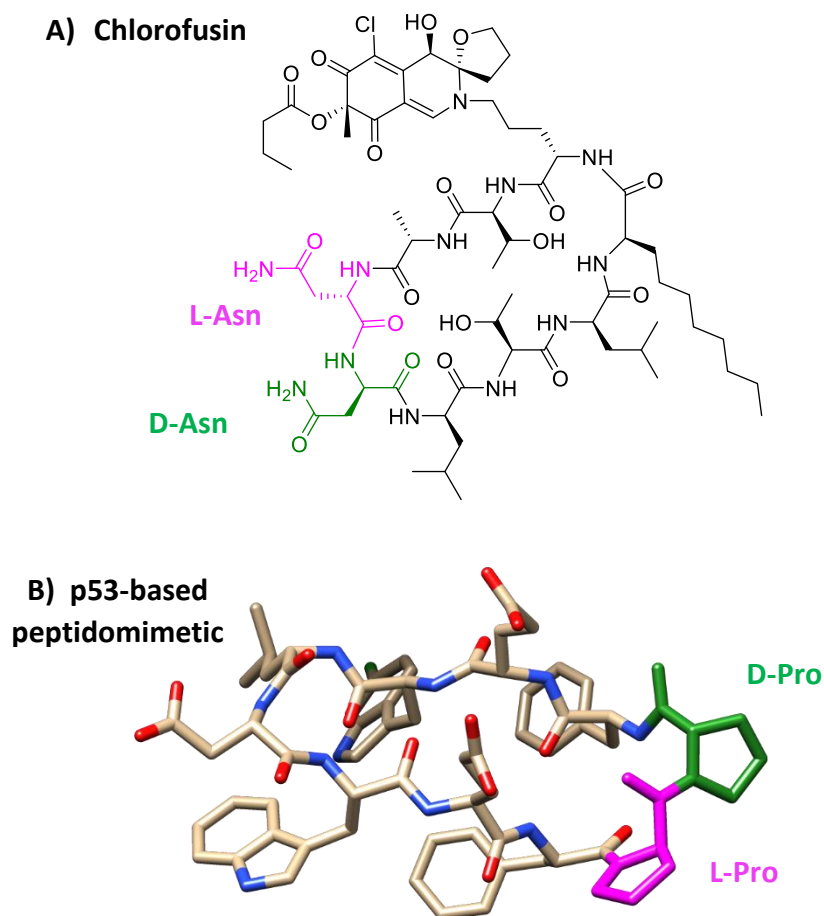


Figure 3-6: **(A)** The chemical structure of chlorofusin,¹⁶⁷ with the D-Asn-L-Asn dipeptide highlighted. **(B)** The crystal structure of the p53-based peptidomimetic (PDB code: 2AXI), with the D-Pro-L-Pro dipeptide highlighted.

3.2 Chapter aims

In this chapter, a series of peptide analogues will be synthesised, based on the 4WJ-binding dodeca-peptide **2**. Different methods for creating peptidomimetics will be used, in an attempt to improve the drug-like properties of **2**, including the use of non-canonical amino acids and peptide cyclisation. A structure-activity relationship

(SAR) study will be carried out to probe the binding of these peptides to 4WJs. SAR studies look at how small changes to chemical structure effect biological activity and seeks to use this information to optimise the drug-like properties of compounds.¹⁶⁹ After the peptide analogues have been synthesised, using a combination of SPPS and solution phase chemistry, they will be tested in the same biological assays as used during **Chapter 2**. This will allow their activity to be directly compared to the original peptide **2**. The knowledge gained from this study will be used to suggest possible next steps for the development of the next generation of 4WJ-binding molecules.

3.3 Results and discussion

3.3.1 The design of peptide analogues

As stated in **section 1.4.1.3**, WRWYRGGRYWRW **2** is a single-chain analogue of the peptide dimer formed by WRWYCR **1**.⁴⁵ The main difference between these two peptides is the replacement of the disulphide bridge-forming cysteine residues of dimerised **1** with the two glycine amino acids in **2**. The second arginine in the WRWYCR sequence is also switched from the 6th position to the 5th position in WRWYRG, one half of the palindromic single-chain analogue. The final difference is that the dimer of **1** contains 2 sets of carboxylic acid and amide termini, on arginine residues and tryptophan residues respectively, whereas **2** contains only one set that are both on tryptophan residues (**Figure 3-7**). This means that the second half of the single-chain analogue has the WRWYRG sequence going from C- to N-termini, whereas the first half WRWYRG sequence and both halves of the WRWYCR dimer are going from N- to C-termini. Despite the differences between these peptides, they both contain WRWY tetrapeptides and have similar activity against the 4WJ.⁴⁵

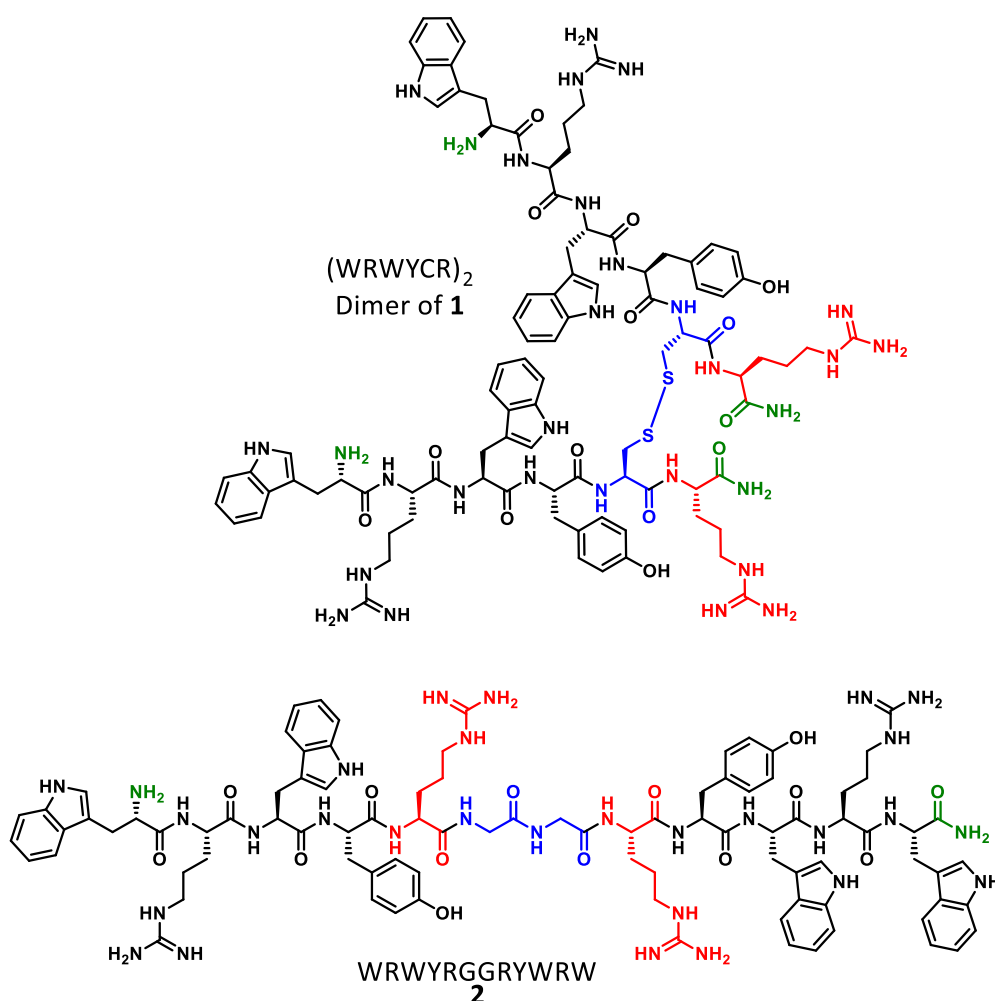


Figure 3-7: Comparison of the chemical structures of a dimer of **1** and peptide **2**. Differences in amino acid identity (blue), the position of Arg residues (red) and positioning and number of C/N-termini (green) are highlighted. Both peptides contain neutral C-terminal amide groups.

With the Gly-Gly dipeptide acting as a replacement for the dimer-forming cysteine disulphide bridge, peptide analogues were designed to explore the effect of substituting the central glycine residues in peptide **2** with alternative linker groups. Using isosteric spacers to create peptidomimetic compounds, SAR studies were carried out into 4WJ binding. By investigating the effects of incorporating flexible or rigid linkers, as well as peptide cyclisation, it may also be possible to predict the active conformations of the peptide analogues when bound to the 4WJ. This information could be used to guide the design of future 4WJ-binding peptides. Some possible active peptide conformations are shown in **Figure 3-8**.

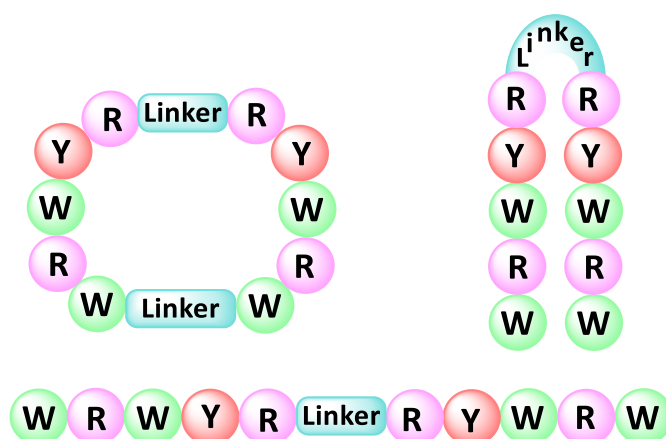


Figure 3-8: Potential different conformations of peptide analogues when bound to the 4WJ.

3.3.2 Choice of linkers

To explore the effects of changing the central Gly-Gly amino acids for different linkers, a wide range of isosteric spacer groups were selected (**Figure 3-9**). When looking at the glycine amino acids, they form a 6-atom linker, when counting from the N-terminal amide to the C-terminal carboxylic acid. Therefore, linkers of a similar length were chosen, between 6- to 7-atoms in length, except for one longer 11-atom linker. The linker groups all contained a terminal carboxylic acid and amine group, to allow for their easy incorporation into SPPS. For the design of the linker groups, synthetic feasibility was considered, with linker groups chosen that could either be purchased commercially as amino acids or synthesised.

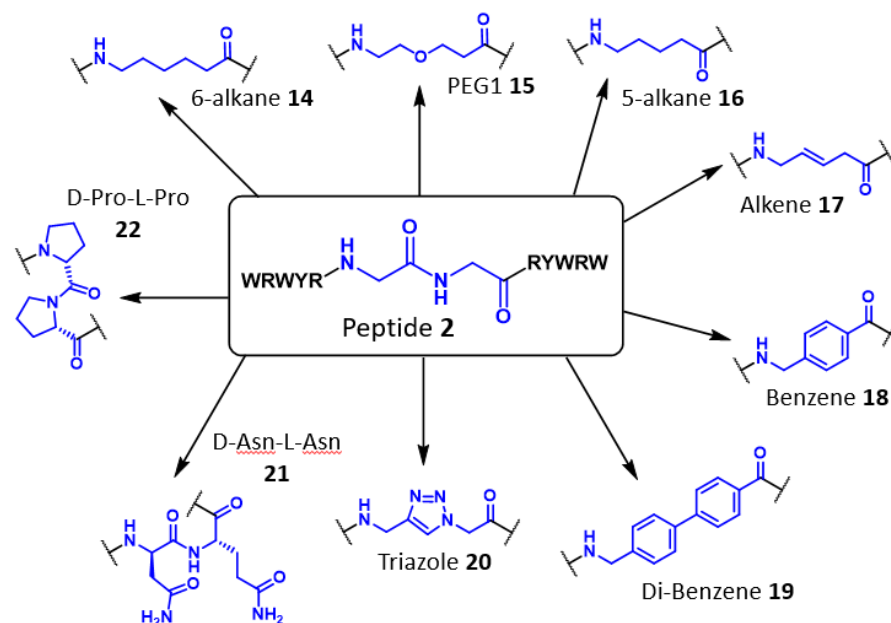


Figure 3-9: Structures of different linker functional groups (blue) used for the replacement of the central Gly-Gly amino acids (blue) of peptide 2.

Each peptide analogue was designed to explore the effects of different functional groups, linker lengths and flexibility on 4WJ binding (**Table 3-1**). Alkane linkers, of 6- or 7-atoms in length, were chosen to determine if a flexible linker helped the peptides adopt more favourable conformations. Additionally, the 6-alkane or Ahx linker placed between two arginine residues was previously shown to improve cell penetration.¹⁴⁴ The hydrogen bond donor groups, ether (PEG1) and alkene, were added into the alkane chains, to explore whether further interactions could be made between the peptide and 4WJ. Alkane chains are highly hydrophobic, therefore adding the polar ether group could also improve aqueous solubility. More rigid linkers were designed to keep the spacer groups in a specific arrangement. The benzene and di-benzene linkers decrease the ability for the spacers to bend and should keep the two halves of the peptide separated. In contrast the D-Asn-L-Asn and D-Pro-L-Pro dipeptide chains should force turns in the spacer and keep the two peptide halves closer together.^{167,168} However, it should be noted that apart from the D-Pro-L-Pro dipeptide, all other linkers contain some flexible bonds, that may affect the formation of the desired linker structures. Finally, a linker containing a triazole ring was designed to act as a more metabolically stable amide bioisoster.¹⁴⁷

Table 3-1: Summary of the properties of each of the spacer groups used in the peptide analogues.

No.	Linker group	Number of atoms	Flexible functional groups	Rigid functional groups
2	Gly-Gly	6	Alkane	Amide
14	6-alkane	7	Alkane	-
15	PEG1	7	Alkane and ether	-
16	5-alkane	6	Alkane	-
17	Alkene	6	Alkane	Alkene
18	Benzene	7	Alkane	Aromatic
19	Dibenzene	11	Alkane	Aromatic
20	Triazole	7	Alkane	Aromatic
21	D-Asn-L-Asn	6	Amino α -carbon and alkane	Amide
22	D-Pro-L-Pro	6	-	Amide and heterocycle

3.3.3 Cyclic peptide analogues

In addition to the linear peptide analogues, two cyclic peptides were designed. This will further rigidify the structure of the peptides and most likely force the spacer groups to act as turns (**Figure 3-8**). As mentioned in **section 1.4.1.4**, cyclic peptides have previously been designed to target the DNA 4WJ.^{73,74} In these studies, the cyclic peptides contained mostly hydrophobic residues, that prevented the compounds from being able to cross bacterial cell walls. By creating a cyclic peptide based on the WRWYRGGRYWRW sequence, it may be possible to produce a more hydrophilic molecule, due to the presence of multiple charged arginine residues.

The design of the cyclic peptide was based on a method developed by Searcey *et al.* for the synthesis of the natural cyclic peptide chlorofusin.¹⁷⁰ Searcey's group used on-resin head-to tail cyclisation, *via* side chain immobilisation. For this, an Asp residue was immobilised, that, when cleaved from a Rink amide resin, produced an Asn amino acid. To use this method, a D-Asn residue was also attached to the N-terminal of the sequence to give a 14 amino acid sequence of ^DN-WRWYRGGRYWRW-^LN. Cyclisation could then be carried out between the two Asn residues, to give cyclic peptide **23**, containing the D-Asn-L-Asn dipeptide. In addition, the 6-alkane linker was used as an alternative linker for the Gly-Gly dipeptide, to give

cyclic peptide **24** (Figure 3-10), to allow the cyclic peptide to have more flexibility if required for adopting an active conformation.

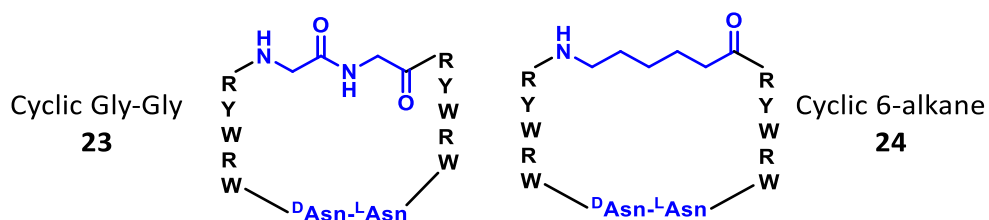


Figure 3-10: Structures of both cyclic peptides **23** and **24**, with linker groups (blue) shown.

3.3.4 Solution phase synthesis and SPPS of peptide analogues

Synthesis of the peptide analogues involved a combination of SPPS and multistep solution phase chemistry. To allow for comparison to peptide **2** in biological assays, all linear peptide analogues were synthesised with amide C-terminal functionalisation. The various linker groups were added as Fmoc amino acids to allow for their easy addition during SPPS. For 6 of the 9 linear peptide analogues, the linker group could be purchased with Fmoc-protected amino acid functionalisation, for direct use in SPPS (red, **Table 3-2**). For triazole **20**, alkene **17** and dibenzene **19** linkers were synthesised prior to use and will be discussed in the next section (black, **Table 3-2**). The 6 peptide analogues, synthesised solely using SPPS, were obtained with purities >80% and low yields of 4 – 11%. This could have been caused by the low solubility of the peptide analogues in aqueous solutions, used during purification, potentially resulting from the increased hydrophobicity of the linkers compared to the original Gly-Gly linker.

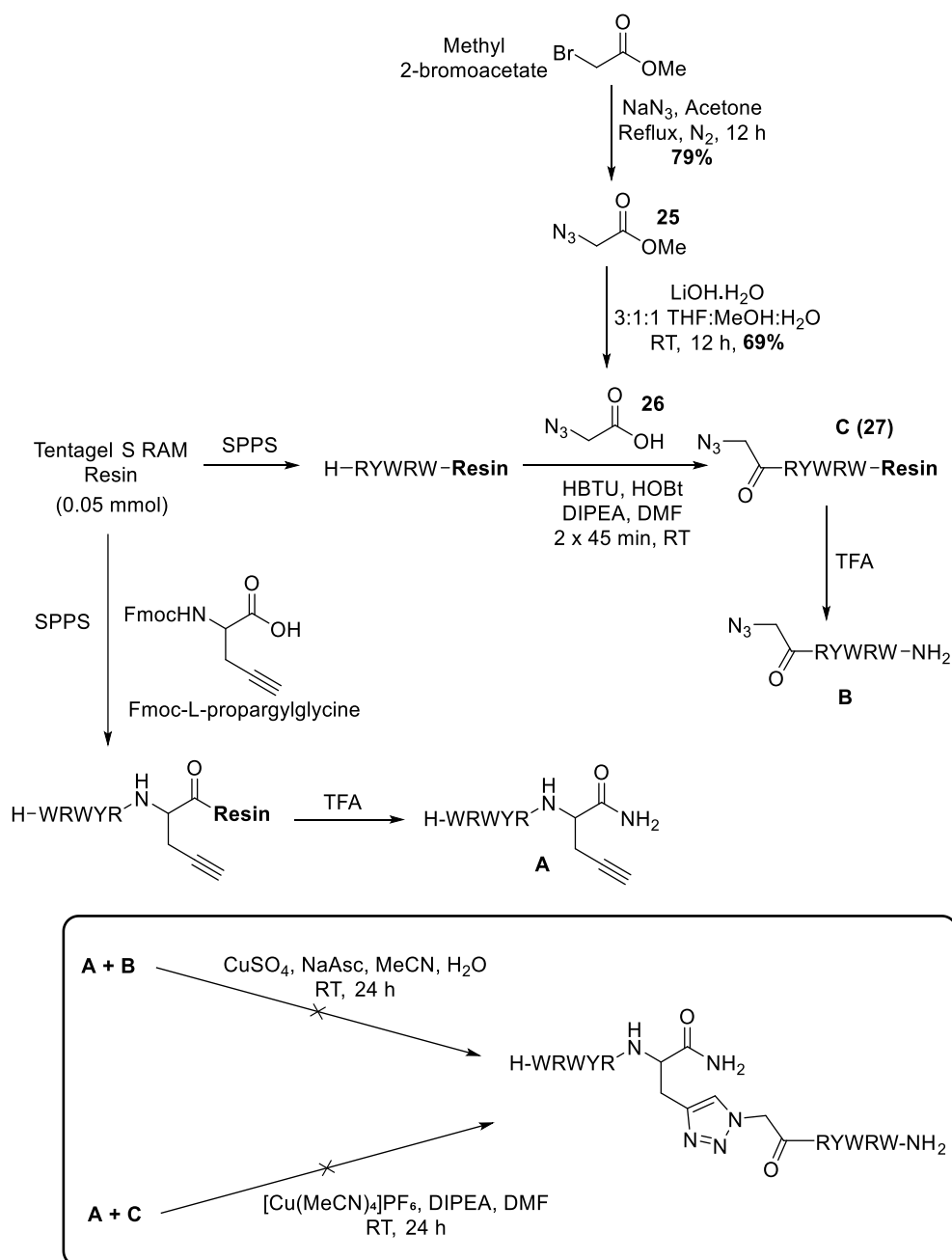
Table 3-2: Peptide analogues based on peptide **2**, with the central Gly-Gly residues replaced with various linker groups, to give WRWYR-Linker-RYWYW analogues. Compounds were synthesised using purchased Fmoc amino acids (red), pre-synthesised amino acids (black) or via peptide cyclisation (green). *Mass provided for [M+2H]⁺.

No.	Linker group	Exact mass (g/mol)	Observed mass [M+H] ⁺	Purity (%)	Yield (%)
14	6-alkane	1824.9590	1825.9686	95	9
15	PEG1	1826.9382	1827.9438	92	4
16	5-alkane	1810.9433	1811.9489	95	11
18	Benzene	1844.9277	1845.9340	90	4
21	D-Asn-L-Asn	1939.9608	1940.9686	89	4
22	D-Pro-L-Pro	1905.9804	1906.9871	80	6
20	Triazole	1849.9291	1850.9369	100	3
17	Alkene	1808.9277	1809.9355	98	8
19	Dibenzene	1920.9590	1921.9668	99	4
23	Cyclic Gly-Gly*	1018.4958	1019.4927	90	3
24	Cyclic 6-alkane	2036.0183	2037.0261	95	8

Triazole-linked peptide (20)

The synthesis of the triazole **20** was initially attempted *via* the ligation of two peptide fragments (**Scheme 3-3**). This involved separating the two WRWYR peptide fragments, then reacting them together *via* a CuAAC click reaction, to form the desired triazole.¹⁷¹ For the synthesis of alkyne fragment **A**, the commercially available amino acid Fmoc-L-propargylglycine was first added to Tentagel S RAM resin, before the subsequent SPPS of the RYWYW chain. For the synthesis of the azide fragment **B**, azidoacetic acid was first made in 2-steps. Commercially available methyl 2-bromoacetate was subjected to an S_N2 reaction using sodium azide, resulting in the replacement of the bromide group for an azide, to afford **25**. Subsequent hydrolysis reduced the methyl ester to a carboxylic acid to give **26**.¹⁷² After the SPPS of WRWYR, the final tryptophan was reacted with azidoacetic acid, under standard amide coupling conditions to afford resin-bound azide fragment **C** (**27**, **Scheme 3-3**). To confirm the completion of the amide coupling, the Kaiser test was carried out (**Section 6.2.2**) and gave a negative result, confirming that no primary amine remained.

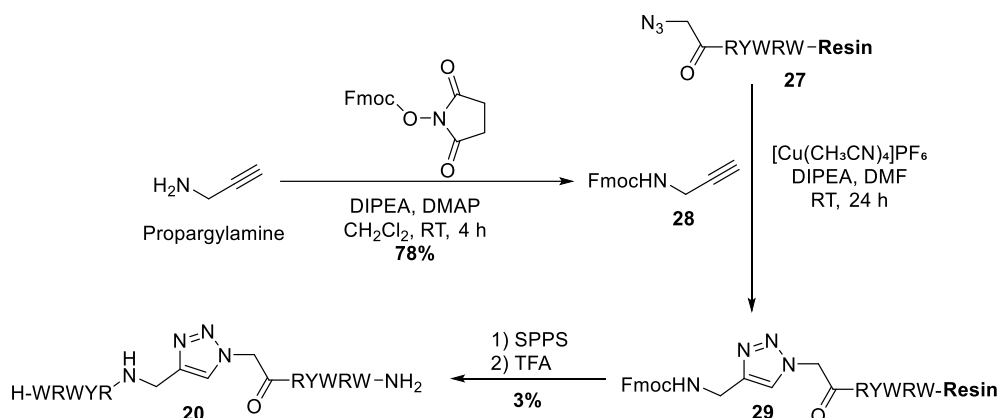
An attempt was made to carry out the click reaction between the peptide fragments, with both halves cleaved from the resin (**A+B**). For this, copper sulphate, reduced by sodium ascorbate, was used as a catalyst¹⁷¹ and resulted in the formation of traces of the desired product, with mostly unreacted SM remaining, as confirmed by MALDI. Next, an attempt was made where the azide fragment was kept on resin and reacted with the cleaved alkyne fragment (**A+C**), using $[\text{Cu}(\text{CH}_3\text{CN})_4]\text{PF}_6$ as a catalyst.¹⁷¹ This gave similar results, with only traces of product formation observed (**Scheme 3-3**). The lack of reactivity of the peptide fragments may have been caused by the low temperatures required during peptide ligation, to prevent possible degradation and side reactions caused by unprotected peptides. Also, the C-terminal amide of the alkyne fragment could have caused steric issues that slowed the formation of the desired triazole.



Scheme 3-3: Synthesis of alkyne peptide **A**, azide peptide **B** and resin-bound azide peptide **C**. Initial attempts to synthesis the triazole-linker peptide **20**, using peptide ligation are shown in the black box.

To avoid the issues encountered with peptide ligation, the next approach was to use a small alkyne molecule and react directly with the resin-attached azide peptide, before finishing synthesis using SPPS (**Scheme 3-4**). Synthesis began with the Fmoc protection of propargylamine to afford **28**. A click reaction was then carried out between alkyne **28** and resin-bound azide peptide **27**, using $[\text{Cu}(\text{CH}_3\text{CN})_4]\text{PF}_6$ as a catalyst, to give **29**.¹⁷¹ Cleavage and HPLC/MALDI analysis of a small aliquot of resin showed that the desired triazole **29** was successfully formed. The resin was then resubjected to SPPS, to complete the addition of the remaining amino acids and TFA

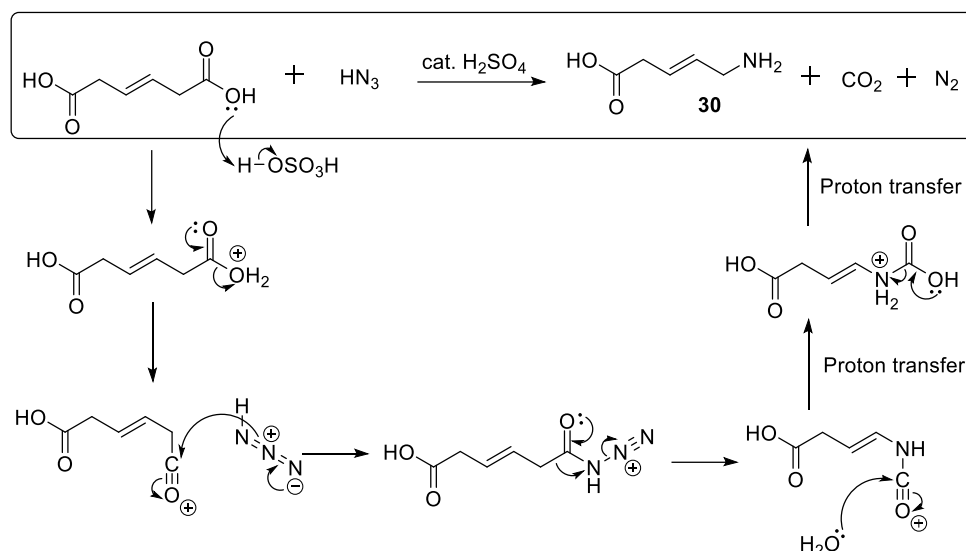
cleavage resulted in the desired peptide **20** being obtained, with 100% purity and 3% yield (**Table 3-2**). This demonstrates how much more easily a small alkyne fragment reacted in the CuAAC reaction compared to the bulky alkyne peptide fragment.



Scheme 3-4: Synthesis of the triazole-linked peptide **20**, using a combination of solution phase chemistry and SPPS.

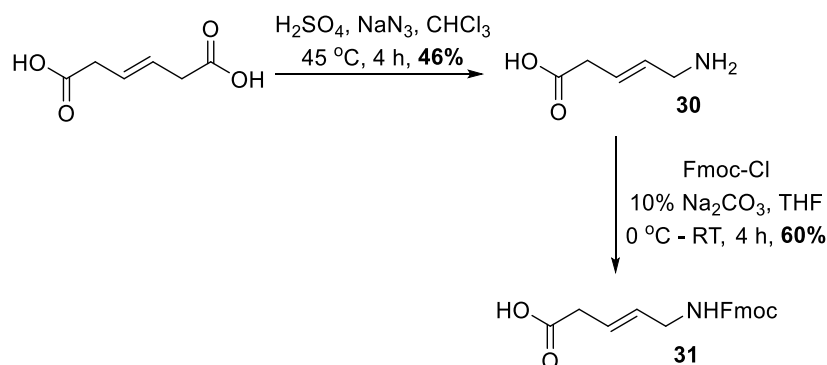
Alkene-linked peptide (**17**)

The 2-step synthesis of alkene **17** (**Scheme 3-6**) was taken from the literature.¹⁷³ Firstly the Schmidt reaction was used, to convert one of the carboxylic acids of commercially available trans- β -hydromuconic acid into an amine and obtain **30**. For this, sodium azide and a catalytic amount of conc. sulfuric acid were added to the dicarboxylic acid, and heated at 45 °C for 4 h, to afford amine **30**. The reaction mechanism (**Scheme 3-5**) proceeds *via* the protonation of a carboxylic acid, leading to the formation of an acylium ion, after the removal of a H₂O molecule. Hydrazoic acid, formed from sodium azide in the presence of conc. sulfuric acid, then reacts with the acylium ion to afford a protonated azido ketone. Subsequent rearrangement provides a protonated isocyanate and the removal of an N₂ group. Attack by water yields a carbamate, that after proton transfer is decarboxylated, to give the desired amine **30** and CO₂. When isolating **30**, due to its high polarity, the product remained in the aqueous phase after work-up. Therefore, to remove any excess conc. sulphuric acid, cation exchange chromatography was carried out using the Dowex 50WX8-100 resin.¹⁷³



Scheme 3-5: Mechanism for the Schmidt reaction, used for the synthesis of amine **30**.¹⁷⁴

Next, Fmoc protection was carried out on the amine, by dissolving **30** in 10% sodium carbonate, before adding Fmoc-Cl in THF at 0 °C. The reaction was then allowed to warm to RT and stirred for 4 h to afford the Fmoc-protected amino acid **31**. This could be used directly in microwave assisted SPPS, allowing for straight-forward synthesis of the desired alkene-linked peptide **17**, obtained with 98% purity and 8% yield (**Table 3-2**).

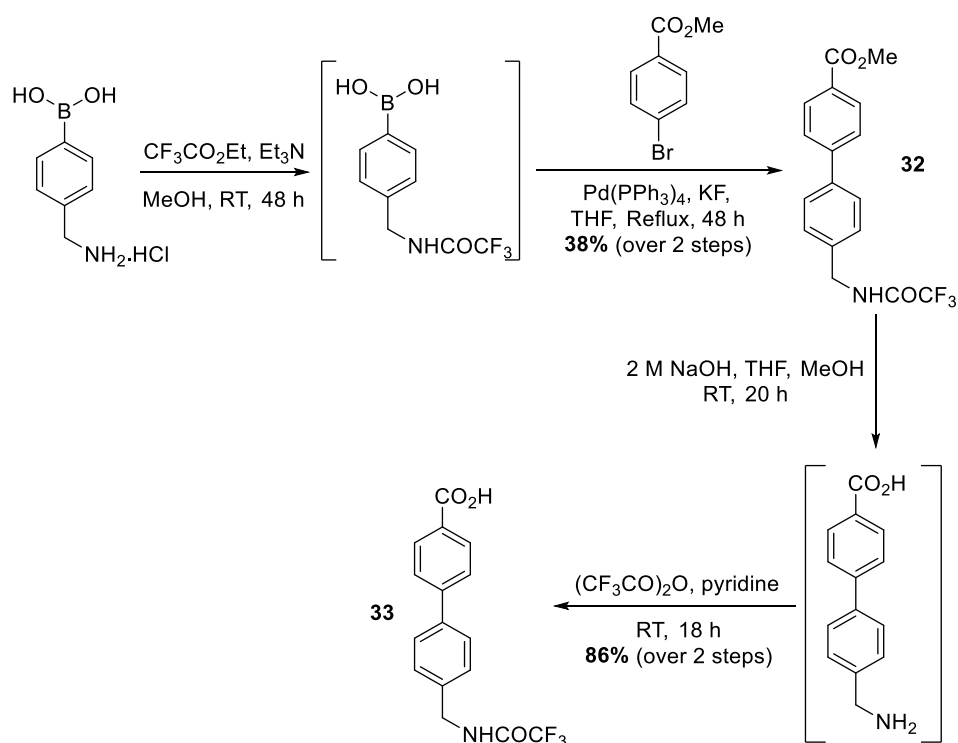


Scheme 3-6: Synthesis of the alkene-containing Fmoc-protected amino acid **31**.

Dibenzene-linked peptide (**19**)

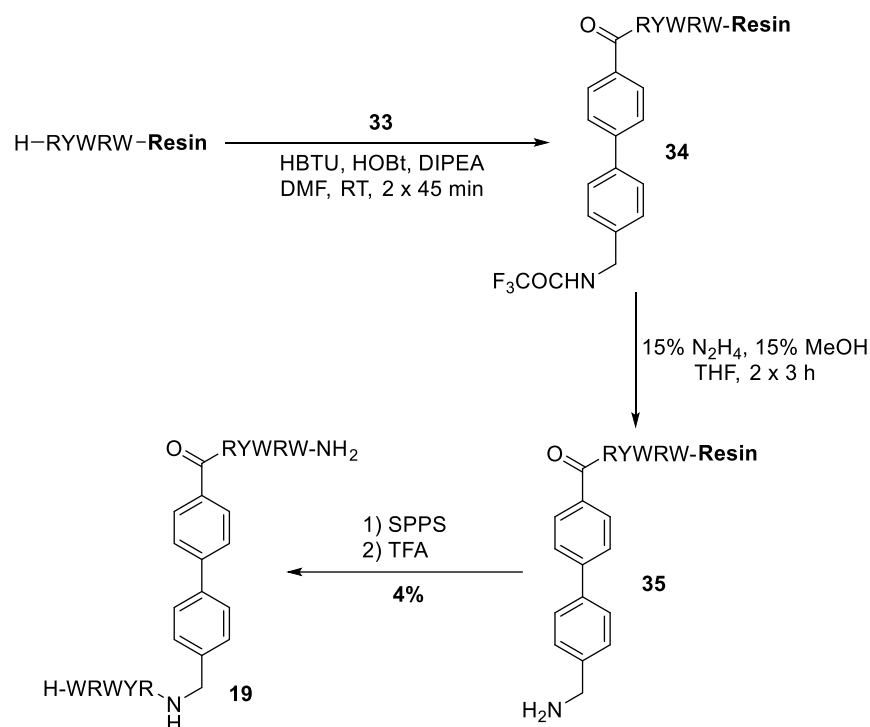
The synthesis of the dibenzene linker was based on a study into developing rigid linkers for peptide ligation, as discussed in **section 3.1.1**.¹⁵² Instead of the usual *N*-Fmoc, *N*-TFA was used as the protecting group for the dibenzene linker amino acid, to provide adequate solubility in DMF required during SPPS. This highlights the high hydrophobicity of this linker. Unlike Fmoc, the TFA protecting group cannot be removed by treatment with piperidine and requires the use of hydrazine instead.

Synthesis began with the TFA protection of commercially available 4-aminomethylphenylboronic acid hydrochloride, using ethyl trifluoroacetate and triethylamine. After work-up, the crude product was immediately used in a Suzuki-Miyaura cross-coupling reaction with methyl 4-bromobenzoate, using $\text{Pd}(\text{PPh}_3)_4$ as a catalyst and refluxing in THF, to afford **32**. Next, a hydrolysis reaction resulted in the production of an amino acid, from the reduction of ester and amide groups. After work-up the crude product was used directly in the next reaction, that involved reattaching the TFA protecting group to the primary amine, in the presence of trifluoroacetic anhydride and pyridine, to afford the final protected amino acid linker **33**.



Scheme 3-7: Synthesis of the dibenzene-containing TFA-protected amino acid **33**.¹⁵²

Peptide **19** was made using both manual and automated peptide synthesis (**Scheme 3-8**). The TFA-protected amino acid **33** was manually added to the peptide fragment **Resin-WRWYR-H** using standard coupling conditions (**Section 6.2.1**), with completion of the amide coupling confirmed by a negative Kaiser test (**Section 6.2.2**), to afford **34**. The TFA protecting group was removed using 15% hydrazine (2 x 3 h)¹⁵² to give **35**, before the resultant peptide was resubjected to SPPS. After TFA cleavage the dibenzene-linked peptide **19** was obtained, with 99% purity and 4% yield (**Table 3-2**).



Scheme 3-8: Synthesis of dibenzene-linker peptide **19**.

When carrying out the TFA deprotection of the dibenzene amino acid **33** on resin, optimisation of reaction length was required. After 3 h of treatment with 15% hydrazine, in 15% MeOH and THF, HPLC analysis of an aliquot of cleaved resin showed that ~25% removal of the TFA group was achieved. To increase the yield of desired product, the reaction mixture was subjected to the hydrazine mixture for a further 3 h, with ~50% TFA removal observed. However, after an additional 3 h both the starting material and product peaks disappeared, with degradation indicated by the formation of several smaller peaks (**Figure 3-11**). Based on these results, the optimum reaction time for TFA removal was 2 x 3 h.

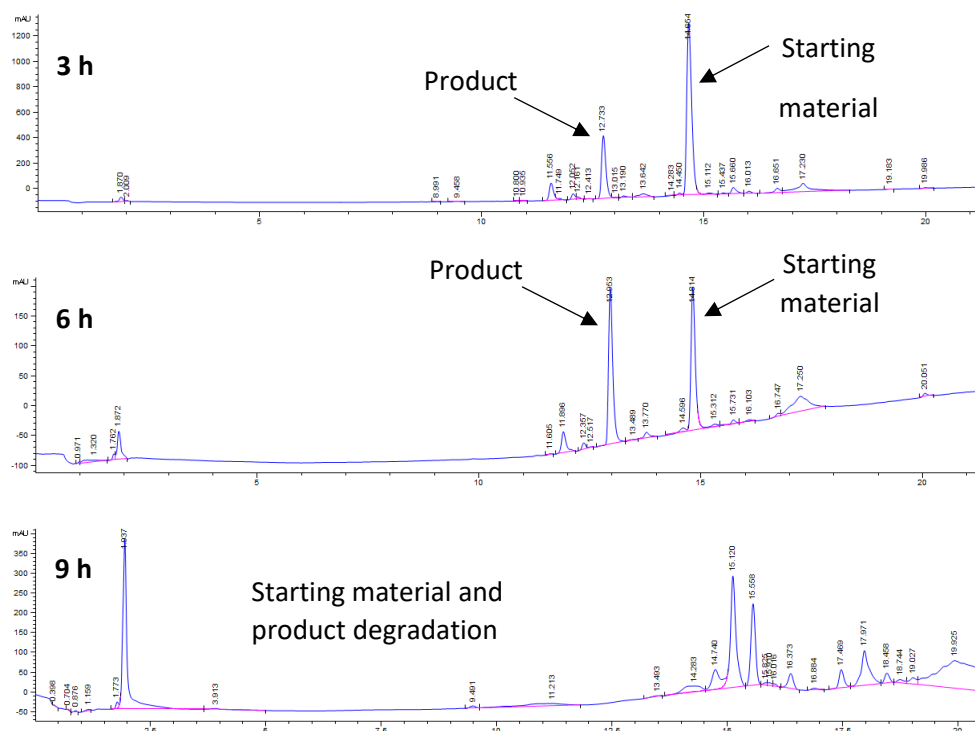
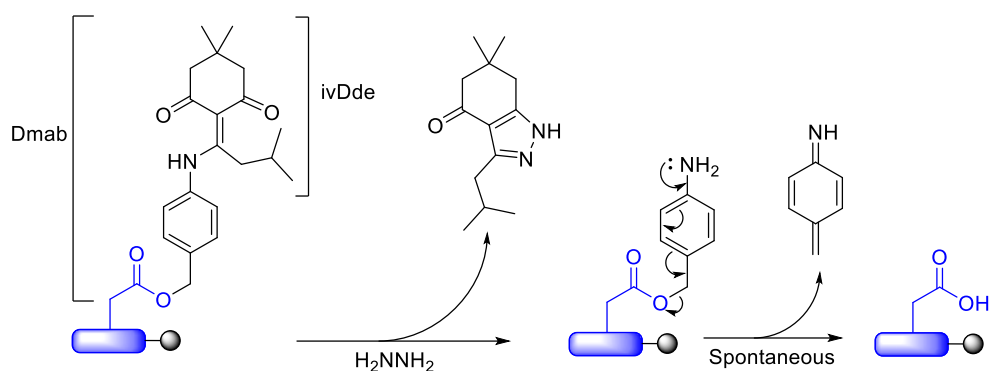


Figure 3-11: HPLC traces (214 nm, 5-95% MeOH in H₂O + 0.05% TFA over 15 min, then 5 min at 95%) showing the conversion of H-WRWYR-Dibenzene-TFA (starting material) to H-WRWYR-Dibenzene-H (product) after treatment with hydrazine for 3 h, 6 h and 9 h.

3.3.5 Synthesis of cyclic peptides

The method for the synthesis of the Gly-Gly (**23**) and 6-alkane (**24**) linked cyclic peptides was based on the prior synthesis of chlorofusin.¹⁷⁰ For this, an Asp amino acid was used, with the carboxylic acid from the backbone protected instead of side chain. This meant that when reacted with resin, the Asp side chain was immobilised rather than the backbone. The protecting group used for the Asp carboxylic side chain needed to stay in place during the synthesis of the linear peptide, before being selectively removed before head-to-tail cyclisation could occur. For this purpose, the 4-*N*-[1-(4,4-dimethyl-2,6-dioxocyclohexylidene)-3-methylbutyl]amino}benzyl ester (Dmab) protecting group was used. Dmab is stable in both 20% piperidine in DMF and TFA, making it an excellent protecting group for orthogonal use during Fmoc-SPPS. Removal of Dmab occurs within minutes of treatment with 2% hydrazine in DMF at RT and results in quantitative deprotection.¹⁷⁵ This occurs *via* a 2-step process, with the initial removal of the 1-(4,4-dimethyl-2,6-dioxocyclohexylidene)-3-methylbutyl (ivDde) group using hydrazine, followed by the spontaneous 1,6-elimination of the aminobenzyl moiety (**Scheme 3-9**).¹⁷⁶



Scheme 3-9: Dmab deprotection mechanism using hydrazine. Resin is represented as a black ball and the peptide as a blue square, with the Asp side chain also shown in blue.¹⁷⁶

Synthesis of the cyclic peptides (**Scheme 3-10**) began with the immobilisation of Fmoc-Asp-ODmab onto Tentagel S RAM resin. Initially this was attempted using microwave assisted SPPS. However, degradation of the Dmab protected peptide was observed under these conditions, shown by the deletion of this amino acid from the linear peptide sequence during MALDI-ToF MS analysis. Therefore, Fmoc-Asp-ODmab was added to the resin at 0 °C with shaking (2 x 30 min) and activated with HATU. A negative Kaiser test confirmed that coupling was complete (**Section 6.2.2**). The resin was moved back onto the automated peptide synthesiser, and SPPS used to complete the linear peptide sequences, using either the Gly-Gly or 6-alkane linker groups. Dmab deprotection was carried out using 2% (v/v) hydrazine hydrate in DMF (2 mL, 3 x 3 min), with reaction completion confirmed by HPLC analysis. To cyclise the peptide *via* an amide coupling, DIC and HOBt in DMF were added to the resin, followed by shaking at RT (2 x 24 h). Although this was a slow cyclisation, a base-free amide coupling reaction was used to reduce levels of oligomerisation and epimerisation.¹⁷⁷ After cyclisation was complete, as shown by a negative Kaiser test, cleavage from the resin using TFA afforded Gly-Gly **23**, with 90% purity and 3% yield, and 6-alkane **24**, with 95% purity and 8% yield (**Table 3-2**).

showed a drop-in activity, with incomplete trapping of the 4WJ with 20 μM . This may indicate a preference for the 7-atom flexible linkers.

Table 3-3: Results from peptide analogues testing using the PAGE assay, using 1.25 μM FAM-labelled 4WJ DNA.

No.	Linker group	Full 4WJ trapping (μM)	Band disappearance (μM)
2	Gly-Gly	12	~50
23	Cyclic Gly-Gly	8	10
14	6-alkane	12	16
21	D-Asn-L-Asn	12	16
20	Triazole	12	14
19	Di-benzene	12	14
15	PEG1	12	16
24	Cyclic 6-alkane	12	14
22	D-Pro-L-Pro	14	16
18	Benzene	14	16
16	5-alkane	>20	>20
17	Alkene	>20	>20

When peptides **1** and **2** were tested at high concentrations (**section 2.3.3**), at a certain point the DNA became trapped at the top of the gel wells, causing the bands to become weaker and eventually completely disappear. For **2**, a concentration of ~50 μM was reached before band disappearance. This is much higher than the concentrations reached by the peptide analogues (**Table 3-3**), indicating that these compounds may be more prone to issues with aggregation and solubility. Out of the peptide analogues, the shorter linkers, 5-alkane **16** and alkene **17**, had the highest concentrations before band disappearance and the cyclic Gly-Gly peptide **23** had the lowest (**Figure 3-12**). This implies a trend, where the lower the concentration required for 4WJ trapping the lower the concentration where band disappearance occurs. This may show that the complex, formed between the 4WJ and peptide analogues, has poor solubility and gets trapped before it can travel through the gel. This could also indicate aggregation, with the peptide analogues binding to multiple 4WJs and causing larger complexes to form.

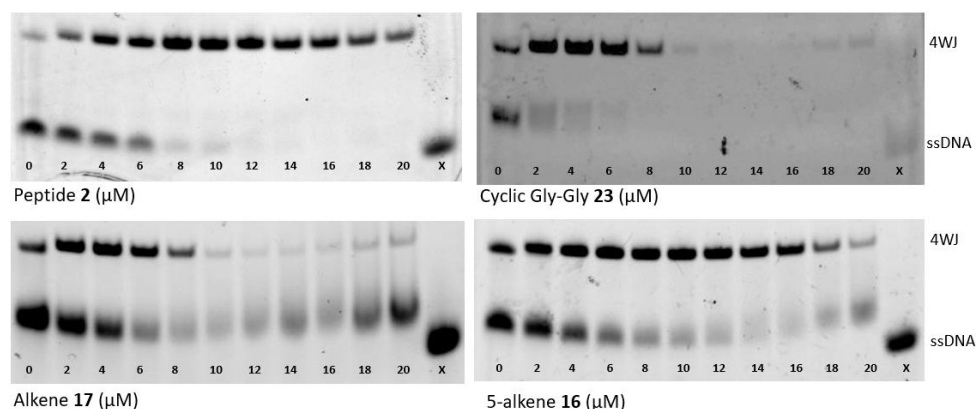


Figure 3-12: PAGE image showing the effect on 4WJ formation from increasing concentrations of peptide **2**, cyclic Gly-Gly **23**, alkene **17** and 5-alkane **16**, using a DNA concentration of 1.25 μM . A solution containing only the labelled X-strand was used as a control for ssDNA. FAM-labelled DNA was used and visualised at 460 nm.

3.3.7 Fluorescence polarisation

As an orthogonal assay, to determine the binding affinity of the peptide analogues to 4WJ DNA, the FP assay was repeated following methods used in **section 2.3.5**. The overlapping of dose-response curves showed the peptide analogues gave relatively similar results (**Figure 3-13**). Overall, peptide **2** had the strongest binding affinity, although the weakest binder, D-Asn-L-Asn **21**, showed a <2-fold reduction in activity (**Table 3-4**). This contrasts the results from the PAGE assay, where peptide **21** and **2** both trapped the 4WJ at the same concentration. In addition, unlike in the PAGE assay, alkene **17** showed slightly higher activity than cyclic Gly-Gly **23**. Together these results may indicate that this ~2-fold difference in 4WJ activities in both the FP and PAGE assays is insignificant. Alternatively, the differences in relative activities in the PAGE assay could have been caused by the solubility/aggregation issues discussed. Additionally, in the FP assay all peptides caused a similar change in polarisation, within 40 mP of each other, most likely due to the similar structure and size of the peptides. These results indicate that the change in linker had modest effects on the binding of the peptide analogues to the FAM-J3 4WJ in the FP assay.

Table 3-4: Results the FP assay testing the peptide analogues, using 10 nM FAM-J3 4WJ. Error in the EC_{50} and polarisation change values are from the mean $\pm$ standard deviation from triplicate readings.

No.	Linker group	EC_{50} (nM)	Polarisation change (mP)
2	Gly-Gly	310 $\pm$ 2.54	249 $\pm$ 0.939
19	Di-benzene	367 $\pm$ 7.79	278 $\pm$ 4.70
20	Triazole	409 $\pm$ 1.92	263 $\pm$ 1.72
22	D-Pro-L-Pro	439 $\pm$ 4.40	256 $\pm$ 0.700
14	6-alkane	455 $\pm$ 4.57	268 $\pm$ 2.01
15	PEG1	484 $\pm$ 17.7	237 $\pm$ 8.38
24	Cyclic 6-alkane	506 $\pm$ 43.9	265 $\pm$ 11.5
18	Benzene	514 $\pm$ 12.6	258 $\pm$ 0.698
17	Alkene	534 $\pm$ 7.47	265 $\pm$ 2.81
23	Cyclic Gly-Gly	540 $\pm$ 21.0	267 $\pm$ 10.2
16	5-alkane	594 $\pm$ 6.81	254 $\pm$ 0.736
21	D-Asn-L-Asn	605 $\pm$ 8.13	258 $\pm$ 5.00

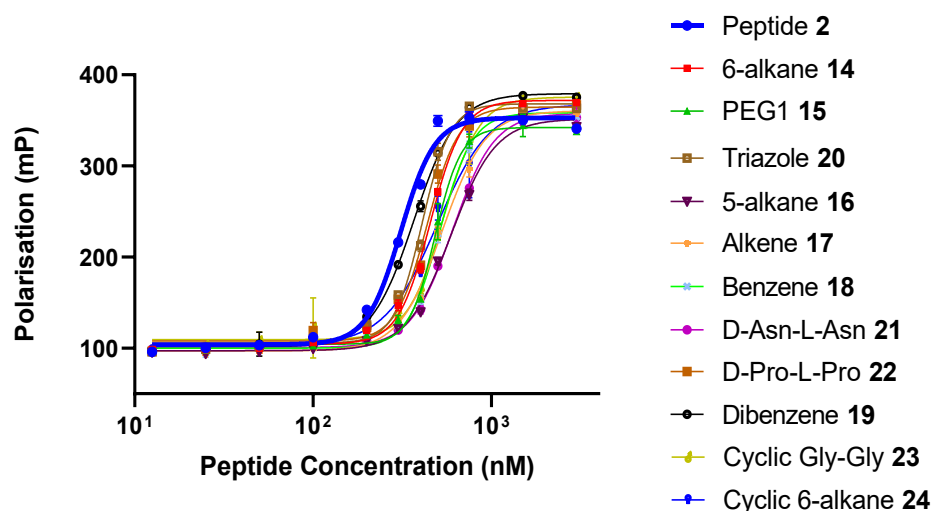


Figure 3-13: Dose-response curves from the FP analysis of the peptide analogues and peptide 2, using 10 nM FAM-J3 4WJ. Error bars represent the mean $\pm$ standard deviation from triplicate readings.

3.3.8 Ethidium bromide displacement assay

To test the binding affinity for the peptide analogues against dsDNA, the EtBr displacement assay was repeated, following methods from **section 2.3.6**. Overlapping dose-response curves showed the similarity in activity between the different peptide analogues (**Figure 3-14**). As with the FP and PAGE assays, all results

were within 2-fold of each other. Unlike in the FP assay, peptide **2** did not have the greatest activity, indicating its higher selectivity for the 4WJ compared to some of the peptide analogues. For example, alkene **17** showed the strongest binding to dsDNA but some of the lowest activities in both the PAGE and FP assays. This may show a stronger preference for binding to the duplex arms of the 4WJ than the centre of the junction. This contrasts with the results for 5-alkane **16**, that had low activity in both the 4WJ and dsDNA binding assays. The only difference between these 2 peptides is the presence of a single or double bond in the centre of the linkers and could show that the alkene group directly interacts with dsDNA. However, the results from all peptide analogues from across the three different binding assays show only slight differences in relative activities, with no obvious overall trends. This points to potential insignificance in the variations in results.

Table 3-5: C_{50} values for the peptide analogues from the EtBr displacement assay. Error in the C_{50} values is from the mean $\pm$ standard deviation from triplicate readings.

No.	Linker group	C_{50} (μ M)
2	Gly-Gly	1.48 $\pm$ 0.0453
17	Alkene	1.01 $\pm$ 0.0542
18	Benzene	1.41 $\pm$ 0.0234
20	Triazole	1.54 $\pm$ 0.0025
21	D-Asn-L-Asn	1.60 $\pm$ 0.0418
14	6-alkane	1.69 $\pm$ 0.0306
22	D-Pro-L-Pro	1.80 $\pm$ 0.123
19	Di-benzene	2.15 $\pm$ 0.393
24	Cyclic 6-alkane	2.17 $\pm$ 0.0714
15	PEG1	2.49 $\pm$ 0.0478
16	5-alkane	3.01 $\pm$ 0.0179
23	Cyclic Gly-Gly	3.02 $\pm$ 0.1179

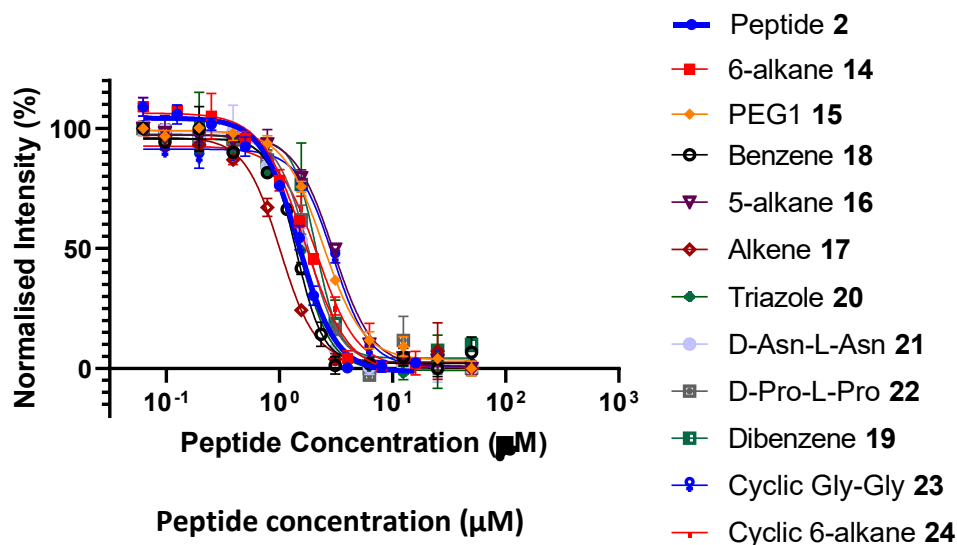


Figure 3-14: EtBr displacement assay dose-response curves for the peptide analogues and peptide **2**. Error bars represent the mean $\pm$ standard deviation from triplicate readings.

3.3.9 FRET

To test the effect of peptide analogue binding on the 4WJ conformation, the FRET assay was used, following methods from **section 2.3.4**. As preceding assays had demonstrated little difference in the binding attributes of the various analogues, for the FRET assay, it was decided to focus on the cyclic versus the linear analogues, to look for any substantial variations. As outlined in **section 2.3.5**, the results from this experiment cannot be conclusively attributed to FRET without further validation and should therefore be interpreted with a level of caution.

When comparing the cyclic and linear versions of the Gly-Gly peptides, the cyclic version **23** showed slightly higher activity (**Table 3-6**). The shape of the binding curves for both peptides were very similar (**Figure 3-15**), with FRET reductions observed straight away, from the first addition of peptide at 5 nM. This indicates that even low concentrations of the peptides start opening the 4WJ. Cyclic Gly-Gly **23** also caused a larger change in FRET efficiency (**Table 3-6**), that may have been caused by the more rigid structure and increased size of the cyclised peptide opening up the junction more. These results show that cyclisation of peptide **22** is well tolerated for 4WJ binding.

Moving onto the 6-alkane linked linear and cyclic peptides, both showed lower levels of 4WJ opening $<1.7 \mu\text{M}$ (**Table 3-6**) compared to the Gly-Gly linked peptides. Cyclic

6-alkane **24** caused FRET to lower from a concentration of ~5 nM, although to a smaller extent than the Gly-Gly linked peptides. In contrast linear 6-alkane **14** caused no change in FRET until concentrations >300 nM were reached. Based on the similarities in activities in the PAGE and FRET assays, this could indicate that 6-alkane **14** has a different 4WJ-binding mode. Perhaps this peptide binds to the X-stacked conformation as a groove binder, rather than forming a turn at the junction's centre. Only when the DNA grooves are saturated does 6-alkane **14** bind to the 4WJ in the same way as peptide **2**. As mentioned in **section 2.4**, additional control experiments are required to confirm these findings and could be further validated using time-resolved FRET and single-molecule FRET assays.

Table 3-6: FRET results for cyclic and associated linear peptides, using 50 nM FAM/TAMRA J3 DNA. Error in K_D values were calculated from the mean $\pm$ standard deviation, from triplicate readings.

No.	Linker group	K_D (nM)	Change in FRET efficiency
2	Gly-Gly	96.4 ± 22.8	0.154
23	Cyclic Gly-Gly	66.7 ± 21.2	0.168
14	6-alkane	N/A	0.136
24	Cyclic 6-alkane	N/A	0.116

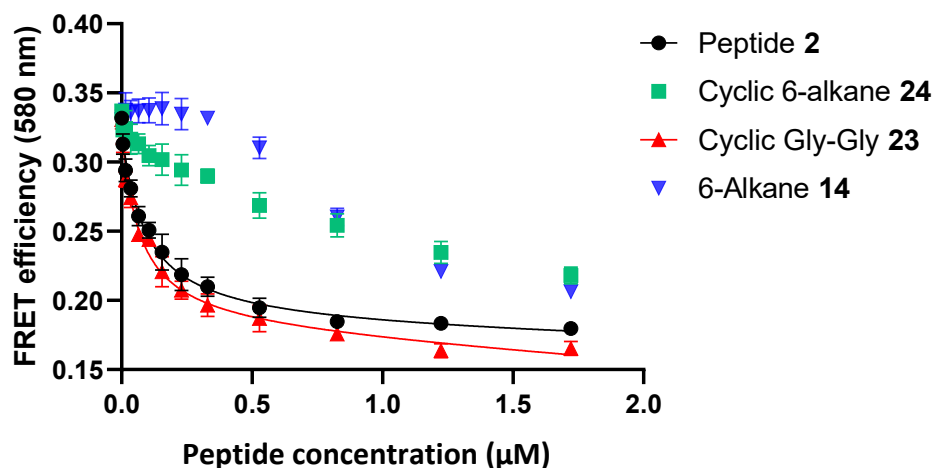


Figure 3-15: Dose-response curve from FRET analysis of cyclic and associated linear peptides, using 50 nM FAM/TAMRA J3 DNA. Error bars represent the mean $\pm$ standard deviation from triplicate readings.

3.3.10 MTS assay

Finally, the peptide analogues were tested in the MTS assay, to look at their levels of cytotoxicity in HL-60 cancer cells, following methods from **section 2.3.7**. Screening the peptide analogues at 25 μ M showed that none caused a >10% decrease in cell

viability (**Figure 3-16**). Comparing the results from the peptide analogues to the DMSO negative control, using a two-tailed unpaired t-test, showed none caused a significant change in cell viability (all *p*-values were >0.05) (**Table 3-7**). This demonstrates that none of the linkers tested improved the cytotoxicity of the peptide analogues against HL-60 cells, when at a 25 μ M concentration.

Table 3-7: Results from the MTS assay using peptide analogues. Errors in the cell viability (%) values are from the mean $\pm$ standard deviation from triplicate readings. P-/T-values were obtained from a two-tailed unpaired t-test, comparing samples to cells treated with DMSO.

No.	Name/linker group	Cell viability (%)	P-value	T-value
2	Gly-Gly	101 $\pm$ 3.73	0.768	-0.315
14	6-alkane	104 $\pm$ 5.77	0.504	-0.733
17	Alkene	95.7 $\pm$ 5.04	0.286	1.23
21	D-Asn-L-Asn	99.4 $\pm$ 2.88	0.829	0.231
22	D-Pro-L-Pro	106 $\pm$ 5.83	0.223	-1.53
20	Triazole	95.1 $\pm$ 3.05	0.145	1.81
15	PEG1	101 $\pm$ 4.49	0.882	-0.158
18	Benzene	96.1 $\pm$ 4.82	0.317	1.14
24	Cyclic 6-alkane	91.3 $\pm$ 4.28	0.0763	2.38
19	Dibenzene	90.1 $\pm$ 0.249	0.0404	4.82
23	Cyclic Gly-Gly	94.4 $\pm$ 0.657	0.117	2.66

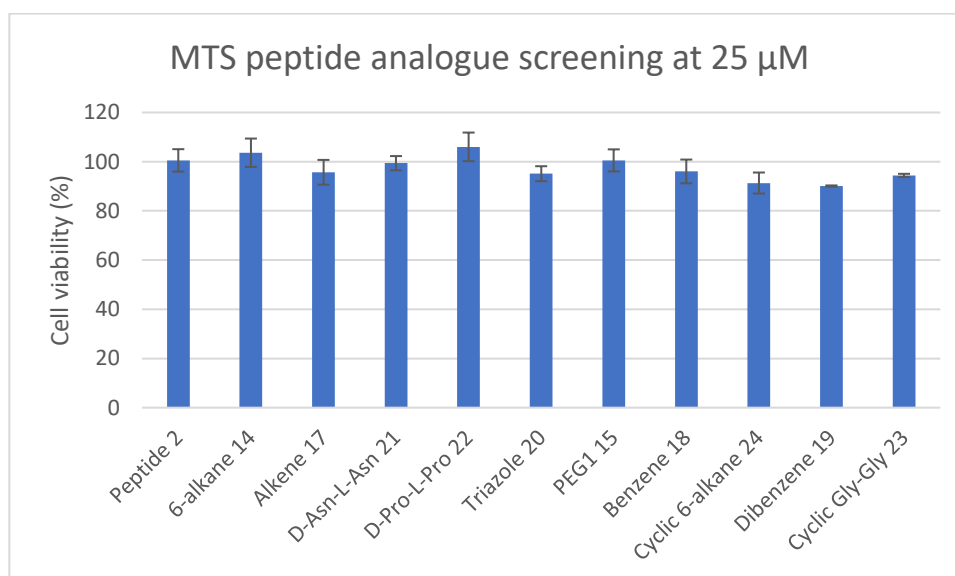


Figure 3-16: MTS assay cell viability assay with peptide **2** or peptide analogue (25 μ M) against HL-60 cells (3×10^5 cells/mL). Error bars represent the mean $\pm$ standard deviation from triplicate readings.

Based on the similar activities of the peptide analogues, it was decided that full binding curves for cell viability (%), using the MTS assay, would only be conducted with the cyclic/linear peptides tested during FRET. IC₅₀ values could be calculated for peptide **2** and cyclic 6-alkane **24** (**Figure 3-17**), however could only be predicted by visual interpretation for cyclic Gly-Gly **23** and 6-alkane **14**, as a result of incomplete binding curves (**Figure 3-18**). Due to the predicted low activity of these two peptides (IC₅₀ values = ~100 µM), the decision was made not to repeat the MTS assay to obtain exact IC₅₀ values.

As with the PAGE, FP and EtBr displacement assays, the activities of the peptides had a <2-fold difference (**Table 3-8**). Cyclic 6-alkane **24** had the highest activity but did not show overall higher activity in the other assays, indicating that this may be caused by a mechanism other than 4WJ or dsDNA binding.

Table 3-8: IC₅₀ values of linear and cyclic peptide analogues in the MTS assay, showing the concentration required to obtain a 50% reduction in cell viability. Errors in the IC₅₀ values are from the mean ± standard deviation from triplicate readings

No.	Linker group	IC ₅₀ (µM)
2	Gly-Gly	96.9 ± 15.8
23	Cyclic Gly-Gly	~100
14	6-alkane	~100
24	Cyclic 6-alkane	65.5 ± 2.37

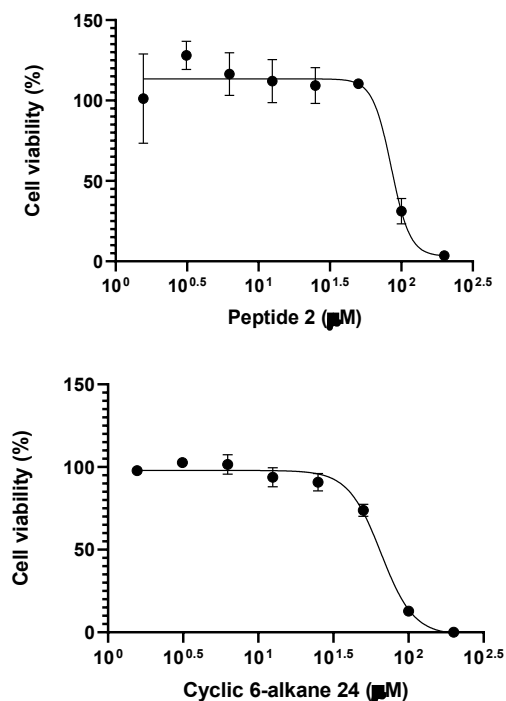


Figure 3-17: Dose-response curves in the MTS cell viability assay for peptide **2** and cyclic 6-alkane **24**, using HL-60 cells (3×10^5 cells/mL). Error bars represent the mean $\pm$ standard deviation from triplicate readings.

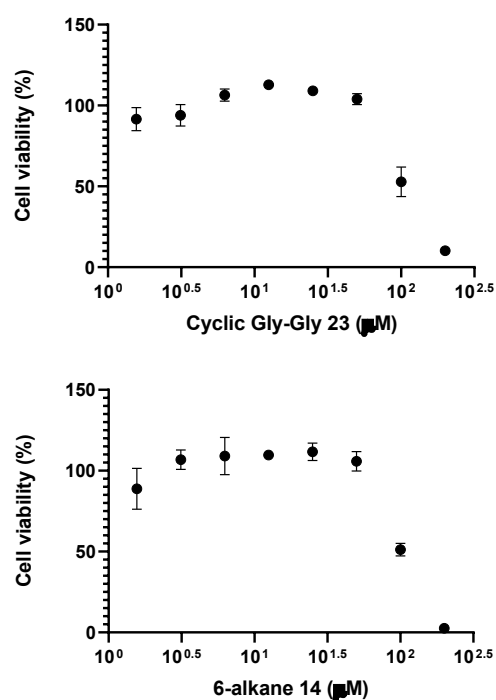


Figure 3-18: Dose-response curves in the MTS cell viability assay for cyclic Gly-Gly **23** and 6-alkane **14**, using HL-60 cells (3×10^5 cells/mL). Error bars represent the mean $\pm$ standard deviation from triplicate readings.

3.4 Conclusion and future work

To summarise, although peptides are promising therapeutic molecules, they often have issues with *in vivo* instability. To improve the drug-like properties of peptides, peptidomimetics can be used, to replace parts of the peptides structure with more stable chemical groups whilst retaining biological activity. Techniques for designing peptidomimetic molecules include adding D-isomer or non-canonical amino acids or cyclising the peptide sequence. In **chapter 2**, peptide **2** was shown to be a strong 4WJ-binder in various biological assays but showed no cytotoxicity at 25 μ M in the MTS assay against HL-60 cancer cells. To improve the 4WJ activity and *in vivo* stability of peptide **2**, peptidomimetic molecules were designed. This involved replacing the central Gly-Gly amino acids with different linkers, of various lengths and functionalities. Additionally, cyclic peptides were designed to see how this effected 4WJ binding. These peptide analogues were synthesised using a combination of multi-step solution phase chemistry and SPPS. Nine linear peptide analogues and two cyclic peptides were successfully synthesised, although with low yields, potentially resulting from poor aqueous solubility and unoptimised synthetic routes, with only $\sim$ 1 mg of ligand being required for initial biological testing.

The peptide analogues were tested for their biological activity, using the assays from **chapter 2**, to allow for the direct comparison of activity to peptide **2**. Binding to the 4WJ was measured in the PAGE/FP assays and dsDNA binding in the EtBr displacement assay. Results showed that all peptide analogues had similar activities to peptide **2**. A $\sim$ 2-fold difference in activities was observed across the ligands, but the relative ordering of activities changed between the different assays, indicating that these variations may have been insignificant. Additionally, testing of all peptide analogues at a concentration of 25 μ M in the MTS assay showed that none of the ligands had substantial cytotoxic effects in HL-60 cancer cells at this concentration. Nonetheless, further cytotoxicity testing across a broader panel of cancer cell lines, such as PC3, Du145, LnCAP, PPC-1, DuPro-1 (prostate cancer models), and A549 (lung carcinoma) previously used by the Segall group,⁷¹ could be used to look for any cell specific susceptibilities to any of these ligands. Minimal inhibitory concentrations of the peptide analogues against various gram positive and negative bacterial strains could also be investigated, with antibacterial activity previously reported for peptide **2**.⁴⁵

Overall, the peptide analogues had similar activity against DNA 4WJs, showing that a range of changes can be made to peptide **2** whilst maintaining 4WJ binding. Further ideas for peptide analogues could include synthesising the equivalent peptide analogues using only D-amino acids, to see if this would improve *in vivo* activity. Another idea could be to use more hydrophilic linker groups to help resolve potential solubility issues, such as extended PEG chains¹⁷⁸ and linkers containing charged amino acids. To further develop 4WJ-binding peptidomimetics, obtaining a crystal structure of peptide **2** bound to 4WJ DNA would provide crucial insight into binding interactions. From this, the active conformation of the peptide could be determined and used to guide the design of future peptide analogues. Additionally, an alanine scan of peptide **2** could be carried out, to find out the amino acids that are non-crucial to 4WJ-binding and replace them with non-canonical amino acids, to improve 4WJ-binding and drug-like properties.

Chapter 4 – Developing 4WJ-binding peptides from protein-4WJ crystal structures

4.1 Introduction

4.1.1 Identifying therapeutically active peptides

As well as developing peptides based on the peptide **2** WRWYRGGRYWRW sequence, another approach could be to discover novel 4WJ-binding peptides. Two common methods for identify therapeutically active peptides are:

1. Randomised high-throughput screening of diverse combinatorial peptide libraries, against a target of interest. This method is often used when structural information is limited.¹⁷⁹
2. The rational design of peptides, based on information about the structure and function of a protein of interest interacting with its target, looking at interactions within protein-target complexes.¹⁷⁹

The first 4WJ-binding peptides, WKHYN Y and peptide **1** WRWYCR, were discovered by Segall *et al.*^{63,65} using the approach of randomised high-throughput screening. The peptides were tested against different bacteriophage λ site-specific recombination pathways, with inhibition resulting in the accumulation of 4WJ DNA, as discussed in **sections 1.4.1.1 – 1.4.1.2**. This involved using a technique known as positional scanning, to screen a hexapeptide combinatorial library.^{63,65} This method involved creating large mixtures of peptides using SPPS and testing these in dose response assays. The procedure for the discovery of the first 4WJ-binding peptide, WKHYN Y, will now be described (**Figure 4-1**).^{63,180} Firstly, a position in the hexapeptide was fixed as a specific residue, out of the 20 canonical amino acids. With there being 6 amino acids in the chain, this created 120 possible fixed positions. To the remaining 5 positions were added a mixture of 19 canonical amino acids (with cysteine excluded). This gave a total of 19^5 different peptides for each of the 120 different fixed positions, creating 300 million different sequences. These peptides were synthesised using SPPS, by coupling 19 equimolar equivalents of Boc-protected L-amino acids to solid resin beads for the mixed positions and carrying out a general amide coupling, using just one amino acid, for the fixed position.¹⁸⁰ After cleaving the peptide mixtures from the resin, the resultant 120 mixtures were directly tested against different bacteriophage λ site-specific recombination pathways. By keeping one of the positions fixed, the effect of a single amino acid could be determined, independent of other residues in the sequence, to find the preferred residue at each position of the peptide.¹⁸¹ This allows for the easy deconvolution of the mixtures of

peptides. The fixed amino acid, in the mixtures that caused the highest accumulation of 4WJ, were identified and chosen for a second round of screening of the top 50 most promising mixtures. This involved fixing two positions with specific amino acids and having a mixture in the remaining 4 positions. This gave a total of 19^4 amino acids per 50 mixtures to give a total of 6.5 million peptides. After repeating the testing with the new mixtures, the best pairs of amino acids were chosen for each position, and 12 peptides were synthesised and tested individually.⁶³

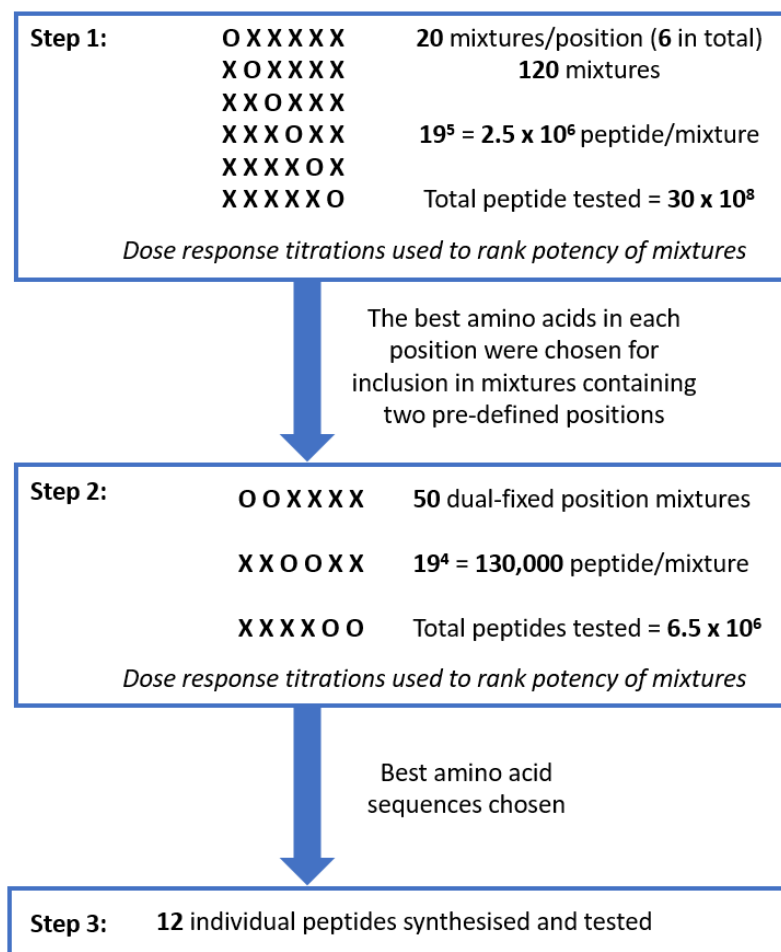


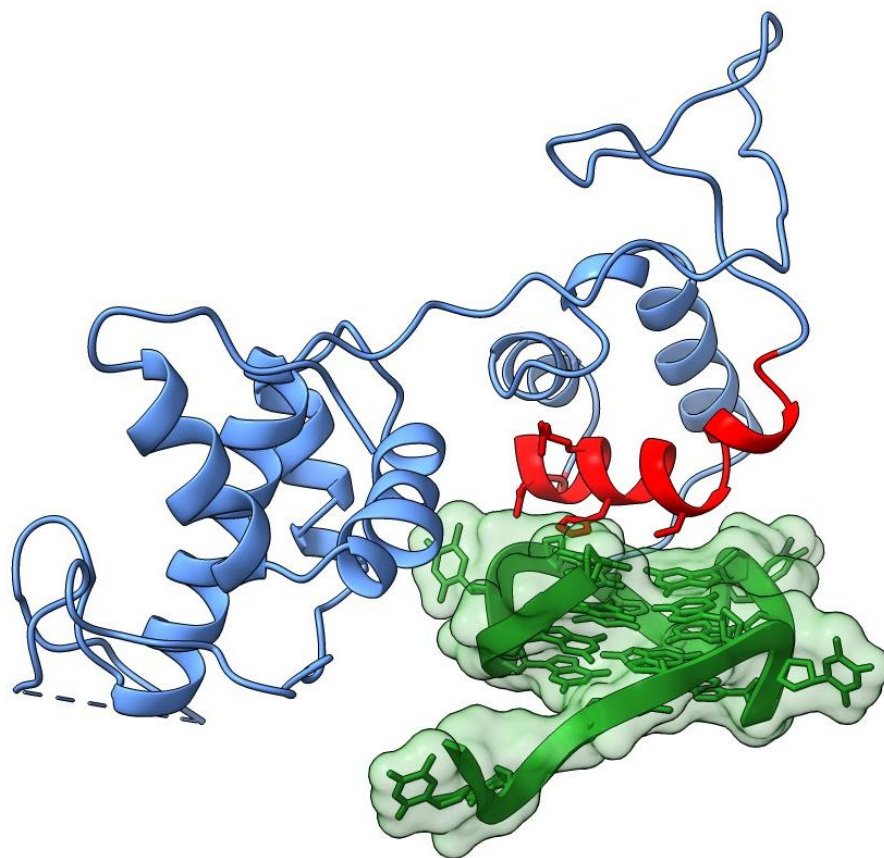
Figure 4-1: Positional scanning of a synthetic combinatorial hexapeptide library used for the discovery of WKHYN^Y.¹ **Step 1:** A single position of the hexapeptide was fixed (O) as one of the 20 canonical amino acids, whilst the remaining 5 positions are mixtures of 19 different residues (cysteine excluded). Peptide mixtures were then tested in dose response assays. **Step 2:** The most potent amino acid combinations were used for hexapeptides containing two fixed and 4 mixed positions. **Step 3:** The best candidates were chosen for individual synthesis and testing.^{63,180}

Using this combinatorial library screening method, WKHYN^Y was found to be able to trap 40 – 60% of 4WJs, by blocking the action of the bacteriophage λ site-specific recombination bent-L pathway. However, a second round of screening was required, to discover peptide **1** that was able to fully trap the 4WJ in all four bacteriophage λ

site-specific recombination pathways.⁶⁵ Per screening iteration, >300 million peptides were synthesised, although by using a combinatorial approach only ~170 mixtures needed to be prepared. This is a substantial amount of work and a time-consuming process. Although a wide variety of hexapeptides were tested, these combinatorial libraries were limited to a single peptide length and the use of only canonical amino acids, leaving a vast amount of chemical space unexplored.

A different approach is to rationally design 4WJ-binding peptides based on structural information. This method has recently been used for the discovery of peptides that bind to G-quadruplex (G4) DNA,¹⁸² a different higher-order DNA structure discussed in **section 1.2**. Rap1 is a protein derived from yeast, that has been found to bind to G4s and promote their formation *in vitro*.¹⁸³ This protein is found at telomeres, with G4 sequences highly evolutionarily conserved between yeast and human telomeric G4 DNA.¹⁸⁴ Therefore, by basing a peptide on interactions between Rap1 and G4 DNA, it was hoped these peptides would be able to bind to human telomeric G4 sequences as well. Analysis of the Rap1-G4 crystal structure showed that a 19 amino acid sequence (397 – 415) acted as the DNA recognition motif, forming hydrophobic, stacking and polar interactions with the G4 DNA (**Figure 4-2**). Biochemical testing of this 19-amino acid peptide, against various human G4 sequences, showed the ability of the peptide to stabilise the G4 conformations. Furthermore, the peptide was shown to form minimal interactions with duplex DNA, demonstrating how this approach can be used to specifically target high-order DNA structures.¹⁸²

A)



B)

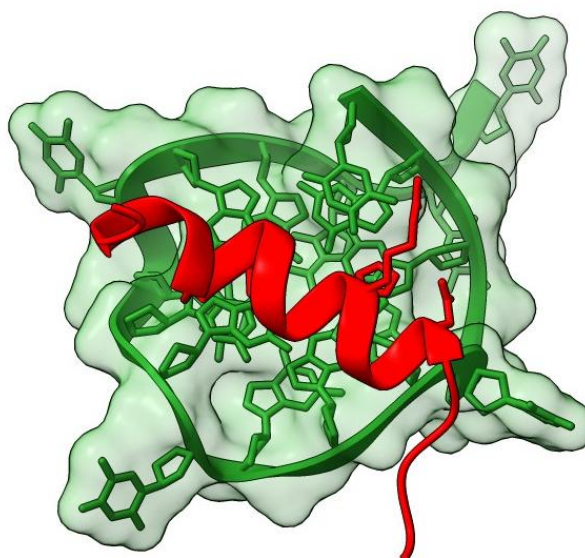


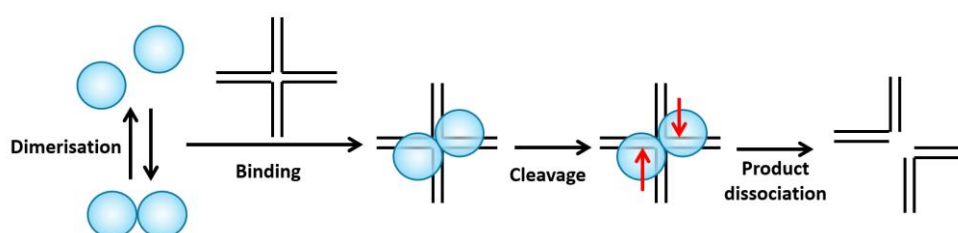
Figure 4-2: Crystal structure of Rap1-G4 complex (PDB code: 6LDM), showing the Rap1 protein (blue), G4 DNA (green) and the Rap 1 DNA recognition motif (red). **(A)** Full Rap1-G4 DNA complex. **(B)** Top view of the Rap 1 DNA recognition motif interacting with G4 DNA.¹⁸²

To use rational design for the discovery of 4WJ-binding peptides, a biologically relevant protein-4WJ complex, with significant structural information available,

must first be chosen. To make a ligand that is selective for 4WJ DNA over duplex DNA, it is important to bind to the centre of the junction instead of the duplex arms.

4.1.2 4WJ resolvases

As mentioned in **section 1.3**, 4WJs are important intermediates in homologous recombination, a process required for DNA repair.²³ During homologous recombination, 4WJs must be cleaved, to allow for the ligation of different DNA strands to form newly recombined DNA products. This process is carried out by structurally selective endonuclease enzymes, known as 4WJ resolvases. These proteins bind as dimers, before symmetrically cleaving the 4WJ around the junction centre, to form two nicked duplex DNA segments (**Scheme 4-1**).¹⁸⁵ The mechanism of 4WJ cleavage by resolvase enzymes is conserved across a range of other organisms, including bacteriophages,¹⁸⁶ yeasts¹⁸⁷ and mammals.¹⁸⁸ Some 4WJ resolvases exhibit no sequence specificity for their DNA cleavage sites, such as T4 and T7 endonucleases.¹⁸⁹ Other proteins in the family form sequence specific 4WJ cleavages ($\downarrow$), such as RuvC ($5'\text{-A/TTT}\downarrow\text{G/C-3}'$)¹⁹⁰ and MOC1 ($5'\text{-C}\downarrow\text{C-3}'$).¹⁹¹



Scheme 4-1: Mechanism of 4WJ resolution by 4WJ resolvase. The resolvase (blue) is in a monomer-dimer equilibrium, with binding to the 4WJ (black) occurring in its dimeric form. The resolvase then makes bilateral cleavages on the 4WJ (red arrows) and dissociates, leaving two nicked duplex DNA segments.¹⁹²

4.1.3 The MOC1-4WJ complex

The monokaryotic chloroplast 1 protein (MOC1) is a sequence selective 4WJ resolvase that is found in green plants.¹⁹³ This protein is present in chloroplasts, where it plays a crucial role in chloroplast nucleoid segregation.¹⁹¹ The crystal structures of *Nicotiana tabacum* MOC1-4WJ-Mg²⁺ (PDB code: 6KVO)¹⁹⁴ and *Zea mays* MOC1-4WJ-Ca²⁺ (PDB code: 6JRF)¹⁹³ have recently been solved, with resolutions of 2.5 Å and 2.35 Å respectively. Superimposition of these two structures showed high levels of similarity, especially for residues crucial to DNA interactions.¹⁹⁴ MOC1 binds as a dimer to the 4WJ, when in an open conformation with 2-fold symmetry (**A and B, Figure 4-3**).

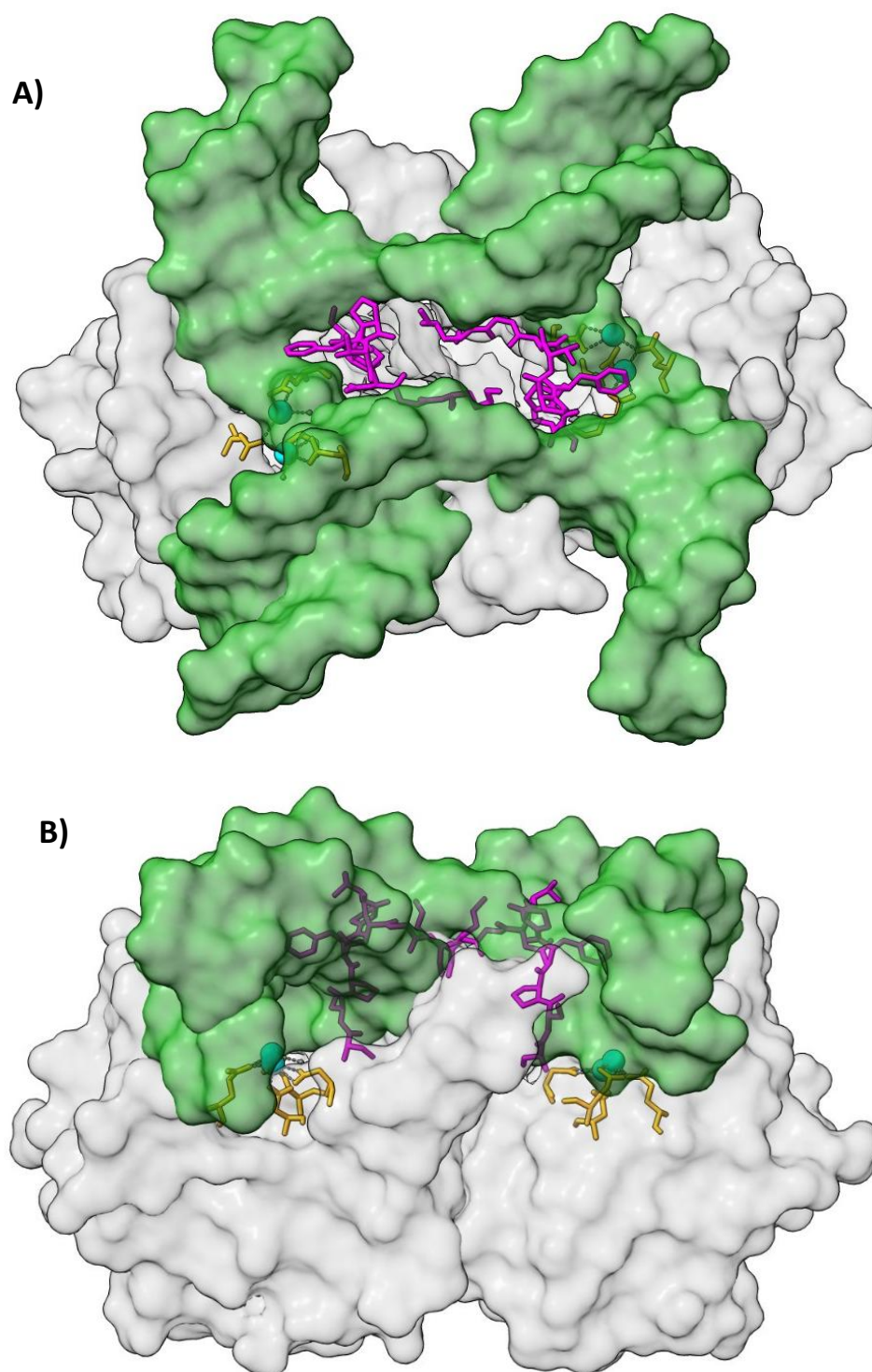


Figure 4-3: Crystal structure of *Zea mays* MOC1-4WJ- Ca^{2+} (PDB code: 6JRF), showing DNA (green), calcium (blue), Moc1 (grey), the active site (yellow) and the BRM (magenta). **(A)** Full complex top view. **(B)** Full complex side view.¹⁹³

The active site contains four amino acids with carboxylic acid side chains (E174, E257, D115 and D117) coordinated to two catalytic Ca^{2+} metal ions. The mutation of any of these residues prevents MOC1 from being able to cleave the 4WJ, showing their importance to the proteins function.¹⁹³ The Ca^{2+} metal ions then form interactions with the scissile phosphate group between the two cytosine bases, allowing cleavage to proceed (**A, Figure 4-4**).¹⁹³

Separate from the active site is another sequence of interest, known as the base-recognition motif (BRM), that consists of residues 176 – 184 in a loop-type conformation. The BRM interacts with DNA bases, allowing MOC1 to recognise the 4WJ in a sequence specific manner. By projecting into the open centre of the junction (**B, Figure 4-4**), the BRM interrupts base stacking interactions.¹⁹³ The aromatic residue F179 forms π - π interactions with two cytosine bases, perfectly placed by the specific loop distortion caused by the two proline residues that straddle this phenylalanine. Additionally, the carboxylic acid group of D182 stabilises the unpaired cytosine, caused by flipping out of guanosine base at the junction's centre, through a pseudo-base-pair interaction. This recognition of cytosine contributes to the selective 4WJ resolution of MOC1 between two consecutive cytosines. The displaced guanosine is further stabilised by a H-bond interaction between the side chain amine of K184 and the DNA base phosphate group.¹⁹⁴

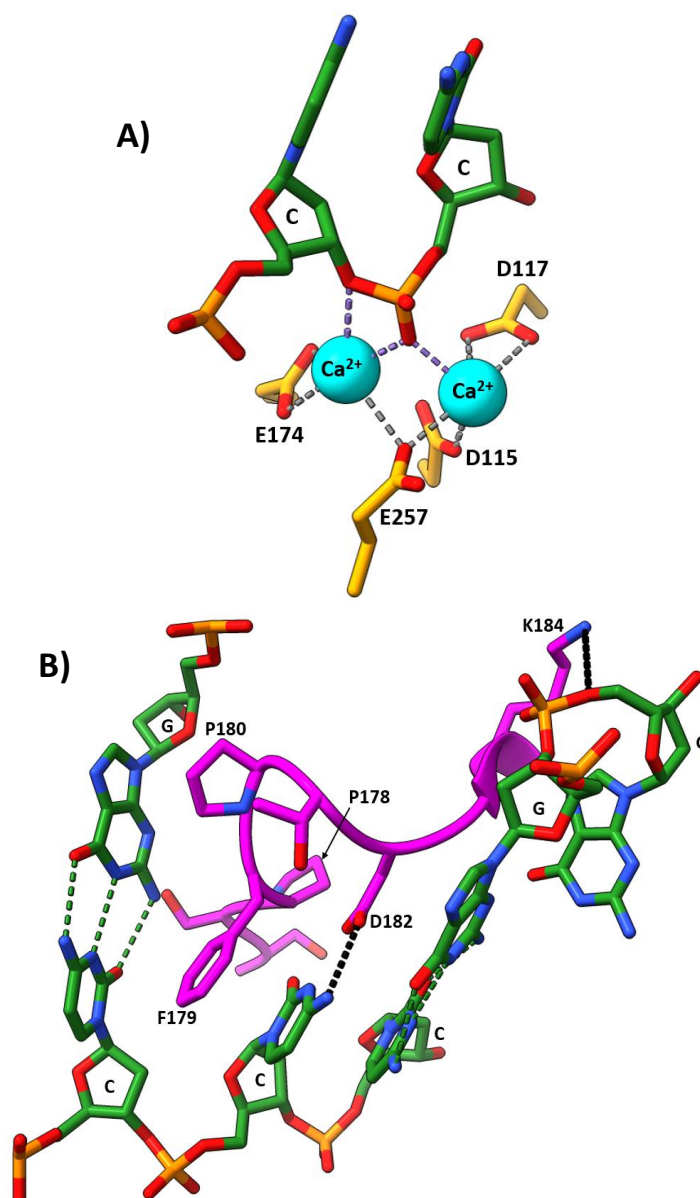


Figure 4-4: Crystal structure of *Zea mays* MOC1-4WJ-Ca²⁺ (PDB code: 6JRF), showing DNA (green), calcium (blue), the active site (yellow) and the BRM (magenta). **(A)** Active site interactions with Ca²⁺ metal ions and cytosine (C) bases. **(B)** BRM interactions with 4WJ DNA, with H-bonds between DNA bases (green) and between the BRM and DNA (black).¹⁹³

When comparing the BRM in the MOC1-4WJ complex and the unbound MOC1, a significant change to the BRM conformation is observed.¹⁹³ Additionally, the BRM and active site are close enough to each other to communicate, with the BRM residue S176 and active site residue E174 directly interacting. It has been observed that when MOC1 binds to the 4WJ, the BRM conformation is changed and subsequently alters the conformation of E174. This most likely regulates the metal catalysed action of the enzyme of cleaving the 4WJ.¹⁹³

4.2 Chapter aims

The aim of this chapter is to design peptides that bind to the open conformation 4WJ, based on the crystal structure of the *Zea mays* MOC1-4WJ-Ca²⁺ complex. This offers a different approach to previous methods, involving the screening of large combinatorial libraries. These peptides will be synthesised using SPPS, before being tested against 4WJ DNA using the PAGE assay previously used in **chapters 2 and 3**.

4.3 Results and discussion

4.3.1 Design of crystal structure-based 4WJ peptides

From analysis of the *Zea mays* MOC1-4WJ-Ca²⁺ crystal structure, discussed in **section 4.1.3**, the base recognition motif (BRM) was found to be crucial for sequence specific 4WJ recognition, through direct contacts with DNA bases, and interacted with the MOC1 active site.^{193,194} The BRM also sits in the centre of the 4WJ, the position peptide **1** was predicted to occupy through molecular modelling.⁶⁶ The BRM (176 – 184) is nine amino acids long and was synthesised first, with N-terminal amidation to follow the synthesis of previous peptides in **chapters 2 and 3**, to give the sequence of H-SSPFPTDGK-NH₂ (peptide **34**, **A**, **Figure 4-5**). This follows the precedent set for designing G-quadruplex binding peptides, where the full DNA recognition motif was synthesised and tested.¹⁸²

The MOC1 resolvase binds to the 4WJ as a dimer, therefore two BRM are present in the junction's centre.¹⁹³ In **chapter 2**, peptide **1** bound to the 4WJ as a dimer whilst peptide **2** was a palindromic peptide, with two glycine residues replacing the cysteine disulphide bridge of (WRWYCR)₂. In a similar manner, two monomers of peptide **34** could be linked using glycine residues, to create another palindromic peptide. The distance between the two monomers is 5.10 Å, when measured from the two α-carbon atoms of the C-terminal lysine residues (**B**, **Figure 4-5**). The average length of a single amino acid in a crystal structure is 3.6 Å,¹⁹⁵ meaning that only one glycine residue is predicted to fit between the two MOC1 monomers. The carbonyl groups would be twisted when fixed in place by a glycine residue and shorten the distance between two lysine residues, potentially allowing the single glycine to stretch across the full distance between the peptide **34** monomers. This gives the overall 19 amino acid sequence H-SSPFPTDGKGKDTPFPSS-NH₂ (peptide **35**, **A**, **Figure 4-5**).

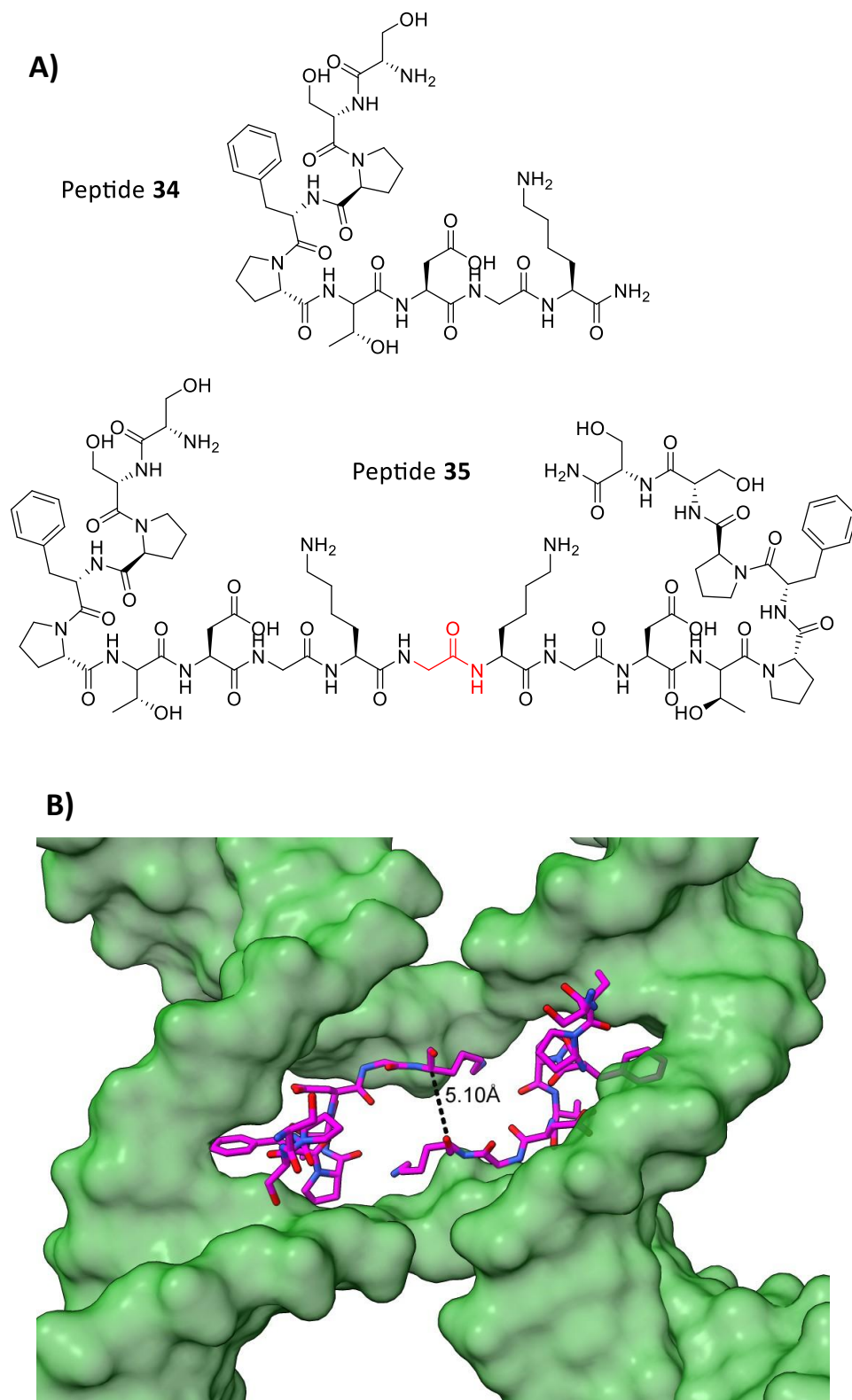
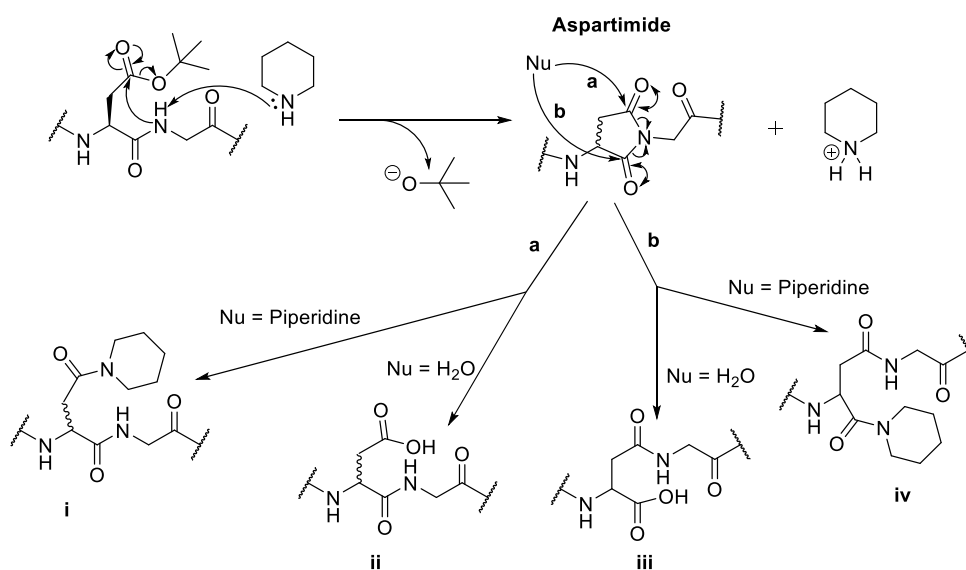


Figure 4-5: (A) The chemical structures of peptides **34** and **35**, with the Gly linker of peptide **35** in red. Both peptides contain neutral C-terminal amide groups. **(B)** Crystal structure of Zea mays MOC1-4WJ- Ca^{2+} (PDB code: 6JRF), showing the 4WJ DNA (green) and BRM sequences (magenta). The distance between the lysine residues is measured between the α -carbon atoms.

4.3.2 Peptide synthesis

Peptides **34** and **35** were synthesised using SPPS, following conditions from **chapters 2 and 3**, except for the addition of 0.1 M oxyma pure to the deprotection solution (20% piperidine in DMF). This was added to avoid side reactions that can occur with Asp-Gly dipeptides in basic solutions. Oxyma pure was also used as an additive during amino acid couplings, as discussed in **section 2.1.1**. This *N*-hydroxyamine acts as a pH regulator due to its low pKa of 4.60.¹⁰²

Both peptides contain Asp-Gly dipeptides, that are prone to the formation of aspartimide *via* intramolecular cyclisation upon treatment with piperidine (**Scheme 4-2**).¹⁹⁶ Firstly the piperidine nucleophile extracts the hydrogen from the backbone amide between both residues. The resulting negatively charged nitrogen then attacks the β -carboxyl group of the Asp residue, with subsequent loss of the *tert*-butyl protecting group, to form an aspartimide. This results in a scrambling of the chiral centre of the Asp residue. Subsequent nucleophilic attack causes ring-opening and several side products to be formed. Hydrolysis of the aspartimide can either restore the original α -Asp-Gly dipeptide **ii**, with the Asp protecting group removed, or cause the peptide backbone to be modified to give the isoaspartyl β -peptide **iii**. Attack on the aspartimide by piperidine gives the equivalent α -peptide **i** or β -peptide **iv**, containing piperidine groups.¹⁹⁷ As well as reducing the yield of the desired Asp-Gly containing peptide, these side products can be difficult to remove, due to the similarities of α , β and epimerised isomers.¹⁹⁸ By adding oxyma pure to the piperidine deprotection solution, hydrogen extraction from the backbone amide is competed against and aspartimide formation prevented.¹⁹⁷



Scheme 4-2: Intramolecular aspartimide formation after treatment of Asp-Gly with piperidine, followed by attack by nucleophiles to produce various side products.¹⁹⁷

Using these modified conditions, both monomer **34** (95% purity, 5% yield) and dimer **35** (95% purity, 11% yield) were successfully synthesised. Both peptides were initially purified by prep-HPLC using 0 – 100% MeOH in H₂O with 0.05% TFA. However, close running impurities co-eluted and were removed by a further prep-HPLC purification using a shallower gradient of 0 – 50% MeOH in H₂O with 0.05% TFA. This may indicate that some of the side products, mentioned in **Scheme 4-2**, were still present in solution and may have caused lower peptide yields. To increase yields, it could be worth increasing the concentration of oxyrna pure, to 1 M in the deprotection solution, as this has previously been shown to increase crude purity from ~60% to ~85%.¹⁹⁷

4.3.3 PAGE assay

To test the ability of peptides **34** and **35** to trap DNA as the 4WJ, the PAGE assay was used, following methods used in **chapter 2 and 3**. The 22 bp FAM-labelled 4WJ was used at a concentration of 1.25 μM and annealed in the presence of different peptide concentrations, dissolved in DMSO (2 μM), at 50-fold the desired final concentration (313 μM – 5 mM). Results from testing both peptides showed no change in band intensities (**Figure 4-6**), indicating that the peptides were unable to induce the formation of the 4WJ upon binding.

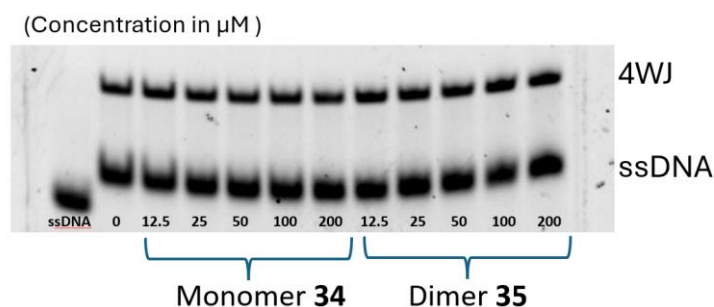


Figure 4-6: PAGE image showing the effect on 4WJ formation from increasing concentrations of peptides **34** and **35**, using a DNA concentration of 1.25 μM . A solution containing only the labelled X-strand was used as a control for ssDNA. Fam-labelled DNA was used and visualised at 460 nm.

The lack of activity could be caused by the 22 bp oligonucleotides used, that did not contain two consecutive cytosine when annealed into a 4WJ. As discussed in **section 4.1.3**, the BRM forms selective interactions with these cytosine bases, meaning BRM-based peptides may be unable to bind effectively to 4WJ centres where these crucial bases are absent. To evaluate this theory, two oligonucleotides Y and Z (**Figure 4-7**) were used in the PAGE assay, that matched the 4WJ sequences from the *Zea mays* MOC1-4WJ- Ca^{2+} crystal structure (PDB code: 6RJF). These oligonucleotides are 33 bp long, containing a pair of 9 bp long arms and a pair of 6 bp arms, with the shorter arms stabilised by three thymine bases that form a turn (**Figure 4-7**).¹⁹³ A FAM label was attached to the Z oligonucleotide, to allow for DNA visualisation.

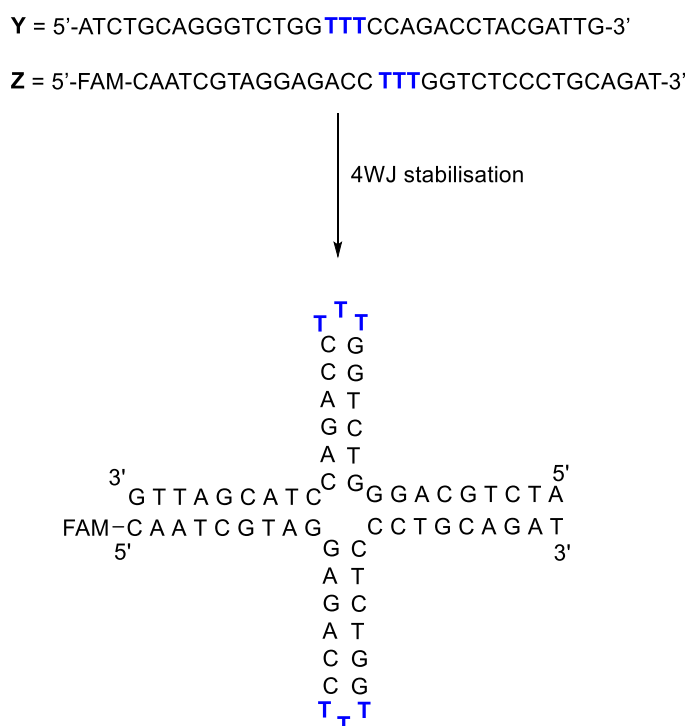


Figure 4-7: Oligonucleotide sequences capable of forming a 4WJ, based on the *Zea mays* MOC1-4WJ- Ca^{2+} crystal structure (PDB code: 6RJ).¹⁹³ The three base thymine turns are shown in blue.

To assess whether the Y/Z DNA was able to form a 4WJ during the PAGE assay, an MgCl_2 titration was carried out. For the 22 bp DNA, previous studies shows that an MgCl_2 concentration ≥ 1 mM resulted in the DNA being predominantly trapped as a 4WJ.¹²⁷ However, for the Y/Z DNA using MgCl_2 concentrations as high as 16 mM resulted in no 4WJ formation being observed (**A, Figure 4-8**). This could indicate that the Y/Z 4WJ is unstable, when attempting to form during annealing in 1 x TBE buffer. These conditions strongly differ to those used during crystallisation, where 20% PEG3350 was used to stabilise the MOC1-4WJ complex and crystals were isolated,¹⁹³ instead of trying to form the 4WJ in solution. Alternatively, this could indicate that the 4WJ is only stable in the open conformation, in the presence of MOC1, instead of the X-stacked conformation induced in a high salt environment. To check whether the peptides were able to induce the formation of the 4WJ, single 100 μM concentrations were tested for both peptides **34** and **35** and showed no change from the ssDNA. A faint second band in the Z-only control could show that this oligonucleotide partially anneals with itself, that may be interfering with 4WJ formation (**B, Figure 4-8**).

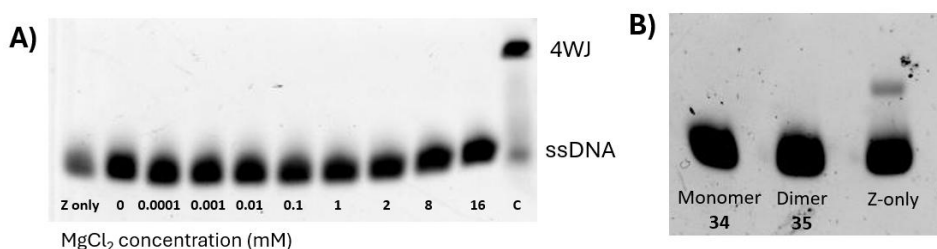


Figure 4-8: PAGE images using 1.25 μM Fam-labelled Y/Z DNA sequences. **(A)** The effect on DNA from increasing concentrations of MgCl_2 , with lane C containing the 22 bp FAM-labelled DNA, annealed in the presence of 2 mM MgCl_2 . **(B)** the effect on 4WJ formation from the addition of 100 μM of peptides **34** and **35**. Fam-labelled DNA visualised at 460 nm.

4.4 Conclusions and future work

This chapter focussed on the rational design of 4WJ-binding peptides, based on the crystal structure of the *Zea mays* MOC1-4WJ- Ca^{2+} complex. The base recognition motif (BRM) is a sequence from the MOC1 protein, that binds across the centre of the 4WJ. The 9 amino acid BRM, peptide **34**, was synthesised using SPPS, followed by the synthesis of a BRM dimer-like peptide **35**. These two peptides were tested for their ability to trap DNA as the 4WJ, using the PAGE assay. Neither peptide showed any activity against the original 4WJ sequences, used in the PAGE assay during

chapters 2 and 3, or against the Y/Z 4WJ, using oligonucleotides that matched those from the crystal structure complex. However, this could result from the original 4WJ sequences not containing two consecutive cytosine bases at the centre of the junction, required for BRM binding, and the potential instability of the Y/Z 4WJ in solution.

An alternative method for test the crystal structure-based peptides, could involve using the PAGE assay in a competition assay. Previous studies have shown the ability of MOC1 to cleave 4WJ DNA, containing central cytosine residues. This was visualised by the disappearance of the 4WJ band and appearance of nicked duplex DNA segments, upon addition of the protein.^{193,194} Peptides **34** and **35** could be added to a solution containing 4WJ DNA at the same time as MOC1 protein, to test the ability of the peptides to compete against MOC1 for 4WJ binding and prevent 4WJ cleavage from occurring. Additionally, the peptides could be tested in the FP or FRET assays used in **chapters 2 and 3**, using 4WJ DNA with the required central cytosine bases.

Different crystal structure-based peptides could also be designed, based on the MOC1 BRM. These could include extending the BRM, to include residues involved in the active site, such as E174 that is just two residues away from S176 in the BRM sequence. This could allow for more interactions to be formed between the peptide and 4WJ, although metal ions may need to be added due to their presence in the MOC1 active site (**C, Figure 4-4**). Alternatively, analogues based in the dimer-like peptide **35** could be designed using the peptidomimetic approaches from **chapter 3**, with different linkers used to bind the two BRM monomers and peptide cyclisation attempted. Another approach could be to crystallise peptide **34** or **35** with 4WJ DNA, to show if these peptides are capable of binding to the junctions' centre and potentially lead to ideas for the rational design of further peptide analogues.

Chapter 5 – General conclusions and future perspectives

5.1 General conclusions

This thesis began with a literature review of known DNA 4WJ-binding molecules, covering peptides, organometallic and small molecule ligands (**Chapter 1**). The molecular features required for 4WJ-binding were explored, as well as the differences in ligands that bind to either the X-stacked or open conformation of the 4WJ. A number of ligands discussed showed promise as anti-cancer and antibacterial therapeutic agents, although their selectivity over other higher-order DNA structures requires further exploration.

In subsequent chapters, the focus moved to systematically exploring the design, synthesis, and biological evaluation of peptides and peptidomimetics that specifically target the open conformation of 4WJ DNA. In **Chapter 2**, five peptides and two acridines compounds, previously shown to have 4WJ-binding activity, were synthesised and their biological activity evaluated in a series of assays. Among these, peptide **2** emerged as the strongest binder to 4WJ DNA across PAGE, FRET, and FP biophysical assays. Notably, peptide **2** also exhibited strong dsDNA binding, as demonstrated in an EtBr displacement assay, highlighting a potential lack of selectivity for 4WJ structures. However, competition experiments with PAGE assays suggested preferential binding to 4WJ despite excess dsDNA, pointing towards selectivity being affected by DNA sequences as well as the 4WJ structure. Despite its promising DNA-binding profile, peptide **2** showed no detectable cytotoxicity against HL-60 cancer cells at concentrations up to 25 μ M. This inactivity could arise from limited cellular uptake or susceptibility to enzymatic degradation in cellular environments, emphasising the therapeutic challenges posed by peptide instability and bioavailability.

Chapter 3 focused on overcoming these limitations by designing peptidomimetic analogues based on peptide **2**. The central Gly-Gly dipeptide was replaced with various linkers differing in length, flexibility, and functionality, while cyclic peptides were also synthesised to explore enhanced conformational stability. Nine linear analogues and two cyclic peptides were prepared via a hybrid approach involving multi-step solution-phase chemistry and SPPS. Biological evaluation of these analogues revealed comparable 4WJ binding to peptide **2**, with only minor variations in affinity across different assays. Unfortunately, none exhibited appreciable cytotoxicity against HL-60 cancer cells, showing that the structural changes made to

the peptide analogues were insufficient to impart cellular activity at the concentrations tested. The data suggested that peptide **2** is tolerant to chemical modification without loss of 4WJ affinity, providing a flexible platform for further optimisation.

In **Chapter 4**, the approach shifted towards rational peptide design informed by the crystal structure of the *Zea mays* MOC1 protein bound to 4WJ DNA. The base recognition motif (BRM), a nine-residue peptide sequence from MOC1 responsible for binding across the 4WJ centre, was synthesised as peptide **34**, with a dimeric variant (peptide **35**) also prepared to mimic protein multivalency. These peptides were tested using the PAGE assay with both 4WJ sequences used in previous chapters and the Y/Z 4WJ matching the crystal structure. Neither peptide displayed binding in these assays, likely due to the absence of consecutive cytosine bases critical for BRM binding and potential instability of the Y/Z 4WJ in solution. To overcome this, development of a competitive PAGE assay was proposed, to test the ability of these peptides to inhibit MOC1-mediated cleavage of 4WJ DNA containing central cytosine bases.

This PhD builds directly on the foundational work of Segall's group, who first identified the 4WJ-binding peptides WKHYNV and WRWYCR (peptide **1**) by employing large-scale combinatorial peptide library screening.^{63,65} This was followed by the rational design of WRWYRGGRYWRW (peptide **2**), that was able to retain the 4WJ-binding activity of peptide **1**, without the need to form a dimer.⁴⁵ Segall's research demonstrated the ability of these peptides to trap 4WJs, inhibit 4WJ junction-processing enzymes, and display antibacterial and anticancer activity.^{45,63,65} While Segall's research established the biological relevance and binding properties of these peptides, the present thesis focussed on chemically modifying the structure of a 4WJ-binding peptide to explore subsequent effects on activity. Additionally, an alternative approach to discovering 4WJ-binding peptides was investigated, using a 4WJ-protein complex, to compliment the large-scale screening methods employed by Segall.^{63,65}

The work described in this thesis provides an expansion to the toolkit of assays available for evaluating ligands for 4WJ-binding activity, with the development of PAGE, FP and FRET assays. These complementary assay formats will enable more robust structure-activity assessments of future potential 4WJ-binding ligands. This

thesis also provides a blueprint to approaching the development of drug-like peptides that target DNA structures, either using peptidomimetics to improve on known peptide binders or utilising structure-based design from known protein-DNA structures. Taken together, this work can be used to progress to the design of more drug-like peptide ligands that target 4WJ DNA.

5.2 Future perspectives

The advancement of 4WJ-binding peptides towards clinical application depends on rigorously confirming their selectivity for 4WJs over other DNA structures. Ensuring that these peptides preferentially engage 4WJs is essential for attributing any observed therapeutic effects to actual 4WJ targeting, rather than non-specific or off-target DNA interactions. This could be achieved by expanding the use of the PAGE competition assay, used to examine selectivity of peptide **2** between 4WJ DNA and dsDNA in **section 2.3.3**, to include other higher order structures, such as G-quadruplexes and 3WJs. Additionally, a panel of 4WJs containing various DNA sequences could be tested with the different peptide ligands, to look for any specific sequence selectivity as well as structural selectivity.

An important question arising from this thesis is whether the peptide ligands bind exclusively to the centre of the 4WJ or also along its dsDNA arms. Obtaining high-resolution structural information on peptide-4WJ complexes, using techniques such as X-ray crystallography, cryo-electron microscopy and nuclear magnetic resonance spectroscopy, would clarify where and how these interactions occur. For junction-centre binding peptides, this structural information could reveal key interactions and conformations that inform ligand optimisation, to further improve 4WJ binding affinity and specificity.

Another essential focus in the development of four-way junction (4WJ)-binding peptides is enhancing their ability to enter cells effectively. Research efforts should prioritize strategies such as conjugating these peptides to cell-penetrating peptides,¹⁹⁹ substituting L-amino acids with protease-resistant D-amino acids¹⁴² and incorporating lipophilic modifications like elongated PEG chains.¹⁷⁸ These approaches could help to membrane permeability and intracellular delivery, which

are critical for ensuring that high-affinity 4WJ-binding peptides reach their DNA targets within cells.

Future development of 4WJ-binding peptides will require confirmed selectivity over other DNA structures, detailed structural insight into binding modes, and effective strategies to improve delivery and stability. Building on the assays and design approaches developed in this work, these peptides could be developed into promising therapeutic agents.

Chapter 6 – Experimental

6.1 Materials and analysis techniques

6.1.1 Solvents and reagents

The solvents and chemicals were purchased at reagent grade from the suppliers: Sigma-Aldrich, Fluorochem, Fisher Scientific, AGTC bioproducts and Novabiochem. Peptide grade piperidine, DMF and DIPEA were acquired from Cambridge Reagents. Distilled water was used throughout the syntheses. MilliQ water was used for the biological assays.

The following Fmoc amino acids were purchased for use in peptide synthesis:

- Fmoc-L-Arg(Pbf)-OH
- Fmoc-L-Asn(Trt)-OH
- Fmoc-D-Asn(Trt)-OH
- Fmoc-L-Asp(O^tBu)-OH
- Fmoc-L-Cys(Acm)-OH
- Fmoc-L-Cys(Trt)-OH
- Fmoc-L-Gly-OH
- Fmoc-L-Lys(Boc)-OH
- Fmoc-L-Phe-OH
- Fmoc-L-Pro-OH
- Fmoc-D-Pro-OH
- Fmoc-L-Ser(^tBu)-OH
- Fmoc-L-Thr(O^tBu)-OH
- Fmoc-L-Trp(Boc)-OH
- Fmoc-L-Tyr(^tBu)-OH

6.1.2 NMR and spectroscopic techniques

¹H-, ¹³C- and ¹⁹F-NMR analysis was carried out a Bruker spectrometer, operating at 400 MHz (¹H), 101 MHz (¹³C) or 376 MHz (¹⁹F) using specified deuterated solvents. Chemical shifts were measured in ppm and splitting patterns were recorded in Hz. NMR multiplicities were labelled as s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet and br s = broad singlet.

MALDI analysis was performed on the Shimadzu AXIMA Performance MALDI ToF/TOF in a positive ionisation setting.

HRMS analysis was performed using the LC-TOF: Waters SYNAPT XS with ACQUITY H-Class PLUS UPLC machine in a positive ionisation setting.

LCMS analysis was performed on the Agilent 1200 Series, equipped with a 6460-triple quadrupole.

IR analysis was performed using a Perkin-Elmer Spectrum BX FT-IR spectrometer and analysed using Spectrum v5.3.1 software.

Fluorescence analysis of 96-well plates were performed on the BMG Labtech CLARIOstar microplate reader. For titrations into a 3.5 mL cuvette with a 1 cm path length, the Edinburgh instruments FS5 fluorimeter was used. Visualisations of electrophoresis gels were performed on the ImageQuant™ LAS 4000 biomolecular imager.

6.1.3 Chromatography techniques

Analytical Reverse Phase (RP) HPLC was performed on an Agilent 1200 Series HPLC, with a 4.6 x 150 mm, 5 µm, Agilent Eclipse XDB-C18 column at a 1 mL/min flow rate. The runs began with a gradient of 5 – 95% (MeOH or MeCN (as specified) containing 0.05% TFA) in (H₂O containing 0.05% TFA) over 15 min, followed by 5 min held at 95%. Peaks were detected at 214 nm and 254 nm wavelengths.

Preparative RP-HPLC was performed on an Agilent 1200 HPLC, with a 21.2 x 150 mm, 5 µm, Agilent Eclipse XDB-C18 column at a 20 mL/min flow rate. Runs were carried out at a gradient of 5% - 95% MeOH/MeCN (as specified) in H₂O with 0.05% TFA, over 20 min.

Flash chromatography was carried out on silica gel (particle size 60 µm) on a Biotage® Isolera Four. For RP column chromatography pre-packed Biotage® KP-C18-HS Flash Cartridges were used.

6.1.4 MALDI sample preparation

MALDI matrix: α-Cyano-4-hydroxycinnamic acid (CHCA) (20 mg/mL) in 70:30 MeCN:H₂O.

Peptide samples: 0.1 mg peptide dissolved in 1 mL MeOH.

Sample preparation: CHCA matrix (5 µL) and H₂O (5 µL) were added to the sample dissolved in MeOH (5 µL). The resultant solution (2 µL) was spotted onto the MALDI

plate and allowed to dry. Subsequently, 1:1 TFA:H₂O (2 µL) was spotted over the matrix and allowed to dry.

The MALDI standard ProteomassTM ACTH Fragment 18-39 from Sigma-Aldrich, with a mass of 2464.1980 Da, was used for calibration. An [M+H]⁺ value of 2465.10 Da was obtained, however some variation was observed over time of approximately 5 Da from the expected value. Mass values were corrected to match the variation observed with the standard.

6.2 Peptide synthesis

6.2.1 General peptide synthesis procedure

Peptides were synthesised using either a Multisynthtech Syro I or a CEM Liberty automated peptide synthesiser.

The following procedure was used for peptide synthesis using a Multisynthtech Syro I: 100 mg of the specified resin 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 the coupling of the first amino acid (4 equiv) using HOBt (4 equiv), HBTU (3.9 equiv) and DIPEA (8 equiv) in DMF for 45 min with shaking. The column was drained, and the reaction was repeated, to ensure completion of the amino acid coupling. After this, the resin was washed with DMF (3 x 2 mL). Subsequent deprotection and coupling reactions were carried out in the same way.

The procedure used for the CEM Liberty automated peptide synthesiser was the same except for the following differences:

- 0.05 mmol of the specified resin was used.
- Couplings used amino acid (5 equiv), oxyma pure (5 equiv) and DIC (10 equiv).
- Couplings of all amino acids except Arg were carried out with microwave heating at 75 °C (15 sec) followed by 90 °C (110 sec).
- For the coupling of the Arg residues, the microwave was heated at 75 °C (300 sec).

All resins were washed with MeOH (3 x 2 mL), CH₂Cl₂ (3 x 2 mL) to remove residual DMF. For cleavage of peptides containing cysteines, a solution of 92.5:2.5:2.5:2.5 TFA:2,2'-(Ethylenedioxy)diethanethiol (DODT):H₂O:TIPS (5 mL) was used. For peptides not containing cysteines, a solution of 96:2:2 TFA:H₂O:TIPS (5 mL) was used. The cleavage cocktail was added to the column before shaking for 3 h. The cleavage solution was collected and concentrated *in vacuo*. The peptide was precipitated using Et₂O (3 x 2 mL) and collected using centrifugation. The peptides were purified using RP-HPLC or RP column chromatography and analysed using HPLC, MALDI or LCMS, and HRMS.

6.2.2 General procedure for Kaiser test

This method followed Arora *et al.* with modification.²⁰⁰ The Kaiser test was used to qualitatively confirm the completion of amide coupling reactions. The presence of unreacted primary amines resulted in a positive test and a colour change to dark blue. The absence of primary amines signified reaction completion and resulted in a negative test, with no significant colour change observed.

The following 3 solutions were prepared:

- 1) Ninhydrin (5 g) in EtOH (100 mL)
- 2) Phenol (80 g) in EtOH (20 mL)
- 3) 0.0001 M KCN (aq, 2 mL) in pyridine (98 mL)

Procedure:

Approximately 5 – 10 resin beads were transferred into a small glass test tube, before 2 drops of solutions **1**, **2** and **3** were added. The resultant solution was heated using a heat gun until bubbling was observed. The solution was examined for any significant colour change.

6.2.3 Peptide data

6.2.3.1 H-WRWYCR-NH₂(1)

Peptide **1** was synthesised and purified using the general procedure in **section 6.2.1**, using H-rink amide ChemMatrix® resin (0.6 mmol/g) to give a white solid (24 mg, 91% purity, 50% yield).

HPLC: *t_R* = 10.1 min in (MeCN and 0.05% TFA)/(H₂O and 0.05% TFA).

MALDI-TOF $[M+H]^+$ m/z calculated for $[C_{46}H_{61}N_{15}O_7S]^+$ 968.46, found 968.80.

HRMS (ESI) calculated for $[C_{46}H_{61}N_{15}O_7S]^+$ 968.4599, found 968.4677.

6.2.3.2 H-WRWYRGGRYWRW-NH₂ (2)

Peptide **2** was synthesised and purified using the general procedure in **section 6.2.1**, using NovaPEG rink amide resin (0.49 mmol/g) to give a brown oil (23 mg, 95% purity, 26% yield).

HPLC: t_R = 10.3 min in (MeCN and 0.05% TFA)/(H₂O and 0.05% TFA).

MALDI-TOF $[M+H]^+$ m/z calculated for $[C_{90}H_{115}N_{29}O_{14}]^+$ 1826.92, found 1827.17.

HRMS (ESI) calculated for $[C_{90}H_{115}N_{29}O_{14}]^+$ 1826.9178, found 1826.9235.

6.2.3.3 H-WRWYCR-OH (4)

Peptide **4** was synthesised and purified using the general procedure in **section 6.2.1**, using H-Arg(Pbf)-HMPB NovaPEG resin (0.49 mmol/g) to give a yellow oil (24 mg, 95% purity, 51% yield).

HPLC: t_R = 9.84 min in (MeCN and 0.05% TFA)/(H₂O and 0.05% TFA).

MALDI-TOF $[M+H]^+$ m/z calculated for $[C_{46}H_{60}N_{14}O_8S]^+$ 969.44, found 969.10.

HRMS (ESI) calculated for $[C_{46}H_{60}N_{14}O_8S]^+$ 969.4439 found 969.4432.

6.2.3.4 H-WRWYC(Acm)R-NH₂ (5)

Peptide **5** was synthesised and purified using the general procedure in **section 6.2.1**, using H-rink amide ChemMatrix® resin (0.6 mmol/g) to give an off-white solid (16 mg, 93% purity, 31% yield).

HPLC: t_R = 12.5 min in (MeCN and 0.05% TFA)/(H₂O and 0.05% TFA).

MALDI-TOF $[M+H]^+$ m/z calculated for $[C_{49}H_{66}N_{16}O_8S]^+$ 1039.50, found 1039.21.

HRMS (ESI) calculated for $[C_{49}H_{66}N_{16}O_8S]^+$ 1039.4970, found 1039.5049.

6.2.3.5 *H-WRWYR-NH₂* (**6**)

Peptide **6** was synthesised and purified using the general procedure in **section 6.2.1**, using Tentagel® S RAM (0.23 mmol/g) to give a white solid (4 mg, 100% purity, 9% yield).

HPLC: t_R = 9.12 min in (MeCN and 0.05% TFA)/(H₂O and 0.05% TFA).

MALDI-TOF [M+H]⁺ m/z calculated for [C₄₃H₅₆N₁₄O₆]⁺ 865.45, found 865.06.

HRMS (ESI) calculated for [C₄₃H₅₆N₁₄O₆]⁺ 865.4507, found 865.4597.

6.2.3.6 *H-SSPFPTDGK-NH₂* (**34**)

Peptide **34** was synthesised and purified using the general procedure in **section 6.2.1**, with oxyma pure (0.1 M in DMF) added to the deprotection solution. H-rink amide ChemMatrix® resin (0.6 mmol/g) was used and give a pale yellow solid (5 mg, 95% purity, 11% yield). All Fmoc amino acids were coupled twice as standard, except for the Fmoc-Pro-OH residues that were coupled once.

HPLC: t_R = 10.4 min in (MeOH and 0.05% TFA)/(H₂O and 0.05% TFA).

MALDI-TOF [M+H]⁺ m/z calculated for [C₄₁H₆₃N₁₁O₁₄]⁺ 934.46, found 934.65.

HRMS (ESI) calculated for [C₄₁H₆₃N₁₁O₁₄]⁺ 934.4556, found 934.4654.

6.2.3.7 *H-SSPFPTDGKGKGDTPFPSS-NH₂* (**35**)

Peptide **35** was synthesised and purified using the general procedure in **section 6.2.1**, with oxyma pure (0.1 M in DMF) added to the deprotection solution. H-rink amide ChemMatrix® resin (0.6 mmol/g) was used and give a pale brown solid (5 mg, 95% purity, 5% yield). All Fmoc amino acids were coupled twice as standard, except for the Fmoc-Pro-OH residues that were coupled once.

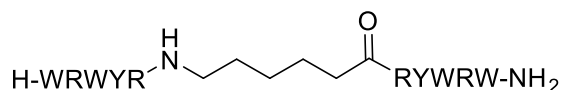
HPLC: t_R = 12.1 min in (MeOH and 0.05% TFA)/(H₂O and 0.05% TFA).

MALDI-TOF [M+H]⁺ m/z calculated for [C₈₄H₁₂₆N₂₂O₂₉]⁺ 1907.91, found 1907.59.

HRMS (ESI) calculated for [C₈₄H₁₂₆N₂₂O₂₉]⁺ 1907.9061, found 1907.9139.

6.2.4 Peptide analogue data

6.2.4.1 H-WRWYR-6-Alkane-RYWRW-NH₂ (14)

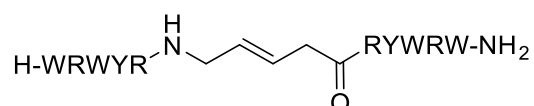


HPLC: $t_R = 13.5$ min in (MeOH and 0.05% TFA)/(H₂O and 0.05% TFA).

MALDI-TOF [M+H]⁺ m/z calculated for [C₉₁H₁₁₈N₂₈O₁₃]⁺ 1811.94, found 1811.60.

HRMS (ESI) calculated for [C₉₁H₁₁₈N₂₈O₁₃]⁺ 1811.9433, found 1811.9489.

6.2.4.4 H-WRWYR-Alkene-RYWRW-NH₂ (17)



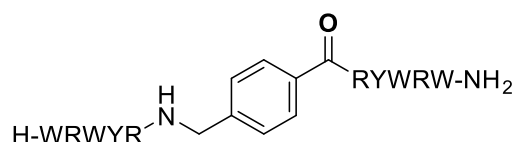
Peptide **17** was synthesised and purified using the general procedure in **section 6.2.1**, using H-rink amide ChemMatrix® resin (0.6 mmol/g) and the Fmoc-protected amino acid **31** (**section 5.3.3.2**), to give a pale brown solid (7 mg, 98% purity, 8% yield).

HPLC: $t_R = 13.2$ min in (MeOH and 0.05% TFA)/(H₂O and 0.05% TFA).

LCMS (ESI-QQQ) m/z calculated for [C₉₁H₁₁₆N₂₈O₁₃]⁺ 1809.92, found 1810.00.

HRMS (ESI) calculated for [C₉₁H₁₁₆N₂₈O₁₃]⁺ 1809.9277, found 1809.9355.

6.2.4.5 H-WRWYR-Benzene-RYWRW-NH₂ (18)



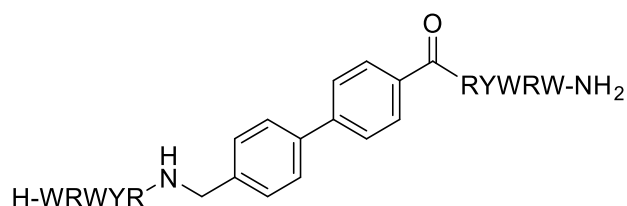
Peptide **18** was synthesised and purified using the general procedure in **section 6.2.1**, using H-rink amide ChemMatrix® resin (0.6 mmol/g) to give a pale brown solid (4 mg, 90% purity, 4% yield).

HPLC: $t_R = 13.385$ min in (MeOH and 0.05% TFA)/(H₂O and 0.05% TFA).

MALDI-TOF [M+H]⁺ m/z calculated for [C₉₄H₁₁₆N₂₈O₁₃]⁺ 1845.93, found 1845.61.

HRMS (ESI) calculated for [C₉₄H₁₁₆N₂₈O₁₃]⁺ 1845.9277, found 1845.9340.

6.2.4.6 H-WRWYR-Dibenzene-RYWRW-NH₂ (**19**)



H-RYWRW-**Resin** was synthesised using the CEM Liberty automated peptide synthesiser, using the general procedure in section 6.2.1 and Tentagel® S RAM (0.23 mmol/g), with the peptide left on the resin. H-RYWRW-**Resin** was treated with a solution containing dibenzene amino acid **33** (section 6.3.4.2, 20 mg, 0.2 mmol), HBTU (74 mg, 0.195 mmol), HOBt (27 mg, 0.2 mmol) and DIPEA (70 µL, 0.4 mmol) in DMF (2mL). The resultant solution was shaken at RT for 45 min before the column was drained, and the reaction repeated. The completion of the coupling was confirmed by a negative Kaiser test. After this, the resin was washed with DMF (3 x 2 mL). To the resin was added 15% N₂H₄, 15% MeOH in THF (2 mL) and the reaction was shaken at RT (2 x 3 h).¹⁵² Full conversion to the TFA deprotected product was confirmed by HPLC, with disappearance of the starting material peak. The resin was placed back on the CEM Liberty automated peptide synthesiser, and the remaining amino acids added, before the desired product was cleaved from the resin and purified using RP-HPLC to afford the desired product **19** as a pale-yellow solid (4 mg, 99% purity, 4% yield).

HPLC: t_R = 13.8 min in (MeOH and 0.05% TFA)/(H₂O and 0.05% TFA).

LCMS (ESI-QQQ) m/z calculated for [C₁₀₀H₁₂₀N₂₈O₁₃]⁺ 1921.96, found 1921.90.

HRMS (ESI) calculated for [C₁₀₀H₁₂₀N₂₈O₁₃]⁺ 1921.9590, found 1921.9668.

6.2.4.7 Azide-RYWRW-Resin (**27**)



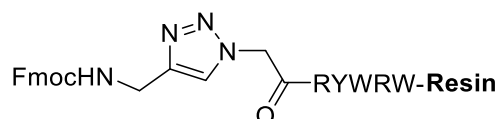
H-RYWRW-**Resin** was synthesised using the CEM Liberty automated peptide synthesiser, using the general procedure in section 6.2.1 and Tentagel® S RAM (0.23 mmol/g), with the peptide left on the resin. The intermediate was treated with a solution containing **26** (section 6.3.2.2, 20 mg, 0.2 mmol), HBTU (74 mg, 0.195 mmol), HOBt (27 mg, 0.2 mmol) and DIPEA (70 µL, 0.4 mmol) in DMF (2mL). The resultant solution was shaken at RT for 45 min before the column was drained,

and the reaction repeated. After this, the resin was washed with DMF (3 x 2 mL).¹⁷¹ Synthesis of intermediate **27** was confirmed by cleaving a few beads of resin and analysing the crude product:

HPLC: t_R = 10.6 min in (MeCN and 0.05% TFA)/(H₂O and 0.05% TFA).

MALDI-TOF [M+H]⁺ m/z calculated for [C₄₅H₅₇N₁₇O₇]⁺ 948.46, found 947.99.

6.2.4.8 Triazole-RYWRW-Resin (**29**)

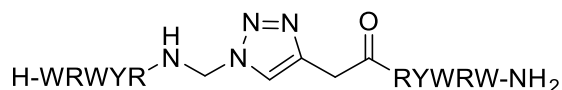


To intermediate **27** was added a solution containing **28** (section 6.3.2.3, 22 mg, 5 μmol), [Cu(CH₃CN)₄]PF₆ (0.2 mg, 0.5 μmol), DIPEA (1.7 μL, 10 μmol) and DMF (100 μL). The resultant solution was shaken for 16 h at RT, then washed with DMF (3 x 2 mL).¹⁷¹ Synthesis of intermediate **29** was confirmed by cleaving a few beads of resin and analysing the crude product:

HPLC: t_R = broad peak at ~13.0 min in (MeCN and 0.05% TFA)/(H₂O and 0.05% TFA).

MALDI-TOF [M+H]⁺ m/z calculated for [C₆₃H₇₂N₁₈O₉]⁺ 1225.57, found 1225.66.

6.2.4.9 H-WRWYR-Triazole-RYWRW-NH₂ (**20**)



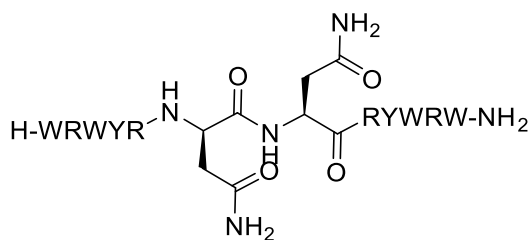
Intermediate **29** was placed back on the CEM Liberty automated peptide synthesiser and the remaining amino acids added, the product was cleaved from the resin and purified using RP-HPLC to afford the desired product **20** as a white solid (3 mg, 100% purity, 3% yield).

HPLC: t_R = 10.9 min in (MeCN and 0.05% TFA)/(H₂O and 0.05% TFA).

MALDI-TOF [M+H]⁺ m/z calculated for [C₉₁H₁₁₅N₃₁O₁₃]⁺ 1850.92, found 1851.23.

HRMS (ESI) calculated for [C₉₁H₁₁₅N₃₁O₁₃]⁺ 1850.9291, found 1850.9369.

6.2.4.10 H-WRWYR-D-Asn-L-Asn-RYWRW-NH₂ (21)



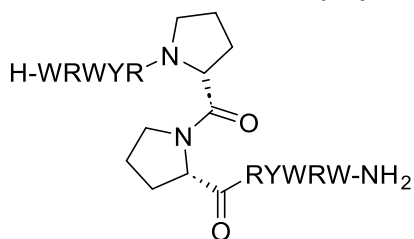
Peptide **21** was synthesised and purified using the general procedure in **section 6.2.1**, using H-rink amide ChemMatrix® resin (0.6 mmol/g) to give a pale brown solid (4 mg, 89% purity, 4% yield).

HPLC: t_R = 13.1 min in (MeOH and 0.05% TFA)/(H₂O and 0.05% TFA).

MALDI-TOF [M+H]⁺ m/z calculated for [C₉₄H₁₂₁N₃₁O₁₆]⁺ 1940.96, found 1940.63.

HRMS (ESI) calculated for [C₉₄H₁₂₁N₃₁O₁₆]⁺ 1940.9608, found 1940.9686.

6.2.4.11 H-WRWYR-D-Pro-L-Pro-RYWRW-NH₂ (22)



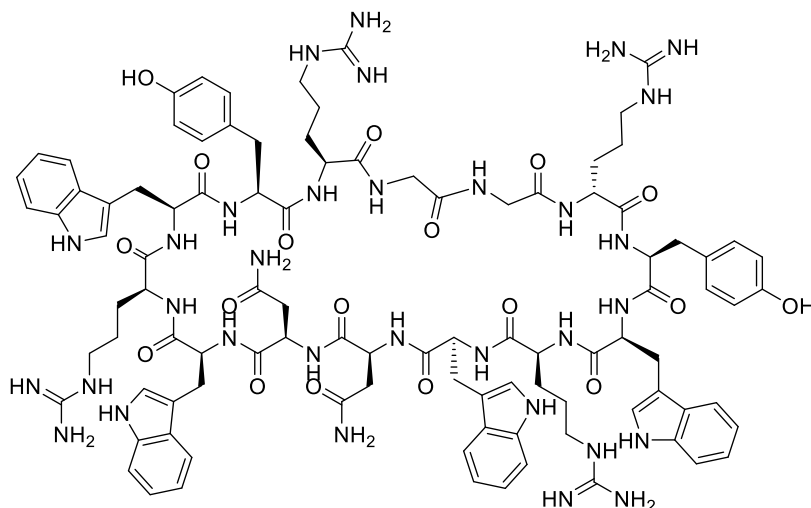
Peptide **22** was synthesised and purified using the general procedure in **section 6.2.1**, using H-rink amide ChemMatrix® resin (0.6 mmol/g) to give a pale brown solid (6 mg, 80% purity, 6% yield).

HPLC: t_R = 13.7 min in (MeOH and 0.05% TFA)/(H₂O and 0.05% TFA).

MALDI-TOF [M+H]⁺ m/z calculated for [C₉₆H₁₂₃N₂₉O₁₄]⁺ 1906.98, found 1907.47.

HRMS (ESI) calculated for [C₉₆H₁₂₃N₂₉O₁₄]⁺ 1906.9804, found 1906.9871.

6.2.4.12 Cyclic Gly-Gly (**23**)



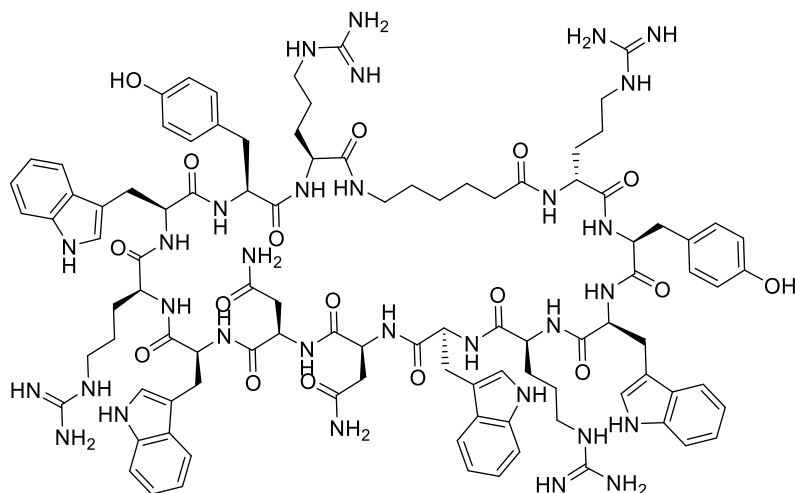
This method followed the method of Searcey *et al.* with modification.¹⁷⁰ Tentagel® S RAM (0.23 mmol/g) resin was Fmoc deprotected by the addition of 20% piperidine in DMF (2 x 2 mL) for 10 min with shaking. The resin was washed with DMF (6 x 2 mL) before a solution containing Fmoc-Asp-ODmab (67 mg, 0.1 mmol), HATU (36 mg, 0.095 mmol) and DIPEA (34 μ L, 0.2 mmol) dissolved in DMF (2 mL) was added. The resultant solution was shaken at 0 °C for 30 min before the column was drained, and the reaction repeated. Completion of the coupling was confirmed by a negative Kaiser test. The remaining amino acids were added using the CEM Liberty automated peptide synthesiser, following the general procedure in **section 6.2.1**. The Dmab group was removed using 2% (v/v) hydrazine hydrate in DMF (2 mL, 3 x 3 min), and the resin washed with DMF (6 x 2 mL). On-resin cyclisation was carried out using DIC (23 μ L, 0.15 mmol) and HOBt (20 mg, 0.15 mmol) in DMF (2 mL) for 24 h before the column was drained, and the reaction repeated. Completion of the reaction was confirmed by a negative Kaiser test. The peptide was cleaved from the resin and purified, following procedures from **section 6.2.1**, to afford the desired product **23** as a white solid (3 mg, 90% purity, 3% yield).

HPLC: t_R = 10.0 min in (MeCN and 0.05% TFA)/(H₂O and 0.05% TFA).

LCMS (ESI-QQQ) m/z calculated for [C₉₈H₁₂₄N₃₂O₁₈]^(M+2H) 1019.5, found 1019.7.

HRMS (ESI) calculated for [C₉₈H₁₂₄N₃₂O₁₈]^(M+2H) 1019.4958, found 1019.4927.

6.2.4.13 Cyclic 6-Alkane (24)



This method followed the method of Searcey *et al.* with modification.¹⁷⁰ Cyclic peptide **24** was synthesised and purified using the procedure from **section 6.2.4.12**, with 6-(Fmoc-amino)hexanoic acid replacing the two glycine amino acids. Tentagel® S RAM (0.23 mmol/g) resin was used to afford the desired product as a white solid (8 mg, 95% purity, 8% yield).

HPLC: $t_R = 10.3$ min in (MeCN and 0.05% TFA)/(H₂O and 0.05% TFA).

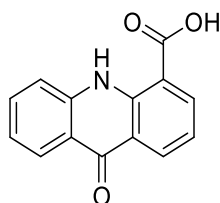
MALDI-TOF $[M+H]^+$ m/z calculated for $[C_{100}H_{129}N_{31}O_{17}]^+$ 2037.018, found 2037.74.

HRMS (ESI) calculated for $[\text{C}_{100}\text{H}_{129}\text{N}_{31}\text{O}_{17}]^+$ 2037.0183, found 2037.0261.

6.3 Chemical synthesis and characterisation

6.3.1 Synthesis of acridines

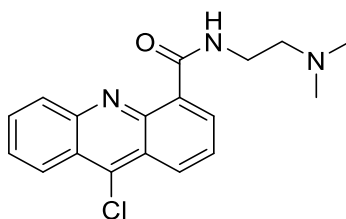
6.3.1.1 9-Oxo-9,10-dihydroacridine-3-carboxylic acid (7)



This method followed the method of Howell with modification.¹²⁷ 2,2-Iminodibenzoic acid (3 g, 11.7 mmol) was dissolved in conc. H₂SO₄ (30 mL) and refluxed at 100 °C for 1 h. On cooling, the solution was poured over ice water to give

a fluorescent yellow precipitate. The solid was washed with water (150 mL) then dissolved in MeOH (150 mL) and 2 M NaOH (aq) (150 mL). The solution was acidified with glacial acetic acid (500 mL) then concentrated *in vacuo*. The resultant precipitate was collected by filtration and washed with H₂O (150 mL). After drying under vacuum **7** was obtained as a yellow solid (2.4 g, >95% purity, 86% yield). IR ν_{max} (neat) / cm⁻¹ 3226 (NH), 2888 (OH), 1692 (C=O carboxylic acid), 1616 (C=O aromatic). ¹H NMR (DMSO-d₆, 400 MHz) δ_{H} ppm: 12.02 (s, 1H), 8.54 (dd, *J* = 7.9, 1.7 Hz, 1H), 8.45 (dd, *J* = 7.5, 1.7 Hz, 1H), 8.27 – 8.20 (m, 1H), 7.83 – 7.71 (m, 2H), 7.43 – 7.27 (m, 2H). ¹³C NMR (DMSO-d₆, 101 MHz) δ_{C} ppm: 176.95, 169.59, 141.64, 140.34, 137.28, 134.49, 132.78, 126.32, 122.70, 122.05, 121.04, 120.62, 119.01, 115.51.

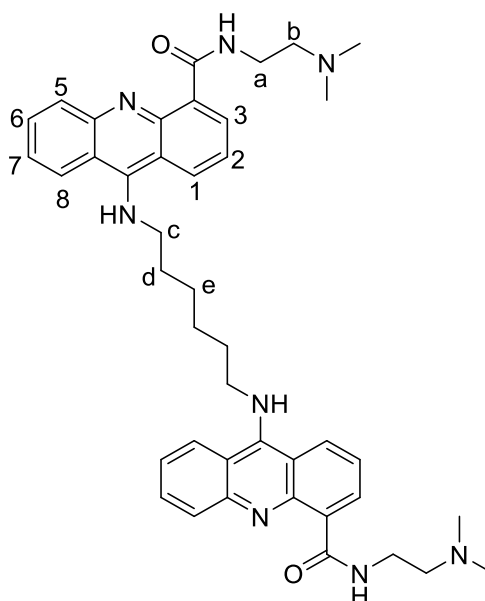
6.3.1.2 9-Chloro-*N*-(2-(dimethylamino)ethyl)acridine-4-carboxamide (**8**)



This method followed the method of L. A. Howell with modification.¹²⁷ **7** (2.4 g, 10 mmol) was dissolved in a solution of SOCl₂ (24 mL) containing 10 drops of DMF. The resultant solution heated under reflux at 90 °C and stirred for 1 h. On cooling, the solution was concentrated *in vacuo* and residual traces of SOCl₂ removed by azeotrope with dry toluene (24 mL). The intermediate 9-chloroacridine-4-carbonyl chloride was cooled to 0 °C and a cold solution of *N,N*-dimethylethylenediamine (4.31 mL, 40 mmol) in anhydrous CH₂Cl₂ (48 mL) was added quickly in one aliquot. The reaction mixture was allowed to warm to RT, stirred until homogenous and then for a further 30 min. The reaction mixture was extracted with brine (3 x 50 mL) and the organic extract was dried over MgSO₄, filtered, and the solvent removed *in vacuo*. Column chromatography (0-10% MeOH in CH₂Cl₂) afforded **8** as a pale orange solid (1.58 g, 95% purity, 48% yield). IR ν_{max} (neat) / cm⁻¹ 3184 (NH), 2943 (CH), 2821 (CH₂/CH₃), 2769 (CH₂/CH₃), 1655 (C=O). ¹H NMR (CDCl₃, 400 MHz) δ_{H} ppm: 11.90 (br s, 1H), 8.92 (dd, *J* = 7.1, 1.5 Hz, 1H), 8.54 (dd, *J* = 8.8, 1.5 Hz, 1H), 8.41 – 8.32 (m, 1H), 8.25 – 8.16 (m, 1H), 7.81 (ddd, *J* = 8.8, 6.6, 1.4 Hz, 1H), 7.68 (dd, *J* = 8.7, 7.1 Hz, 1H),

7.63 (ddd, $J = 8.8, 6.6, 1.2$ Hz, 1H), 3.88 – 3.72 (m, 2H), 2.75 (t, $J = 6.2$ Hz, 2H). ^{13}C NMR (CDCl_3 , 101 MHz) δ_{C} ppm: 165.43, 147.40, 146.32, 142.78, 135.46, 131.40, 129.39, 128.94, 128.43, 127.46, 126.56, 124.64, 124.46, 123.79, 58.15, 45.42, 37.91.

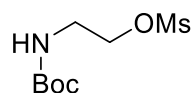
6.3.1.3 9,9'-(Hexane-1,6-diylbis(azanediy))bis(N-(2-(dimethylamino)ethyl)acridine-4-carboxamide (3)



This method followed the method of L. A. Howell with modification.¹²⁷ **8** (50 mg, 0.153 mmol) and excess dry phenol (250 mg) were stirred and heated at 110 °C for 15 min. On cooling to 55 °C, 1,6-hexane diamine (9 mg, 0.0765 mmol) was added and the reaction stirred for 2 h. On cooling to RT, 2 M NaOH(aq) (20 mL) was added and the aqueous phase was extracted with CH_2Cl_2 (3 x 20 mL). The organic extracts were combined, dried over MgSO_4 , filtered and the solvent was removed *in vacuo*. Column chromatography (0-20% (MeOH in EtOAc + 1% Et_3N) in CH_2Cl_2) yielded **3** as a bright orange solid (25 mg, 95% purity, 47% yield). ^1H NMR (MeOD, 400 MHz) δ_{H} ppm: 8.49 (2H, dd, $J = 7.1, 1.4$ Hz, 2ArH(3)), 8.25 (2H, dd, $J = 8.7, 1.5$ Hz, 2ArH(1)), 8.12 (2H, d, $J = 8.8$ Hz, 2ArH(8)), 7.81 (2H, d, $J = 8.6$ Hz, 2ArH(5)), 7.65 – 7.59 (2H, m, 2ArH(6)), 7.30 – 7.19 (4H, m, 4ArH(2+7)), 3.71 – 3.63 (8H, m, $\text{NHCH}_2(\text{a+c})$), 2.71 (4H, t, $J = 6.5$ Hz, $\text{CH}_2\text{N}(\text{CH}_3)_2(\text{b})$), 2.41 (12H, s, $\text{N}(\text{CH}_3)_2$), 1.71 – 1.61 (m, 4H, $\text{NHCH}_2\text{CH}_2(\text{d})$), 1.31 – 1.24 (4H, m, $\text{NHCH}_2\text{CH}_2\text{CH}_2(\text{e})$). ^{13}C NMR (MeOD, 101 MHz) δ_{C} ppm: 167.62 (CONH), 154.03 (ArC), 146.69 (ArC), 145.83 (ArC), 133.79 (ArC), 131.12 (ArC), 128.23 (ArC), 127.14 (ArC), 125.28 (ArC), 123.43 (ArC), 122.77 (ArC), 120.58 (ArC), 115.65 (ArC),

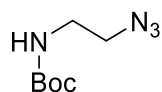
114.81 (ArC), 57.84 (CH₂N(CH₃)₂), 49.60 (CONHCH₂), 44.12 (N(CH₃)₂), 37.04 (NHCH₂), 30.16 (NHCH₂CH₂), 25.87 (NHCH₂CH₂CH₂). HRMS (ES⁺) calculated for C₄₂H₅₁N₈O₂ (M+H)⁺ 699.41 found 699.41.

6.3.1.4 2-(*Tert*-butoxycarbonylamino)ethyl methane sulfonate (**10**)



This method followed the method of L. A. Howell with modification.¹²⁷ To a solution of Boc-ethanolamine (1 g, 6.2 mmol) and Et₃N (2.6 mL, 18.6 mmol) in CH₂Cl₂ (6 mL) at 0 °C, whilst stirring vigorously, was added MsCl (480 µL, 6.2 mmol) *via* a dropping funnel. The reaction was allowed to warm to RT and stirred for 12 h. Water (20 mL) was added to the solution and the product extracted with CH₂Cl₂ (3 x 20 mL). The combined organic phases were dried (MgSO₄), filtered, and concentration *in vacuo* to give **10** as a pale-yellow liquid (1.41 g, 90% purity, 95% yield). This compound was stored in the freezer due to instability. IR ν_{max} (neat) / cm⁻¹ 3343 (NH), 2976 (CH₂/CH₃), 2935 (CH₂/CH₃), 1695 (C=O). ¹H NMR (CDCl₃, 400 MHz) δ_{H} ppm: 4.97 (1H, br s, NH), 4.27 (2H, t, *J* = 5.2 Hz, CH₂OMs), 3.45 (2H, q, *J* = 5.5 Hz, CH₂NH), 3.02 (3H, s, CH₃), 1.43 (9H, s, C(CH₃)₃). ¹³C NMR (CDCl₃, 101 MHz) δ_{C} ppm: 155.79 (C=O), 79.90 (C(CH₃)₃), 69.80 (CH₂OMs), 39.95 (CH₂NH), 37.38 (CH₃), 28.31 (C(CH₃)₃).

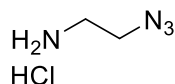
6.3.1.5 *Tert*-butyl 2-azidoethylcarbamate (**11**)



This method followed the method of L. A. Howell with modification.¹²⁷ **10** (800 mg, 3.3 mmol) and sodium azide (540 mg, 8.4 mmol) were dissolved in DMF (10 mL) and heated at 67 °C for 12 h. On cooling, water was added (20 mL) and the reaction mixture extracted with Et₂O (3 x 20 mL). The organic phases were dried (MgSO₄), filtered and the solvent removed *in vacuo* to yield **11** as a colourless liquid (366 mg, 95% purity, 59% yield). IR ν_{max} (neat) / cm⁻¹ 3341 (NH), 2978 (CH₂/CH₃), 2932 (CH₂/CH₃), 2096 (N₃), 1688 (C=O). ¹H NMR (CDCl₃, 500 MHz) δ_{H} ppm: δ 4.87 (1H, br s, NH), 3.43 – 3.37 (2H, m, CH₂N₃), 3.32 – 3.25 (2H, m, NHCH₂), 1.43 (9H, s, CH₃). ¹³C

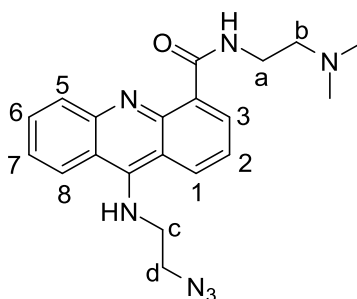
NMR (CDCl₃, 101 MHz) δ_c ppm: 155.71 (C=O), 79.81 (C(CH₃)₃), 51.27 (CH₂N₃), 40.06 (CH₂NH), 28.35 (C(CH₃)₃).

6.3.1.6 2-Azidoethanamine hydrochloride (**12**)



This method followed the method of L. A. Howell with modification.¹²⁷ **11** (366 mg, 2.0 mmol) was stirred at RT in a mixture of 1 M HCl/EtOAc (4 mL) for 12 h. 4 M HCl/EtOAc (4 mL) was added, and the reaction was stirred for a further 12 h. The solvent was removed under vacuum and the product dried using a freeze-drier (3 days) to yield **12** as an off-white solid (210 mg, 80% purity, 87% yield). IR $\nu_{\max}$ (neat) / cm⁻¹ 2882 (NH), 2099 (N₃). ¹H NMR (D₂O, 400 MHz) δ_H ppm: 3.68 – 3.61 (2H, m, CH₂N₃), 3.11 (2H, t, *J* = 5.6 Hz, CH₂NH₂). ¹³C NMR (D₂O, 101 MHz) δ_c ppm: 47.92 (CH₂N₃), 38.61 (CH₂NH₂).

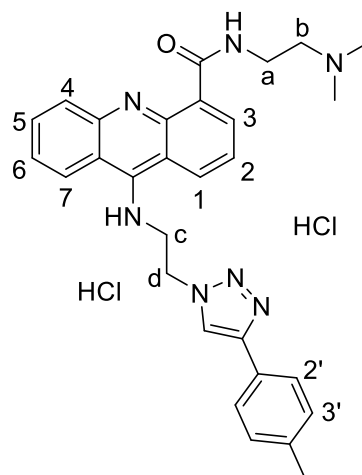
6.3.1.7 9-(2-Azidoethylamino)-N-(2-(dimethylamino)ethyl)acridine-4-carboxamide (**13**)



This method followed the method of L. A. Howell with modification.¹²⁷ **8** (100 mg, 0.31 mmol) was stirred in excess phenol (292 mg, 3.1) at 100 °C for 15 min. On cooling to 55 °C, **12** (42 mg, 0.34 mmol) was added and the reaction stirred at 55 °C overnight. On cooling, NaOH (2 M, 20 mL) was added and the reaction mixture extracted with CH₂Cl₂ (3 x 20 mL). The combined organic phases were dried (MgSO₄), filtered and the solvent removed *in vacuo*. Column chromatography (0-5% MeOH in EtOAc + 1% Et₃N) yielded the **13** as a yellow oil (77 mg, 95% purity, 66% yield). IR $\nu_{\max}$ (neat) / cm⁻¹ 3355 (NH), 2951 (CH₂/CH₃), 2862 (CH₂/CH₃), 2083 (N₃), 1638 (C=O). ¹H NMR (MeOD, 400 MHz) δ_H ppm: 8.71 – 8.46 (1H, m, ArH(1)), 8.33 (1H, d, *J* = 8.6 Hz, ArH(3)), 8.18 (1H, d, *J* = 8.7 Hz, ArH(8)), 7.96 – 7.79 (1H, m, ArH(5)), 7.64 (1H, t, *J* =

7.8 Hz, ArH(6)), 7.43 – 7.27 (2H, m, ArH(2+7)), 3.89 (2H, t, $J = 5.7$ Hz, CH_2N_3 (d)), 3.67 (2H, t, $J = 6.6$ Hz, CONHCH_2 (a)), 3.57 (2H, t, $J = 5.7$ Hz, NHCH_2 (c)), 2.67 (2H, t, $J = 6.6$ Hz, $\text{CH}_2\text{N}(\text{CH}_3)_2$ (b)), 2.36 (6H, s, $\text{N}(\text{CH}_3)_2$). HRMS (ES+) calculated for $\text{C}_{20}\text{H}_{24}\text{N}_7\text{O}$ ($\text{M}+\text{H}$)⁺ 378.2037 found 378.2042.

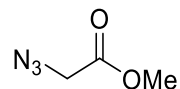
6.3.1.8 *N*-(2-(dimethylamino)ethyl)-9-(2-(4-phenyl-1H-1,2,3-triazol-1-yl)ethylamino)acridine-4-carboxamide (9)



This method followed the method of L. A. Howell with modification.¹²⁷ **13** (67 mg, 0.18 mmol) was dissolved in 1:1 *t*-BuOH:H₂O (890 μL) and to this was added 4-ethynyltoluene. The reaction mixture was stirred at RT for 2 min followed by the addition of sodium ascorbate (3.5 mg, 0.018 mmol) and copper sulphate (0.28 mg, 0.0018 mmol). This was stirred at RT overnight before the solvents were removed *in vacuo*. The product was purified by column chromatography (10-15% MeOH in EtOAc + 1% Et₃N) and converted into the dihydrochloride salt, by stirring the compound in a solution of 1.25 M HCl in MeOH (2 mL) for 10 min followed by precipitation with EtOAc. The salt was freeze dried from water to give **9** as a dark yellow solid (66 mg, 95% purity, 75% yield). IR ν_{max} (neat) / cm^{-1} 3233 (NH), 3041.72 (CH_2/CH_3), 2968 (CH_2/CH_3), 1622 (C=O). ¹H NMR (MeOD, 400 MHz) δ_{H} ppm: 8.72 (1H, d, $J = 8.6$ Hz, ArH(1)), 8.62 (1H, s, triazole CH), 8.58 (1H, dd, $J = 7.5, 1.0$ Hz, ArH(3)), 8.52 (1H, d, $J = 8.7$ Hz, ArH(8)), 7.99 – 7.93 (1H, m, ArH(6)), 7.85 (1H, d, $J = 8.6$ Hz, ArH(5)), 7.68 – 7.56 (2H, m, ArH(2+7)), 7.52 – 7.48 (2H, m, ArH(2')), 7.25 (2H, d, $J = 7.9$ Hz, ArH(3')), 5.21 – 5.13 (2H, m, CH_2N_3 (d)), 4.88 – 4.80 (2H, m, NHCH_2), 3.93 – 3.88 (2H, m, CONHCH_2), 3.54 (2H, t, $J = 5.7$ Hz, $\text{CH}_2\text{N}(\text{CH}_3)_2$), 3.06 (6H, s, $\text{N}(\text{CH}_3)_2$), 2.37 (3H, s, CH₃). HRMS (ES+) calculated for $\text{C}_{29}\text{H}_{32}\text{N}_7\text{O}$ ($\text{M}+\text{H}$)⁺ 494.2663 found 494.2668.

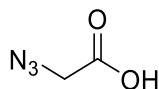
6.3.2 Synthesis of triazole-linked peptide analogue (20)

6.3.2.1 Methyl 2-azidoacetate (25)



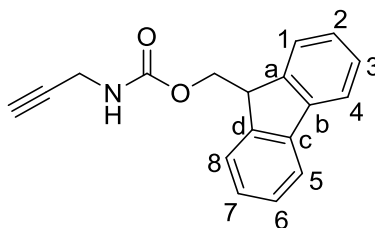
This method followed the method of Spring *et al.* with modification.¹⁷² A solution of methyl 2-bromoacetate (617 μ L, 6.54 mmol) and NaN₃ (1.7 g, 26.15 mmol) in acetone was refluxed under a nitrogen atmosphere and stirred for 12 h. The solvent was removed *in vacuo* and water (20 mL) added to the residue. The product was extracted with Et₂O (3 x 20 mL) and the combined organic phases were dried (MgSO₄), filtered and concentrated *in vacuo* to afford the desired product as a clear, colourless liquid (594 mg, 79%). ¹H NMR (CDCl₃, 400 MHz) δ_{H} ppm: 3.88 (2H s, CH₂), 3.79 (3H, s, CH₃). ¹³C NMR (CDCl₃, 101 MHz) δ_{C} ppm: 169.84 (C=O), 53.21 (OCH₃), 50.18 (CH₂). IR ν_{max} (neat) / cm⁻¹ 2103 (N₃), 1744 (C=O).

6.3.2.2 2-Azidoacetic acid (26)



This method followed the method of Spring *et al.* with modification.¹⁷² To a solution of **25** (594 mg, 5.16 mmol) in THF:MeOH:H₂O (15 mL, 3:1:1) was added LiOH.H₂O (1.08 g, 25.8 mmol) and the reaction was stirred for 12 h at RT. The reaction mixture was diluted with THF (5 mL) and acidified to pH 3 using 3 M HCl (aq). The product was extracted with Et₂O (3 x 20 mL) and the combined organic fractions were dried (MgSO₄), filtered and concentrated *in vacuo* to afford the desired product as a yellow oil (359 mg, 69%). ¹H NMR (CDCl₃, 400 MHz) δ_{H} ppm: 10.18 (1H, br s, OH), 4.00 (2H, s, CH₂). ¹³C NMR (CDCl₃, 101 MHz) δ_{C} ppm: 174.32 (C=O), 50.04 (CH₂). IR ν_{max} (neat) / cm⁻¹ 2108 (N₃), 1721 (C=O).

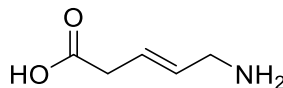
6.3.2.3 (9H-Fluoren-9-yl)methyl prop-2-yn-1-ylcarbamate (**28**)



This method followed the method of Wagner *et al.* with modification.²⁰¹ To a solution of propargylamine (690 μ L, 10 mmol) and DIPEA (1.74 mL, 10 mmol), in CH_2Cl_2 (30 mL) at 0 $^\circ\text{C}$, was added (9H-fluoren-9-yl)methyl 2,5-dioxopyrrolidine-1-carboxylate (3.37 g, 10 mmol) and DMAP (120 mg, 1 mmol). The reaction was allowed to warm to RT and stirred for 24 h. The resultant solution was successively washed with 1 M HCl (aq) (30 mL), sat. NaHCO_3 (30 mL) and brine (30 mL). The organic layer was dried (MgSO_4), filtered and concentrated *in vacuo*. The residue was purified by column chromatography (0-20% EtOAc in hexane) to afford the desired product **28** as a white solid (2.16 g, 78%). ^1H NMR (CDCl_3 , 400 MHz) δ_{H} ppm: 7.80 – 7.73 (2H, m, ArH(4+5)), 7.62 – 7.56 (2H, m, ArH(2+7)), 7.41 (2H, tt, J = 7.5, 1.0 Hz, ArH(1+8)), 7.32 (2H, td, J = 7.4, 1.2 Hz, ArH(3+6)), 4.96 (1H, br s, CH_2CCH), 4.43 (2H, d, J = 6.9 Hz, OCH_2), 4.23 (1H, t, J = 6.9 Hz, OCH_2CH), 4.04 – 3.97 (2H, m, NHCH_2), 2.26 (1H, t, J = 2.5 Hz, NH). ^{13}C NMR (CDCl_3 , 101 MHz) δ_{C} ppm: 156.16 (C=O), 143.65 (Ar(a+d)), 141.55 (Ar(b+c)), 127.76 (ArC(1+8)), 127.09 (ArC(3+6)), 125.04 (ArC(2+7)), 120.03 (ArC(4+5)), 79.28 (CH_2CCH), 72.18 (CH_2CCH), 67.08 (OCH_2), 47.16 (OCH_2CH), 30.88 (NHCH_2). IR ν_{max} (neat) / cm^{-1} 3335 (NH), 3300 (CH), 2966 (CH_2/CH_3), 1687 (C=O), 1529 (CH_2).

6.3.3 Synthesis of alkene-linked peptide analogue (**17**)

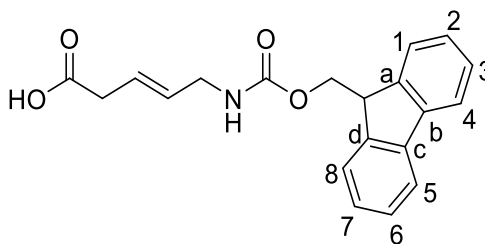
6.3.3.1 (3E)-5-Amino-3-pentenoic acid (**30**)



This method followed the method of Gendron *et al.* with modification.¹⁷³ Trans- β -hydromuconic acid (1.4 g, 9.70 mmol) was dissolved in CHCl_3 (40 mL) at 45 $^\circ\text{C}$ under stirring. Excess conc. H_2SO_4 (4 mL) was added, followed by small portions of NaN_3 (650 mg, 10.0 mmol) over a period of 30 min. After 4 h the CHCl_3 layer was decanted, and the viscous residue washed with CHCl_3 (30 mL). The organic layers were washed with water (40 mL). The aqueous layer was combined with the viscous residue and

diluted with water (110 mL). Meanwhile, Dowex™ resin 50WX8-100 (50 mL) was washed with water (100 mL) and 0.1 M HCl (aq) (80 mL). The resulting resin was then loaded with the aqueous extract and rinsed with water until pH 7. The product was eluted with 1 M pyridine (aq). The resultant precipitate was filtered, rinsed with iPrOH and dried under vacuum to afford the desired product **30** as a white solid (518 mg, 46%). ¹H NMR (D₂O, 400 MHz) δ_{H} ppm: 6.07 – 5.95 (1H, m, CHCH₂NH₂), 5.71 – 5.58 (1H, m, CHCH₂CO), 3.60 (2H, d, J = 6.7 Hz, CH₂NH₂), 3.01 (2H, d, J = 7.1 Hz, CH₂CO). ¹³C NMR (D₂O, 101 MHz) δ_{C} ppm: 180.20 (C=O), 133.01 (CHCH₂NH₂), 122.98 (CHCH₂CO), 40.94 (CH₂NH₂), 40.89 (CH₂CO). IR ν_{max} (neat) / cm⁻¹ 2752 (br, NH), 1560 (C=O), 1365 (OH), 976 (C=C).

6.3.3.2 (3E)-5-[[[(9H-Fluoren-9-ylmethoxy)carbonyl]amino]-3-pentenoic acid (**31**)

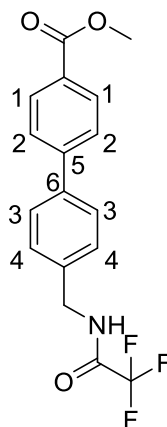


This method followed the method of Gendron *et al.* with modification.¹⁷³ Amino acid **30** (518 mg, 4.5 mmol) was dissolved in aq Na₂CO₃ 10% (w/w) (14.3 mL, 13.5 mmol). Fmoc-Cl (1.16 g, 4.5 mmol) was dissolved in THF (30 mL) and slowly added to the aqueous solution under stirring at 0°C. The reaction was allowed to warm and stirred at RT for 4 h. The mixture was diluted in water (30 mL) and washed with Et₂O (2 x 30 mL). The aqueous phase was acidified to pH 4 with 1 M HCl (aq) and extracted with EtOAc (3 x 30 mL). The organic phases were combined, dried (MgSO₄), filtered and concentrated under vacuum to afford the desired product **31** as a white solid (904 mg, 60%). ¹H NMR (CDCl₃, 400 MHz) δ_{H} ppm: 7.73 – 7.66 (2H, m, ArH(4+5)), 7.52 (2H, d, J = 7.5 Hz, ArH(2+7)), 7.37 – 7.29 (2H, m, ArH(1+8)), 7.24 (2H, td, J = 7.5, 1.2 Hz, ArH(3+6)), 5.70 – 5.53 (2H, m, 2CH), 4.78 (1H, s, CCH), 4.36 (2H, d, J = 7.0 Hz, OCH₂), 4.19 – 4.13 (1H, m, NH), 3.78 – 3.73 (2H, m, CH₂NH), 3.09 – 3.03 (2H, m, CH₂CO). ¹³C NMR (CDCl₃, 101 MHz) δ_{C} ppm: 175.66 (C=OOH), 156.43 (C=OCH₂), 143.85 (Ar(a+d)), 141.33 (Ar(b+c)), 130.48 (CHCH₂NH), 127.75 (ArC(1+8)), 127.08 (ArC(3+6)), 125.05 (ArC(2+7)), 124.15 (CHCH₂CO), 120.01 (ArC(4+5)), 66.75 (OCH₂),

51.96 (CH₂O), 47.04 (CH₂NH), 37.34 (CH₂CO). IR $\nu_{\max}$ (neat) / cm⁻¹ 3315 (NH), 2915 (CH₂), 1697 (NHC=O), 1542 (C=OOH), 1255 (CN), 975 (C=C), 733 (CH).

6.3.4 Synthesis of dibenzene-linked peptide analogue (19)

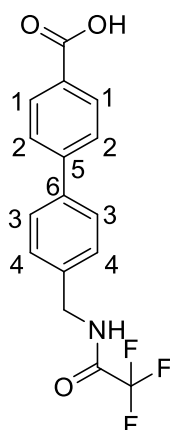
6.3.4.1 Methyl 4'-(N-Trifluoroacetyl)aminomethyl-[1,1'-biphenyl]-4-carboxylate (32)



This method followed the method of Mash *et al.* with modification.¹⁵² To a solution of 4-aminomethylphenylboronic acid hydrochloride (1 g, 5.34 mmol) in MeOH (10 mL) were successively added Et₃N (744 μ M, 5.34 mmol) and CF₃CO₂Et (1.28 mL, 10.7 mmol). After stirring at RT for 2 days, the reaction mixture was concentrated *in vacuo*. EtOAc (60 mL) and H₂O (10 mL) were added, and the layers separated. The organic layer was washed with brine (10 mL) and the combined organic layers extracted with EtOAc (3 x 60 mL). The organic layers were combined, dried (MgSO₄), filtered and concentration *in vacuo* to afford an intermediate boronic acid, that was used immediately in the next step. To a solution of methyl 4-bromobenzoate (733 mg, 3.41 mmol) in THF (6 mL) was added Pd(PPh₃)₄ (394 mg, 0.341 mmol) under a nitrogen atmosphere and the mixture was stirred at RT for 20 min. To the solution were successively added the boronic acid intermediate (1.009 g, crude) and KF (832 mg, 14.3 mmol) and the mixture was heated to reflux. After 48 h the mixture was cooled to 0 °C and EtOAc (20 mL) and brine (20 mL) were added. The layers were separated, and the organic layer was washed with brine (20 mL x 2). The aqueous layers were combined and extracted with EtOAc (60 mL). The combined organic layers were dried (MgSO₄), filtered and concentration *in vacuo*. The product was purified by column chromatography (0-50% EtOAc in hexane) to afford **32** as a white solid (683 mg, 38%). ¹H NMR (DMSO-*d*₆, 400 MHz) δ_{H} ppm: 10.10 – 10.05 (1H, br s,

NH), 8.04 (2H, d, $J = 8.5$, ArH(2)), 7.82 (2H, d, $J = 8.4$, ArH(1)), 7.74 (2H, d, $J = 8.3$, ArH(3)), 7.42 (2H, d, $J = 8.3$, ArH(4)), 4.50 – 4.42 (2H, m, CH₂), 3.88 (3H, s, CH₃). ¹³C NMR (DMSO-*d*₆, 101 MHz) δ_c ppm: 166.04 (COO), 156.45 (q, $J = 36.4$ Hz, COCF₃), 144.28 (ArCCO), 137.90 (ArCCH₂), 137.78 (ArC(5)), 129.81 (ArC(2)), 128.46 (ArC(6)), 128.15 (ArC(1)), 127.15 (ArC(3)), 126.87 (ArC(4)), 116.11 (q, $J = 287.7$ Hz, CF₃), 52.14 (CH₃), 42.29 (CH₂). ¹⁹F NMR (376 MHz, DMSO-*d*₆) δ_f ppm: -74.30 (CF₃). HRMS (ES+) calculated for C₁₇H₁₄F₃NO₃ (M+H)⁺ 338.0926 found 338.0990.

6.3.4.2 4'-(*N*-Trifluoroacetyl)aminomethyl-[1,1'-biphenyl]-4-carboxylic acid (**33**)



This method followed the method of Mash *et al.* with modification.¹⁵² To a solution of ester **32** (520 mg, 1.54 mmol) in THF:MeOH (20 mL 1:1) was added 2 M NaOH (10 mL, aq.) dropwise with stirring. After 20 h the reaction mixture was concentrated under reduced pressure until ~3 mL remained, and H₂O (5 mL) was added. The mixture was cooled to 0 °C and acidified to pH 4 with 2 M HCl (aq). The precipitate was collected on a Büchner funnel, washed successively with H₂O (5 mL x 3), MeOH (2 mL x 2), acetone (2 mL), CH₂Cl₂ (2 mL), and Et₂O (5 mL), then dried in air to afford an amine intermediate, that was immediately used in the next step. To a suspension of the amine intermediate (230 mg, crude) in pyridine (35 mL) was added trifluoroacetic anhydride (352 μ M, 2.53 mmol) dropwise. The mixture was stirred at RT for 18 h and concentrated *in vacuo*, co-evaporating with toluene (50 mL x 3). The residue was purified by column chromatography (0-60% EtOAc in hexane) to afford **33** as a white solid (306 mg, 86%). ¹H NMR (DMSO-*d*₆, 400 MHz) δ_H ppm: 13.07 – 12.85 (1H, br s, OH), 10.06 (1H, t, $J = 6.0$ Hz, NH), 8.02 (2H, d, $J = 8.2$ Hz, ArH(2)), 7.79 (2H, d, $J = 8.1$ Hz, ArH(1)), 7.73 (2H, d, $J = 8.1$ Hz, ArH(3)), 7.40 (2H, d, $J = 8.1$ Hz, ArH(4)), 4.45 (2H, d, $J = 6.0$ Hz, CH₂). ¹³C NMR (DMSO-*d*₆, 101 MHz) δ_c ppm: 167.13 (COO), 156.46 (q, $J = 36.3$ Hz, COCF₃), 143.92 (ArCCO), 138.12 (ArCCH₂),

137.64 (ArC(5)), 130.00 (ArC(2)), 129.66 (ArC(6)), 128.15 (ArC(1)), 127.15 (ArC(3)), 126.75 (ArC(4)), 116.06 (q, J = 288.1 Hz, CF₃), 42.31 (CH₂). ¹⁹F NMR (376 MHz, DMSO-*d*₆) δ_F ppm: -74.28 (CF₃). HRMS (ES+) calculated for C₁₆H₁₂F₃NO₃ (M+H)⁺ 324.0769 found 324.0701.

6.4 Biological Assays

6.4.1 PAGE assay

This method followed Howell with modification.¹²⁷

DNA sequences:

B = 5'-GCCTAGCATGATACTGCTACCG-3'

H = 5'-CGGTAGCAGTACCGTTGGTGGC-3'

R = 5'-GCCACCAACGGCGTCAACTGCC-3'

X (unlabelled) = 5'-GGCAGTTGACGTCATGCTAGGC-3'

X (fluorescently labelled) = 5'-FAM-GGCAGTTGACGTCATGCTAGGC-TAMRA-3'

dsDNA 1 = 5'-CGGAATTGCTGAGTCACTGTTACTC-3'

dsDNA 2 = 5'-GAGTAACAGTGACTCAGCAATTCCG-3'

Y = 5'-ATCTGCAGGGTCTGGTTTCCAGACCTACGATTG-3'

Z = 5'-FAM-CAATCGTAGGAGACCTTTGGTCTCCCTGCAGAT-3'

General gel electrophoresis procedure:

Gel electrophoresis was carried out with the PowerEase™ Touch 120 W Power Supply and Mini Gel Tank using a 10% TBE native polyacrylamide gel for 2 h at 50 V, with 1 x TBE as the running buffer. PAGE analysis was carried out once for each compound, meaning reproducibility cannot be discussed.

Fluorescently labelled B/H/R/X DNA:

Oligonucleotides B, H, R and X (fluorescently labelled) were diluted to 5 μM in 1 x Tris-Borate-EDTA (TBE) pH 8.3 buffer. Stoichiometric amounts of each oligonucleotide were combined, to give a final 4WJ concentration of 1.25 μM. The DNA was annealed in the presence of 2 μL of drug (50-fold dilution in DMSO), at 90 °C for 5 min followed by gradual cooling to 20 °C over 12 h. After electrophoresis gels

were visualised using the ImageQuant™ LAS 4000 biomolecular imager, detecting FAM (460 nm).

Unlabelled B/H/R/X DNA:

Oligonucleotides B, H, R and X (unlabelled) were diluted to 1 μ M in 1 x TBE. Stoichiometric amounts of each oligonucleotide were combined, to give a final 4WJ concentration of 250 nM. The DNA was annealed at 90 °C for 5 min followed by gradual cooling to 20 °C over 12 h. 2 μ L of H-WRWYRGGRYWRW-NH₂ (50-fold dilution in DMSO) was added to the DNA solution and incubated at RT for 1 h. After electrophoresis the gel was stained for 40 min at RT in a solution of 1 x SYBR™ Gold Nucleic Acid Gel Stain in 1 x TBE. Gels were then visualised using the ImageQuant™ LAS 4000 biomolecular imager at 460 nm.

4WJ/dsDNA competition:

Oligonucleotides B, H, R and X (unlabelled) were diluted to 2 μ M in 1 x TBE. Stoichiometric amounts of each oligonucleotide were combined, to give a final 4WJ concentration of 500 nM. The DNA was annealed at 90 °C for 5 min followed by gradual cooling to 20 °C over 12 h. The 4WJ DNA (50 μ L) was incubated, for 1 h at RT, in the presence of 2 μ L H-WRWYRGGRYWRW-NH₂ (12 molar equiv final concentration). 50 μ L solutions of pre-annealed dsDNA (0.195 – 25 final molar equiv), dissolved in 1 x TBE, were added to give a final 4WJ-DNA concentration of 250 nM. The resultant solution was incubated for 1 h at RT, before gel electrophoresis and SYBR™ Gold staining/visualisation were carried out, using the same method as for the unlabelled DNA.

Fluorescently labelled Y/Z DNA:

Oligonucleotides Y and Z (labelled with FAM) were diluted to 2.5 μ M in 1 x Tris-Borate-EDTA (TBE) pH 8.3 buffer. Stoichiometric amounts of each oligonucleotide were combined, to give a final 4WJ concentration of 1.25 μ M. The DNA was annealed in the presence of 2 μ L of drug (50-fold dilution in DMSO), at 90 °C for 5 min followed by gradual cooling to 20 °C over 12 h. Electrophoresis was carried out and visualised using the same method as for the fluorescently labelled B/H/R/X DNA.

6.4.2 FRET

This method followed the method of Vitoc and Mukerji with modification.¹²⁵

Junction 3 DNA sequences:

B = 5'-CCTCCGTCCTAGCAAGGGGCTGCTACCGGAAGGG-3'

H = 5'-CCCTTCCGGTAGCAGCCTGAGCGGTGGTTGAAGG-3'

R = 5'-CCTTCAACCACCGCTCAACTCAACTGCAGTCTGG-3'

X = 5'-CCAGACTGCAGTTGAGTCCTTGCTAGGACGGAGG-3'

X(TAMRA) = 5'-TAMRA-CCAGACTGCAGTTGAGTCCTTGCTAGGACGGAGG-3'

R(FAM) = 5'-FAM-CCTTCAACCACCGCTCAACTCAACTGCAGTCTGG-3'

4WJs:

(F+T) FAM and TAMRA = X(Tamra)/R(Fam)/B/H

(F) FAM-only = X/R(FAM)/B/H

(T) TAMRA-only = X(TAMRA)/R/B/H

Procedure:

The 3 4WJs, containing 17 bp arms and differing fluorescent labels, were constructed *via* annealing at 90 °C for 5 min followed by gradual cooling to 20 °C over 12 h in 10 mM Tris, 1 mM EDTA and 300 mM NaCl (pH 7.6). Each 4WJ was buffer exchanged into a low salt buffer containing 10 mM Tris-HCl and 0.1 mM EDTA (pH 7.6), using Amicon® Ultra Centrifugal Filters with a 3 kDa molecular weight cut-off. Experiments were performed using a 3.5 mL cuvette with a 1 cm path length in an Edinburgh instruments FS5 fluorimeter, with an excitation bandwidth of 4 nm and an emission bandwidth of 8 nm. Measurements were taken at 20 °C, with the polarisers set to 0° for excitation and 55° for emission. The **(F+T)** FAM and TAMRA and **(F)** FAM-only junctions were excited at 490 nm and the **(T)** TAMRA-only junction was excited at 565 nm. All emission spectra recorded from 500 – 650 nm at 0.2 nm/s. For all FRET experiments, the DNA concentration and volume was kept consistent (50 nM, 2 mL) and ligands (1 µL per addition) were titrated into the solution. Dilution effects were corrected for in calculations. The titration was repeated in triplicate for each of the 3 4WJs. Control experiments of ligands titrated into buffer were also carried out.¹²⁵

FRET efficiency (E) was calculated using the following equation:

$$E = F_{DA} \div F_A$$

Where F_{DA} is the emission of the acceptor (TAMRA) in the presence of the donor (FAM) at 580 nm and F_A is the emission of the acceptor only at 580 nm. The contribution of FAM to the F_{DA} value was removed by first calculating the conversion

factor required to match the spectra of the **F+T** 4WJ to the **F** 4WJ between 500 – 530 nm (where only FAM emission is observed). This conversion factor was used to multiply each wavelength of **F+T**. The **F** 4WJ values were subtracted from the converted **F+T** 4WJ spectrum at each wavelength to provide the F_{DA} value, also known as the “extracted acceptor emission”.¹²³ The average values of triplicates from the **F** and **T** singly labelled 4WJs were used for the calculations and each **F+T** value was used separately.

6.4.3 Fluorescence polarisation assay

This method followed the method of Beekman *et al.* with modification.²⁰² Fluorescence Polarisation (FP) was performed on a BMG Labtech CLARIOstar microplate reader, with an FP optic measured at 482 ± 16 nm/ 530 ± 40 nm. Corning™ 96-well black polystyrene plates were used. The singly labelled Junction 3 4WJ (**B**), containing X/R(FAM)/B/H oligonucleotides, from the FRET procedure (**section 6.4.2**) was used and annealed in the same conditions. The FAM-labelled 4WJ was buffer exchanged into a low salt buffer containing 1 mM PBS, 10 mM NaCl and 0.05% Tween at pH 6.5, using Amicon® Ultra Centrifugal Filters with a 3 kDa molecular weight cut-off. Each binding well contained 40 μ L of 20 nM FAM-labelled 4WJ (10 nM final concentration) and 40 μ L peptide in increasing concentrations (25 nM to 3 μ M). Peptides were dissolved in the same low salt buffer as the DNA. Reagents were incubated at RT for 20 min and bubbles removed prior to read out in triplicate experiments. The data was processed using nonlinear regression to give EC_{50} values, using GraphPad Prism Version 6.0 software.

6.4.4 Ethidium bromide displacement assay

This method followed Howell with modification.¹²⁷ The ethidium bromide displacement assay was performed on a BMG Labtech CLARIOstar microplate reader, with emission intensity recorded at 610 ± 40 nm after excitation at 555 ± 20 nm. A 14.1 mg/mL stock solution Calf thymus DNA was diluted using MilliQ water to a concentration of 40.5 μ g/mL, measured using a Nanodrop microvolume spectrophotometer. Using the average molecular weight of a DNA base pair of 650 g/mol, this is equivalent to 62.3 μ M (bp). Corning™ 96-well non-binding black polystyrene plates were used. Each well contained 88 μ L of Tris buffer containing ethidium bromide (0.1 M Tris, 0.1 M NaCl, pH 7 and 4.4 μ M ethidium bromide) and 10 μ L of CT-DNA (4.05 μ g/mL final concentration). To each well was added 2 μ L of

compound in DMSO, in increasing concentrations (10 nM – 400 μ M final concentrations). After incubation at RT, for 2 h in the dark, the plate read in triplicate experiments. Control wells containing no drug (100% fluorescence) and no DNA (0% fluorescence) were also carried out in triplicate. The data was processed using nonlinear regression to give C_{50} values, using GraphPad Prism Version 6.0 software.

6.4.5 MTS assay

This method followed Howell with modification.¹²⁷ Compound cytotoxicity was measured using an MTS assay with the HL-60 promyelocytic cell line. The outer wells of a 96-well plate were filled 100 μ L of MilliQ water, to prevent the evaporation of the central wells. The remaining wells were seeded with HL-60 cells (100 μ L, 3×10^5 cells/mL) and incubated overnight at 37 °C, 5% CO₂. Cells were treated with 1 μ L of compound dissolved in DMSO in increasing concentrations (1.56 nM – 200 μ M) in triplicate for 72 h at 37 °C, 5% CO₂. MTS (10 μ L) was added, and the cells were incubated for a further 3 h. Absorbance was measured at 492 nm using the BMG Labtech Clariostar microplate reader. Cell viability (%) values were obtained from triplicate readings, looking at the relative fluorescence of the treated cells, compared to untreated cells containing 1 μ L DMSO (100%) and wells containing MilliQ water with MTS added (0%). The data for the full binding curves was processed using nonlinear regression to give IC_{50} values, using GraphPad Prism Version 6.0 software.

References

- 1 J. D. Watson and F. H. C. Crick, Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid, *Nature*, 1953, **171**, 737–738.
- 2 R. E. Franklin and R. G. Gosling, Evidence for 2-Chain Helix in Crystalline Structure of Sodium Deoxyribonucleate, *Nature*, 1953, **172**, 156–157.
- 3 A. Klug, The Discovery of the DNA Double Helix, *J. Mol. Biol.*, 2004, **335**, 3–26.
- 4 B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts and P. Walter, *Molecular biology of the cell*, Garland Science, New York, 4th edn., 2002.
- 5 J. B. Krall, P. J. Nichols, M. A. Henen, Q. Vicens and B. Vögeli, Structure and Formation of Z-DNA and Z-RNA, *Molecules*, 2023, **28**, 843.
- 6 W. Fuller, W. H. F. Wilkins, H. R. Wilson and L. D. Hamilton, The molecular configuration of deoxyribonucleic acid: IV. X-ray diffraction study of the A form, *J. Mol. Biol.*, 1965, **12**, 60–76.
- 7 Y. Tanaka, S. Fujii, H. Hiroaki, T. Sakata, T. Tanaka, S. Uesugi, K. Tomita and Y. Kyogoku, A'-form RNA double helix in the single crystal structure of r(UGAGCUUCGGCUC), *Nucleic Acids Res.*, 1999, **27**, 949–955.
- 8 Z. Luo, M. Dauter and Z. Dauter, Phosphates in the Z-DNA dodecamer are flexible, but their P-SAD signal is sufficient for structure solution, *Acta Crystallogr. D Biol. Crystallogr.*, 2014, **70**, 1790–1800.
- 9 M. Kaushik, S. Kaushik, K. Roy, A. Singh, S. Mahendru, M. Kumar, S. Chaudhary, S. Ahmed and S. Kukreti, A bouquet of DNA structures: Emerging diversity, *Biochem. Biophys. Rep.*, 2016, **5**, 388–395.
- 10 Y. Hashimoto, S. Shil, M. Tsuruta, K. Kawauchi and D. Miyoshi, Three- and four-stranded nucleic acid structures and their ligands, *RSC Chem. Biol.*, 2025, **6**, 466–491.
- 11 M. Dalla Pozza, A. Abdullrahman, C. J. Cardin, G. Gasser and J. P. Hall, Three's a crowd – stabilisation, structure, and applications of DNA triplexes, *Chem. Sci.*, 2022, **13**, 10193–10215.
- 12 A. S. Ricciardi, N. A. McNeer, K. K. Anandalingam, W. M. Saltzman and P. M. Glazer, in *Cancer Genomics and Proteomics: Methods and Protocols*, ed. N. Wajapeyee, Springer New York, New York, NY, 2014, pp. 89–106.
- 13 C. Li, Z. Zhou, C. Ren, Y. Deng, F. Peng, Q. Wang, H. Zhang and Y. Jiang, Triplex-forming oligonucleotides as an anti-gene technique for cancer therapy, *Front. Pharmacol.*, 2022, **13**, 1007723.
- 14 D. Sun, B. Thompson, B. E. Cathers, M. Salazar, S. M. Kerwin, J. O. Trent, T. C. Jenkins, S. Neidle and L. H. Hurley, Inhibition of Human Telomerase by a G-Quadruplex-Interactive Compound, *J. Med. Chem.*, 1997, **40**, 2113–2116.
- 15 J. Figueiredo, J.-L. Mergny and C. Cruz, G-quadruplex ligands in cancer therapy: Progress, challenges, and clinical perspectives, *Life Sci.*, 2024, **340**, 122481.

- 16 J. Hilton, K. Gelmon, P. L. Bedard, D. Tu, H. Xu, A. V Tinker, R. Goodwin, S. A. Laurie, D. Jonker, A. R. Hansen, Z. W. Veitch, D. J. Renouf, L. Hagerman, H. Lui, B. Chen, D. Kellar, I. Li, S.-E. Lee, T. Kono, B. Y. C. Cheng, D. Yap, D. Lai, S. Beatty, J. Soong, K. I. Pritchard, I. Soria-Bretones, E. Chen, H. Feilotter, M. Rushton, L. Seymour, S. Aparicio and D. W. Cescon, Results of the phase I CCTG IND.231 trial of CX-5461 in patients with advanced solid tumors enriched for DNA-repair deficiencies, *Nat. Commun.*, 2022, **13**, 3607.
- 17 A. A. Ahmed, W. Greenhalf, D. H. Palmer, N. Williams, J. Worthington, T. Arshad, S. Haider, E. Alexandrou, D. Guneri, Z. A. E. Waller and S. Neidle, The Potent G-Quadruplex-Binding Compound QN-302 Downregulates S100P Gene Expression in Cells and in an In Vivo Model of Pancreatic Cancer, *Molecules*, 2023, **28**, 2452.
- 18 J. A. Capra, K. Paeschke, M. Singh and V. A. Zakian, G-Quadruplex DNA Sequences Are Evolutionarily Conserved and Associated with Distinct Genomic Features in *Saccharomyces cerevisiae*, *PLoS Comput. Biol.*, 2010, **6**, 1000861.
- 19 S. Sedghi Masoud and K. Nagasawa, i-Motif-Binding Ligands and Their Effects on the Structure and Biological Functions of i-Motif, *Chem. Pharm. Bull.*, 2018, **66**, 1091–1103.
- 20 H. A. Day, P. Pavlou and Z. A. E. Waller, i-Motif DNA: Structure, stability and targeting with ligands, *Bioorg. Med. Chem.*, 2014, **22**, 4407–4418.
- 21 C. Roxo and A. Pasternak, Switching off cancer – An overview of G-quadruplex and i-motif functional role in oncogene expression, *Bioorg. Med. Chem. Lett.*, 2025, **116**, 130038.
- 22 E. Alexandrou, D. Guneri, S. Neidle and Z. A. E. Waller, QN-302 demonstrates opposing effects between i-motif and G-quadruplex DNA structures in the promoter of the S100P gene, *Org. Biomol. Chem.*, 2024, **22**, 55–58.
- 23 D. M. J. Lilley, in *Mathematics of DNA Structure, Function and Interactions*, eds. C. J. Benham, S. Harvey, W. K. Olson, D. W. Sumners and D. Swigon, Springer New York, New York, NY, 2009, pp. 213–224.
- 24 K. Hiom, Recombination: Homologous recombination branches out, *Curr. Biol.*, 2001, **11**, R278–R280.
- 25 J. W. Szostak, T. L. Orr-Weaver, R. J. Rothstein and F. W. Stahl, The double-strand-break repair model for recombination, *Cell*, 1983, **33**, 25–35.
- 26 N. Hirano, T. Muroi, H. Takahashi and M. Haruki, Site-specific recombinases as tools for heterologous gene integration, *Appl. Microbiol. Biotechnol.*, 2011, **92**, 227–239.
- 27 G. Meinke, A. Bohm, J. Hauber, M. T. Pisabarro and F. Buchholz, Cre Recombinase and Other Tyrosine Recombinases, *Chem. Rev.*, 2016, **116**, 12785–12820.
- 28 A. Upadhyay, Cancer: An unknown territory; rethinking before going ahead, *Genes Dis.*, 2021, **8**, 655–661.
- 29 M. G. Vander Heiden and R. J. DeBerardinis, Understanding the Intersections between Metabolism and Cancer Biology, *Cell*, 2017, **168**, 657–669.

- 30 A. W. Lambert, D. R. Pattabiraman and R. A. Weinberg, Emerging Biological Principles of Metastasis, *Cell*, 2017, **168**, 670–691.
- 31 F. Bray, M. Laversanne, H. Sung, J. Ferlay, R. L. Siegel, I. Soerjomataram and A. Jemal, Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, *CA Cancer J. Clin.*, 2024, **74**, 229–263.
- 32 K. Cheung-Ong, G. Giaever and C. Nislow, DNA-Damaging Agents in Cancer Chemotherapy: Serendipity and Chemical Biology, *Chem. Biol.*, 2013, **20**, 648–659.
- 33 L. R. Ferguson, H. Chen, A. R. Collins, M. Connell, G. Damia, S. Dasgupta, M. Malhotra, A. K. Meeker, A. Amedei, A. Amin, S. S. Ashraf, K. Aquilano, A. S. Azmi, D. Bhakta, A. Bilsland, C. S. Boosani, S. Chen, M. R. Ciriolo, H. Fujii, G. Guha, D. Halicka, W. G. Helferich, W. N. Keith, S. I. Mohammed, E. Niccolai, X. Yang, K. Honoki, V. R. Parslow, S. Prakash, S. Rezazadeh, R. E. Shackelford, D. Sidransky, P. T. Tran, E. S. Yang and C. A. Maxwell, Genomic instability in human cancer: Molecular insights and opportunities for therapeutic attack and prevention through diet and nutrition, *Semin. Cancer Biol.*, 2015, **35**, S5–S24.
- 34 W. D. Wright, S. S. Shah and W.-D. Heyer, Homologous recombination and the repair of DNA double-strand breaks, *J. Biol. Chem.*, 2018, **293**, 10524–10535.
- 35 S. B. Chernikova, J. C. Game and J. M. Brown, Inhibiting homologous recombination for cancer therapy, *Cancer Biol. Ther.*, 2012, **13**, 61–68.
- 36 Q. Yin, X. Liu, L. Hu, Q. Song, S. Liu, Q. Huang, Z. Geng, Y. Zhu, X. Li, F. Fu and H. Wang, VE-822, a novel DNA Holliday junction stabilizer, inhibits homologous recombination repair and triggers DNA damage response in osteogenic sarcomas, *Biochem. Pharmacol.*, 2021, **193**, 114767.
- 37 L. S. Cox and D. P. Lane, Tumour suppressors, kinases and clamps: How p53 regulates the cell cycle in response to DNA damage, *BioEssays*, 1995, **17**, 501–508.
- 38 S. Lee, L. Cavallo and J. Griffith, Human p53 Binds Holliday Junctions Strongly and Facilitates Their Cleavage, *J. Biol. Chem.*, 1997, **272**, 7532–7539.
- 39 E. Ivens, M. M. D. Cominetti and M. Searcey, Junctions in DNA: underexplored targets for therapeutic intervention, *Bioorg. Med. Chem.*, 2022, **69**, 116897.
- 40 Q. Song, Y. Hu, A. Yin, H. Wang and Q. Yin, DNA Holliday Junction: History, Regulation and Bioactivity, *Int. J. Mol. Sci.*, 2022, **23**, 9730.
- 41 W. C. Reygaert, An overview of the antimicrobial resistance mechanisms of bacteria, *AIMS Microbiol.*, 2018, **4**, 482–501.
- 42 E. M. Darby, E. Trampari, P. Siasat, M. S. Gaya, I. Alav, M. A. Webber and J. M. A. Blair, Molecular mechanisms of antibiotic resistance revisited, *Nat. Rev. Microbiol.*, 2023, **21**, 280–295.
- 43 B. Spellberg and D. N. Gilbert, The Future of Antibiotics and Resistance: A Tribute to a Career of Leadership by John Bartlett, *Clin. Infect. Dis.*, 2014, **59**, S71–S75.
- 44 S. Saikia and P. Chetia, Antibiotics: From Mechanism of Action to Resistance and Beyond, *Indian J. Microbiol.*, 2024, **64**, 821–845.

- 45 M. C. Rideout, I. Naili, J. L. Boldt, A. Flores-Fujimoto, S. Patra, J. E. Rostron and A. M. Segall, wrwryggrywrw is a single-chain functional analog of the Holliday junction-binding homodimer, (wrwycr)₂, *Peptides (N.Y.)*, 2013, **40**, 112–122.
- 46 S. C. West, RuvA Gets X-Rayed on Holliday, *Cell*, 1998, **94**, 699–701.
- 47 R. Holliday, A mechanism for gene conversion in fungi, *Genet. Res. (Camb.)*, 2007, **89**, 285–307.
- 48 N. C. Seeman, M. F. Maestre, R. I. Ma and N. R. Kallenbach, Physical characterization of a nucleic acid junction, *Prog. Clin. Biol. Res.*, 1985, **172A**, 99–108.
- 49 N. R. Kallenbach, R.-I. Ma, A. J. Wand, G. H. Veeneman, J. H. van Boom and N. C. Seeman, Fourth Rank Immobile Nucleic Acid Junctions, *J. Biomol. Struct. Dyn.*, 1983, **1**, 159–168.
- 50 J. P. Cooper and P. J. Hagerman, Gel electrophoretic analysis of the geometry of a DNA four-way junction, *J. Mol. Biol.*, 1987, **198**, 711–719.
- 51 J. P. Cooper and P. J. Hagerman, Geometry of a Branched DNA Structure in Solution, *Proc. Natl. Acad. Sci. U. S. A.*, 1989, **86**, 7336–7340.
- 52 N. C. Seeman and N. R. Kallenbach, Design of immobile nucleic acid junctions, *Biophys. J.*, 1983, **44**, 201–209.
- 53 D. R. Duckett, A. I. H. Murchie, S. Diekmann, E. von Kitzing, B. Kemper and D. M. J. Lilley, The structure of the Holliday junction, and its resolution, *Cell*, 1988, **55**, 79–89.
- 54 P. A. Khuu, A. R. Voth, F. A. Hays and P. S. Ho, The stacked-X DNA Holliday junction and protein recognition, *J. Mol. Recognit.*, 2006, **19**, 234–242.
- 55 C. Joo, S. A. McKinney, D. M. J. Lilley and T. Ha, Exploring Rare Conformational Species and Ionic Effects in DNA Holliday Junctions Using Single-molecule Spectroscopy, *J. Mol. Biol.*, 2004, **341**, 739–751.
- 56 S. M. Roe, T. Barlow, T. Brown, M. Oram, A. Keeley, I. R. Tsaneva and L. H. Pearl, Crystal Structure of an Octameric RuvA–Holliday Junction Complex, *Mol. Cell*, 1998, **2**, 361–372.
- 57 D. Hargreaves, D. W. Rice, S. E. Sedelnikova, P. J. Artymiuk, R. G. Lloyd and J. B. Rafferty, Crystal structure of *E. coli* RuvA with bound DNA Holliday junction at 6 Å resolution, *Nat. Struct. Biol.*, 1998, **5**, 441–446.
- 58 R. J. Bennett and S. C. West, Structural Analysis of the RuvC–Holliday Junction Complex Reveals an Unfolded Junction, *J. Mol. Biol.*, 1995, **252**, 213–226.
- 59 F. Guo, D. N. Gopaul and G. D. Van Duyne, Structure of Cre recombinase complexed with DNA in a site-specific recombination synapse, *Nature*, 1997, **389**, 40–46.
- 60 B. F. Eichman, J. M. Vargason, B. H. M. Mooers and P. S. Ho, The Holliday junction in an inverted repeat DNA sequence: Sequence effects on the structure of four-way junctions, *Proc. Natl. Acad. Sci. U.S.A.*, 2000, **97**, 3971–3976.
- 61 M. Ortiz-Lombardía, A. González, R. Eritja, J. Aymamí, F. Azorín and M. Coll, Crystal structure of a DNA Holliday junction, *Nat. Struct. Mol. Biol.*, 1999, **6**, 913–917.

- 62 E. Logerot, G. Cazals, A. Memboeuf and C. Enjalbal, Revealing C-terminal peptide amidation by the use of the survival yield technique, *Anal. Biochem.*, 2022, **655**, 114823.
- 63 G. Cassell, M. Klemm, C. Pinilla and A. Segall, Dissection of bacteriophage λ site-specific recombination using synthetic peptide combinatorial libraries, *J. Mol. Biol.*, 2000, **299**, 1193–1202.
- 64 K. Ghosh, C. K. Lau, F. Guo, A. M. Segall and G. D. Van Duyne, Peptide Trapping of the Holliday Junction Intermediate in Cre-*loxP* Site-specific Recombination, *J. Biol. Chem.*, 2005, **280**, 8290–8299.
- 65 J. L. Boldt, C. Pinilla and A. M. Segall, Reversible Inhibitors of λ Integrase-mediated Recombination Efficiently Trap Holliday Junction Intermediates and Form the Basis of a Novel Assay for Junction Resolution, *J. Biol. Chem.*, 2004, **279**, 3472–3483.
- 66 K. V. Kepple, N. Patel, P. Salamon and A. M. Segall, Interactions between branched DNAs and peptide inhibitors of DNA repair, *Nucleic Acids Res.*, 2008, **36**, 5319–5334.
- 67 K. V. Kepple, J. L. Boldt and A. M. Segall, Holliday junction-binding peptides inhibit distinct junction-processing enzymes, *Proc. Natl. Acad. Sci. U.S.A.*, 2005, **102**, 6867–6872.
- 68 M. C. Rideout, J. L. Boldt, G. Vahi-Ferguson, P. Salamon, A. Nefzi, J. M. Ostresh, M. Giulianotti, C. Pinilla and A. M. Segall, Potent antimicrobial small molecules screened as inhibitors of tyrosine recombinases and Holliday junction-resolving enzymes, *Mol. Divers.*, 2011, **15**, 989.
- 69 L. Y. Su, D. L. Willner and A. M. Segall, An Antimicrobial Peptide That Targets DNA Repair Intermediates In Vitro Inhibits *Salmonella* Growth within Murine Macrophages, *Antimicrob. Agents Chemother.*, 2010, **54**, 1888–1899.
- 70 C. W. Gunderson and A. M. Segall, DNA repair, a novel antibacterial target: Holliday junction-trapping peptides induce DNA damage and chromosome segregation defects, *Mol. Microbiol.*, 2006, **59**, 1129–1148.
- 71 M. Dey, S. Patra, L. Y. Su and A. M. Segall, Tumor Cell Death Mediated by Peptides That Recognize Branched Intermediates of DNA Replication and Repair, *PLoS One*, 2013, **8**, 78751.
- 72 A. A. Baxter, F. T. Lay, I. K. H. Poon, M. Kvensakul and M. D. Hulett, Tumor cell membrane-targeting cationic antimicrobial peptides: novel insights into mechanisms of action and therapeutic prospects, *Cell. Mol. Life Sci.*, 2017, **74**, 3809–3825.
- 73 M. L. Bolla, E. V. Azevedo, J. M. Smith, R. E. Taylor, D. K. Ranjit, A. M. Segall and S. R. McAlpine, Novel Antibiotics: Macrocyclic Peptides Designed to Trap Holliday Junctions, *Org. Lett.*, 2003, **5**, 109–112.
- 74 L. A. Liotta, I. Medina, J. L. Robinson, C. L. Carroll, P.-S. Pan, R. Corral, J. V. C. Johnston, K. M. Cook, F. A. Curtis, G. J. Sharples and S. R. McAlpine, Novel antibiotics: second generation macrocyclic peptides designed to trap Holliday junctions, *Tetrahedron Lett.*, 2004, **45**, 8447–8450.

- 75 P.-S. Pan, F. A. Curtis, C. L. Carroll, I. Medina, L. A. Liotta, G. J. Sharples and S. R. McAlpine, Novel antibiotics: C-2 symmetrical macrocycles inhibiting Holliday junction DNA binding by *E. coli* RuvC, *Bioorg. Med. Chem.*, 2006, **14**, 4731–4739.
- 76 Q. Guo, N. C. Seeman and N. R. Kallenbach, Site-specific interaction of intercalating drugs with a branched DNA molecule, *Biochemistry*, 1989, **28**, 2355–2359.
- 77 M. Lu, Q. Guo, N. C. Seeman and N. R. Kallenbach, Drug binding by branched DNA: selective interaction of the dye Stains-All with an immobile junction, *Biochemistry*, 1990, **29**, 3407–3412.
- 78 M. Lu, Q. Guo, R. F. Pasternack, D. J. Wink, N. C. Seeman and N. R. Kallenbach, Drug binding by branched DNA: selective interaction of tetrapyrrolyl porphyrins with an immobile junction, *Biochemistry*, 1990, **29**, 1614–1624.
- 79 Q. Guo, M. Lu, N. C. Seeman and N. R. Kallenbach, Drug binding by branched DNA molecules: analysis by chemical footprinting of intercalation into an immobile junction, *Biochemistry*, 1990, **29**, 570–578.
- 80 A. Oleksi, A. G. Blanco, R. Boer, I. Usón, J. Aymamí, A. Rodger, M. J. Hannon and M. Coll, Molecular Recognition of a Three-Way DNA Junction by a Metallosupramolecular Helicate, *Angew. Chem. Int. Ed.*, 2006, **45**, 1227–1231.
- 81 L. Cardo, I. Nawroth, P. J. Cail, J. A. McKeating and M. J. Hannon, Metallo supramolecular cylinders inhibit HIV-1 TAR-TAT complex formation and viral replication *in cellulo*, *Sci. Rep.*, 2018, **8**, 13342.
- 82 V. H. S. van Rixel, A. Busemann, M. F. Wissingh, S. L. Hopkins, B. Siewert, C. van de Griend, M. A. Siegler, T. Marzo, F. Papi, M. Ferraroni, P. Gratteri, C. Bazzicalupi, L. Messori and S. Bonnet, Induction of a Four-Way Junction Structure in the DNA Palindromic Hexanucleotide 5'-d(CGTACG)-3' by a Mononuclear Platinum Complex, *Angew. Chem. Int. Ed.*, 2019, **58**, 9378–9382.
- 83 J. S. Craig, L. Melidis, H. D. Williams, S. J. Dettmer, A. A. Heidecker, P. J. Altmann, S. Guan, C. Campbell, D. F. Browning, R. K. O. Sigel, S. Johannsen, R. T. Egan, B. Aikman, A. Casini, A. Pöthig and M. J. Hannon, Organometallic Pillarplexes That Bind DNA 4-Way Holliday Junctions and Forks, *J. Am. Chem. Soc.*, 2023, **145**, 13570–13580.
- 84 C.-M. Chien, P.-C. Wu, R. Satange, C.-C. Chang, Z.-L. Lai, L. D. Hagler, S. C. Zimmerman and M.-H. Hou, Structural Basis for Targeting T:T Mismatch with Triaminotriazine-Acridine Conjugate Induces a U-Shaped Head-to-Head Four-Way Junction in CTG Repeat DNA, *J. Am. Chem. Soc.*, 2020, **142**, 11165–11172.
- 85 M. L. Carpenter, G. Lowe and P. R. Cook, The structure of 4-way DNA junctions: specific binding of bis-intercalators with rigid linkers, *Nucleic acids res.*, 1996, **24**, 1594–1601.
- 86 N. H. Hopcroft, A. L. Brogden, M. Searcey and C. J. Cardin, X-ray crystallographic study of DNA duplex cross-linking: simultaneous binding to two d(CGTACG)₂ molecules by a bis(9-aminoacridine-4-carboxamide) derivative, *Nucleic Acids Res.*, 2006, **34**, 6663–6672.

- 87 M. Kozurkova, Acridine derivatives as inhibitors/poisons of topoisomerase II, *J. Appl. Toxicol.*, 2022, **42**, 544–552.
- 88 A. L. Brogden, N. H. Hopcroft, M. Searcey and C. J. Cardin, Ligand Bridging of the DNA Holliday Junction: Molecular Recognition of a Stacked-X Four-Way Junction by a Small Molecule, *Angew. Chem. Int. Ed.*, 2007, **46**, 3850–3854.
- 89 L. A. Howell, R. A. Bowater, M. A. O’Connell, A. P. Reszka, S. Neidle and M. Searcey, Synthesis of Small Molecules Targeting Multiple DNA Structures using Click Chemistry, *ChemMedChem*, 2012, **7**, 792–804.
- 90 L. A. Howell, Z. A. E. Waller, R. Bowater, M. O’Connell and M. Searcey, A small molecule that induces assembly of a four-way DNA junction at low temperature, *Chem. Commun.*, 2011, **47**, 8262–8264.
- 91 D. K. Ranjit, M. C. Rideout, A. Nefzi, J. M. Ostresh, C. Pinilla and A. M. Segall, Small molecule functional analogs of peptides that inhibit λ site-specific recombination and bind Holliday junctions, *Bioorg. Med. Chem. Lett.*, 2010, **20**, 4531–4534.
- 92 B. F. Eichman, B. H. M. Mooers, M. Alberti, J. E. Hearst and P. S. Ho, The crystal structures of psoralen cross-linked DNAs: drug-dependent formation of Holliday junctions, *J. Mol. Biol.*, 2001, **308**, 15–26.
- 93 L. Wang, N. Wang, W. Zhang, X. Cheng, Z. Yan, G. Shao, X. Wang, R. Wang and C. Fu, Therapeutic peptides: current applications and future directions, *Signal Transduct. Target. Ther.*, 2022, **7**, 48.
- 94 M. Stawikowski and G. B. Fields, Introduction to Peptide Synthesis, *Curr. Protoc. Protein Sci.*, 2012, **69**, 18.1.1-18.1.13.
- 95 R. B. Merrifield, Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide, *J. Am. Chem. Soc.*, 1963, **85**, 2149–2154.
- 96 J. M. Palomo, Solid-phase peptide synthesis: an overview focused on the preparation of biologically relevant peptides, *RSC Adv.*, 2014, **4**, 32658–32672.
- 97 T. Tatsumi, K. Sasamoto, T. Matsumoto, R. Hirano, K. Oikawa, M. Nakano, M. Yoshida, K. Oisaki and M. Kanai, Practical N-to-C peptide synthesis with minimal protecting groups, *Commun. Chem.*, 2023, **6**, 231.
- 98 W. Chan and P. White, *Fmoc Solid Phase Peptide Synthesis: A Practical Approach*, Oxford University Press, 1999.
- 99 C.-D. Chang and J. Meienhofer, Solid-phase peptide synthesis using mild base cleavage of N α -fluorenylmethyloxycarbonylamino acids, exemplified by a synthesis of dihydrosomatostatin, *Int. J. Pept. Protein Res.*, 1978, **11**, 246–249.
- 100 R. B. Merrifield, Solid Phase Synthesis (Nobel Lecture), *Angew. Chem. Int. Ed. Engl.*, 1985, **24**, 799–810.
- 101 S. Duengo, M. I. Muhajir, A. T. Hidayat, W. J. A. Musa and R. Maharani, Epimerisation in Peptide Synthesis, *Molecules*, 2023, **28**, 8017.

- 102 R. Subirós-Funosas, R. Prohens, R. Barbas, A. El-Faham and F. Albericio, Oxyma: An Efficient Additive for Peptide Synthesis to Replace the Benzotriazole-Based HOBt and HOAt with a Lower Risk of Explosion, *Chem. Eur. J.*, 2009, **15**, 9394–9403.
- 103 E. Kaiser, R. L. Colescott, C. D. Bossinger and P. I. Cook, Color test for detection of free terminal amino groups in the solid-phase synthesis of peptides, *Anal. Biochem.*, 1970, **34**, 595–598.
- 104 V. K. Sarin, S. B. H. Kent, J. P. Tam and R. B. Merrifield, Quantitative monitoring of solid-phase peptide synthesis by the ninhydrin reaction, *Anal. Biochem.*, 1981, **117**, 147–157.
- 105 R. Behrendt, P. White and J. Offer, Advances in Fmoc solid-phase peptide synthesis, *J. Pept. Sci.*, 2016, **22**, 4–27.
- 106 J. R. Lakowicz, *Principles of Fluorescence Spectroscopy*, Springer US, 2007.
- 107 D. Frackowiak, The Jablonski diagram, *J. Photochem. Photobiol. B: Biol.*, 1988, **2**, 399.
- 108 G. Anand, T. Thyagarajan and S. Ramakrishnan, Fluorescence Nano Particle Detection in a Liquid Sample Using the Smartphone for Biomedical Application, *J. Fluoresc.*, 2022, **32**, 135–143.
- 109 T. Beppu, K. Tomiguchi, A. Masuhara, Y.-J. Pu and H. Katagiri, Single Benzene Green Fluorophore: Solid-State Emissive, Water-Soluble, and Solvent- and pH-Independent Fluorescence with Large Stokes Shifts, *Angew. Chem. Int. Ed.*, 2015, **54**, 7332–7335.
- 110 K. Rurack and M. Spieles, Fluorescence Quantum Yields of a Series of Red and Near-Infrared Dyes Emitting at 600–1000 nm, *Anal. Chem.*, 2011, **83**, 1232–1242.
- 111 S. V Sciuto and R. A. Jockusch, The intrinsic photophysics of gaseous ethidium ions, *J. Photochem. Photobiol. A.*, 2015, **311**, 186–192.
- 112 J. R. Lakowicz, B. Shen, Z. Gryczynski, S. D'Auria and I. Gryczynski, Intrinsic Fluorescence from DNA Can Be Enhanced by Metallic Particles, *Biochem. Biophys. Res. Commun.*, 2001, **286**, 875–879.
- 113 M. D. Hall, A. Yasgar, T. Peryea, J. C. Braisted, A. Jadhav, A. Simeonov and N. P. Coussens, Fluorescence polarization assays in high-throughput screening and drug discovery: a review, *Methods Appl. Fluoresc.*, 2016, **4**, 022001.
- 114 W. C. Tse and D. L. Boger, A Fluorescent Intercalator Displacement Assay for Establishing DNA Binding Selectivity and Affinity, *Acc. Chem. Res.*, 2004, **37**, 61–69.
- 115 J. Olmsted and D. R. Kearns, Mechanism of ethidium bromide fluorescence enhancement on binding to nucleic acids, *Biochemistry*, 1977, **16**, 3647–3654.
- 116 B. R. Madhanagopal, S. Chen, C.-D. Platt and A. R. Chandrasekaran, Caffeine-induced release of small molecules from DNA nanostructures, *iScience*, 2023, **26**, 106564.
- 117 Y. Zhang, H. Tang, W. Chen and J. Zhang, Nanomaterials Used in Fluorescence Polarization Based Biosensors, *Int. J. Mol. Sci.*, 2022, **23**, 8625.
- 118 B. T. Bajar, E. S. Wang, S. Zhang, M. Z. Lin and J. Chu, A Guide to Fluorescent Protein FRET Pairs, *Sensors*, 2016, **16**, 1488.

- 119 H. Sahoo, Förster resonance energy transfer – A spectroscopic nanoruler: Principle and applications, *J. Photoch. Photobio. C.: Photoch. Rev.*, 2011, **12**, 20–30.
- 120 B. Rana and G. K. Joshi, in *Basic Biotechniques for Bioprocess and Bioentrepreneurship*, eds. A. K. Bhatt, R. K. Bhatia and T. C. Bhalla, Academic Press, 2023, pp. 183–193.
- 121 S. D. Levene and B. H. Zimm, Understanding the Anomalous Electrophoresis of Bent DNA Molecules: a Reptation Model, *Science (1979)*, 1989, **245**, 396–399.
- 122 D. M. J. Lilley, Analysis of branched nucleic acid structure using comparative gel electrophoresis, *Q. Rev. Biophys.*, 2008, **41**, 1–39.
- 123 R. M. Clegg, A. I. H. Murchie, A. Zechel, C. Carlberg, S. Diekmann and D. M. J. Lilley, Fluorescence resonance energy transfer analysis of the structure of the four-way DNA junction, *Biochemistry*, 1992, **31**, 4846–4856.
- 124 S. A. McKinney, A.-C. Déclais, D. M. J. Lilley and T. Ha, Structural dynamics of individual Holliday junctions, *Nat. Struct. Biol.*, 2003, **10**, 93–97.
- 125 C. I. Vitoc and I. Mukerji, HU Binding to a DNA Four-Way Junction Probed by Förster Resonance Energy Transfer, *Biochemistry*, 2011, **50**, 1432–1441.
- 126 M. D’Hondt, N. Bracke, L. Taevernier, B. Gevaert, F. Verbeke, E. Wynendaele and B. De Spiegeleer, Related impurities in peptide medicines, *J. Pharm. Biomed. Anal.*, 2014, **101**, 2–30.
- 127 L. A. Howell, Design, synthesis and biological evaluation of 9-aminoacridines that target the Holliday Junction, *PhD Thesis*, University of East Anglia, 2009.
- 128 J. Yu, T. Ha and K. Schulten, Conformational model of the Holliday junction transition deduced from molecular dynamics simulations, *Nucleic Acids Res.*, 2004, **32**, 6683–6695.
- 129 S. J. Dettmer, H. D. Williams, R. Napier, J. M. Beames, S. Walker-Griffiths, T. D. Craggs and M. J. Hannon, Supramolecular Recognition of a DNA Four-Way Junction by an M2L4 Metallo-Cage, Inspired by a Simulation-Guided Design Approach, *Angew. Chem.*, 2025, **64**, e202504866.
- 130 Guillen Danielle, M. Schievelbein, Patel Kushkumar, Jose Davis and Ouellet Jonathan, A simple and affordable kinetic assay of nucleic acids with SYBR Gold gel staining, *PLoS One*, 2020, **15**, 1–16.
- 131 B. Nickel, On the elimination of the polarization bias of the luminescence or transient absorption of photoexcited isotropic solutions, *J. Lumin.*, 1989, **44**, 1–18.
- 132 J. M. Fogg, M. Kvaratskhelia, M. F. White and D. M. J. Lilley, Distortion of DNA junctions imposed by the binding of resolving enzymes: a fluorescence study, *J. Mol. Biol.*, 2001, **313**, 751–764.
- 133 S. H. Lin, D. Zhao, V. Deng, V. K. Birdsall, S. Ho, O. Buzovetsky, C. M. Etson and I. Mukerji, Integration Host Factor Binds DNA Holliday Junctions, *Int. J. Mol. Sci.*, 2023, **24**, 580.

- 134 N. J. Moerke, Fluorescence Polarization (FP) Assays for Monitoring Peptide-Protein or Nucleic Acid-Protein Binding, *Curr. Protoc. Chem. Biol.*, 2009, **1**, 1–15.
- 135 A.-C. Déclais, J. Liu, A. D. J. Freeman and D. M. J. Lilley, Structural Recognition between a Four-way DNA Junction and a Resolving Enzyme, *J. Mol. Biol.*, 2006, **359**, 1261–1276.
- 136 S. Kamiloglu, G. Sari, T. Ozdal and E. Capanoglu, Guidelines for cell viability assays, *Food Front.*, 2020, **1**, 332–349.
- 137 L. Di, Strategic Approaches to Optimizing Peptide ADME Properties, *AAPS J.*, 2015, **17**, 134–143.
- 138 Y. Han, Z. Gao, L. Chen, L. Kang, W. Huang, M. Jin, Q. Wang and Y. H. Bae, Multifunctional oral delivery systems for enhanced bioavailability of therapeutic peptides/proteins, *Acta Pharm. Sin. B*, 2019, **9**, 902–922.
- 139 M. J. Humphrey, in *Delivery Systems for Peptide Drugs*, eds. S. S. Davis, L. Illum and E. Tomlinson, Springer US, Boston, MA, 1986, pp. 139–151.
- 140 E. Lenci and A. Trabocchi, Peptidomimetic toolbox for drug discovery, *Chem. Soc. Rev.*, 2020, **49**, 3262–3277.
- 141 L.-M. Rečnik, W. Kandjoller and T. L. Mindt, 1,4-Disubstituted 1,2,3-Triazoles as Amide Bond Surrogates for the Stabilisation of Linear Peptides with Biological Activity, *Molecules*, 2020, **25**, 3576.
- 142 H. M. Werner, C. C. Cabalteja and W. S. Horne, Peptide Backbone Composition and Protease Susceptibility: Impact of Modification Type, Position, and Tandem Substitution, *ChemBioChem*, 2016, **17**, 712–718.
- 143 V. P. Reddy Chichili, V. Kumar and J. Sivaraman, Linkers in the structural biology of protein–protein interactions, *Protein Sci.*, 2013, **22**, 153–167.
- 144 A. Markowska, A. R. Markowski and I. Jarocka-Karpowicz, The Importance of 6-Aminohexanoic Acid as a Hydrophobic, Flexible Structural Element, *Int. J. Mol. Sci.*, 2021, **22**, 12122.
- 145 H.-K. Lee, M. D. Smith, B. J. Smith, J. Grussendorf, L. Xu, R. J. Gillies, H. S. White and G. Bulaj, Anticonvulsant Met-Enkephalin Analogues Containing Backbone Spacers Reveal Alternative Non-Opioid Signaling in the Brain, *ACS Chem. Biol.*, 2009, **4**, 659–671.
- 146 R. Abes, A. Arzumanov, H. Moulton, S. Abes, G. Ivanova, M. J. Gait, P. Iversen and B. Lebleu, Arginine-rich cell penetrating peptides: Design, structure–activity, and applications to alter pre-mRNA splicing by steric-block oligonucleotides, *J. Pept. Sci.*, 2008, **14**, 455–460.
- 147 N. Agouram, E. M. El Hadrami and A. Bentama, 1,2,3-Triazoles as Biomimetics in Peptide Science, *Molecules*, 2021, **26**, 2937.
- 148 H. C. Kolb, M. G. Finn and K. B. Sharpless, Click Chemistry: Diverse Chemical Function from a Few Good Reactions, *Angew. Chem.*, 2001, **40**, 2004–2021.

- 149 M. Meldal and F. Diness, Recent Fascinating Aspects of the CuAAC Click Reaction, *Trends Chem.*, 2020, **2**, 569–584.
- 150 C. W. Tornøe, S. J. Sanderson, J. C. Mottram, G. H. Coombs and M. Meldal, Combinatorial Library of Peptidotriazoles: Identification of [1,2,3]-Triazole Inhibitors against a Recombinant *Leishmania mexicana* Cysteine Protease, *J. Comb. Chem.*, 2004, **6**, 312–324.
- 151 J. E. Hein and V. V Fokin, Copper-catalyzed azide–alkyne cycloaddition (CuAAC) and beyond: new reactivity of copper(i) acetylides, *Chem. Soc. Rev.*, 2010, **39**, 1302–1315.
- 152 Y. Monguchi, J. Vagner, H. L. Handl, U. Jana, L. J. Begay, V. J. Hruby, R. J. Gillies and E. A. Mash, Design, synthesis, and validation of rigid linkers for bioactive peptides, *Tetrahedron Lett.*, 2005, **46**, 7589–7592.
- 153 L. Dall’Olmo, N. Papa, N. C. Surdo, I. Marigo and S. Mocellin, Alpha-melanocyte stimulating hormone (α -MSH): biology, clinical relevance and implication in melanoma, *J. Transl. Med.*, 2023, **21**, 562.
- 154 R. M. Slominski, T. Sarna, P. M. Płonka, C. Raman, A. A. Brożyna and A. T. Slominski, Melanoma, Melanin, and Melanogenesis: The Yin and Yang Relationship, *Front. Oncol.*, 2022, **12**, 842496.
- 155 V. Martí-Centelles, M. D. Pandey, M. I. Burguete and S. V Luis, Macrocyclization Reactions: The Importance of Conformational, Configurational, and Template-Induced Preorganization, *Chem. Rev.*, 2015, **115**, 8736–8834.
- 156 H. C. Hayes, L. Y. P. Luk and Y.-H. Tsai, Approaches for peptide and protein cyclisation, *Org. Biomol. Chem.*, 2021, **19**, 3983–4001.
- 157 A. Zorzi, K. Deyle and C. Heinis, Cyclic peptide therapeutics: past, present and future, *Curr. Opin. Chem. Biol.*, 2017, **38**, 24–29.
- 158 S. Caulier, C. Nannan, A. Gillis, F. Licciardi, C. Bragard and J. Mahillon, Overview of the Antimicrobial Compounds Produced by Members of the *Bacillus subtilis* Group, *Front. Microbiol.*, 2019, **10**, 302.
- 159 B. Behsaz, H. Mohimani, A. Gurevich, A. Prjibelski, M. Fisher, F. Vargas, L. Smarr, P. C. Dorrestein, J. S. Mylne and P. A. Pevzner, *De Novo* Peptide Sequencing Reveals Many Cyclopeptides in the Human Gut and Other Environments, *Cell Syst.*, 2020, **10**, 99–108.
- 160 M. Malesevic, U. Strijowski, D. Bächle and N. Sewald, An improved method for the solution cyclization of peptides under pseudo-high dilution conditions, *J. Biotech.*, 2004, **112**, 73–77.
- 161 L. T. Scott, J. Rebek, L. Ovsyanko and C. L. Sims, Organic chemistry on the solid phase. Site-site interactions on functionalized polystyrene, *J. Am. Chem. Soc.*, 1977, **99**, 625–626.
- 162 J. S. McMurray, Solid phase synthesis of a cyclic peptide using Fmoc chemistry, *Tetrahedron Lett.*, 1991, **32**, 7679–7682.

- 163 A. B. Yongye, Y. Li, M. A. Giulianotti, Y. Yu, R. A. Houghten and K. Martínez-Mayorga, Modeling of peptides containing D-amino acids: implications on cyclization, *J. Comput. Aided Mol. Des.*, 2009, **23**, 677–689.
- 164 P. Hainaut and M. Hollstein, in *Adv. Cancer Res.*, eds. G. F. Vande Woude and G. Klein, Academic Press, 1999, vol. 77, pp. 81–137.
- 165 Y. Haupt, R. Maya, A. Kazaz and M. Oren, Mdm2 promotes the rapid degradation of p53, *Nature*, 1997, **387**, 296–299.
- 166 P. Chène, Inhibiting the p53–MDM2 interaction: an important target for cancer therapy, *Nat. Rev. Cancer*, 2003, **3**, 102–109.
- 167 S. J. Duncan, S. Grüşchow, D. H. Williams, C. McNicholas, R. Purewal, M. Hajek, M. Gerlitz, S. Martin, S. K. Wrigley and M. Moore, Isolation and Structure Elucidation of Chlorofusin, a Novel p53-MDM2 Antagonist from a *Fusarium sp.*, *J. Am. Chem. Soc.*, 2002, **124**, 14503.
- 168 J. A. Robinson, β -Hairpin Peptidomimetics: Design, Structures and Biological Activities, *Acc. Chem. Res.*, 2008, **41**, 1278–1288.
- 169 R. Guha, in *In Silico Models for Drug Discovery*, ed. S. Kortagere, Humana Press, Totowa, NJ, 2013, pp. 81–94.
- 170 J. P. Malkinson, M. Zloh, M. Kadom, R. Errington, P. J. Smith and M. Searcey, Solid-Phase Synthesis of the Cyclic Peptide Portion of Chlorofusin, an Inhibitor of p53-MDM2 Interactions, *Org. Lett.*, 2003, **5**, 5051–5054.
- 171 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.
- 172 F. G. Glansdorp, R. J. Spandl, J. E. Swatton, O. Loiseleur, M. Welch and D. R. Spring, Using chemical probes to investigate the sub-inhibitory effects of azithromycin, *Org. Biomol. Chem.*, 2008, **6**, 4120–4124.
- 173 A. Proteau-Gagné, V. Bournival, K. Rochon, Y. L. Dory and L. Gendron, Exploring the Backbone of Enkephalins To Adjust Their Pharmacological Profile for the δ -Opioid Receptor, *ACS Chem. Neurosci.*, 2010, **1**, 757–769.
- 174 G. I. Koldobskii, V. A. Ostrovskii and B. V. Gidasov, Schmidt Reaction with Aldehydes and Carboxylic Acids, *Russ. Chem. Rev.*, 1978, **47**, 1084–1094.
- 175 W. C. Chan, B. W. Bycroft, D. J. Evans and P. D. White, A novel 4-aminobenzyl ester-based carboxy-protecting group for synthesis of atypical peptides by Fmoc-But solid-phase chemistry, *J. Chem. Soc., Chem. Commun.*, 1995, 2209–2210.
- 176 T. Conroy, K. A. Jolliffe and R. J. Payne, Efficient use of the Dmab protecting group: applications for the solid-phase synthesis of N-linked glycopeptides, *Org. Biomol. Chem.*, 2009, **7**, 2255–2258.
- 177 G. B. Fields and R. L. Noble, Solid phase peptide synthesis utilizing 9-fluorenylmethoxycarbonyl amino acids, *Int. J. Pept. Protein Res.*, 1990, **35**, 161–214.

- 178 C. Li, T. Li, X. Tian, W. An, Z. Wang, B. Han, H. Tao, J. Wang and X. Wang, Research progress on the PEGylation of therapeutic proteins and peptides (TPPs), *Front. Pharmacol.*, 2024, **15**, 1353626.
- 179 G. Castel, M. Chtéoui, B. Heyd and N. Tordo, Phage Display of Combinatorial Peptide Libraries: Application to Antiviral Research, *Molecules*, 2011, **16**, 3499–3518.
- 180 C. Pinilla, J. R. Appel, P. Blanc and R. A. Houghten, Rapid identification of high affinity peptide ligands using positional scanning synthetic peptide combinatorial libraries, *Biotechniques*, 1992, **13**, 901–905.
- 181 E. L. Schneider and C. S. Craik, in *Proteases and Cancer: Methods and Protocols*, eds. T. H. Bugge and T. M. Antalis, Humana Press, Totowa, NJ, 2009, pp. 59–78.
- 182 F. Merlino, S. Marzano, P. Zizza, F. D’Aria, N. Grasso, A. Carachino, S. Iachettini, A. Biroccio, S. Di Fonzo, P. Grieco, A. Randazzo, J. Amato and B. Pagano, Unlocking the potential of protein-derived peptides to target G-quadruplex DNA: from recognition to anticancer activity, *Nucleic Acids Res.*, 2024, **52**, 6748–6762.
- 183 R. Giraldo and D. Rhodes, The yeast telomere-binding protein RAP1 binds to and promotes the formation of DNA quadruplexes in telomeric DNA., *EMBO J.*, 1994, **13**, 2411–2420–2420.
- 184 L. Lototska, J. Yue, J. Li, M. Giraud-Panis, Z. Songyang, N. J. Royle, G. Liti, J. Ye, E. Gilson and A. Mendez-Bermudez, Human RAP1 specifically protects telomeres of senescent cells from DNA damage, *EMBO rep.*, 2020, **21**, 49076.
- 185 D. M. J. Lilley and M. F. White, The junction-resolving enzymes, *Nat. Rev. Mol. Cell Biol.*, 2001, **2**, 433–443.
- 186 K. Mizuuchi, B. Kemper, J. Hays and R. A. Weisberg, T4 endonuclease VII cleaves holliday structures, *Cell*, 1982, **29**, 357–365.
- 187 S. Kleff, B. Kemper and R. Sternglanz, Identification and characterization of yeast mutants and the gene for a cruciform cutting endonuclease., *EMBO J.*, 1992, **11**, 699–704.
- 188 A. Constantinou, A. A. Davies and S. C. West, Branch Migration and Holliday Junction Resolution Catalyzed by Activities from Mammalian Cells, *Cell*, 2001, **104**, 259–268.
- 189 J. M. Hadden, A.-C. Déclais, S. B. Carr, D. M. J. Lilley and S. E. V Phillips, The structural basis of Holliday junction resolution by T7 endonuclease I, *Nature*, 2007, **449**, 621–624.
- 190 J. M. Fogg, M. J. Schofield, M. F. White and D. M. J. Lilley, Sequence and Functional-Group Specificity for Cleavage of DNA Junctions by RuvC of *Escherichia coli*, *Biochemistry*, 1999, **38**, 11349–11358.
- 191 Y. Kobayashi, O. Misumi, M. Odahara, K. Ishibashi, M. Hirono, K. Hidaka, M. Endo, H. Sugiyama, H. Iwasaki, T. Kuroiwa, T. Shikanai and Y. Nishimura, Holliday junction resolvases mediate chloroplast nucleoid segregation, *Science (N.Y.)*, 2017, **356**, 631–634.

- 192 D. M. J. Lilley, Holliday junction-resolving enzymes: Structures and mechanisms, *FEBS Lett.*, 2017, **591**, 1073–1082.
- 193 H. Lin, D. Zhang, K. Zuo, C. Yuan, J. Li, M. Huang and Z. Lin, Structural basis of sequence-specific Holliday junction cleavage by MOC1, *Nat. Chem. Biol.*, 2019, **15**, 1241–1248.
- 194 J. Yan, S. Hong, Z. Guan, W. He, D. Zhang and P. Yin, Structural insights into sequence-dependent Holliday junction resolution by the chloroplast resolvase MOC1, *Nat. Commun.*, 2020, **11**, 1417.
- 195 S. R. K. Ainavarapu, J. Brujić, H. H. Huang, A. P. Wiita, H. Lu, L. Li, K. A. Walther, M. Carrion-Vazquez, H. Li and J. M. Fernandez, Contour Length and Refolding Rate of a Small Protein Controlled by Engineered Disulfide Bonds, *Biophys. J.*, 2007, **92**, 225–233.
- 196 R. Behrendt, S. Huber and P. White, Preventing aspartimide formation in Fmoc SPPS of Asp-Gly containing peptides — practical aspects of new trialkylcarbinol based protecting groups, *J. Pept. Sci.*, 2016, **22**, 92–97.
- 197 R. Subirós-Funosas, A. El-Faham and F. Albericio, Use of Oxyma as pH modulatory agent to be used in the prevention of base-driven side reactions and its effect on 2-chlorotriyl chloride resin, *Pept. Sci.*, 2012, **98**, 89–97.
- 198 P. Stathopoulos, S. Papas, S. Kostidis and V. Tsikaris, α - and β - aspartyl peptide ester formation via aspartimide ring opening, *J. Pept. Sci.*, 2005, **11**, 658–664.
- 199 G. Guidotti, L. Brambilla and D. Rossi, Cell-Penetrating Peptides: From Basic Research to Clinics, *Trends Pharmacol Sci.*, 2017, **38**, 406–424.
- 200 A. Patgiri, M. Z. Menzenski, A. B. Mahon and P. S. Arora, Solid-phase synthesis of short α -helices stabilized by the hydrogen bond surrogate approach, *Nat Protoc*, 2010, **5**, 1857–1865.
- 201 G. Leriche, G. Budin, L. Brino and A. Wagner, Optimization of the Azobenzene Scaffold for Reductive Cleavage by Dithionite; Development of an Azobenzene Cleavable Linker for Proteomic Applications, *European J Org Chem*, 2010, **2010**, 4360–4364.
- 202 A. M. Beekman, M. A. O’Connell and L. A. Howell, Identification of Small-Molecule Inhibitors of the Antiapoptotic Protein Myeloid Cell Leukaemia-1 (Mcl-1), *ChemMedChem*, 2016, **11**, 840–844.