Targeting PCNA with Cell and Nuclear Permeable p21-derived Peptides

Submitted in partial fulfilment of the degree

Master of Philosophy (Chemical Sciences)

School of Physical Sciences

Department of Chemistry

Theresa Chav *(B.Sc.)*

April 2021

Supervisors: Professor Andrew Abell

& Dr John Bruning

Table of Contents

Abstract

The Proliferating Cell Nuclear Antigen (PCNA) is a sliding clamp protein essential for DNA replication and repair and is upregulated in a large number of cancers. This work centres on targeting PCNA with peptides derived from a segment of the p21$^{WAF/CIP1}$ protein, termed **p21$_{139\text{-}160}$** (^{139}GRKRRQTSMTDFYHSKRRLIFS^{160}), that is known to inhibit PCNA. A potential therapeutic must be cell permeable and translocate to the nucleus to reach PCNA, thus defining the aims of Chapter Two and Three. Furthermore, the **p21$_{139\text{-}160}$** sequence was used to develop a peptide-based fluorescent sensor for PCNA which is discussed in Chapter Four.

Previous work in our lab identified the truncated p21-derived peptide, p21$_{143\text{-}155}$ (^{143}RQTSMTDFYHSK^{155}), termed **P1** as a lead inhibitor of PCNA due to its short length and high affinity for PCNA. However, **P1** was found to be impermeable to breast cancer cells (T47D). In Chapter Two, peptides **P2** (^{140}RKRRQTSMTDFYHSK^{155}), **P3** (^{143}RQTSMTDFYHSKRR^{157}) and **P4** (^{140}RKRRQTSMTDFYHSKRR^{157}) with additional residues from the longer and cell-permeable **p21$_{139\text{-}160}$** were tagged with fluorescein and administered to breast cancer cells to determine if the added residues facilitate cell permeability. This revealed modest cell permeability of **P2** and **P4**, whereas **P3** showed no cell entry. **P4** displayed the most intracellular accumulation, which indicated extension of the **P1** sequence at both termini facilitated cell permeability.

Chapter Three presents studies on conjugating Nuclear Localisation Sequence (NLS) peptides; Tat$_{48\text{-}57}$ (**N1F**), SV40$_{126\text{-}132}$ (**N2F**), cMyc$_{320\text{-}328}$, (**N3F**) and R6W3 (**N4F**) to **P1** to provide a series of linear peptide conjugates termed **P1b-N1F**, **P1b-N2F**, **P1b-N3F** and **P1b-N4F**, respectively. **P1b-N1F** and **P1b-N3F** displayed modest cell permeability to breast cancer cells. **P1b-N2F** displayed cell and nuclear permeability, which suggests the **N2F** imparted both cell and nuclear entry. The NLS peptides were also conjugated to a

macrocyclic bimane analogue of **P1**, **P1c**. This gave a series of macrocyclic bimane peptide conjugates **P1c-N1F**, **P1c-N2F**, **P1c-N3F** and **P1c-N4F**. The SV40$_{126-132}$ tagged macrocyclic peptide, **P1c-N2F**, showed nuclear entry. Additionally, the control peptide, **P1bimF**, which contains a bimane linker and fluorescein tag but no NLS peptide, was also nuclear permeable. In contrast, analogues of **P1bimF** and **P1c-N2F**, without the fluorescein tag (**P1bim** and **P1c-N2)**, were only cell permeable, highlighting the effect the fluorescein tag has in altering nuclear uptake, in this instance. This work presents **P1b-N2F**, **P1bimF** and **P1c-N2F** as three nuclear permeable peptide leads towards the development of a viable pre-clinical anti-cancer therapeutic that targets PCNA.

Chapter Four details the development of a p21-derived fluorescent sensor for PCNA. The solvatochromic fluorophore, 4-dimethylaminophthalimide (4-DMAP) was introduced at positions 147, 150 or 151 in a p21$_{141-155}$ (^{141}KRRQTSMTDFYHSKR^{155}) scaffold to provide three fluorescent sensor peptides termed **B1**, **B2** and **B3**, respectively. The 151-substituted peptide, **B3**, exhibited the largest fluorescence response in the presence of PCNA, with a 7.9-fold change. The binding affinity of all peptides for PCNA were determined by Surface Plasmon Reasonance (SPR) with only **B3** binding specifically to PCNA with a K_D of 1.28 μM. **B1** and **B2** largely interacted non-specifically with PCNA, suggesting insertion of the 4-DMAP fluorophore at 147 or 150 disrupts PCNA binding. This work demonstrates that incorporation of a solvatochromic fluorophore is most favourable at position 151 for a p21$_{141-155}$ scaffold, which facilitates effective PCNA binding and a resultant 'turn on' fluorescence response. **B3** presents a promising lead for further development of a fluorescent PCNA sensor to measure levels of cell proliferation.

Publications

Publications from this work

Title of paper:	A turn-on fluorescent sensor for PCNA
Publication Status:	Published
Publication Details:	(Ref 89) Horsfall, A. J.; Chav, T.; Bruning, J. B.; Abell, A. D., A turn-on fluorescent PCNA sensor. *Bioorg Med Chem Lett* **2021,** *41*, 128031

Title of paper:	A nuclear permeable peptidomimetic to inhibit human PCNA
Publication Status:	manuscript in preparation (see Appendix 4: Publications)
Publication Details:	A. J. Horsfall, T. Chav, Z. Kikhtyak, J. L. Pederick, W. Kowalczyk, D. B. Scanlon, W. D. Tilley, T. E. Hickey, A. D. Abell and J. B. Bruning, Nature Chemical Biology 2021

Other publications contributed to

Title of paper:	Unlocking the PIP-box: Understanding factors for high affinity binding to human PCNA
Contribution to paper:	SPR assays, discussed results and edited manuscript.
Publication Status:	Published
Publication Details:	(Ref 90) Horsfall, A. J.; Vandborg, B. A.; Kowalczyk, W.; Chav, T.; Scanlon, D. B.; Abell, A. D.; Bruning, J. B., Unlocking the PIP-box: A peptide library reveals interactions that drive high-affinity binding to human PCNA. *J Biol Chem* **2021,** *296*, 100773.

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint award of this degree.

I give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Theresa Chav

Acknowledgements

This work was funded with help from the Centre for Nanoscale BioPhotonics (CNBP). To CNBP and their members, I thank you for the opportunities and knowledge provided to me in the last two years. I also acknowledge the Australian Nanoscale Fabrication Facility for the ongoing funding they provide toward our analytical instruments and the Institute for Photonics and Advanced Sensing for their facilities.

First and foremost, I want to thank my supervisor Prof. Andrew Abell for his continuous encouragement and guidance over the last two years. I would not have been able to complete this degree without your support and assistance that truly went above and beyond what was called of you. Thank you to Dr John Burning for being a fantastic co-supervisor as well as Denis Scanlon for the endless amounts of knowledge offered to me.

I owe a huge, huge thanks to Aimee Horsfall for teaching me the ropes and for the endless support over the last couple years. I have learnt so much under your guidance and you continue to be an inspiration to me. To Pat Capon, thank you for the amount of help you have offered me throughout my degree. I don't know what I would have done without the wealth of advice and therapy chats both Pat and Aimee provided.

Whilst my sanity is currently *under review*, I owe a lot of said sanity to my labmates. To Kathryn Palasis, thank you for the never-ending conversations about pop culture, books and **boy/s**. To Nic Schumann, the connoisseur in synthesis, fitness and coffee, thank you for being a laugh. Special thanks to the Siggy gang for providing endless opportunities to waste time and have pretty good coffee with a side of spicy chats. To the rest of the Abell group, thank you for all the help and support I received, it was a pleasure being in the group.

To Lachie, you're a gem. Thank you for being a lovely shoulder to cry on, I appreciate you. To my friends, thank you for the support and the distractions. Also, big thanks for being so **crazy**. Ryan, Clare and Anna, especially. You guys are the best.

Mum and Dad, there aren't enough words to express my gratitude for the both of you. Thank you for the sacrifices you've made and the hard work you've put it to get me where I am today, I hope I make you proud.

The past two years have been incredibly challenging but also extremely rewarding. As a result, I've learnt so much, about science and myself. This experience will undeniably continue to shape who I am for years to come. Above all else, I feel ready.

Abbreviations

4-DAPA: 4-dimethylaminothalimidoalanine; **4-DMAP**: 4-dimethylaminophthalimide; **4-DMN**: 4-*N,N*-dimethylamino-1,8-naphthalimide; **4-DMNA**: 4-*N,N*-dimethylamino-1,8-naphthalimidoalanine; **ACN**: Acetonitrile; **Alloc**: Allyloxycarbonyl; **Boc**: Tert-butyloxycarbonyl; **CPP**: Cell Penentrating Peptide; **DCM**: Dichloromethane; **DIPEA**: *N,N*-Diisopropylethylamine; **DMF**: *N,N'* -Dimethylformamide; **DODT**: 2,2′-(Ethylenedioxy)diethanethiol; **ESI**: Electrospray Ionisation; **FITC**: Fluorescein isothiocyanate; **Fmoc**: 9-Fluorenylmethoxycarbonyl; **HEPES**: 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; **HOBt**: 1-Hydroxybenzotriazole hydrate; **HRMS**: High Resolution Mass Spectrometry; **Mmt**: 4-Methoxytrityl; **MS**: Mass Spectrometry; **NLS**: Nuclear Localisation Sequence; **NMR**: Nuclear Magnetic Resonance; **Pbf**: 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl; **PCNA**: Proliferating Cell Nuclear Antigen; **PPI**: Protein-Protein Interactions; **RP-HPLC**: Reverse-Phase High-Performance Liquid Chromatography; **SPPS**: Solid Phase Peptide Synthesis; **SPR**: Surface Plasmon Resonance; **tBu**: Tert-butyl; **TFA**: Trifluoroacetic acid; **TIPS**: Triisopropylsilane; **TNBS**: 2,4,6-Trinitrobenzenesulfonic acid

Chapter 1 Introduction

1.1   Inhibiting Protein-Protein Interactions

Protein-protein interactions (PPIs) are central mediators in biological processes and many human diseases are a result of abnormal PPIs.[1, 2] Targeting PPIs with traditional small molecule drugs is difficult due to the large binding surfaces involved in PPIs (~1500–3000 $Å^2$).[3, 4] In addition, these surfaces are typically flat and featureless and lack well-defined grooves or pockets for small molecule binding.[2, 5] Such small molecule inhibitors are likely to be displaced from the interface by native protein binding partners in the event of disease-related protein abundance.[4, 6] Furthermore, single mutations on the protein interaction site can eliminate binding of these drugs. Small molecule PPI inhibitors also inherently lack specificity as a result of the small protein-drug interface.[7]

1.2   Peptide therapeutics and their synthesis

Peptides present as ideal candidates for PPI inhibition (see Figure 1). Therapeutic peptides are often <50 residues long and have a MW of 500-5000 Da, offering an ability to target larger sites and achieve higher specificity and selectivity than small molecule inhibitors.[3, 8] These peptides can be designed from known sequences of native proteins involved in PPIs with incorporation of key residues. Additionally, peptide sequences can be easily accessed synthetically and biologically. Fmoc/tBu-Solid Phase Peptide Synthesis (Fmoc/tBu-SPPS) is primarily used for preparation of synthetic peptides because of its orthogonal protection strategy and milder reaction conditions relative to Boc-SPPS. Since the seminal conception of Fmoc/tBu-SPPS, there have been considerable refinements in protecting groups, polymeric solid supports, linkers and activation techniques, however, the synthesis of long and difficult peptides remains a major consideration. [9, 10]   The tendency of some peptide sequences to aggregate and form secondary structures such as β-sheets is well documented to result in failed peptide synthesis, low yields and peptide heterogeneity. [11, 12] Certain amino acids also

exhibit reactivities leading to different side reactions. Of particular interest is reactivity of aspartic acid residues.
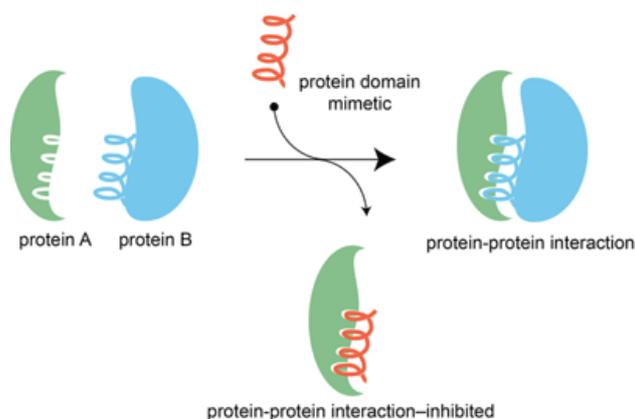
Figure 1: Protein-protein interactions can be inhibited by peptides derived from a native protein. Figure adapted[13].

The first peptide drug was insulin, isolated from canine and bovine pancreas, for diabetes treatment.[14, 15] Around 80 peptide drugs are available in the global market at present.[14] The most notable of which include insulin and analogues for diabetes treatment; Leuprolide and Octreotide for cancer treatment and Teriparatide for Osteoporosis.[14]

1.2.1 Aspartimide formation

Aspartimide formation is a well-documented side reaction of aspartic acid in Fmoc/tBu-SPPS. The sequence-dependent side reaction features in peptides containing aspartic acid, particularly those with glycine, asparagine, serine or alanine as the neighbouring residue (i.e. Asp-Gly, Asp-Asn, Asp-Ser or Asp-Ala), with Asp-Gly being the most susceptible.[16]
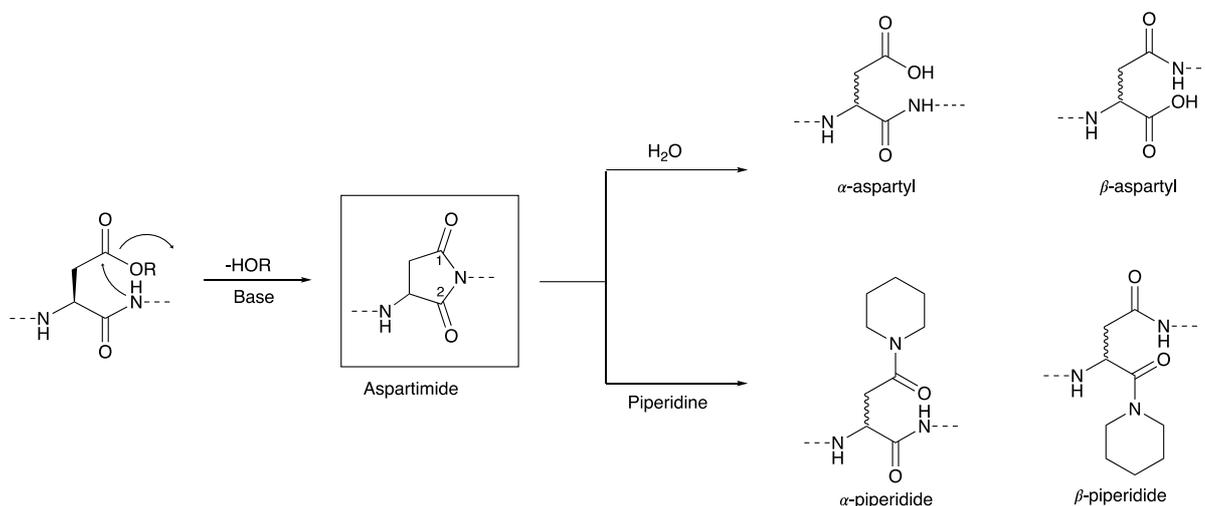
Figure 2: Base-catalysed aspartimide formation. Nucleophilic attack of the aspartimide by water and piperidine results in epimerised α-and β-peptides. Attack at the C1 leads to α-peptides. Attack at C2 leads to β-peptides. R= protecting groups, such as tBu. Figure adapted. [17]

Base-catalysed aspartimide formation occurs during 9-Fluorenylmethoxycarbonyl (Fmoc) deprotection with piperidine (see Figure 2). The ester side chain of the aspartic acid undergoes a cyclization to form the aspartimide structure. Aspartimide formation can give rise to epimerisation via nucleophilic attack and ring-opening of the imide ring. Piperidine can react with the aspartimide ring resulting in α-and β-piperidides. Hydrolysis can lead back to the target peptide, α-aspartyl or to the undesired β-aspartyl peptide, depending on which carbon is nucleophilically attacked. The β-aspartyl is difficult to identify due to its identical mass with the desired peptide and frequently co-elutes with the desired peptide on HPLC. [17-19] In order to decrease the potential formation of such inseparable epimerized side products, our efforts turned to minimising aspartimide formation through optimising synthetic conditions.

Three main strategies have been developed to decrease aspartimide formation, one of which includes maximising the steric hindrance of the ester side chain protecting group of aspartic acid.[17] Such protecting groups from least to most bulky include: Odie, OMpe, OtBu, O-1-adamantyl, trityl-based, OBzl, OAll and O-phenacyl.[9] The highly hydrophobic nature of these aspartic ester moieties, however, lead to low coupling yields.[17]

17
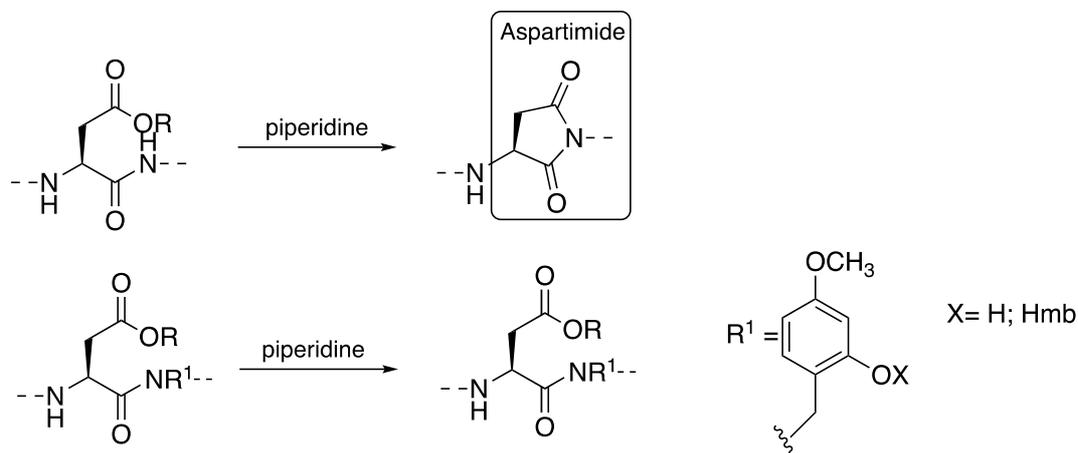
Figure 3: Amide backbone protection with Hmb amino acids that can prevent side reactions such as aspartimide formation.

Another approach, amide backbone protection, involves reversible alkylation of peptide bonds, in this case, the Asp-X bond (where X=amino acid), to prevent attack of the ester carbonyl. This is achieved using secondary amino acid surrogates such as N-alkylamino acids and proline analogues.[11] The 2-hydroxyl-4-methoxybenzyl (Hmb) amino acids developed in Sheppard's group demonstrates the former (see Figure 3) while Mutter's pseudoproline dipeptides, the latter.[20, 21]

While Hmb protection is widely used and can completely eliminate aspartimide in Asp-Gly containing peptides, it presents certain limitations.[18] Coupling of the subsequent residue after Hmb incorporation is challenging, despite being aided by an intramolecular O to N acyl transfer.[21] [10]The low coupling efficiencies necessitates the use of Hmb dipeptides instead, of which only Fmoc-Asp(tBu)-(Hmb)Gly is commercially available.[17, 19] Mutter's pseudoproline dipeptides contain a cysteine or serine/threonine residue that is reversibly protected in a proline-like thiazolidine or oxazolidine, respectively. The native sequence is regenerated upon acid cleavage/deprotection. Both backbone protecting approaches also prevent secondary structure formation during synthesis, reducing in decrease aggregation and minimized side products.

Addition of acidic modifiers is the last of the strategies to suppress aspartimide formation. These include hydroxybenzotriazole (HOBt) and ethylcyano(hydroxyamino)acetate (Oxyma) Pure addition to the deprotection solution.[9, 17, 19] These work by protonating the amide nitrogen adjacent to the aspartic acid residue, decreasing its nucleophilicity and minimising attack at the carbonyl group of aspartic acid.

1.3   PCNA

Proliferating Cell Nuclear Antigen (PCNA), first discovered in 1978, is critical for proliferation of cells. [22] PCNA functions in a number of nuclear processes such as cell cycle control and DNA-replication and -repair. [23] It facilitates these processes by encircling DNA to act as a mobile docking platform, providing access for requisite proteins to bind and interact with DNA (see Figure 4). [24] Due to its central role in DNA-replication and its upregulation in almost all cancer cell lines, PCNA is an attractive inhibition target for development of broad-spectrum cancer therapeutics. PCNA inhibition has shown to be cytotoxic to proliferative cells. [23, 25]

PCNA is a ring-shaped trimeric protein, with each monomer containing two similar domains consisting of two α-helices and nine β-strands (see Figure 4). [24, 26] These two domains are connected by the inter-domain connecting loop which is a strand lacking defined secondary structure. [27] Located near this loop is the binding site for many of the PCNA protein binding partners.
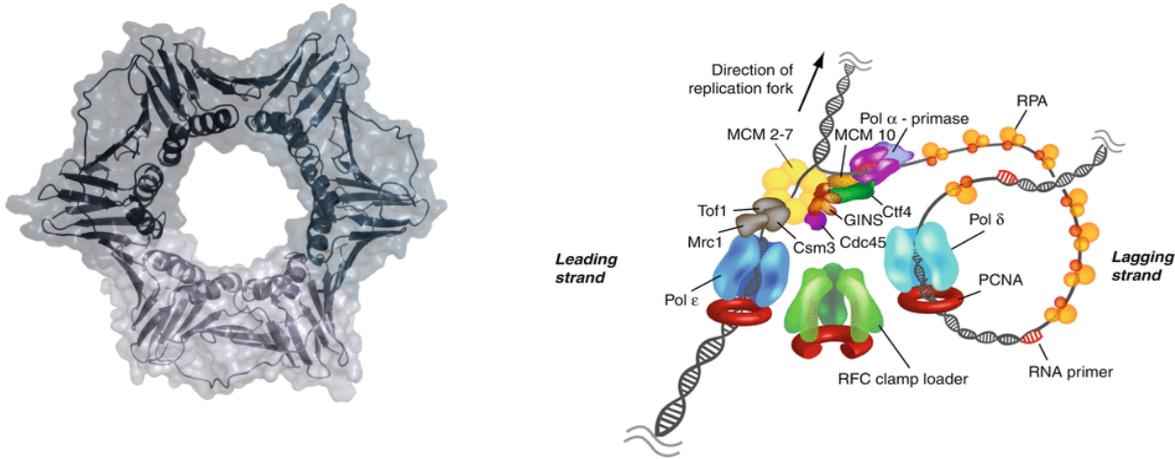
Figure 4: Structure of PCNA (left)(1AXC). PCNA wraps around DNA and facilitates DNA-replication (right, figure adapted[28]) and -repair.

1.3.1 PCNA and p21

PCNA binding partners bind via a conserved motif termed the PCNA-Interacting Protein (PIP) box. [25] Of these proteins, the cell cycle control protein p21[WAF/CIP1] binds PCNA with the highest known affinity and inhibits PCNA-dependent DNA-replication by binding to the PIP-box of PCNA, preventing access to proteins involved in replication. [25, 29] p21[WAF/CIP1] thus presents a structural template for therapeutically relevant peptide inhibitors. p21-peptide derivatives have been shown to maintain the properties of the parent protein, inhibiting malignant cell proliferation selectively by preventing access to the PIP-box motif of PCNA. [25] The PIP-box is defined as Qxxφxxψψ (Q= glutamine, x= any amino acid, φ= hydrophobic residue, ψ= aromatic residue). [30] All PIP-peptides/proteins form a characteristic 3_{10}-helix upon binding and inserts the φ and ψ residues into a hydrophobic pocket of PCNA. [26]

A 22-residue peptide derived from the p21[WAF/CIP1] protein, [139]GRKRRQTSMTDFYHSKRRLIFS[160], was found to bind PCNA and a co-crystal structure was elucidated in 1996 (see Figure 5). [27] This peptide binds with the same affinity as the parent protein and disrupted PCNA-dependent DNA-replication upon binding. [24] This 22mer, herein termed **p21$_{139\text{-}160}$**, is cell permeable, however lacks nuclear localisation capability.

20

Previously (unpublished), the **p21$_{139-160}$** was shortened to a 12mer (^{143}RQTSMTDFYHSK^{154}) which maintained a low nanomolar binding affinity with PCNA but was not cell permeable. Here, the peptide sequence was optimised to achieve permeability while maintaining the shortest length.

p21 PIP- Box
^{144}QTSMTDFY^{151}
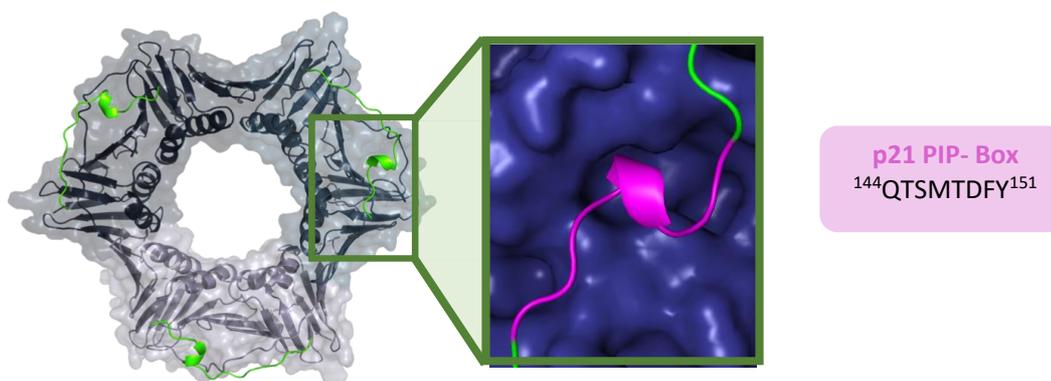
Figure 5: Co-crystal structure of PCNA and bound p21$_{139-160}$ (green) (1AXC). p21$_{139-160}$ binds PCNA via its PIP-box sequence ^{144}QTSMTDFY^{151}.

1.4 Peptide drug permeability

1.4.1 Cell permeability

For therapeutics with an intracellular target, the cell membrane is the first obstacle to overcome. However, peptide therapeutics typically exhibit low cellular permeability. The cellular membrane is a semi-permeable barrier which is integral to the survival and function of cells.[31] It is a highly dynamic structure composed of both lipids and membrane proteins. The membrane forms a bilayer with 3 main lipid classes – glycerophospholipids, sphingolipids and cholesterol.[32] Due to the lipophilic nature of plasma membranes, intracellular entry of compounds is restricted. Small, highly lipophilic molecules are able to traverse the membrane via passive diffusion.[31] Larger, more polar molecules – such as sugars, peptides or amino acids – require active transport and the use of membrane transporters.[32, 33]

Molecular permeation through lipid membranes is an essential biological process and is also vital for intracellular drug delivery. The selective nature of biological membranes thus presents a key challenge for transport of drugs to intracellular targets – whether that be in the cytosol or other organelles.[34] Additionally drug permeability is of critical importance to bioavailability of drug candidates, further demonstrating the necessity of tools that enhance or facilitate cell permeability.

Several approaches have been investigated to facilitate cellular uptake of therapeutics. Chemically altering the structure and functional groups of small molecules has been used extensively to achieve moderate lipophilicity which facilitates passive permeation. Increasing lipophilicity allows interaction with the hydrophobic lipid region of plasma membranes, a concept outlined in Lipinski's 'Rule of 5.'[32] This approach, however, can affect the specificity of the drug by altering the groups and size of the molecule which is especially incompatible with peptide and protein therapeutics.

The interaction of potential therapeutics with proteins located on plasma membranes has also been studied as a means to drug internalisation. Drugs can be developed to target membrane proteins such as transporters or receptors, facilitating transporter-mediated transcytosis and receptor-mediated endo/transcytosis, respectively.[35] While such an approach has advantages of high specificity, it does limit the development of broad spectrum therapeutics and the drug can be made ineffective by the ability of diseases to change receptor binding, levels of binding substrates or receptor sensitivity.[36] Recently, the use of peptides as drug transporters has gained wide attention. Peptides present high internalization efficiency, specificity, low cytotoxicity and compatibility with many drug cargoes.[31, 32]

Cell-penetrating peptides (CPPs) or protein transduction domains are short peptides (5-30 amino acids) that can translocate through the cell membrane. The use of CPPs as carriers for

intracellular drug delivery has been extensively researched in recent decades. CPP-cargo conjugates are able to traverse the cell membrane. The cargo molecules – which are covalently or non-covalently attached – range from drugs, proteins, antibodies, peptides, nanoparticles, oligonucleotides, amongst others.[37] Generally, CPPs exist in 3 classes – cationic, amphipathic or hydrophobic with the large majority carrying a net positive charge.[38] The positive charge facilitates CPP electrostatic interaction with negatively charged cell surfaces.[37]

Arginine residues are commonly found in cationic CPPs and have been found to aid in internalization.[37] The positively charged residue interacts electrostatically with negatively charged structures of the cell membrane such as heparan sulfate proteoglycans, phosphatidylserines and sialoglycolipids.[37] Arginine, particularly, is able to bind to negatively charged molecules such as phosphates in a bidentate fashion which induces membrane curvature, leading to internalization.[37]

The exact mechanisms by which CPPs promote cell internalisation remain unknown.[39] Studies on the mechanism of CPP action indicate both endocytosis (energy dependent) and direct translocation (energy independent) occur (see Figure 6).[37] The mechanism used by a CPP will depend heavily on the nature and size of the CPP and its cargo.[40]
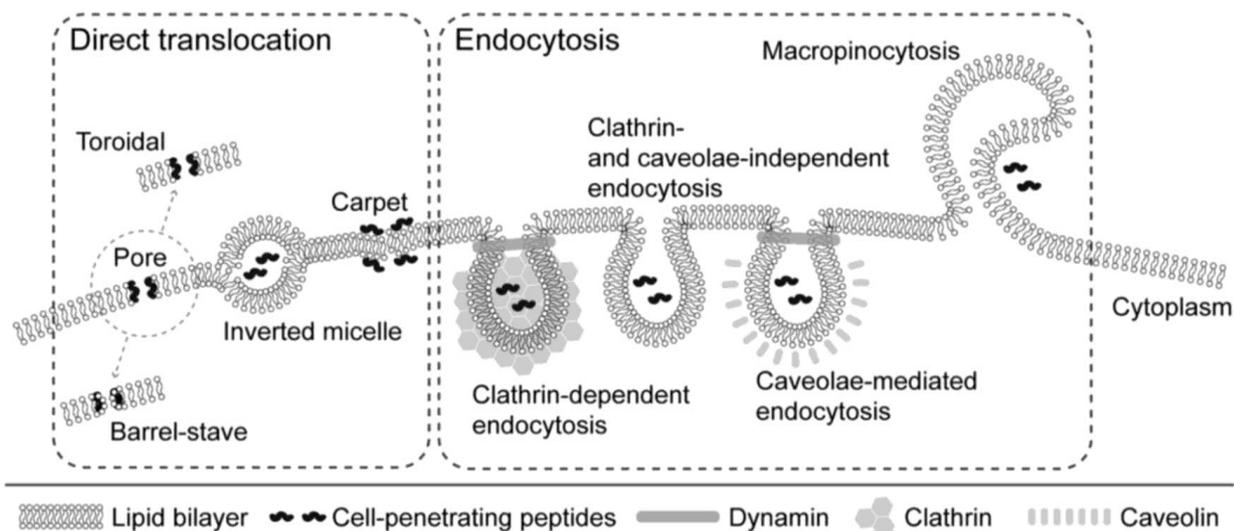
Figure 6: CPP mechanisms of cell entry. These mechanisms can follow direct translocation or endocytosis pathways. (See Ref [41])

Whilst the mechanism of action of CPPs are not fully understood, their utility and potential remains significant as evidenced by the CPP-drug systems that are currently being evaluated in clinical trials.[37]

1.4.2 Nuclear permeability

To affect its therapeutic activity, a nuclear targeted drug needs to permeate the cell membrane, translocate to the nucleus and pass through the nuclear membrane, thus presenting another barrier.[42] The nucleus is the control centre of eukaryotic cells, storing both genetic information and transcription machinery. Separating the nucleus from the cytoplasm is the nuclear envelope, a phospholipid bilayer. Like the plasma membrane, the nuclear envelope presents another selective barrier for molecular entry. Nuclear pore complexes –protein complexes embedded in the nuclear membrane – are the site of transport into the nucleus.[43] Due to the small size of nuclear pore complexes, molecules larger than 5 nm cannot diffuse through.[43] Larger molecules, such as proteins, require specific targeting signals, termed Nuclear Localisation Sequences (NLSs) which are short peptide sequences that facilitate nuclear entry of their respective protein.[44] NLS peptides are a type of cationic CPP and are generally short in length (less than 12 residues) that are rich in basic amino acids.[45] [38]

Numerous strategies have been studied to overcome nuclear permeability, some of which include exploiting these NLSs. NLS peptides have been shown to be effective as drug delivery tools.[46]

Nuclear transport systems were first proposed when the sequence required for nuclear localisation of the simian virus 40 large T antigen protein was elucidated (^{126}PKKKRRV^{132}) and termed the $SV40_{126-132}$ NLS peptide. $SV40_{136-132}$ is now considered the prototypical monopartite NLS sequence which is defined a single stretch of basic amino acids. $cMyc_{320-328}$ (^{320}PAAKRVKLD^{328}) is originally derived from the human cMyc oncoprotein and is also a classical monopartite NLS, despite the sequence containing few basic residues. [47] The difference between the $cMyc_{320-328}$ and $SV40_{136-132}$ sequences highlights the diversity of NLS sequences. The lysine residue in the fourth position of both $cMyc_{320-328}$ and $SV40_{136-132}$ were found to be critical for nuclear localisation.[44]

NLS peptides derived from viral proteins are one of the most well-known and studied classes of NLS sequences.[48] This includes the Tat_{48-57} sequence (^{48}GRKKRRQRRR^{57}) derived from the Tat (Transcription-Activating) protein from the highly infectious virus, HIV-1. Tat_{48-57} was found to facilitate its cell and nuclear entry. Tat_{48-57} is also a widely used CPP of the cationic class. Another NLS of interest is R6W3 (RRWWRRWRR), a synthetic analogue of penetratin (^{43}RQIKIWFGNRRMKWKK^{58}), a well-known CPP. [45] The penetratin CPP is derived from a *Drosophila* Antennapedia homeodomain protein and was found to cross the cell membrane and translocate to the nucleus. [49] R6W3 mimics the structure of the original penetratin CPP, forming an amphipathic, helical secondary structure, which exemplifies all 3 CPP classes – cationic, amphipathic and hydrophobic. [37, 49]. R6W3 forms an amphipathic α-helix, a secondary structural motif that may facilitate permeability.[50] Amphipathic α-helices have two 'faces', one of which is hydrophobic while the other is cationic, anionic or polar. Amphipathic CPPs also tend to accumulate in the nucleus.[51]

1.4.3 Nuclear Import Mechanism

A heterodimeric transport carrier complex composed of importin α and importin β recognises NLSs (see Figure 7). When present, importin β binds to importin α, which allows NLS recognition and subsequent binding.[44, 46] The β-subunit then directs the importin(s)-cargo through the nuclear pore via specific interactions between importin β and nucleoporins which make up the nuclear pore complexes.

Ran (a GTPase) is primarily bound, in the nucleus, to GTP which binds to importin β, triggering release of the importin α– NLS complex cargo into the nucleoplasm. Without importin β bound, the interaction between importin α and NLS cargo is displaced, facilitating the release of the NLS cargo.[44, 46]
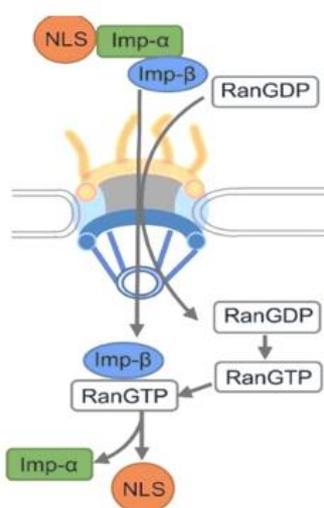
**Figure 7:** Nuclear import mechanism. Figure adapted[52]

Here, NLS attachment to p21-peptides is investigated. The NLS peptides chosen are derived from SV40, cMyc and Tat proteins as well as the synthetic mimetic of penetratin, R6W3.[40] p21-peptides have successfully been directed to the nucleus previously, thus this work will attempt to refine the sequence to ensure cell and nuclear permeability with minimum length

and maximum binding affinity.[25] Whilst other therapeutic peptides are undergoing clinical trials or are on market, no PIP-box inhibitors have reached clinical trials stages.[25]

1.5 Protein fluorescent sensors

Changes in protein levels, structure or function are often indicative of pathological abnormalities and thus, can be used as diagnostic and prognostic biomarkers.[53] Proteins have been used as disease biomarkers for over 150 years.[54]

Many techniques have been developed to investigate biological processes including protein dynamics.[55] However, methods such as analytical ultracentrifugation and photoaffinity labelling lack spatial and/or temporal details of the interactions and are often cell destructive.[56] This necessitates development of tools that can analyse the dynamic protein-protein interactions in real time within living cells.[56] One such example is the application of a biosensor, an analytical tool containing a bio-recognition element, which generates a signal in response to a stimulus. Of particular interest is fluorescent-based sensors that display environment-dependent fluorescent properties. [57]

1.5.1 Sensing PCNA

PCNA levels are a marker for cell proliferation which has implications in cancers.[58] Overexpression of PCNA has been described as a reliable biomarker for many tumour types. The potential use of PCNA as a marker for diseases has been documented for cellular rejection in renal allografts and various cancers including parathyroid carcinoma.[59, 60] The structure of PCNA is also well documented as discussed earlier and thus presents an attractive structural template for development of a protein biosensor to indicate cell proliferation levels. Selectively sensing large biomolecules such as proteins, however, can be difficult as a high number of interactions are needed between the sensing probe and the large biomolecule to

elicit a response. Peptides can be easily modified to introduce fluorophores into their structure while still maintaining interactions and selectivity for their binding target. [57, 61]

1.5.2 Solvatochromic fluorophores

Incorporation of solvatochromic fluorophores into sensing probes has received considerable attention for imaging. [62, 63] Solvatochromic fluorophores are a type of environment-sensitive fluorophore which have emission properties that depend heavily on the local solvent microenvironment.[64] These can be sensitive to a range of changes including solvent relaxation rates, rigidity of local environment and solvent polarity.[56] Particularly useful are the latter species which typically show low fluorescence in polar protic solvent environments such as water (e.g. cytosol) while exhibiting high fluorescence in hydrophobic environments such as protein hydrophobic pockets or membranes (see Figure 8).[62] This change in emission intensity is generally coupled with a shift in maximum emission wavelength.[65] Such probes have been successfully adapted for incorporation into biopolymers of interest including peptides, proteins or DNA.[64] An example of this is the incorporation of novel amino acids that contain solvatochromic fluorophore side chains into peptides. Introducing a solvatochromic amino acid into a peptide which selectively binds a protein, gives rise to peptide which produces a fluorescence signal upon binding, acting as a fluorescent sensor for the target protein.
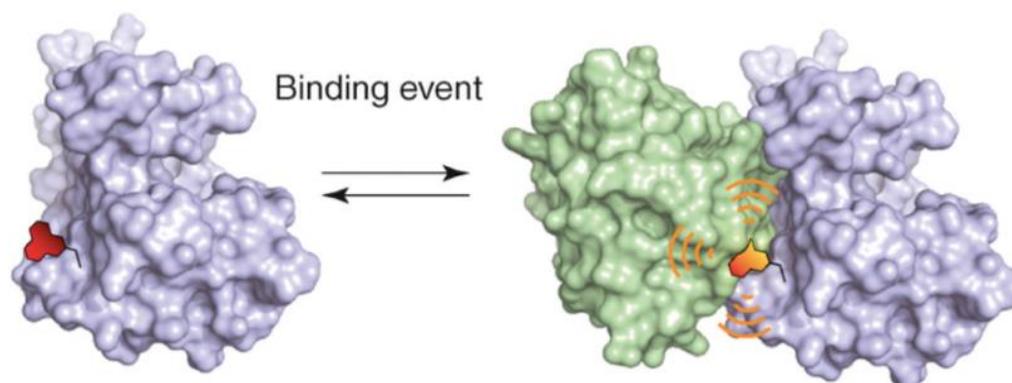
Figure 8: Solvatochromic fluorophores can be used to report protein-protein interactions by sensing changes in solvent polarity. Binding to a hydrophobic pocket will cause a shift in fluorescence intensity or emission wavelength. Figure adapted [62]

Whilst the use of PRODAN (2-propionyl-6-dimethylaminonaphthalene), dansyl (5-dimethylamino-1-naphtha-lenesulfonyl) and NBD (7-nitrobenz-2-oxa-1,3-diazole) as solvatochromic fluorophores is widespread, they are associated with limitations (see Figure 9). NBD has relatively low sensitivity and PRODAN and dansyl derivatives have poor signal to noise and on/off fluorescence ratios.[65] Fluorophores of the dimethylamino-phthalimide and -naphthalimide family hold certain advantages over these more widespread fluorophores, most significantly, is high sensitivity.[56] They exhibit exceptionally low fluorescence in water relative to those in apolar solvents, creating a greater signal-to-noise ratio.[56] The relatively small size of these fluorophores is comparable to tryptophan, offering easy incorporation into the primary sequence of peptides or protein.[56] Thus, they are of particular interest for investigating peptide-or protein-protein interactions.
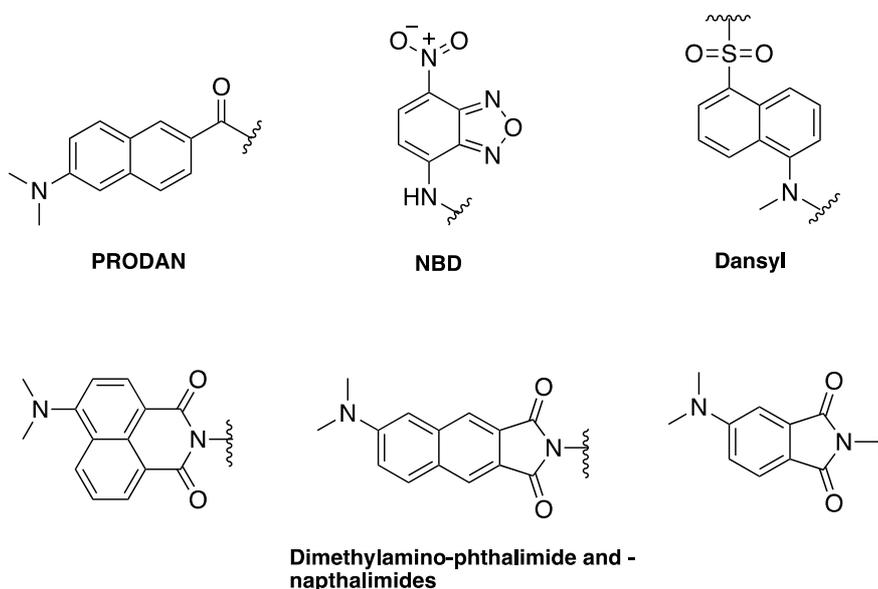
Figure 9: Widely used solvatochromic fluorophores used in peptide/protein studies. Figure adapted [62].

1.5.3 p21-derived peptides as a PCNA sensor

p21[139-160] binds PCNA with a 3_{10}-helix defined by the Met147, Phe150 and Tyr151 residues. Upon binding, these residues are inserted into a hydrophobic pocket on the surface of PCNA and thus, are ideal residues to substitute for a solvatochromic amino acid (see Figure 10).[24, 25] This provides an opportunity to sense PCNA levels through p21-PCNA binding events. Upon binding, the solvatochromic fluorophore is inserted into the PCNA cleft, resulting in a change in fluorescence. Previously in our lab (Aimee Horsfall[64]), three peptides were prepared based on the p21-derived sequence [141]KRRQTSMTDFYHSKR[155] (**P5**), where 4-*N,N*-dimethylamino-1,8-naphthalimidoalanine (4-DMNA) was separately inserted into the 147, 150 and151 positions (see Figure 11). These three peptides showed increased fluorescence upon binding PCNA, however, they bound with significantly lower affinity than the native **P5** peptide. The 4-DMNA fluorescent side chain is relatively large in comparison to the Met147, Phe150 and Tyr151 residues of the native peptide. This likely led to a decrease of binding in the hydrophobic PCNA pocket.
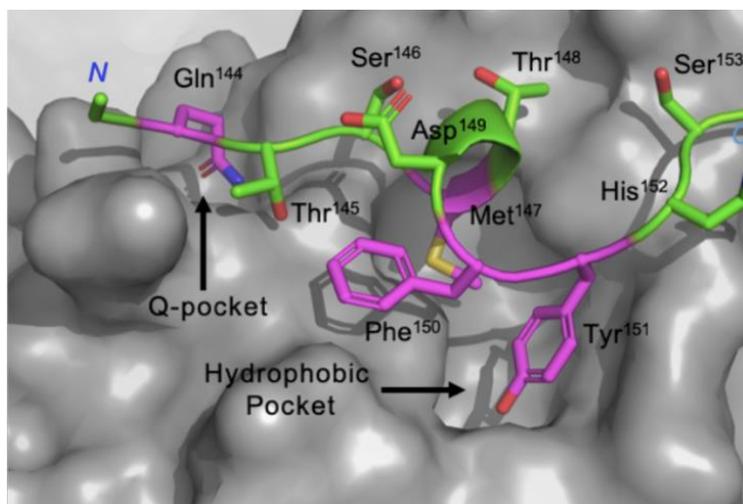
Figure 10: p21 (144-153): QTSMTDFYHS bound too PCNA (grey, 1AXC). M147, F150 and Y151 insert into the hydrophobic pocket of PCNA. Figure adapted [25].

Here, the smaller 4-dimethylaminophthalimidoalanine (4-DAPA) solvatochromic amino acid is incorporated at the same M147, F150, Y151 positions of a p21 derivative (see Figure 11). The smaller size of the fluorescent side chain increases binding affinity by allowing improved insertion into the hydrophobic PCNA pocket. The fluorescence response of these biosensor peptides was investigated as well as their binding affinity for PCNA in an effort to develop a turn-on fluorescent PCNA sensor for insight into PCNA-linked diseases, such as cancer.
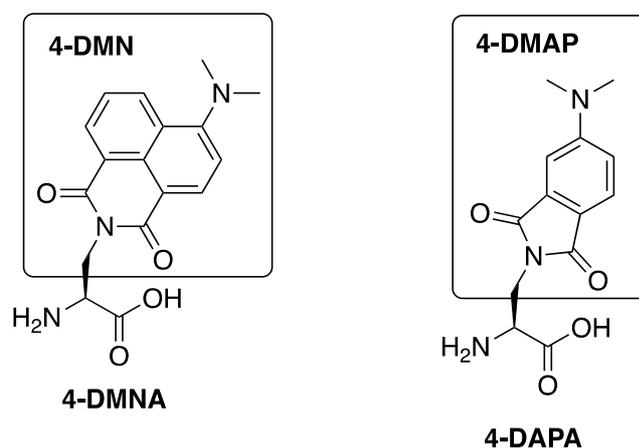
Figure 11: Structures of solvatochromic amino acids, 4-DMNA and 4-DAPA and their corresponding fluorophore, 4-DMN and 4-DMAP. Figure adapted [56]

1.6   Aims for this thesis

p21-derived peptides that inhibit PCNA present an opportunity for the treatment of diseases associated with PCNA upregulation, such as cancer. Systematic modification of the p21-derived sequence provides an opportunity to develop a pre-clinical therapeutic that targets PCNA.

A viable peptide therapeutic must be both cell and nuclear permeable in order to target PCNA. Hence, the aim of the first project (Chapter Two and Three) was to synthesise a cell and nuclear permeable p21-derived peptide. Chapter Two described studies to define the minimum p21-derived peptide scaffold required for cell permeability. The starting template was $p21_{143-155}$, ^{143}RQTSMTDFYHSK^{155} and a series of peptides were prepared with residues reintroduced to the *C*-and *N*-termini, as the longer **p21$_{139-160}$** has been reported to be cell permeable. The cell permeability of the peptides synthesised was determined in a cell uptake assay, the findings of which were used to inform the second aim of this project. Chapter Three explores modifying the p21-derived peptides to facilitate nuclear permeability. Two series of p21-derived peptides were synthesised, linear and macrocyclic, and conjugated to four selected NLS peptides, in order to identify nuclear permeable p21-peptide leads for the development of a viable pre-clinical PCNA inhibitor.

The second project (Chapter Four) describes the development of a p21-derived fluorescent PCNA sensor. A series of three p21-peptides containing a solvatochromic fluorophore, 4-DMAP, at positions 147, 150 or 151 were synthesised, and the optimal position was determined. This work highlighted interesting leads for developing a PCNA-selective turn-on fluorescent sensor to measure proliferation levels and thus investigate diseases such as cancer.

Chapter 2   Cell permeable peptides

2.1   Design of peptides

Work in this chapter is concerned with studies on identifying the shortest p21-derived peptide capable of cellular entry. A 22-residue peptide derived from the p21$^{\text{WAF1/CIP1}}$ protein,^{139}GRKRRQTSMTDFYHSKRRLIFS^{160} (**p21$_{\text{139-160}}$**), has been reported to bind PCNA with high affinity (K$_D$= 2.5 nM) and disrupt DNA-replication upon binding.[27] A fluorescein tagged analogue of this peptide has shown to be cell permeable in a cell permeability assay using T47D breast cancer cells (unpublished). Work in our lab (unpublished) by Aimee Horsfall [66] identified a truncated derivative of **p21$_{\text{139-160}}$**, ^{143}RQTSMTDFYHSK^{155}(**P1**), as the shortest p21-derived peptide needed to effectively bind PCNA (K$_D$= 102 nM). A fluorescein-tagged derivative of **P1**, **P1F**, was shown to lack cell permeability in a cell uptake assay, revealing the truncation of **p21$_{\text{139-160}}$**, to the **P1** sequence, removed residues that facilitate cell permeability. Here, three subsequent peptides; ^{140}RKRRQTSMTDFYHSK^{155} (**P2),** ^{143}RQTSMTDFYHSKRR^{157} (**P3**) and ^{140}RKRRQTSMTDFYHSKRR^{157} (**P4**) (see Table 1) were prepared that extend the **P1** sequence by reintroducing positively charged residues (arginine and lysine) in an attempt to restore the cell permeability seen with **p21$_{\text{139-160}}$**. Cationic residues are known to aid in cell permeability by electrostatically interacting with the negatively charged cell surface.[37]

Fluorescein-tagged analogues **P2F**, **P3F** and **P4F** of **P2, P3** and **P4**, respectively were also prepared for cell permeability studies (see Table 1). These peptides contain an *N*-terminal fluorescein isothiocyanate (FITC) group in order to image the peptides in cell permeability assays. This group is known to be susceptible to an acid-catalysed cyclisation that removes the final *N*-terminal α-amino acid from the peptide (see Figure 12). [67] **P2F**, **P3F** and **P4F** contain a β-alanine residue as a spacer between the rest of the peptide sequence and the fluorescein tag in order to avoid this degradation pathway. [67]

Table 1: p21-derived peptide sequences.

| Peptide	Sequence	Fluorescent tag
p21$_{139-160}$	H-GRKRRQTSMTDFYHSKRRLIFS-NH$_2$	
P1	H-RQTSMTDFYHSK-NH$_2$	
P2	H-RKRRQTSMTDFYHSK-NH$_2$	N/A
P3	H-RQTSMTDFYHSKRR-NH$_2$	
P4	H-RKRRQTSMTDFYHSKRR-NH$_2$	
p21$_{139-160F}$	Fl-GRKRRQTSMTDFYHSKRRLIFS-OH	
P1F	FITC-A$_\beta$-RQTSMTDFYHSK-NH$_2$	
P2F	FITC-A$_\beta$-RKRRQTSMTDFYHSK-NH$_2$	
P3F	FITC-A$_\beta$-RQTSMTDFYHSKRR-NH$_2$	
P4F	FITC-A$_\beta$-RKRRQTSMTDFYHSKRR-NH$_2$	FITC=
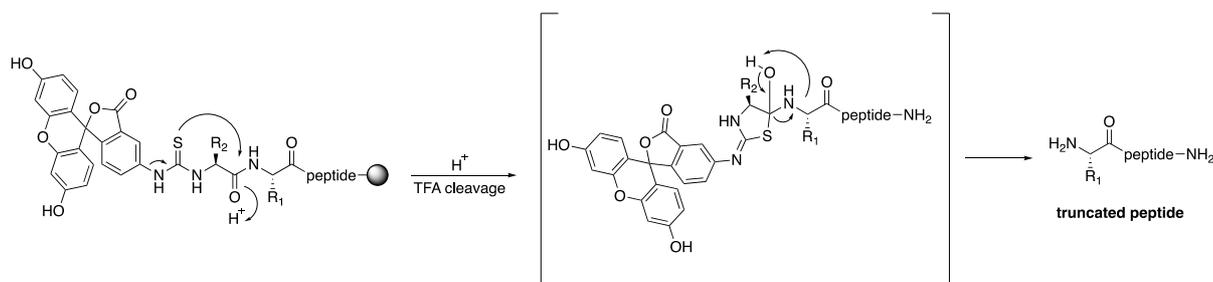
Figure 12: *N*-terminal FITC labelled peptides can undergo a cyclisation in acidic conditions such as during TFA cleavage. This results in removal of the N-terminal α-amino acid and can be avoided by inserting a β-amino acid as a spacer. Figure adapted. [67]

2.2   Synthesis optimisation

P2, P3, **P4**, **P2F, P3F** and **P4F** were synthesised on-resin by Fmoc/tBu-SPPS (described in Experimental). In brief, amino-acids were coupled onto a resin-bound growing peptide chain by addition of Fmoc-protected amino acids (5 equiv), HATU (5 equiv) and DIPEA (10 equiv) in DMF, to the resin for 1 h. On-resin Fmoc-deprotections were achieved on treatment with 20% piperidine and 0.1 M HOBt for 15 min. Successive amino acid couplings and Fmoc-deprotections then gave the desired peptide sequence. **P2F, P3F** and **P4F** were fluorescently labelled following the last Fmoc-deprotection by addition of fluorescein-5-isothiocyanate (5 equiv) and DIPEA (10 equiv) in DMF (4 mL) to the resin for 2 h. All peptides were

simultaneously cleaved from resin and globally deprotected by treatment with a solution of TFA (92.5%), DODT (2.5%), TIPS (2.5%) and water (2.5%) for 2 h. **P2** and **P4** were purified by semi-preparative Reverse Phase-High Performance Liquid Chromatography (RP-HPLC) and identity confirmed by High-resolution Mass Spectrometry (HRMS). The crude **P3** sample contained a major side aspartimide reaction by-product, with further discussion on its preparation and purification included below. **P2F, P3F** and **P4F** were purified by semi-preparative RP-HPLC and identity confirmed by HRMS.
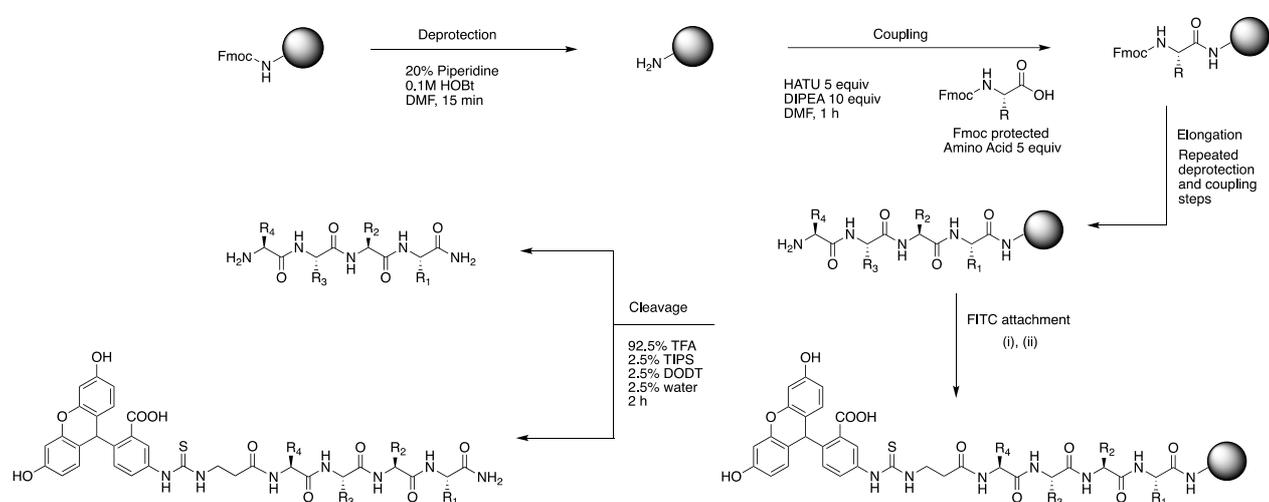
Scheme 1: General scheme of Fmoc/tBu-SPPS and on-resin FITC labelling. *Reagents and conditions*: i) Fmoc-β-Ala-OH, HATU, DIPEA, DMF, 1 h, followed by 20% piperidine, 0.1 M HOBt, 15min; ii) fluorescein-5-isothiocyanate, DIPEA, DMF, 2 h. R= amino acid side chains.

2.2.1 Aspartimide formation

Significant aspartimide formation was observed during the synthesis of **P3**. MS analysis of the crude product revealed the desired product (**P3**) with a mass of 1812 au (M) but also a by-product with a mass of 1794 au (M-18). It was hypothesised that the M-18 signal in the MS spectrum (see Figure 13) corresponds to loss of water as a result of the aspartic acid ester side chain undergoing a base-catalysed cyclisation to form aspartimide (see Figure 15). MS analysis of crude **P2** and **P4** samples obtained under these conditions did not reveal the presence of aspartimide (i.e., M-18 peak) (see Figure 16 and Figure 17). Interestingly, **P2, P3**

and **P4** were synthesised under the same Fmoc/tBu-SPPS conditions, however, aspartimide was only observed in **P3** synthesis, indicating the aspartimide formation was sequence dependence. This is likely a result of enhanced flexibility of the **P3** peptide chain which allows the Phe150 amide in **P3** to attack the Asp149 sidechain, resulting in aspartimide formation (see Figure 14).
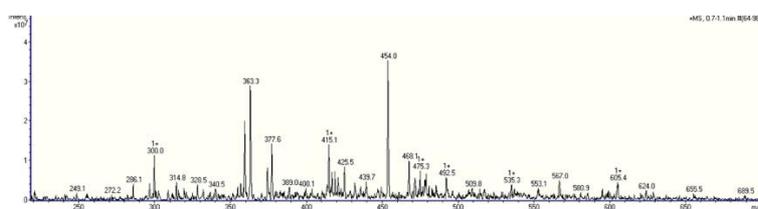

Figure 13: MS of crude **P3.** Target peptide (1812 au): 363 $[M+5H]^{5+}$, 454 $[M+4H]^{4+}$, 605 $[M+3H]^{3+}$ *m/z*. Side-product (1794) (M-18): 300 $[M+6H]^{6+}$, 359 $[M+5H]^{5+}$ *m/z.*
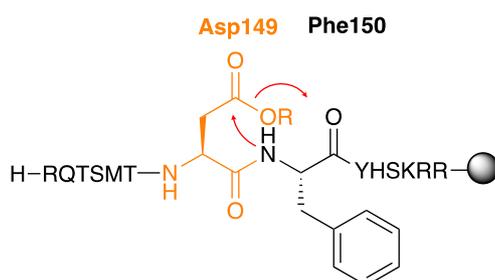
Figure 14: Aspartimide formation mechanism of the P3 sequence.

Aspartimide formation was also observed in an earlier synthesis of **p21₁₃₉₋₁₆₀** in our lab[66], despite the Asp-Phe motif of this sequence not normally being associated with aspartimide formation. Sequence motifs most susceptible to aspartimide formation generally do not contain a large, bulky residue adjacent to the aspartic acid like phenylalanine. Motifs such as Asp-Gly, Asp-Asn, Asp-Se and Asp-Ala are prone to aspartimide formation.[9, 16] The synthesis of **p21₁₃₉₋₁₆₀** by Fmoc/tBu-SPPS was repeated under microwave irradiation conditions in an attempt to minimise aspartimide formation and was found to be successful. Microwave irradiation is reported to provide shorter Fmoc-deprotection times, this limits exposure of the peptide to piperidine. [68]
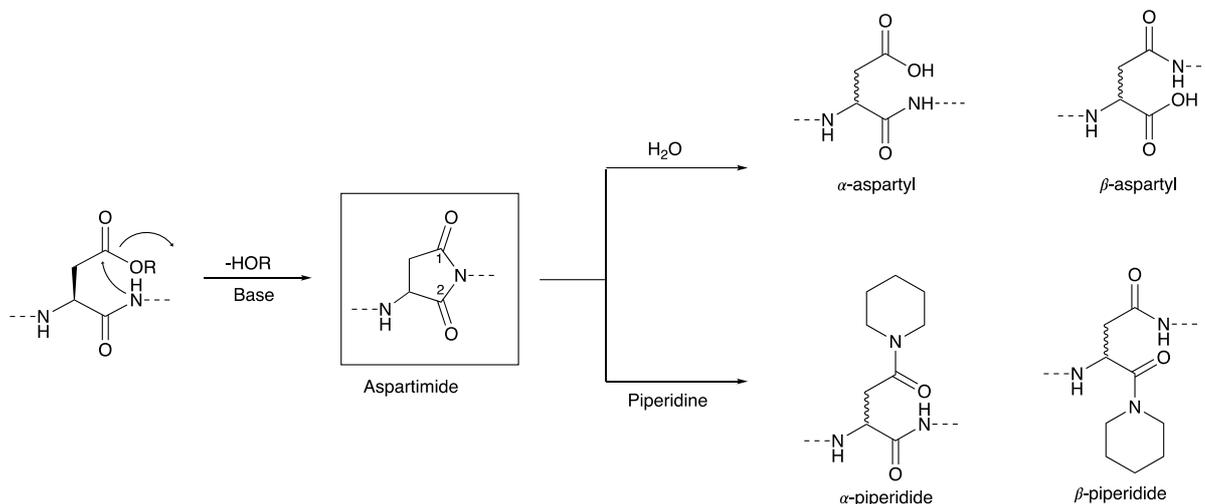
Figure 15: Base-catalysed aspartimide formation. Nucleophilic attack of the aspartimide by water and piperidine results in epimerisised α-and β-peptides. Attack at the C1 leads to α-peptides. Attack at C2 leads tp β-peptides. R= protecting groups, such as tBu. Figure adapted. [17]
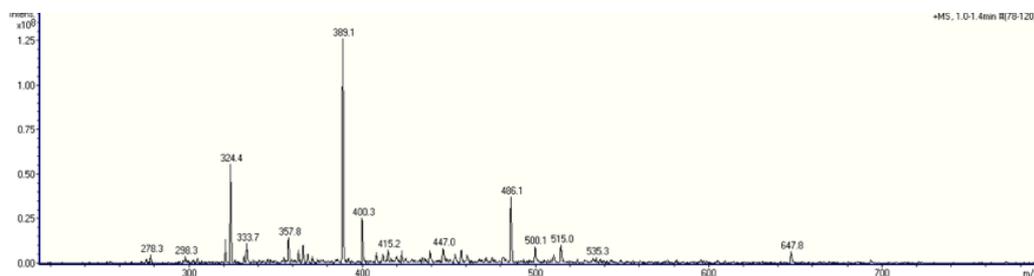
Figure 16: MS of crude **P2.** Target peptide (1940 au): 324 [M+6H]$^{6+}$, 389 [M+5H]$^{5+}$, 486 [M+4H]$^{4+}$, 647 [M+3H]$^{3+}$ *m/z*.
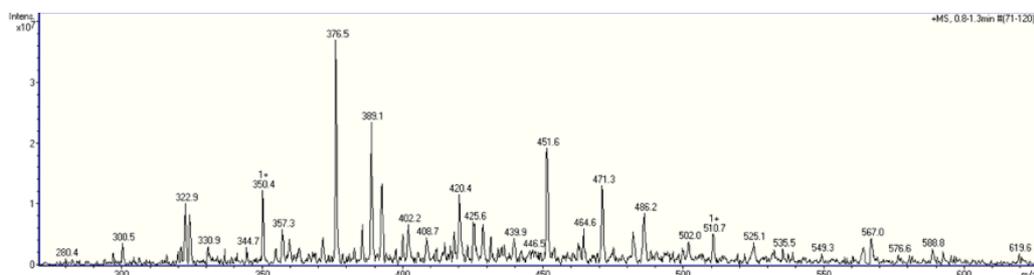
Figure 17: MS of crude **P4**. Target peptide (2252 au): 322 [M+7H]$^{7+}$, 376 [M+6H]$^{6+}$, 451 [M+5H]$^{5+}$ *m/z*.

The aspartimide can undergo epimerisation at both carbonyls on reaction with either water or piperidine, resulting in epimerised by-products (see Figure 15). Ring-opening of aspartimide on reaction with piperidine gives rise to α-and β-piperidides. α-and β-piperidides are characterised in MS as a mass 67 au greater than that of the target peptide. A mass of 1879 au (M+67) was not observed in the MS of crude **P3**, which confirms absence of α-and β-piperidides.

37

Ring-opening of aspartimide on reaction with water can yield the target peptide or the β-aspartyl peptide. The β-aspartyl peptide, unlike α-and β-piperidides, usually co-elutes with the desired peptide and has an identical mass. Purification of the crude **P3** mixture, by semi-preparative RP-HPLC, was initially attempted in order to separate the desired product from peptide with aspartimide. A fraction from the major peak in the trace was collected and analysed by RP-HPLC (see Figure 18) and MS (see Figure 19). A sharp peak was observed in the RP-HPLC trace and MS analysis revealed peptide masses of 1812 au (**P3**, M) and 1798 au (M-18), indicating **P3** and **P3** with aspartimide formation were co-eluting in the RP-HPLC and their separation was unsuccessful. We next looked to minimising aspartimide formation during the preparation of **P3** by altering the synthetic conditions in the deprotection step.
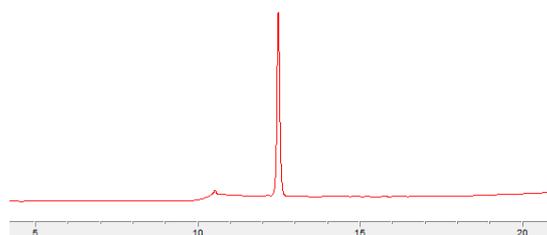
Figure 18: Analytical RP-HPLC spectrum of **P3**, 0-100% aq. ACN gradient over 15 minutes on a C18 column visualised at 220nm.
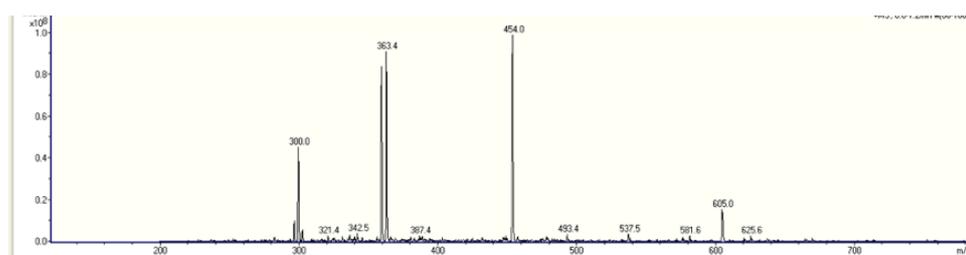
Figure 19: MS of purified fraction of **P3**

A range of deprotection conditions were investigated for the synthesis of **P3** in order to combat aspartimide formation and possible formation of β-aspartyl peptides (see Table 2). These include addition of different acidic modifiers and also using an alternative base and decreasing the deprotection time, the results of which are summarised in Table 3.

Firstly, the addition of acidic modifiers HOBt (**D1**) and formic acid (**D2**) to the deprotection solution was investigated as both acids have been reported to suppress aspartimide formation.[19, 69] **P3** was synthesised using the Fmoc-deprotection conditions **D1** and **D2**, defined in Table 2. RP-HPLC analysis of the peptide sample prepared by **D1** revealed a sharp peak at 8.9 min (see Figure 20) and MS confirmed the major product to be the target peptide (**P3**) with a mass of 1812 au (see Figure 21). This suggests addition of 0.5 M HOBt as an acidic modifier, in deprotection condition **D1**, suppressed aspartimide formation. The peptide sample prepared by **D2** eluted as a twinned peak at 8.5 min in RP-HPLC analysis (see Figure S1), that corresponded to unidentified products in the MS (see Figure S2). This indicates that the use of 2% formic acid, in deprotection condition **D2**, hindered **P3** synthesis.

Table 2: Deprotection conditions trialled for minimising aspartimide formation during the synthesis of **P3**

| Deprotection condition	Deprotection agent (in DMF)	Acidic modifiers	Deprotection time
0	20% piperidine	0.1 M HOBt	15 min
1	20% piperidine	0.5 M HOBt	2 x 1 min
2	20% piperidine	2% formic acid	2 x 1 min
3	20% piperidine	8% formic acid	10 min
4	20% piperidine	8% formic acid	2 x 1 min
5	5% piperazine	-	2 x 1 min
6	5% piperazine	0.5 M HOBt	2 x 1 min

Formic acid has been reported to significantly reduce aspartimide formation.[19] Fmoc-deprotection conditions **D3** and **D4** were investigated where formic acid was added to the deprotection solution as an acidic modifier. **D3** and **D4** include 8% formic acid, a higher concentration of formic acid than condition **D2**, to determine if increasing the concentration aids in **P3** synthesis and minimises aspartimide formation. The effect of deprotection time on successful **P3** synthesis was also examined, where condition **D3** tested 10 min Fmoc-deprotections and condition **D4** tested 2 x 1 min Fmoc-deprotections. Longer Fmoc-deprotection times have been previously linked to increase aspartimide formation.[70] RP-HPLC analysis of the peptide sample prepared under condition **D3** revealed broad peaks eluting between 8 and 9 min and a cluster of peaks between 12 and 15 min (see Figure S3).

MS revealed the products at 8 to 9 min as peptides with various amino acid deletions (see Figure S4). MS also identified Fmoc-protected products that correlate to the cluster of peaks between 12-15 min in the RP-HPLC as these peaks exhibit absorption in the 254 nm range, associated with the Fmoc-group. The peptide sample prepared under condition **D4** displayed similar results. RP-HPLC revealed peaks eluting at 8 to 9 min and 12 to 15 min (see Figure S5). MS confirmed the products as peptides with various amino acid deletions and Fmoc-protected peptides. MS analysis also revealed the desired peptide (**P3**) with a mass of 1812 au, however, **P3** was not the major product. These results indicate that increasing the formic acid concentration to 8%, in conditions **D3** and **D4**, did not improve **P3** synthesis. Additionally, a shorter deprotection time, in condition **D4** (2 x 1 min), did not drastically alter the results of the synthesis.

Next, piperazine was investigated as an alternative to piperidine for Fmoc-deprotections, in conditions **D5** and **D6**, to determine if a weaker base may minimise aspartimide formation. [69,71] RP-HPLC analysis of the peptide samples prepared by condition **D6** (see Figure S9) and **D5** (see Figure S7) both revealed broad peaks between 8 and 9 min and a cluster of peaks between 12 and 15 min. MS confirmed the products in the **D6** peptide samples as peptides with various sequence terminations and Fmoc-protected peptides (Figure S10). MS identified the products in the **D5** peptide sample as peptides with various amino acid deletions and Fmoc-protected peptides (see Figure S8). These results indicate the use of piperazine, in deprotection condition **D5** and **D6**, hindered Fmoc-deprotections, leading to Fmoc-protected peptides and sequence terminated peptides.

In summary, MS analysis indicated only two of the six deprotection conditions trialled, **D1** and **D4**, resulted in of the formation of **P3. D1** resulted in a superior yield for **P3** with the major product being **P3**, as confirmed by MS, whereas the **P3** yield from **D4** was significantly lower. MS analysis of the peptide sample prepared by **D1** conditions revealed a smaller, less

intense M-18 peak, relative to that of the peptide sample prepared by **D0**. This indicates that **D1** conditions in the preparation of **P3** was found to suppress aspartimide formation when compared to the original deprotection condition **D0**. **D1** utilised an increased HOBt concentration that was likely required due to commercial HOBt hydrate containing >20% w/w water, thus lowering the concentration of HOBt. The shorter deprotection times, in condition **D1**, compared to **D0**, were found to be sufficient for Fmoc-deprotection and led to decreased aspartimide formation. **D1** conditions were used for the synthesis of **P3**.
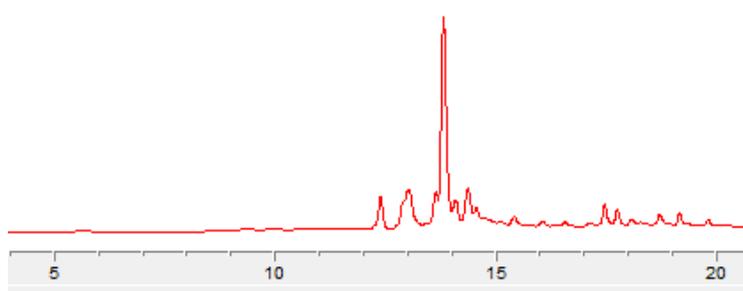

Figure 20: Analytical RP-HPLC of crude **P3** peptide prepared by D1 conditions, 0-50% aq. ACN over 15 min (from 5-20 min), visualised at 220 nm.
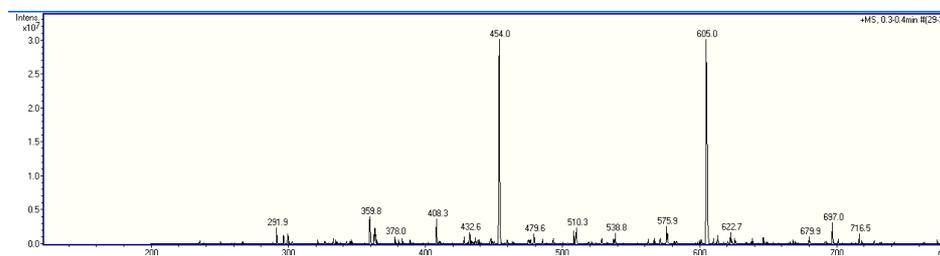

Figure 21: MS of crude peptide prepared by D1 conditions. Target peptide (1812 au): 363 [M+5H]$^{5+}$, 454 [M+4H]$^{4+}$, 605 [M+3H]$^{3+}$ *m/z*.

Table 3: MS and RP-HPLC analysis of crude **P3** peptides synthesised with different deprotection conditions. T_D= deprotection time.

Code	Deprotection Conditions	Mass (au)	*m/z* peaks	Mass difference (M= 1812)	ID	RP-HPLC
D1	20% piperidine + 0.5 M HOBt (T_D=2x1min)	1812	363, 454, 605	M	Molecular ion	Major, sharp peak at 8.9min (see Figure 20)
D2	20% piperidine + 2% formic acid (T_D= 2x1min)	1573	525, 787	M-239	Unidentified	Split peak at 8.5 min (see Figure S1),
		1125	376, 563	M-687	Unidentified	
		1188	397, 595	M-624	Unidentified	
D3	20% piperidine + 8% formic acid (T_D=10min)	1804	361, 454, 602	M-8	Unidentified	Multiple small and broad peaks at 8-9 min (see Figure S3).
		1906	382, 477, 636	M+94	Fmoc PG Lysine deletion	
		1933	387, 484	M+121	Fmoc PG Threonine deletion	
		2034	407, 509, 679	M+222	Fmoc PG	
		1779	445, 593	M-36	Unidentified	
D4	20% piperidine + 8% formic acid (T_D= 2x1min)	2034	407, 509, 679	M+222	Fmoc PG	Multiple small and broad peaks at 8-9 min (see Figure S5).
		1684	337, 422, 562	M-128	Lysine deletion	
		1906	382, 477	M+94	Fmoc PG Lysine deletion	
		1812	303, 363, 454	M	Molecular ion	
		1655	332, 415, 553	M-157	Arginine deletion	
D5	5% piperazine (T_D= 2x1min)	1933	387, 484, 645	M+120	Fmoc PG Threonine deletion	Multiple small and broad peaks at 7-8 min (see Figure S7)
		1947	487, 650	M+135	Fmoc PG Serine deletion	
		2034	407, 609, 679	M+222	Fmoc PG	
		2090	419, 523	M+278	Fmoc PG tBu	
D6	5% piperazine + 0.5 M HOBt (T_D= 2x1min)	1329	333, 444, 665	M-483	Fmoc PG Sequence termination	Multiple small and broad peaks at 8-9 min (see Figure S9)
		1430	358, 477, 716	M-382	Fmoc PG Sequence termination	
		1485	372, 496, 743	M-327	Fmoc PG Sequence termination Arginine deletion	

2.3   Binding affinity

Samples of **P2**, **P3**, **P4**, **P2F**, **P3F** and **P4F** were analysed against PCNA by Surface Plasmon Resonance (SPR) to investigate the effect of extending the **P1** sequence on binding. The binding experiments were performed using a Biacore S200 system (see Experimental for details). PCNA was first immobilised onto a gold sensor chip with a carboxymethylated

dextran matrix surface, by EDC/NHS mediated coupling, with a running buffer of 10 mM HEPES buffer with 150 mM NaCl, 3 mM EDTA and 0.05% Tween 20 (pH 7.4). Binding affinity data was collected for **P2**, **P3**, **P4**, **P2F**, **P3F** and **P4F** by first diluting each peptide with running buffer and flowing over the PCNA-loaded sensor chip. Each peptide was serially diluted (1:2) eight times and run from lowest to highest concentration. Binding events, for example, the peptide to the surface immobilised PCNA, results in a change in refractive index of the sensor chip. This change is measured and converted to a 'response', and once equilibration of binding was reached, the steady state binding at each concentration of peptide was determined. The response versus peptide concentration was plotted and a line fitted to provide the equilibrium dissociation constant K_D.

Peptides **P2**, **P3** and **P4** bound PCNA with a K_D of 41 nM, 45 nM and 38 nM, respectively (see Table 4). **P2**, **P3** and **P4** all displayed higher affinity for PCNA compared to **P1**, which binds with 102.3 nM affinity. This indicates the increased length of **P2**, **P3** and **P4** relative to **P1**, facilitated higher affinity binding to PCNA.

Peptides **P2F**, **P3F** and **P4F** bound PCNA with K_D constants of 202, 191 and 61 nM, respectively (see Table 4). This indicates that *N*-terminal attachment of FITC, to **P2**, **P3** and **P4**, somewhat lowered the binding affinity of all three peptides. **P2F** and **P3F** showed a 5-fold and 4-fold lower binding affinity for PCNA compared to **P2** and **P3**, respectively. **P4F** bound PCNA with a K_D of 61 nM, indicating a 1.6-fold lower affinity than **P4**.

All three peptides, **P2**, **P3** and **P4** exhibited improved binding affinity for PCNA, relative to **P1**, indicating that lengthening the **P1** sequence facilitates PCNA binding. This is significant as p21-derived PCNA inhibitors need to bind to PCNA in order to exhibit their inhibitory effect. Fluorescein labelling the peptides was found to lower PCNA affinity, highlighting a

possible limitation in future activity assays by hindering the ability of the fluorescein tagged

peptides to bind and therefore, inhibit PCNA.

Table 4: Binding affinity of the proposed cell permeable peptides and their fluorescein tagged analogues, determined by SPR. SE = standard error. χ^2 = chi^2 (gives a measure of accuracy of fitting).

Peptide	Sequence	Affinity K_D (nM)	K_D SE (nM)	χ^2
P1	H-RQTSMTDFYHSK-NH$_2$	102.3	5.3	0.0701
P2	H-RKRRQTSMTDFYHSK-NH$_2$	41	5.7	1.01
P3	H-RQTSMTDFYHSKRR-NH$_2$	45	4.3	0.301
P4	H-RKRRQTSMTDFYHSKRR-NH$_2$	38	10	7.06
P2F	FITC-A$_\beta$-RKRRQTSMTDFYHSK-NH$_2$	202	18	0.556
P3F	FITC-A$_\beta$-RQTSMTDFYHSKRR-NH$_2$	191	8.7	0.1
P4F	FITC-A$_\beta$-RKRRQTSMTDFYHSKRR-NH$_2$	61	5.6	1.09

2.4 Cell permeability assay

A fluorescein-tagged derivative of **p21$_{139-160}$, p21$_{139-160}$F,** was tested in a cell permeability

assay, as a control, using T47D breast cancer cells by our collaborator Zoya Kikhtyak[72] (see

Experimental for details). The cells were treated with 5 µM of **p21$_{139-160}$F** and incubated for

24 h. The cells were subsequently fixed with formaldehyde and stained with DAPI to mark

the nucleus, and phalloidin to stain actin filaments, marking the cell cytosol. The treated cells

were imaged by fluorescence microscopy which revealed significant green fluorescence in the

cell cytosol of treated cells, indicating **p21$_{139-160}$F** was accumulating within the cell (see

Figure 22). Interestingly, the **p21$_{139-160}$F** fluorescence within the cell appeared punctate,

suggesting endosomal entrapment of **p21$_{139-160}$F**.[32] The cell permeability assay was repeated

with **P1F** as another control. **P1F** did not exhibit cell permeability (image not shown),

confirming that truncating the **p21$_{139-160}$** sequence, to **P1**, eliminated the cell permeability.

P2F, P3F and **P4F** were then evaluated for cell permeability, to determine whether extending

the **P1** sequence and reintroducing positively charged residues could restore the cell

permeability. T47D breast cancer cells were treated with 5 µM of each peptide and incubated

by 24 h, after which the cells were fixed, stained and imaged by fluorescence microscopy (see Figure 22). **P2F** and **P4F** treated cells both displayed some green fluorescence corresponding to the fluorescein tagged peptide in the cell cytosol, indicating modest intracellular accumulation of both **P2F** and **P4F.** No green fluorescence was observed in the cells treated with **P3F**, indicating that **P3F** was not cell permeable. **P4F** displayed the most cell permeability, suggesting the reintroduction of positively charged residues to both termini facilitated greater cell uptake. **p21$_{139\text{-}160}$F** however, showed significantly more cell permeability compared to **P4F**, suggesting the hydrophobic ^{157}LIFS^{160} segment of the **p21$_{139\text{-}160}$** sequence facilitates cell permeability.

The series of proposed cell permeable peptides proved to be moderately successfully with **P2** and **P4** showing modest cell uptake. This indicates reintroducing positively charged residues to the *C*-terminal of **P1**, in **P2**, or both termini, in **P4**, confers cell permeability. A fluorescein tagged derivative of **P3**, on the other hand, displayed no cell entry. Extending the **P1** sequence at both termini, as in **P4**, was found to facilitate the most cell uptake, further highlighting the role of additional positively charged residues in cell permeability.
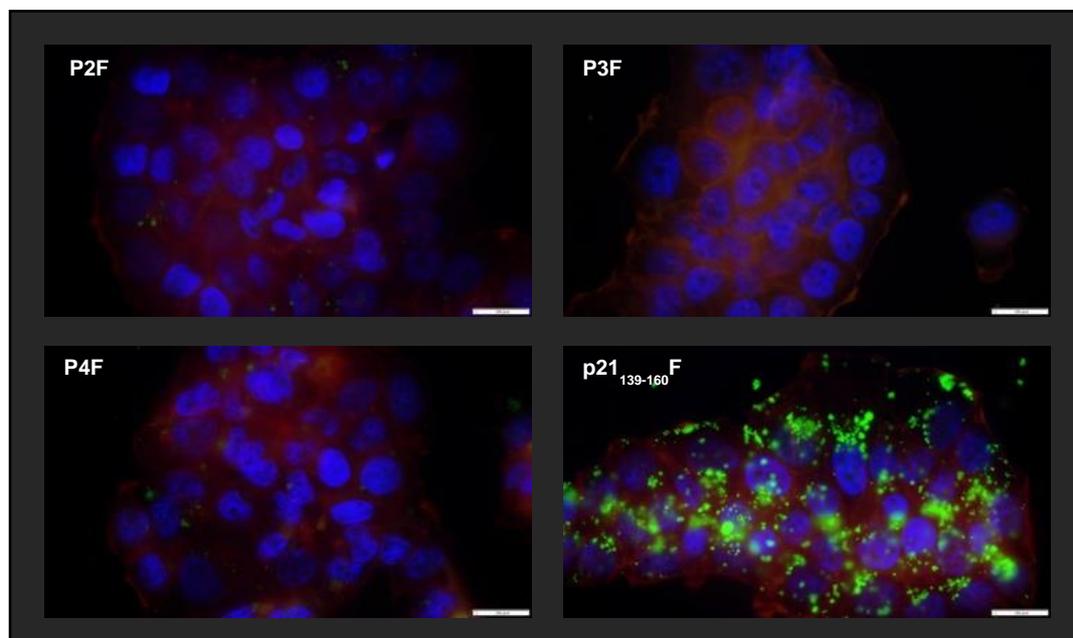
Figure 22: Fluorescence microscopy images of T-47D breast cancer cells treated with 5 µM P2F, P3F, P4F or p21$_{139-160}$F for 24 h. The peptides were covalently labelled with FITC (green), the DNA stained with DAPI (blue) and actin filaments with Phalloidin (red). Images were taken at 100x magnification. Scale bar reads 200 µm.

2.5   Chapter conclusions

The three proposed cell permeable peptides **P2**, **P3** and **P4** were successfully synthesised. **P3** was found to be especially susceptible to aspartimide formation. Various deprotection conditions were investigated to minimise aspartimide formation in the synthesis of **P3**. The susceptibility of **P3** to form aspartimide is likely attributed to aspartimide formation being highly sequence specific.[73]   **P3** was synthesised a number of times, each with different deprotection conditions. Deprotection solutions containing acidic modifiers (HOBt or formic acid) and different deprotecting agents (piperazine) were trialled but few resulted in successful **P3** synthesis. The most optimal deprotection condition for **P3** synthesis was found to be 20% piperidine and 0.5M HOBt with 2 x 1 min deprotection times.

SPR data revealed **P2**, **P3** and **P4** bound PCNA in the nanomolar range with K$_D$ values of 41 nM, 45 nM and 38 nM, respectively. This presented a significant improvement of **P1** which has a K$_D$ of 102.3 nM, indicating that extending the **P1** sequence, enhances binding affinity

for PCNA. Attachment of the fluorescein tag to the peptides decreased the binding affinity of the **P2**, **P3** and **P4** sequences with **P2F**, **P3F** and **P4F** giving K_D values of 202 nM, 191 nM and 61 nM, respectively.

The cell permeability of **P2F**, **P3F** and **P4F** was investigated by fluorescence microscopy using T47D breast cancer cells. **P2F** and **P4F** showed modest cell permeability. **P4F** showed the most intracellular accumulation, demonstrating the extension of the sequence at both *N*- and *C*-termini, relative to **P1**, increased the overall cationic charge of the peptide and enhanced cell entry. **p21$_{139-160}$ F,** however, was significantly more cell permeable than **P4F**, suggesting cell permeability is facilitated by the ^{157}LIFS^{160} region of the **p21$_{139-160}$** sequence. This stretch of hydrophobic residues increases the amphipathic nature of the peptide, a feature common amongst CPPs. [37]

P2, **P3** and **P4** did not confer significant cell permeability, relative to **p21$_{139-160}$,** suggesting that extending the **P1** sequence to **P2**, **P3** and **P4** was not enough to facilitate substantial cell entry. Further study was continued in the following chapter with **P1**, **P2**, **P3** and **P4**, to investigate whether appending a Nuclear Locating Sequence tag could facilitate both significant cell and nuclear penetration.

Chapter 3   Nuclear permeable peptides

This chapter describes studies on the design of two series of nuclear permeable p21-derived peptides for the inhibition of PCNA. Specifically, two series of peptides were prepared and investigated where Nuclear Localisation Sequence (NLS) peptides were chemically linked to linear and macrocyclic p21-derived peptides. Here, NLS peptides were appended to p21-derived peptides and explored as a means to facilitate cell and nuclear entry.

3.1   Selection of NLS peptides

Four NLS peptides, $SV40_{126-132}$, $cMyc_{320-328}$, Tat_{48-57} and R6W3, were proposed for attachment to p21-derived peptides in order to facilitate cell and nuclear entry (see Table 5). These peptides vary in length, origin and properties to allow investigation of a diverse selection of NLS peptides. $SV40_{126-132}$ (^{126}PKKKRRV^{132}) is derived from the SV40 large T antigen protein and is an example of a classical monopartite NLS, in which the sequence consists of a single stretch of basic residues.[44] $SV40_{126-132}$ has previously been appended to a p21-derived peptide (residues 1-79) to promote nuclear localisation of the peptide in transfected cells.[29] $cMyc_{320-328}$ (^{320}PAAKRVKLD^{328}) was derived from the human cMyc oncoprotein and is also a classical monopartite NLS, despite the sequence containing few basic residues. The $cMyc_{320-328}$ NLS sequence was first identified and reported to facilitate nuclear transport of covalently attached human serum albumin protein. [47] The Tat_{48-57} peptide(^{48}GRKKRRQRRR^{57}) is derived from the HIV-1 Tat regulatory protein that was found to enter cells and translocate into the nucleus. [74, 75] The Tat_{48-57} peptide has been reported to promote nuclear accumulation of a p21-derived peptide. [23] The R6W3 sequence (RRWWRRWRR) is a synthetic analogue of penetratin (^{43}RQIKIWFGNRRMKWKK^{58}), which is a well-known Cell Penetrating Peptide (CPP). The penetratin CPP is derived from a *Drosophila* Antennapedia homeodomain protein and is reported to cross the cell membrane and translocate to the nucleus.[40] R6W3 mimics the structure and properties of the original

48

penetratin CPP, where it forms an amphipathic, α-helical secondary structure that is favourable to cell permeability. [37]

Table 5: NLS tags selected to facilitate nuclear entry of p21-derived peptides

NLS	Sequence	Origin	Type
Tat[74, 75]	GRKKRRQRRR	Protein-derived	Cationic CPP
SV40[76]	PKKKRKV	Protein-derived	Classical monopartite[a] NLS
cMyc[47]	PAAKRVKLD	Protein-derived	Classical monopartite[a] NLS
R6W3[40]	RRWWRRWRRC	Synthetic	Cationic CPP Secondary amphipathic CPP

[a] Classical monopartite NLS peptides are defined as a single stretch of basic amino acids

3.2   Design and synthesis of linear peptides

The preparation of NLS tagged p21-peptides by simply combining the NLS and p21-sequences and synthesising the resulting peptide by Fmoc/tBu-SPPS may result in synthetic difficulties of aggregation, racemisation or side reactions. Long peptides are susceptible to aggregation and solubility issues during synthesis due to secondary structure formation (i.e., β-sheet structures). Such formation is considered to be a major source of synthetic problems and leads to difficult preparation and purification.[10, 77]

In the first instance, a peptide termed **P4Tat** (GRKKRRQRRRRKRRQTSMTDFYHSKRR), derived from both the **P4** (see Table 6 and Chapter 2 for a discussion) and **Tat** (see Table 5) sequences, was investigated. Synthesis of **P4Tat** was attempted to test the feasibility of preparing an NLS tagged p21-peptide by simply combining the NLS and p21-sequences and synthesising the resulting peptide by Fmoc/tBu-SPPS.

Table 6: p21-derived peptides prepared and tested for cell permeability (see more in Chapter 2).

Peptide	Sequence
P1	H-RQTSMTDFYHSK-NH$_2$
P2	H-RKRRQTSMTDFYHSK-NH$_2$
P3	H-RQTSMTDFYHSKRR-NH$_2$
P4	H-RKRRQTSMTDFYHSKRR-NH$_2$

P4Tat was synthesised by Fmoc/tBu -SPPS with microwave irradiation in our lab by Aimee Horsfall.[66] Microwave irradiation was used to minimise peptide aggregation and to accelerate the reaction kinetics.[68] In brief, amino acids were coupled to resin-bound peptide with addition of Fmoc-protected amino-acids (5 equiv), OxymaPure® (5 equiv) and DIC (5 equiv) in DMF, to the resin with a maximum temperature of 90°C. Fmoc-deprotections were achieved with 20% piperidine and 0.1 M OxymaPure® in DMF. These couplings and Fmoc-deprotections were repeated to give the desired peptide sequence. The final peptide was simultaneously cleaved from resin and globally deprotected on treatment with a solution of TFA (92.5%), DODT (2.5%), TIPS (2.5%) and water (2.5%) for 2 h. The crude **P4Tat** sample was analysed by RP-HPLC analysis to reveal a broad peak spanning 3 and 12 min (see Figure 23). Purification of the crude sample was attempted by RP-HPLC, but the desired peptide could not be isolated. These results indicate significant difficulty in the synthesis of **P4Tat**, likely attributed to its length. Instead, a conjugation approach was investigated for the synthesis of **P4Tat**.
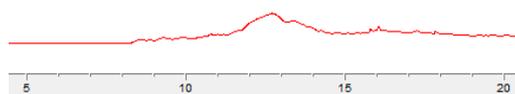

Figure 23: Analytical RP-HPLC spectrum on a C18 column of crude **P4Tat** sample, 0-50% aq. ACN over 15 minutes, visualised at 220 nm.

Next, a two-component synthetic approach was explored for the synthesis of the NLS tagged p21-peptides. This strategy involves individual preparation of the NLS and p21-derived

peptide and reacting them to give the full-length peptide. This allows for a modular approach, where the peptide segments can be mixed and matched to access peptide combinations conveniently and strategically. This also avoids synthetic problems associated with long peptide synthesis via sequential amino acid coupling as seen for the attempted synthesis of **P4Tat**.

Thiol conjugation was investigated in which thiols react with electrophilic groups (e.g. haloacetyls and maleimides) via alkylation to form a thioether bond.[78] The sequences of the four proposed NLS peptides; Tat_{48-57} , $cMyc_{320-328}$, $SV40_{126-132}$ and R6W3 were modified to include a *C*-terminal cysteine residue, to provide a free thiol group for conjugation to a p21-derived peptide. The NLS peptides were also modified to include an *N*-terminal fluorescein to allow fluorescent visualisation of the resulting peptide conjugates in a cell uptake assay. The NLS peptides, with these two design modifications, are referred to herein as **N1F**, **N2F**, **N3F** and **N4F**, respectively (see Table 7).

Table 7: Peptide sequence of four modified NLS tags proposed for conjugation to p21-derived peptides. *C*-terminal cysteine residues are in orange. *N*-terminal fluorescein tag is in green. A β-Alanine residue (A_β) is added to the NLS sequences as a spacer amino acid before fluorescein attachment (see Chapter 2 for discussion).

| Peptide	Sequence
N1F	FITC-A_βGRKKRRQRRR**C**-NH$_2$
N2F	FITC-A_βPKKKRKV**C**-NH$_2$
N3F	FITC-A_βPAAKRVKLD**C**-NH$_2$
N4F	FITC-A_βRRWWRRWRR**C**-NH$_2$

3.2.1 Thiol conjugation via chloroacetyl moiety

Firstly, the thiol conjugation approach was trialled with **P4** and **N1F**. The **P4** sequence was modified to include an *N*-terminal chloroacetyl group, to give **P4a** (Cl-RKRRQTSMTDFYHSKRR) (see Figure 24), to allow conjugation with the thiol containing

N1F peptide using the approach shown in Scheme 2, to give **P4a-N1F** (FITC-AβGRKKRRQRRRCRKRRQTSMTDFYHSKRR) (see Figure 25).

Figure 24: Structure of **P4a**
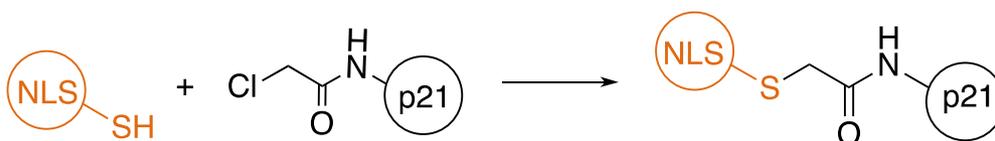
Scheme 2: General scheme of peptide conjugation via NLS tag and chloroacetylated p21 peptide.
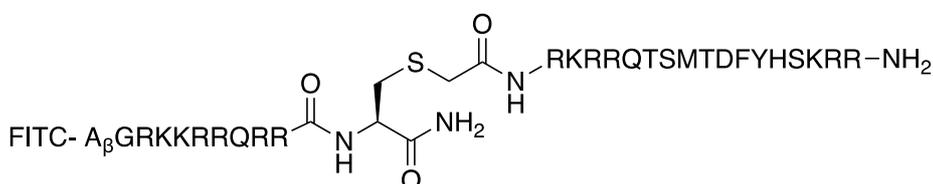

Figure 25: Structure of **P4a-N1F**.

P4a and **N1F** were synthesised by Fmoc/tBu-SPPS (described in Experimental). In brief, Fmoc-protected amino acids were coupled to the growing peptide chain by addition of the Fmoc-protected amino-acid (5 equiv), HATU (5 equiv) and DIPEA (10 equiv) in DMF, to the resin for 1 h. On-resin Fmoc-deprotections were achieved with 15 min treatment with 20% piperidine and 0.5 M HOBt. Successive amino acid couplings and Fmoc-deprotections were repeated to achieve the desired peptide sequence. The *N*-terminal chloroacetyl moiety was installed by reaction of the resin-bound peptide with freshly prepared chloroacetic anhydride. Chloroacetic acid (0.38 g, 4.0 mmol) and DCC (0.41 g, 2.0 mmol) were dissolved in DCM (10 mL) and the mixture stirred at rt for 15 min. The mixture was filtered to remove precipitate and the filtrate was evaporated *in vacuo* to give chloroacetic anhydride as a white solid. The chloroacetic anhydride was reacted with the resin-bound peptide immediately without purification. Following the last Fmoc-deprotection, a solution of chloroacetic anhydride (1 equiv) in DMF to the resin for 1 h (see Scheme 3). **N1F** was assembled using the

same Fmoc/tBu-SPPS procedure and then fluorescently labelled following the last Fmoc-deprotection by treatment of the resin with a solution of fluorescein-5-isothiocyanate (5 equiv) and DIPEA (10 equiv) in DMF for 2 h (see Scheme 1 from Chapter 2). The final peptides were cleaved from resin and simultaneously globally deprotected on treating the resin with a solution of TFA (92.5%), DODT (2.5%), TIPS (2.5%) and water (2.5%) for 2 h to give **P4a** and **NF**. **P4a** and **N1F** were purified by semi-preparative RP-HPLC and identity confirmed by HRMS, with associated data given in the experimental section.
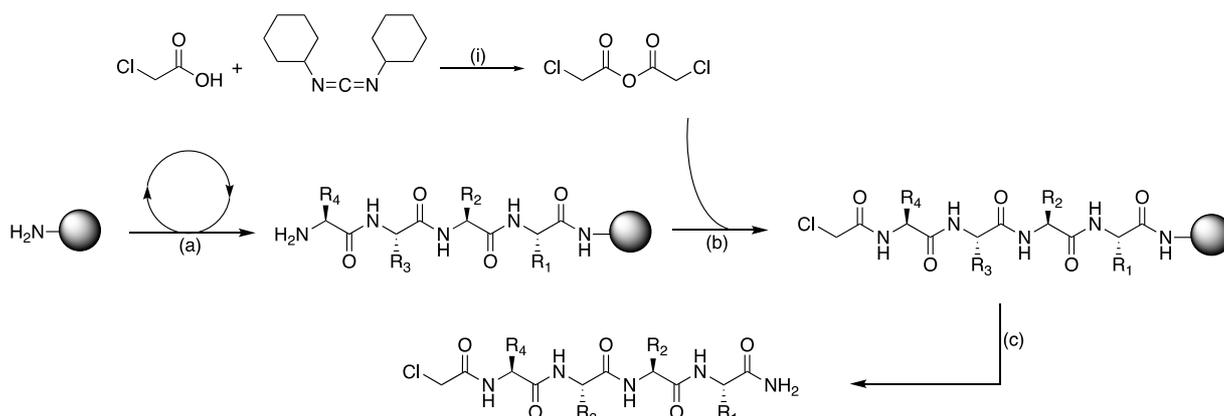
Scheme 3: General scheme for synthesis of a chloroacetylated peptide. *Reagents and conditions:* i) DCM, rt, 15 min; a) Repeated steps of coupling and deprotection; b) chloroacetic anhydride, DMF, 1 h; c) 92.5% TFA, 2.5% DODT, 2.5% TIPS, 2.5% water, 2 h. R= sidechain of amino acid.

Next, trial conjugations of peptides **P4a** and **N1F** peptides were attempted under different solvent conditions in an attempt to prepare **P4a-N1F**, see discussion below. Phosphate buffer was tested in these reactions in order to maintain a physiological pH required for thiols to react with haloacetyl groups.[78] Aqueous sodium bicarbonate was also trialled for the haloacetyl-thiol conjugations.[78-80]

Condition 1: 100 mM phosphate buffer (pH 7.8).

P4a was dissolved in 100 mM phosphate buffer (pH 7.8) and **N1F** in water. The peptide solutions were then combined and stirred at rt. After 24 h, a sample of the crude reaction mixture was removed and analysed by RP-HPLC and MS. Two major peaks were observed in

the RP-HPLC trace (see Figure 26), with MS revealing unidentified masses that did not correspond to the desired peptide **P4a-N1F** (see Figure 27). This indicates conjugation was unsuccessful under these conditions.
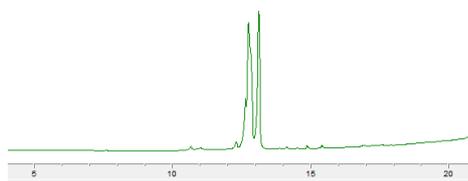

Figure 26: Analytical RP-HPLC C18 spectrum of crude **P4a-N1F** prepared by condition 1, 0-50% aq. ACN over 15 min (from 5-20 min), visualised at 254 nm.
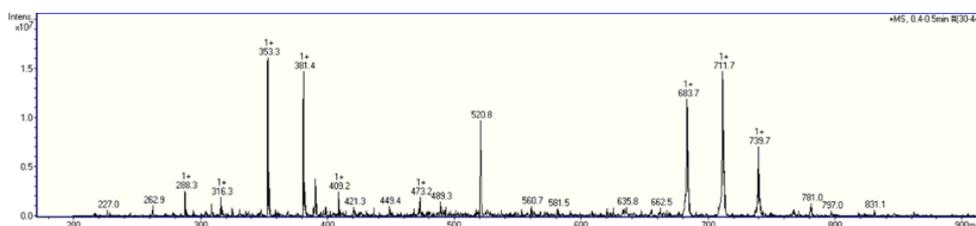

Figure 27: MS of crude **P4a-N1F** prepared by condition 1. Target peptide (4251 au) not isolated.

Condition 2: 100mM phosphate buffer (pH 7.8) + TCEP (2%)

The addition of TCEP to the reaction mixture **P4a** and **N1F** was also investigated. TCEP reduces disulphide bonds that may have formed between the cysteine residue of two **N1F** peptides, ensuring the presence of free thiol groups for conjugation. **N1F** was dissolved in water and **P4a** in 100 mM phosphate buffer (pH 7.8) with 2% TCEP. The peptide solutions were combined, stirred at rt and a sample of the reaction mixture was removed after 24 h and analysed by RP-HPLC and MS. Similarly, to the results of condition 1, two major peaks were observed in the RP-HPLC trace (see Figure 28) with MS revealing masses that did not correspond to the desired peptide **P4a-N1F** (see Figure 29). This suggests that disulphide bond formation in the **N1F** sample was not the limiting factor for reaction of **P4a** and **N1F**.
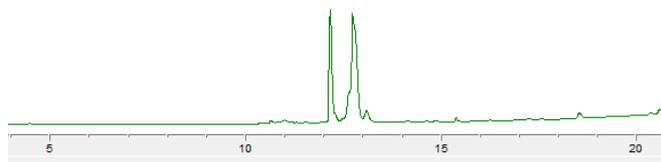
Figure 28: Analytical RP-HPLC C18 spectrum of crude **P4a-N1F** prepared by condition 2, 0-50% aq. ACN over 15 min (from 5-20 min), visualised at 254 nm.
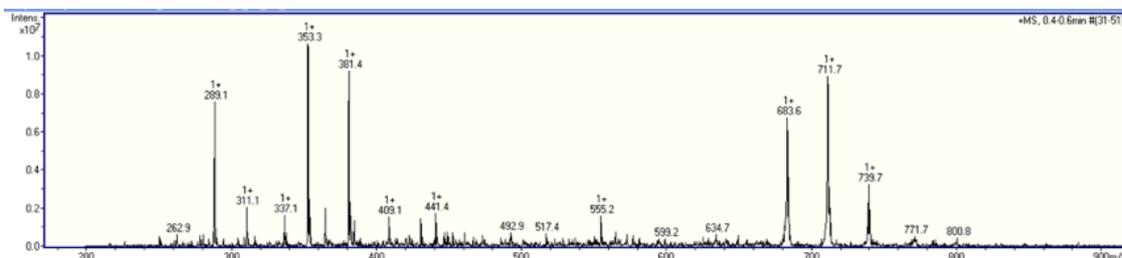

Figure 29: MS of crude **P4a-N1F** prepared by condition 2. Target peptide (4251 au) not isolated.

Condition 3: 0.1M sodium bicarbonate

P4a was next dissolved in 0.1 M sodium bicarbonate and **N1F** in water. The solutions were then combined and stirred at rt. After 24 h, a sample of the reaction mixture was removed and analysed by RP-HPLC and MS to reveal two major peaks (see Figure 30), with MS indicating the major product as **P4a** (M= 2330, with m/z signals= 337, 389, 466, 583) (see Figure 31). Once again peptide conjugation was unsuccessful.
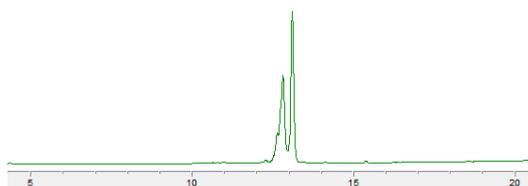

Figure 30: Analytical RP-HPLC C18 spectrum of crude **P4a-N1F** prepared by condition 3, 0-50% aq. ACN over 15 min (from 5-20 min), visualised at 254 nm.
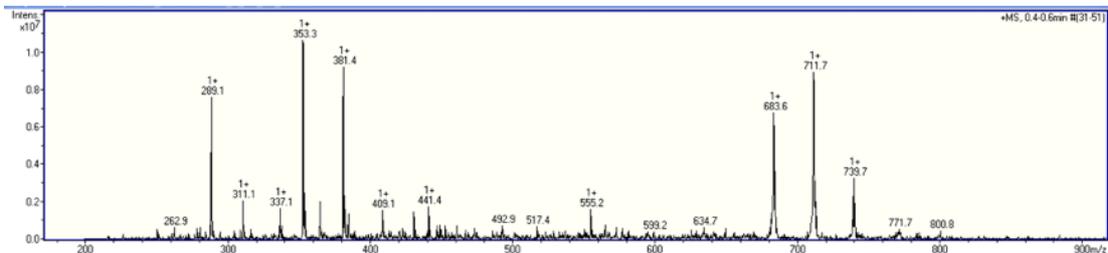
Figure 31: MS of crude **P4a-N1F** prepared by condition 1. Target peptide (4251 au) not isolated.

The three conditions trialled (condition 1, condition 2 and condition 3) for conjugation of **P4a** and **N1F** failed to give **P4a-N1F**. Conditions 1 and 2, which used phosphate buffer as the solvent, resulted in unidentified product by MS. Condition 3, in sodium bicarbonate, returned starting material as identified by MS. In order to ensure these results were not sequence dependent, several other thiol-chloroacetyl conjugations were trialled with different NLS and p21-peptides. These experiments simply gave unidentified products and/or starting material as confirmed by MS, indicating conjugation was not successful (see Table 8 entries 4-7). This suggests that thiol conjugation via a chloroacetyl moiety was incompatible with the p21-peptides and NLS peptides selected. This may be due to chloride being a poor leaving group in comparison to other halogens. [78, 79] Our efforts were then turned to substituting the chloroacetyl moiety with another more reactive thiol-reactive group.

Table 8: Summary of various thiol-chloroacetyl conjugation attempts between chloroacetylated p21-derived peptides and thiol-containing NLS tags. Target peptides: **P4a-N3F** (4194 au), P1a-N1F (3441 au), **P1a-N3F** (3099 au), **P1a-N2F** (3013 au), **P3a-N1F** (3754 au), **P3a-N4F** (3928 au).

Peptides	Conjugation conditions	Mass (au)	*m/z* peaks	ID	RP-HPLC
P4a N1F	100 mM phosphate buffer (pH 7.8), rt, 24 h	4678	391, 520, 781	Unidentified	Split peak spanning 7.5-8.2 min (Figure 26)
P4a N1F	0.1 M sodium bicarbonate, rt, 24 h	2330	337, 389, 466, 583	P4a	Split peak spanning 7.5-8.2 min (Figure 28)
P4a N1F	100 mM phosphate buffer (pH 7.8) + 2% TCEP, rt, 24 h	*	*	Unidentified	Split peak spanning 7-8 min (Figure 30)
P1a N1F	100 mM phosphate buffer (pH 7.8) + 0.02% TCEP, rt, 24 h	1576	316, 526, 788	P1a	Multiple small and broad peaks at 8-14 min, major sharp peak at 10.5 min (Figure S11)
P1a N3F	100 mM phosphate buffer (pH 7.8) + 0.02% TCEP, rt, 24 h	1559	391, 521	N1F	Multiple, small peaks spanning 8.5-13 min (Figure S12)
		1541	514, 771	N1F with aspartimide formation	
		3098	620, 775	Product mass	
P1a N2F	100 mM phosphate buffer (pH 7.8) + 0.02% TCEP, rt, 24 h	1445	482, 724	N2F	Small signals spanning 9-11 min (Figure S13)
		1413	354, 472, 708	Lanthionine formation from cysteine of N2F	
		1789	448, 597, 896	Unidentified	
P3a N1F	0.1 M sodium bicarbonate + DMSO	*	*	Unidentified	Major peak at 9.5 min which overlaps with P3a HPLC peak (Figure S14)
P3a N4F	0.1 M sodium bicarbonate + DMSO	*	*	Unidentified	Major peak at 9.5 min which overlaps with P3a HPLC peak (Figure S15)

*No peptide mass was observed

3.2.2 Thiol conjugation via maleimide moiety

Thiol conjugation via a maleimide moiety was next investigated for the preparation of NLS tagged p21-peptides (see Scheme 4). Maleimide-thiol conjugation is commonly used in functionalisation strategies due to the highly reactive nature of maleimide groups and their high selectivity for thiols.[81]
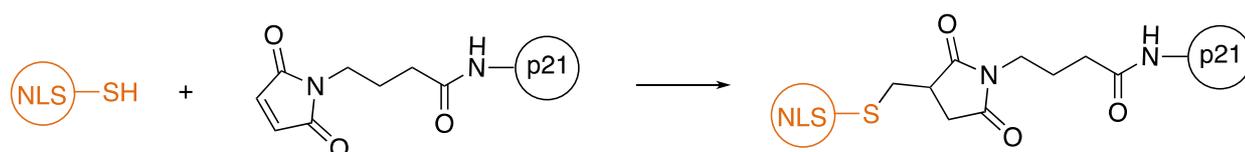
Scheme 4: General scheme of peptide conjugation via a thiol containing NLS peptide and an *N*-alkyl maleimide containing p21-derived peptide.

The shorter **P1** sequence was chosen as the p21-derived peptide segment for these studies. The *N*-terminal of **P1** was functionalised with a maleimide group, giving **P1b** (see Table 9). The NLS peptides; **N1F, N2F, N3F** and **N4F** (see Table 9) were prepared. **P1b** was then conjugated to each of the four NLS peptides to give **P1b-N1F, P1b-N2F**, **P1b-N3F** and **P1b-N4F** (see Table 10).

Table 9: Sequences of the peptide segments prepared for thiol conjugation to give linear P1-NLS peptides. Mal= maleimide.

p21-peptide	Sequence	NLS peptide	Sequence	Conjugate Peptide Code
P1b	mal-RQTSMTDFYHSK-NH$_2$	**N1F**	FITC-A$_\beta$GRKKRRQRRRC-NH$_2$	**P1b-N1F**
		N2F	FITC-A$_\beta$-PKKKRKVC-NH$_2$	**P1b-N2F**
		N3F	FITC-A$_\beta$PAAKRVKLDC -NH$_2$	**P1b-N3F**
		N4F	FITC-A$_\beta$RRWWRRWRRC-NH$_2$	**P1b-N4F**

Table 10: Peptide sequences of linear P1-NLS peptides. Suc= succinimide group.

Peptide	Sequence
P1b-N1F	FITC-A$_\beta$GRKKRRQRRRC[suc-RQTSMTDFYHSK-NH$_2$]
P1b-N2F	FITC-A$_\beta$-PKKKRKVC[suc-RQTSMTDFYHSK-NH$_2$]
P1b-N3F	FITC-A$_\beta$PAAKRVKLDC[suc-RQTSMTDFYHSK-NH$_2$]
P1b-N4F	FITC-A$_\beta$RRWWRRWRRC[suc-RQTSMTDFYHSK-NH$_2$]

P1b was synthesised on-resin by Fmoc/tBu-SPPS (described in Experimental) (see Scheme 1 of Chapter 2). The maleimide group was introduced into **P1b** after the final Fmoc-deprotection by treating the resin with a solution of 4-maleimidobutyric acid (3 equiv), DIPEA (10 equiv) and HATU (3 equiv) in DMF for 1 h (see Scheme 5). The maleimide functionalised peptide was then cleaved from resin and globally deprotected by treatment with a solution of TFA (95%), TIPS (2.5%) and H$_2$O (2.5%) for 2 h. DODT (2,2'-(ethylenedioxy)diethanethiol), a commonly used scavenger of t-butyl cations was excluded from the TFA cocktail in this instance to prevent reaction with the maleimide group of **P1b** as it contains a thiol group. The crude sample of **P1b** was purified by semi-preparatory RP-HPLC and identity confirmed by HRMS. **N1F**, **N2F**, **N3F** and **N4F** were synthesised on-resin by Fmoc/tBu-SPPS (described in Experimental and see Scheme 1 of Chapter 2). After the final Fmoc-deprotection, the fluorescein tag was appended by treating the resin with a solution of fluorescein-5-isothiocyanate (5 equiv) and DIPEA (10 equiv) in DMF for 2 h. **N1F**, **N2F**, **N3F** and **N4F** were cleaved from the resin and globally deprotected by treating the resin with a solution of TFA (92.5%), DODT (2.5%), TIPS (2.5%) and water (2.5%) for 5 h. This extended cleavage time was used to ensure complete removal of Pbf and tBu protecting groups as noted for **N1F**, **N2F**, **N3F** and **N4F** during previous syntheses. The resulting **N1F**, **N2F**, **N3F** and **N4F** peptides were purified semi-preparative RP-HPLC and identities confirmed by HRMS.
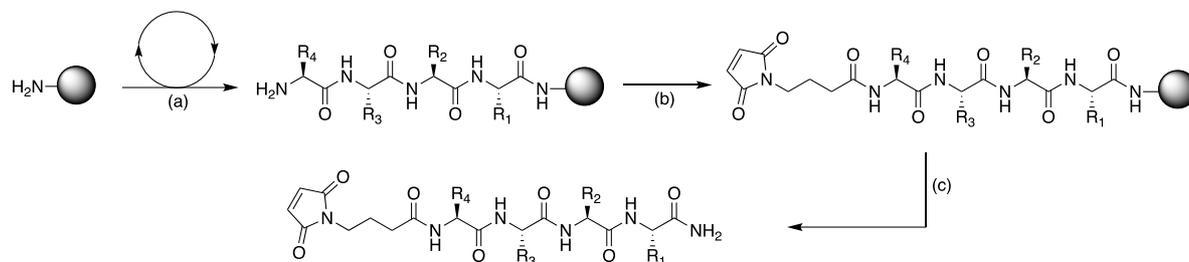
Scheme 5: General scheme for synthesis of maleimide-functionalised peptide. *Reagents and conditions:* a) repeated steps of coupling (Fmoc-protected amino-acid, HATU, DIPEA, DMF, 1 h) and deprotection (20% piperidine, 0.5 M HOBt, DMF, 15 min), b) 4-maleimidobutyric acid, HATU, DIPEA, DMF, 1 h c) TFA, DODT, TIPS, H_2O, 2 h. R= sidechain of amino acids.

N1F (15 mg) and **P1b** (1 equiv) were each dissolved in water (3 mL). The peptide solutions were then combined and stirred at rt. The reaction was monitored by sampling the reaction mixture and analysing by MS and RP-HPLC. After 24 h, the solution was diluted with water and lyophilised to give the crude peptide conjugate **P1b-N1F**. This was repeated with **N2F**, **N3F** and **N4F** to give **P1b-N2F**, **P1b-N3F** and **P1b-N4F**, respectively (see Scheme 6).

N1F	R' = FITC- A$_\beta$GRKKRRQRRRC
N2F	R' = FITC- A$_\beta$PKKKRKVC
N3F	R' = FITC- A$_\beta$PAAKRVKLDC
N4F	R' = FITC- A$_\beta$RRWWRRWRRC

P1b-N1F	R' = FITC- A$_\beta$GRKKRRQRRRC
P1b-N2F	R' = FITC- A$_\beta$PKKKRKVC
P1b-N3F	R' = FITC- A$_\beta$PAAKRVKLDC
P1b-N4F	R' = FITC- A$_\beta$RRWWRRWRRC
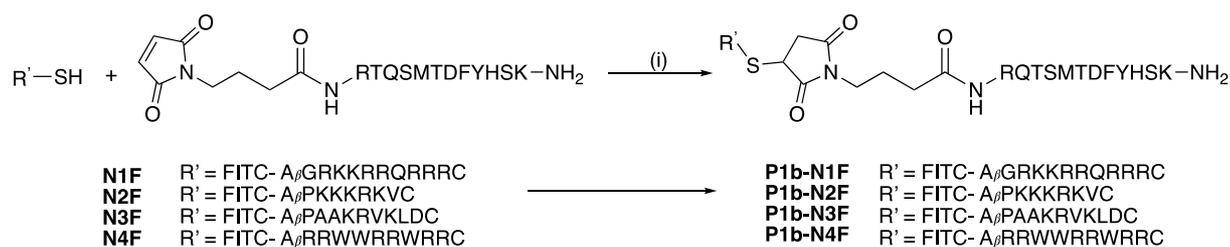
Scheme 6: peptide conjugation via cysteine containing peptide and maleimide-functionalised peptide. *Reagents and Conditions:* i) H_2O, rt, 24 h.

3.2.2.1 Purification of peptide conjugates

Following purification of **P1b-N3F** and **P1b-N2F** by semi-preparatory RP-HPLC, MS of the thus obtained samples of **P1b-N3F** and **P1b-N2F** revealed the desired peptide masses, 3224 au and 2834 au, respectively. A peptide with a mass 390 au less than the target peptide mass was also identified in both MS spectra. We hypothesised that fragmentation of the FITC-peptide bond was occurring in the MS and the mass loss of 390 au corresponded to the cleaved FITC group. FITC-modified peptides and proteins have been reported to undergo a MS-induced fragmentation.[82, 83]

m/z = 390
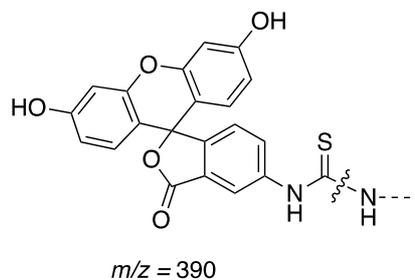
**Figure 32**: FITC-peptide bond fragmentation can lead to a free FITC tag with mass 390 au.
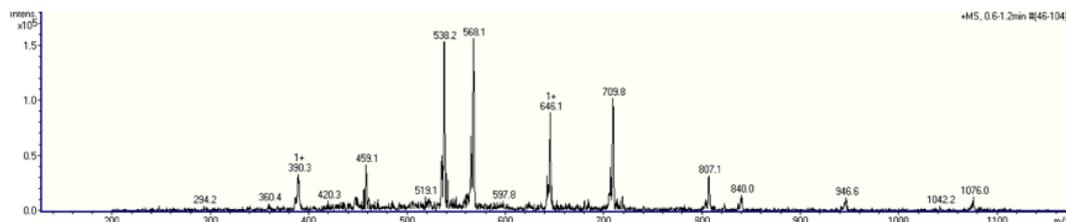
Figure 33: MS of purified **P1b-N3F**. Tagged peptide= 3224: 459 [M+7H]$^{7+}$, 538 [M+6H]$^{6+}$, 646 [M+5H]$^{5+}$, 807 [M+4H]$^{4+}$ *m/z.* Untagged peptide= 2834: 568 [M+5H]$^{5+}$, 709 [M+4H]$^{4+}$, 946 [M+3H]$^{3+}$ *m/z.*
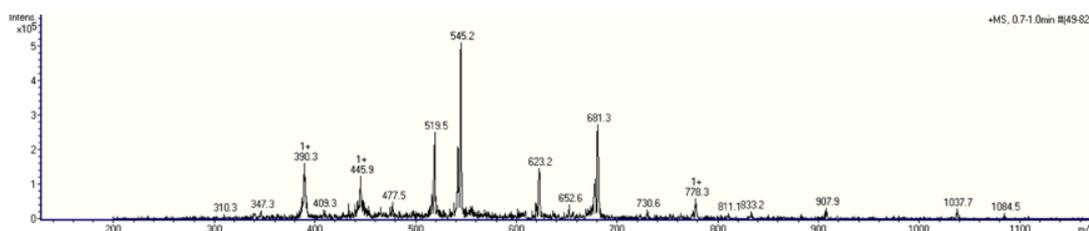
Figure 34: MS of purified **P1b-N2F**. Tagged peptide= 3111: 445 [M+7H]$^{7+}$, 519 [M+6H]$^{6+}$, 623 [M+5H]$^{5+}$, 778 [M+4H]$^{4+}$ *m/z.* Untagged peptide= 2721: 545 [M+5H]$^{5+}$, 681 [M+4H]$^{4+}$, 907 [M+3H]$^{3+}$ *m/z.*

The same **P1b-N3F** and **P1b-N2F** samples were then analysed by RP-HPLC to determine if the FITC tag was indeed attached to the peptide. RP HPLC traces of **P1b-N3F** and **P1b-N2F** each revealed one major peak (see Figure 35). This suggests that either FITC detachment was occurring in the MS and the only product observed in the RP-HPLC was the tagged peptides or the presence of both untagged and tagged peptides, coeluting in the RP-HPLC.

Figure 35: Analytical RP-HPLC spectra, 0-50% aq. ACN gradient over 15 minutes on a C18 column visualised at 254 nm of purified **P1b-N3F** (A) and purified **P1b-N2F** (B).

61

P1b-N3F and **P1b-N2F** were then analysed by MS under mild electrospray ionisation (ESI) conditions, in order to determine whether this would minimise FITC detachment, thus minimising the presence of untagged peptide in the MS. $[M+3H]^{3+}$ molecular ions were observed for both **P1b-N3F** and **P1b-N2F**, indicating FITC detachment could be avoided under gentle ESI conditions (see Figure 36). Thus, the samples were lyophilised to give purified **P1b-N3F** and **P1b-N2F**.
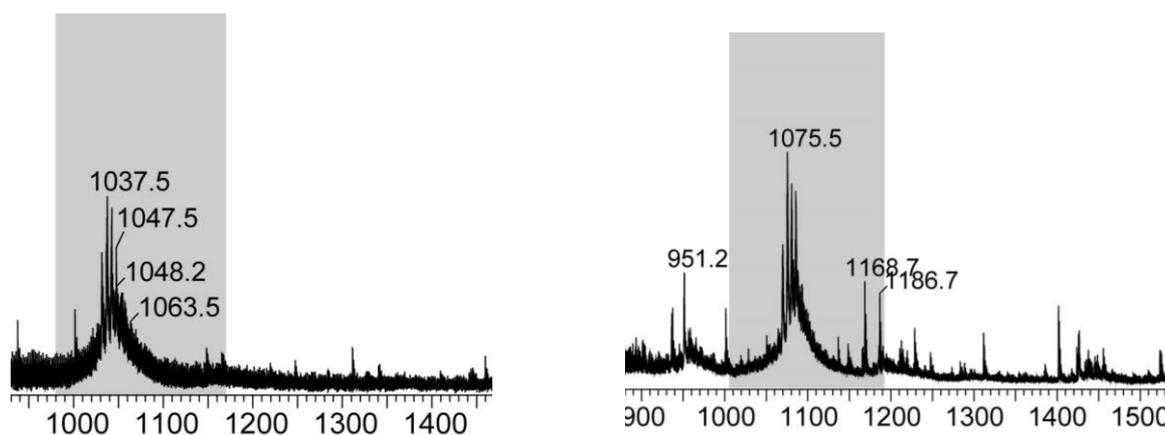

Figure 36: $[M+3H]^{3+}$ molecular ions of **P1b-N2F** (left) and **P1b-N3F** (right) obtained using a nano-electrospray ionisation source.

Interestingly, FITC detachment was not observed with **P1b-N3F** and **P1b-N4F** during MS analysis with standard ESI conditions. This suggests that the FITC-peptide thiourea bond in **P1b-N1F** and **P1b-N2F** is less susceptible to cleavage.

3.3   Binding affinity of linear peptides

K_D values were determined for **P1b-N1F**, **P1b-N2F**, **P1b-N3F** and **P1b-N4F** binding to PCNA by SPR to investigate the influence of the attached NLS tag (see Table 11). SPR experiments were conducted by immobilising PCNA on a gold sensor chip and flowing the peptide solutions over the chip as described in Section 2.3 and Experimental. **P1b-N1F** and **P1b-N3F** gave similar K_D values of 130 nM and 101 nM, respectively. This indicates that high affinity for PCNA is maintained with an NLS peptide appended to **P1** which binds PCNA with 102.3 nM affinity. Interestingly, **P1b-N2F** bound with the highest affinity of the

62

conjugates with a K_D of 33.1 nM, suggesting the **N2F** peptide provides improved PCNA binding, relative to **N1F** and **N3F**. **P1b-N4F**, on the other hand, had the lowest affinity for PCNA with a K_D of 373 nM.

In summary, appending **N1F** and **N3F** to **P1** did not drastically alter the PCNA binding affinity of the resulting peptides, whereas attachment of **N4F** to **P1** greatly reduced PCNA binding affinity. **N2F**, however, improved PCNA binding when attached to **P1**. These results indicate that the choice of NLS peptide for attachment to **P1** is important for maintaining PCNA binding affinity. This suggests the NLS sequences lead to different interactions between the peptide and PCNA, resulting in different binding affinities.

Table 11: Binding affinity of linear P1-NLS conjugates. K_D = binding affinity for PCNA. SE = standard error. χ^2 = chi^2 (gives a measure of accuracy of fitting).

Peptide Code	Affinity K_D (nM)	K_D SE (nM)	χ^2
P1b-N1F	101	9.45	3.60
P1b-N2F	33.1	4.23	2.63
P1b-N3F	130	15.1	0.778
P1b-N4F	373	29.9	3.70

3.4 Nuclear permeability assay of linear peptides

Uptake of **P1b-N1F**, **P1b-N2F**, **P1b-N3F** and **P1b-N4F** into the nucleus of MDA-MB-468 breast cancer cells was evaluated in a permeability assay by collaborator Zoya Kikhtyak[72] in order to determine which NLS peptide facilitates cell and nuclear permeability (see Figure 37) (see Experimental for details). The cells chosen express nuclear mKate, a red-fluorescent monomeric protein that localises in the nucleus, allowing for visualisation of the nucleus when imaged by confocal microscopy. Cells were dosed with 10 μM of peptide and incubated for 24 h before the cells were fixed on treating with formaldehyde and imaged.

Cells dosed with **P1b-N1F** and **P1b-N3F** showed some green fluorescence in the cell cytosol, corresponding to the fluorescein tagged peptide (surrounding the red marked nucleus),

indicating modest cell permeability of **P1b-N1F** and **P1b-N3F**, respectively. No green fluorescence, however, was observed in the nucleus, indicating both peptide conjugates were not nuclear permeable. **P1b-N4F** treated cells also did not show cell or nuclear accumulation of **P1b-N4F**, indicated by lack of green fluorescence corresponding to the peptide in the cell cytosol or nucleus. Interestingly, a decrease in red fluorescence corresponding to the mKate protein at the cell nuclei was observed in **P1b-N4F** treated cells. This indicated significant cell death, suggesting **P1b-N4F** was toxic to cells. Importantly, cells dosed with **P1b-N2F** showed green fluorescence, corresponding to the fluorescein tagged peptide within the nucleus, revealing **P1b-N2F** as nuclear permeable. This suggests the $SV40_{126-132}$ NLS peptide, **N2F**, was able to deliver the peptide into the nucleus.
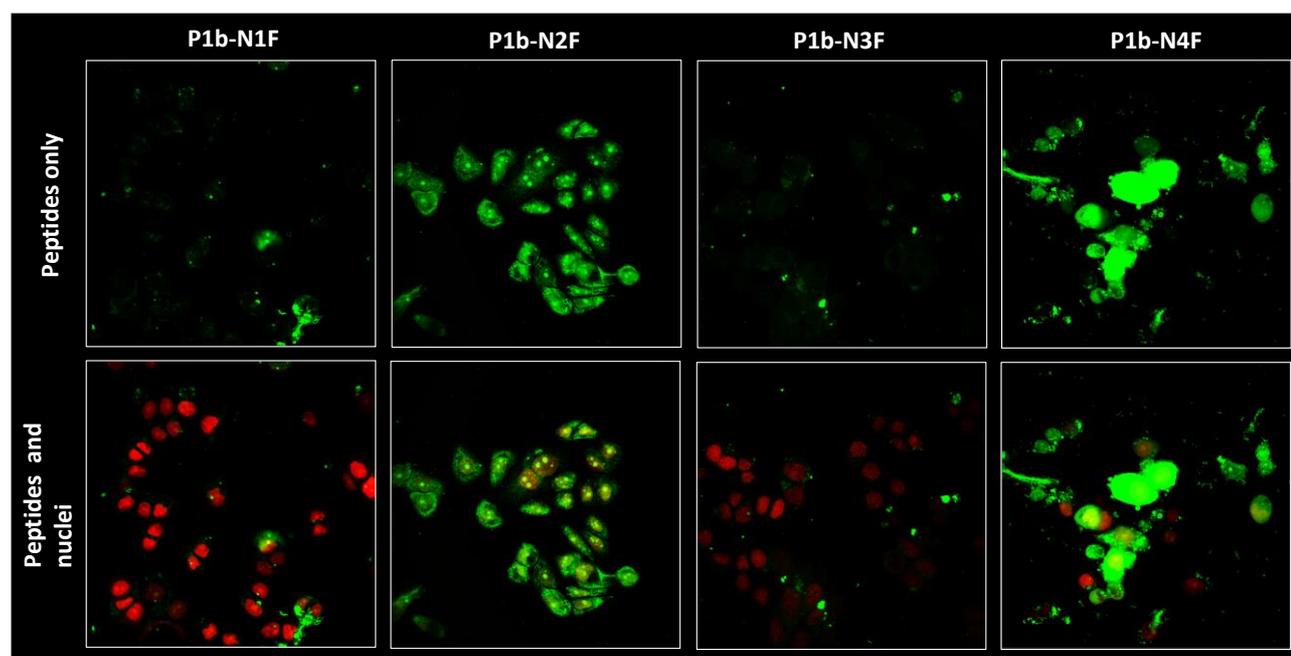

Figure 37: Confocal fluorescence microscopy images of MDA-MB-468 breast cancer cells incubated with 10 μM of peptide for 24 h. The peptides were covalently labelled with FITC (green). Cells are expressing nuclear mKate (red). Overlap of red and green fluorescence indicates the presence of peptide in nucleic of cells.

These results indicate **N1F** and **N3F** were able to facilitate cell entry as the **P1** peptide has previously been evaluated and found to lack cell permeability. **N2F** was able to impart cell entry and subsequent nuclear entry of the peptide conjugate, **P1b-N2F.**

3.5 Design of macrocyclic bimane peptides

A second series of NLS tagged p21-peptides was investigated next. Recent work in our lab by Aimee Horsfall[66] identified a cell permeable cyclised derivative of **P1**. The **P1** sequence was modified to include i and $i+4$ separated cysteines, for reaction with dibromobimane to give a macrocyclic peptide where the cysteine residues are linked via a bimane group (see Figure 38). This gave a **P1bim** with sequence of Ac-RQC(-)SMTC(Bim)FYHSK-NH$_2$ (see

Table **13**) (see structure in Scheme 7). Here, bimane-based linkers, which are thiol-specific, were used to crosslink cysteine-containing p21-peptides, resulting in the bimane linker acting as a structural constraint and a fluorescent tag. Bimane-based linkers have been reported to stabilise secondary structure within short peptides, thus bimane cyclising **P1**, stabilised the 3_{10}-helical turn that p21-peptides adopt upon binding PCNA (see Figure 39). [84] Additionally, **P1bim** was cell permeable in MDA-MB-468 cells. The bimane is inherently fluorescent and **P1bim** was imaged by fluorescence microscopy, without an auxiliary fluorescent tag (e.g. fluorescein). [84] Here, **P1bim** was conjugated to the each of the four previously explored NLS tags; Tat$_{48-57}$, cMyc$_{320-328}$, SV40$_{126-132}$ and R6W3.
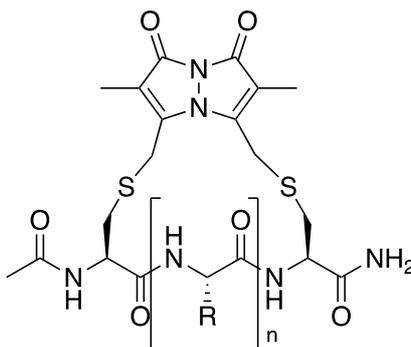

Figure 38: General structure of a bimane linked macrocyclic peptide.
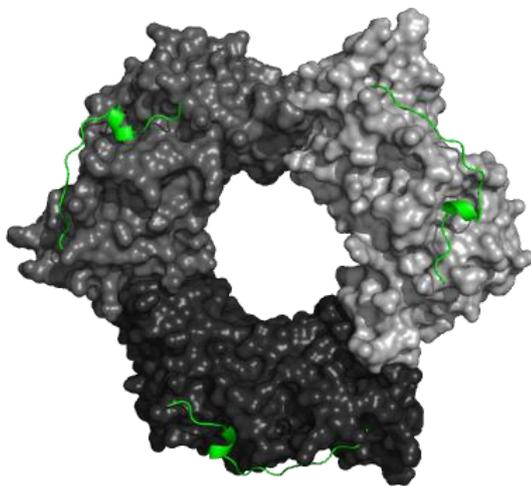
Figure 39: PCNA bound to three p21$_{139-160}$ peptides (green). The p21$_{139-160}$ peptide forms a 3$_{10}$-helix upon binding PCNA. 1AXC

In this work, a derivative of **P1bim** with an *N*-terminal maleimide, **P1c,** was first prepared (see Table 12). Samples of **P1c** were then separately conjugated to all four NLS peptides **N1F**, **N2F**, **N3F** and **N4F** by thiol conjugation (see Table **12**) to give **P1c-N1F**, **P1c-N2F**, **P1c-N3F** and **P1c-N4F** (see Table 14). These conjugates were then evaluated in a cell uptake assay to investigate the impact of the NLS peptides to the cell permeable **P1bim** on nuclear permeability. Additionally, a fluorescein tagged analogue of **P1bim**, **P1bimF**, was prepared in order to determine if a fluorescein tag influences uptake (see Table 13).

3.6   Synthesis of macrocyclic bimane peptides

P1bimF and **P1c** were synthesised on-resin by Fmoc/tBu-SPPS (described in Experimental). In brief, Fmoc-protected amino acids were coupled to resin-bound peptide with addition of the Fmoc-protected amino-acids (5 equiv), HATU (5 equiv) and DIPEA (10 equiv) in DMF, to the resin for 1 h. On-resin Fmoc-deprotections were achieved with 15 min treatment with 20% piperidine and 0.5 M HOBt. The synthesis of **P1bimF** and **P1c** required introduction of 2-methoxytrityl (Mmt)-protected cysteines. This was achieved by coupling Fmoc-L-Cys (Mmt)-OH into desired position by standard coupling conditions. Successive amino acid couplings and Fmoc-deprotections were carried out and the desired peptide sequence was

achieved, **1** (see Scheme 7). Prior to the final Fmoc-deprotection, Mmt-protecting groups were first removed by successive treatments of the resin with 2% TFA in DCM, to liberate the thiol sidechains, to give **2**. On-resin cyclisation was then achieved by adding a solution of dibromobimane (1.5 equiv) and DIPEA (3 equiv) in DMF to the resin for 3 h to give **3**. Final Fmoc-deprotection was performed, to give **4**, an *N*-deprotected, bimane cyclised resin-bound peptide. At this point in the synthesis of **P1c**, the maleimide group was then introduced by addition of a solution with 4-maleimidobutyric acid (3 equiv), DIPEA (10 equiv) and HATU (3 equiv) in DMF to the resin for 1 h. The maleimide functionalised peptide was then cleaved from the resin and globally deprotected by treatment with a solution of TFA (95%), TIPS (2.5%) and water (2.5%) for 2 h, to give **P1c**.

In the case of **P1bimF**, upon reaching the peptide **4** and coupling a β-Alanine residue, a solution of fluorescein-5-isothiocyanate (5 equiv) and DIPEA (10 equiv) in DMF was added to the resin for 2 h. The fluorescein tagged peptide was cleaved from the resin and globally deprotected, by treatment of the resin with TFA (92.5%), TIPS (2.5%), DODT (2.5%) and water (2.5%) for 2 h, to give **P1bimF**. **P1c** and **P1bimF** were purified by semi-preparatory RP-HPLC and identity confirmed by HRMS.
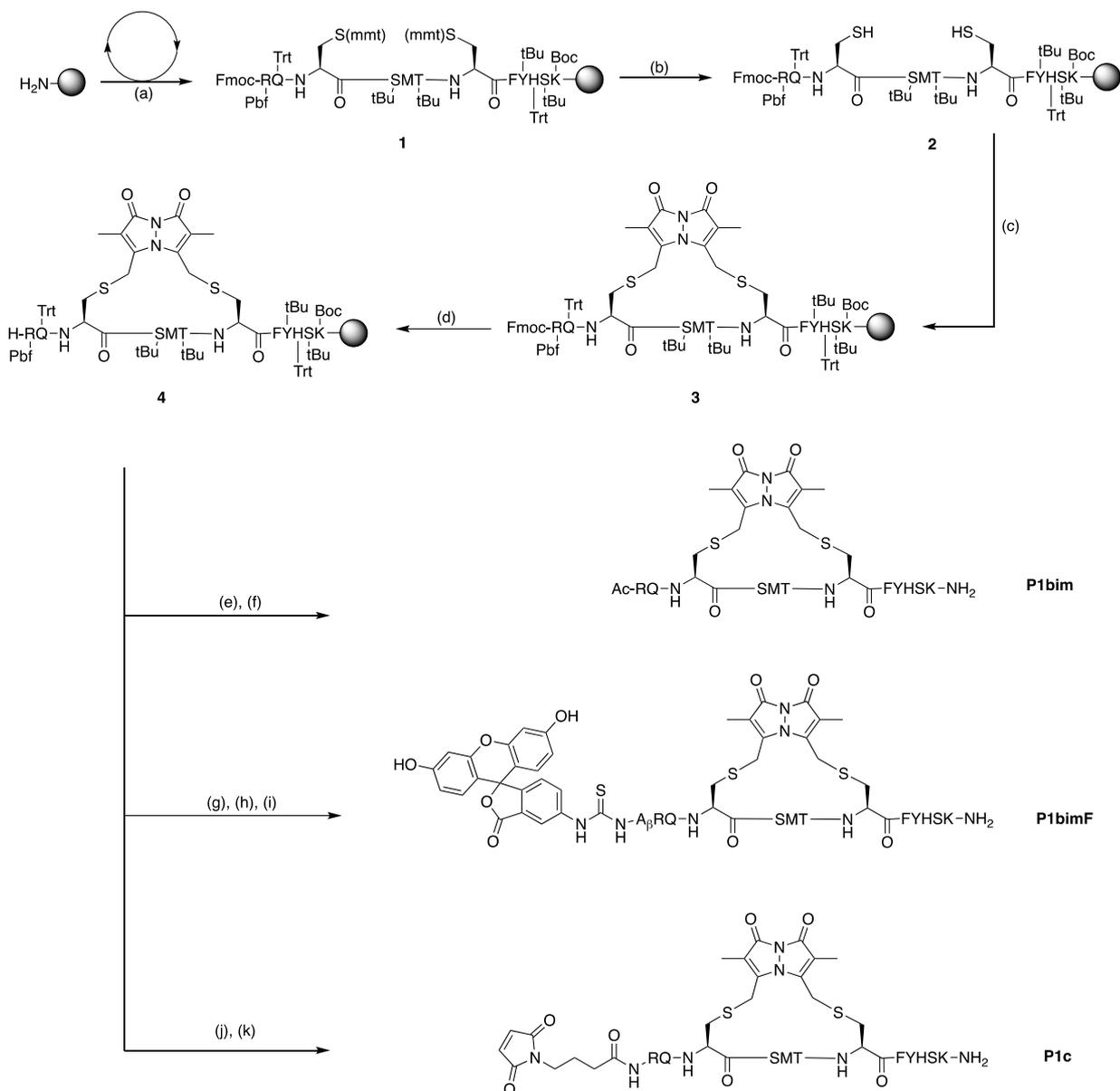
Scheme 7: Synthesis of **P1c** and **P1bimF**. *Reagents and conditions:* a) repeated steps of coupling (Fmoc-protected amino-acid, HATU, DIPEA, DMF, 1 h) and deprotection (20% piperidine, 0.5 M HOBt, DMF, 15 min); b) 2% TFA in DCM; repeat until colourless upon addition; c) dibromobimane, DIPEA, DMF, 3 h; d) 20% piperidine, DMF, 15 min; e) acetic anhydride, DIPEA, DMF, 15 min; f) 92.5% TFA, 2.5% DODT, 2.5% TIPS, 2.5% water, 2 h; g) Fmoc-βAla-OH, HATU, DIPEA, DMF, 1h, followed by 20% piperidine, DMF, 15 min; h) fluorescein-5-isothiocyanate, DIPEA, DMF, 2 h; i) 92.5% TFA, 2.5% DODT, 2.5% TIPS, 2.5% water, 2 h; j) 4-maleimidobutyric acid, HATU, DIPEA, DMF, 1 h; k) 95% TFA, , 2.5% TIPS, 2.5% water, 2 h.

Table 12: Sequences of the peptide segments prepared for thiol conjugation to give bimane cyclic P1-NLS peptides.

p21-peptide	Sequence	NLS peptide	Sequence	Conjugate Peptide Code
P1c	mal-RQC(-)SMTC(Bim)FYHSK-NH₂	N1F	FITC-A$_\beta$GRKKRRQRRRC-NH₂	**P1c-N1F**
		N2F	FITC-A$_\beta$-PKKKRKVC-NH₂	**P1c-N2F**
		N3F	FITC-A$_\beta$PAAKRVKLDC -NH₂	**P1c-N3F**
		N4F	FITC-A$_\beta$RRWWRRWRRC-NH₂	**P1c-N4F**

Table 13: Peptide sequence of macrocyclic bimane peptides. See structures in Scheme 7.

Peptide	Sequence
P1bim	(3,7-cyclo)-Ac-RQC(-)SMTC(Bim)FYHSK-NH$_2$
P1bimF	(4,8-cyclo)-FITC-A$_\beta$RQC(-)SMTC(Bim)FYHSK-NH$_2$

Table 14: Peptide sequences of macrocyclic bimane conjugate peptides. Suc= succinimide group.

Peptide Code	Sequence
P1c-N1F	FITC-A$_\beta$GRKKRRQRRRC[suc-RQC(-)SMTC(Bim)FYHSK-NH$_2$]
P1c-N2F	FITC-A$_\beta$PKKKRKVC[suc-RQC(-)SMTC(Bim)FYHSK-NH$_2$]
P1c-N3F	FITC-A$_\beta$PAAKRVKLDC[suc-RQC(-)SMTC(Bim)FYHSK-NH$_2$]
P1c-N4F	FITC-A$_\beta$RRWWRRWRRC[suc-RQC(-)SMTC(Bim)FYHSK-NH$_2$]
P1c-N2	H-PKKKRKVC[suc-RQC(-)SMTC(Bim)FYHSK-NH$_2$]

N1F (15 mg) and **P1c** (1 equiv) were each dissolved in water (3 mL). The peptide solutions were then combined and stirred at rt. The reaction was monitored by sampling the reaction mixture and analysing by MS and RP-HPLC. After 24 h, the reaction mixture was diluted with water and lyophilised to give the crude peptide conjugate **P1c-N1F**. This was repeated with **N2F**, **N3F** and **N4F** to give **P1c-N2F**, **P1c-N3F** and **P1c-N4F**, respectively (see Scheme 8). The peptides were purified by semi-preparative RP-HPLC and identity verified by HRMS.
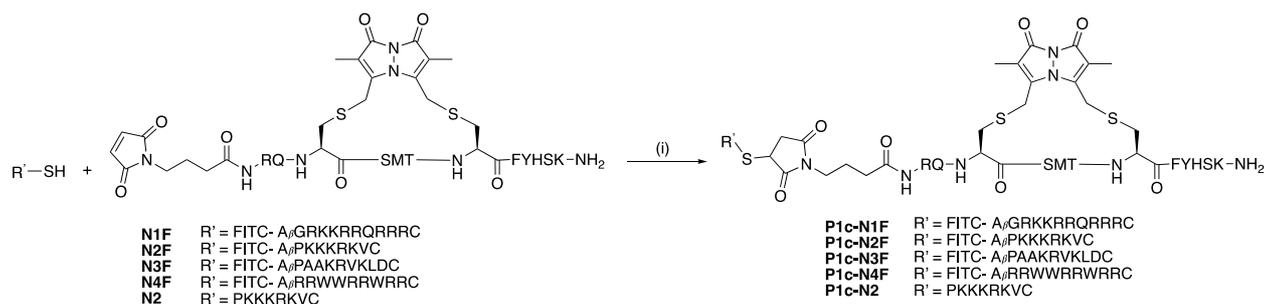
Scheme 8: Peptide conjugation scheme of bimane cyclised P1c and NLS tags, to give **P1c-N1F**, **P1c-N2**, **P1c-N2F**, **P1c-N3F**, **P1c-N4F**. *Reagents and conditions:* i) H$_2$O, 24 h, rt.

3.7   Nuclear permeability assay of macrocyclic bimane peptides

Firstly, as a control, uptake of **P1bim** into MDA-MB-468 breast cancer cells was evaluated in a permeability assay by collaborator Zoya Kikhtyak[72], as previously described in Section 3.4

(see Experimental for details). **P1bim** treated cells displayed blue fluorescence, corresponding to the bimane linked peptide, within the cell cytosol, confirming **P1bim** is cell permeable and not nuclear permeable (see Figure 40).
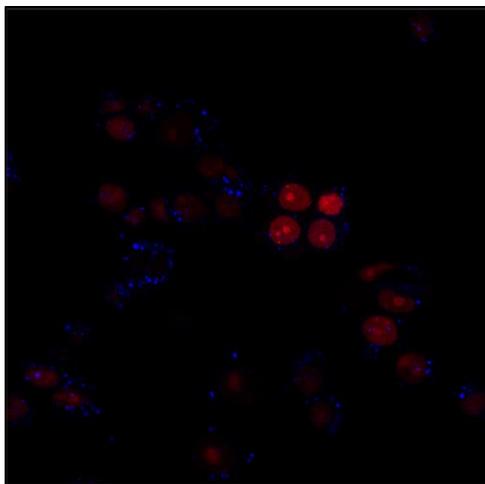

Figure 40: Confocal fluorescence microscopy images of MDA-MB-468 breast cancer cells treated with 10 µM of **P1bim**, incubated for 24 h, fixed and imaged. Blue fluorescence (bimane). Red fluorescence (nucleus). Overlap of blue and red fluorescence indicates presence of peptide in nucleus.

P1c-N1F, **P1c-N2F**, **P1c-N3F**, **P1c-N4F** and **P1bimF** were then tested for nuclear permeability in the same cell uptake assay (see Figure 41). **P1c-N1F** and **P1c-N3F** treated cells both displayed green fluorescence, corresponding to the fluorescein tagged peptide, in the cell cytosol (around the red marked nucleus). This indicated modest cellular accumulation of **P1c-N1F** and **P1c-N3F** but no nuclear accumulation, which was also seen with the linear analogues, **P1b-N1F** and **P1b-N3F**, respectively. These results suggest **N1F** and **N3F** were unable to impart nuclear entry of a linear (**P1b**) or cyclic (**P1c**) p21-peptide.

P1c-N4F, like its linear analogue **P1b-N4F**, displayed toxicity to the treated cells, with few cells remaining viable post treatment as evidenced by decreased red fluorescence corresponding to the nucleus. This indicates that the **P1** sequence, linear and bimane cyclised is toxic to cells when conjugated to **N4F**.

Importantly, cells dosed with **P1c-N2F** displayed green fluorescence, corresponding to the fluorescein tagged peptide in the nucleus, indicating **P1c-N2F** was nuclear permeable. This suggests the fluorescein tagged $SV40_{126-132}$ NLS imparted nuclear permeability to **P1bim**. Cells dosed with **P1bimF** also displayed green fluorescence in the nucleus, indicating nuclear accumulation of **P1bimF**. This suggests that the attachment of just the fluorescein tag, in **P1bimF**, was able to confer nuclear permeability, indicating the fluorescein tag does impact the cell uptake of the peptide.

A derivative of **P1c-N2F** without the fluorescein tag was prepared to further investigate how fluorescein labelling may affect nuclear permeability. **P1c** and **N2** were synthesised and conjugated as described in Section 3.6 to give **P1c-N2** (see Table 12 and Table 14) which was then subjected to the same cell uptake assay. Interestingly, cells treated with **P1c-N2** displayed blue fluorescence, corresponding to macrocyclic bimane peptide, in the cell cytoplasm but not the nucleus, indicating **P1c-N2** is not nuclear permeable. This suggests that attachment of the untagged $SV40_{126-132}$ NLS to **P1bim**, in **P1c-N2**, was not able to confer nuclear permeability, however, attachment of fluorescein tagged $SV40_{126-132}$ NLS, in **P1c-N2F**, was able to facilitate nuclear entry. The permeability results of both **P1bim** and **P1c-N2**, compared to the corresponding fluorescein tagged analogues, **P1bimF** and **P1c-N2F**, indicate incorporating a fluorescein tag may promote nuclear uptake in this instance. This demonstrates that fluorescent tags can drastically impact the properties of a peptide. Future studies should be conducted with the peptide itself, without an appended auxiliary tag, to accurately measure the *in vitro* activity of the peptide itself.
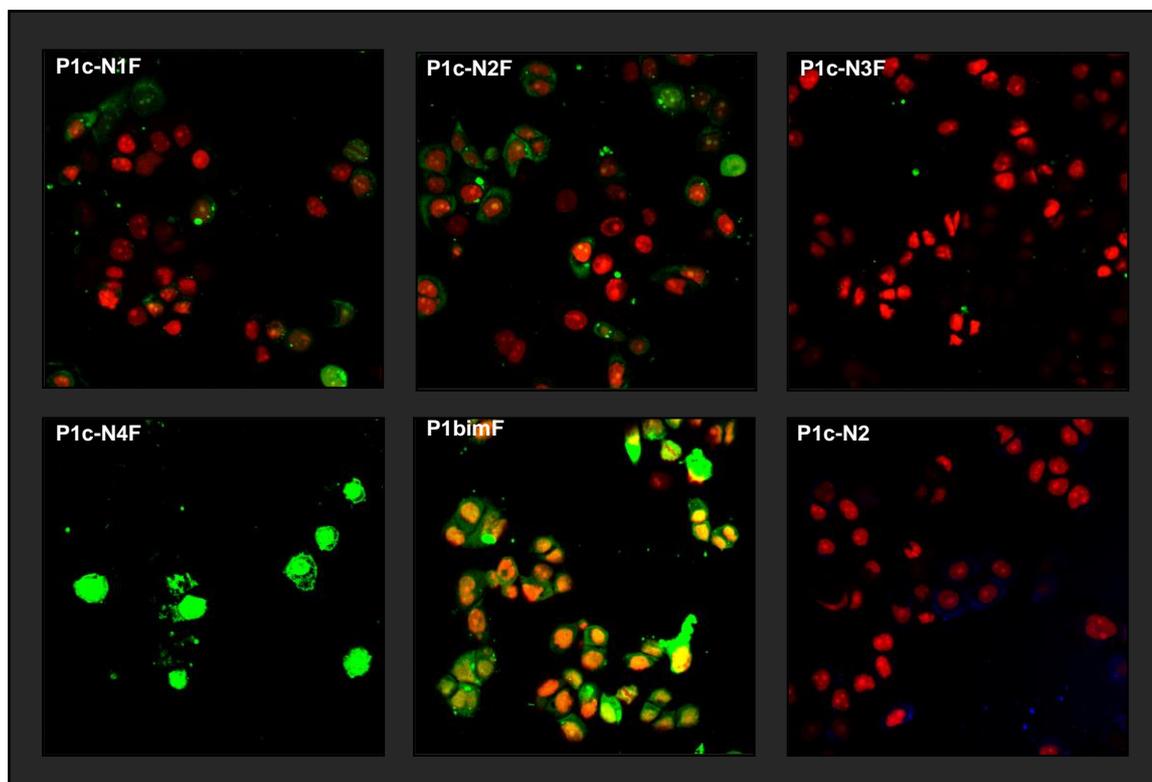
Figure 41: confocal fluorescence microscopy images of MDA-MB-468 breast cancer cells treated with 10 μM of **P1c-N1F**, **P1c-N2F**, **P1c-N3F**, **P1c-N4F**, **P1bimF** or **P1c-N2** for 24 h. All peptides excluding **P1c-N2** were covalently labelled with FITC (green fluorescence). **P1c-N2** was tagged with a bimane staple (blue fluorescence). Cells are expressing nuclear mKate (red fluorescence). Overlap of green or blue fluorescence (peptide) and red fluorescence (nuclear mKate protein) indicates presence of peptide in nucleus.

The same cell uptake experiment was then repeated with **P1bimF** and **P1c-N2F**, where treated cells were imaged under the fluorescein channel (490-534 nm), as well as a bimane channel (410-485 nm) and overlaid with images of the cell nuclei (600-700 nm) (see Figure 42). This was done to verify the peptide conjugates were intact such that the fluorescein tag or fluorescein tagged peptide remained attached to the macrocyclic bimane peptide. **P1bimF** and **P1c-N2F** treated cells both displayed substantial colocalization of the bimane and fluorescein fluorescence suggesting the peptide conjugate remained intact and was accumulating within the nucleus. Both samples display blue fluorescence corresponding to the macrocyclic bimane peptide surrounding the nucleus, in the cell cytosol. These results may suggest the fluorescein tag or fluorescein tagged peptide have been cleaved and the bimane peptide cannot enter the nucleus. One explanation for this may be proteases within the cell samples breaking down the fluorescein-peptide bond. In the case of **P1c-N2F**, the

72

thioether bond formed by conjugation could have undergone hydrolysis converting the peptide conjugate back to the starting **P1c** and **N2F** peptides. [85, 86]

Whilst the thioether linkage between the NLS peptide and p21-peptide is important to transport the p21-peptide to the nucleus, once in the nucleus, only the p21-peptide is needed for subsequent targeting of PCNA, thus detachment of the NLS peptide is not necessarily unfavourable. Studying the *in vitro* activity of these conjugate peptides against cancer cells would provide insight into whether the p21-derived segment of the peptide conjugate is entering the nucleus where it can inhibit cell proliferation.
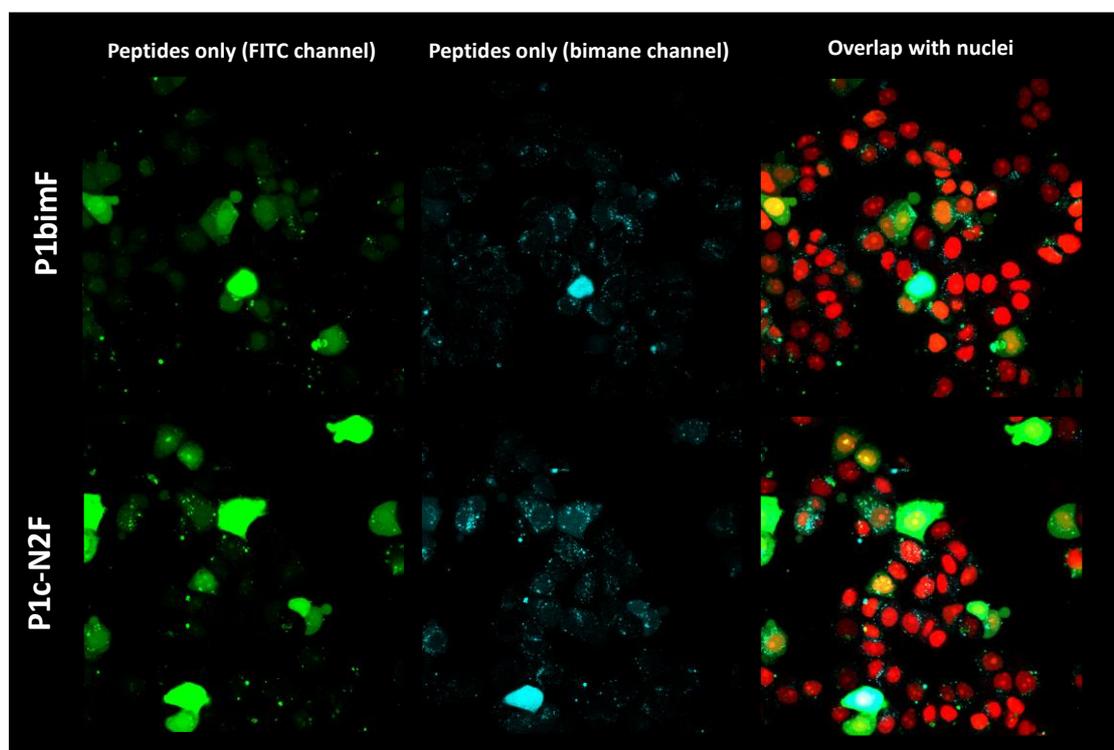

Figure 42: confocal fluorescence microscopy images of MDA-MB-468 breast cancer cells treated with 10 μM of **P1bimF** or **P1c-N2F** for 24 h.

Importantly, **P1bimF** and **P1c-N2F** have been identified as nuclear permeable, highlighting interesting leads to developing a cell and nuclear permeable PCNA inhibitor. In contrast to other reported NLS tagged p21-peptides, [23, 29, 87] **P1bimF** and **P1c-N2F** contain a much shorter p21-derived scaffold, which enhances the druglikeness of both. The cyclic nature of **P1bimF** and **P1c-N2F** is also desirable as cyclic peptides are reported to generally have

improved proteolytic, *in vitro* stability and structural stability, relative to linear peptides.[84, 88] **P1bim** and **P1c-N2F** are both promising candidates for developing a viable pre-clinical peptide-based cancer therapeutic that targets and inhibits PCNA.

3.8 Binding affinity of macrocyclic bimane peptides

The K_D constants were determined for **P1bim**, **P1bimF**, **P1c-N2** and **P1c-N2F** binding to PCNA by SPR to investigate how cyclising the p21-derived peptide might affect binding affinity (see Table 15). SPR experiments were conducted as described in Section 2.3 (see Experimental for details). **P1bimF** and **P1c-N2F** were selected for SPR as they were found to be nuclear permeable, thus potential candidates for *in vitro* activity studies, which would require PCNA binding to exhibit inhibitory activity. The untagged analogues, **P1bim** and **P1c-N2**, were also evaluated by SPR, to determine how a fluorescein tag may affect PCNA binding.

The K_D value for **P1bim** was determined to be 570 nM affinity, whereas the linear **P1** bound PCNA with 102 nM affinity. This indicates that a macrocyclic bimane peptide, as in **P1bim**, reduced PCNA binding affinity. **P1bimF** bound PCNA with a K_D of 25 μM, indicating attachment of a fluorescein tag resulted in a severely reduced affinity for PCNA, relative to **P1bim**. The reduced binding affinity of **P1bimF** suggests the PCNA binding surface was not able to accommodate the large fluorescein tag at that position.

P1c-N2 bound with the highest affinity of the macrocyclic bimane peptides with a K_D of 176 nM, suggesting attachment of the $SV40_{126-132}$ NLS to **P1bim** greatly improved the PCNA binding affinity. Attachment of a fluorescein tagged $SV40_{126-132}$ to **P1bim**, as in **P1c-N2F**, however, lowered the binding affinity to 535 nM, a trend also seen with **P1bim** and **P1bimF**. This reinforces that attachment of a fluorescein tag reduces PCNA binding affinity. **P1c-N2F**

displayed a higher affinity for PCNA than **P1bimF**. One explanation for this may be that attachment of a fluorescein tagged NLS to **P1bim** placed the fluorescein in a more accommodating position in the PCNA binding surface, as opposed to attachment of the tag on **P1bim** directly, which led to an increased binding affinity.

The linear **P1b-N2F** showed the highest affinity for PCNA of all peptide conjugates (linear and macrocyclic bimane conjugates) with a K_D of 33.1 nM. **P1c-N2** displayed the highest affinity for PCNA of the bimane cyclised peptides, at 176 nM. Attachment of the $SV40_{126-132}$ peptide improved PCNA binding affinity in both linear and macrocyclic peptide set. This indicates that the $SV40_{126-132}$ NLS is the most favourable of the four NLS peptides chosen, to impart high affinity PCNA binding. This is important in an *in vitro* setting where the peptide must translocate to the nucleus and still bind PCNA tightly to exhibit its therapeutic effect.

Table 15: Binding affinity of macrocyclic bimane peptides for PCNA. K_D = binding affinity for PCNA. SE = standard error. χ^2 = chi^2 (gives a measure of accuracy of fitting).

Peptide Code	Affinity K_D (nM)	K_D SE (nM)	χ^2
P1bim	570	30.2	0.118
P1bimF	25.2 µM	1.95 µM	0.0547
P1c-N2F	535	73.8	4.92
P1c-N2	176	27.1	5.58

3.9   Chapter conclusions

A series of linear NLS tagged **P1** peptides were successfully synthesised via a thiol conjugation approach to give **P1b-N1F**, **P1b-N2F**, **P1b-N3F** and **P1b-N4F**.

SPR analysis of each peptide conjugate found **P1b-N2F** displayed the highest affinity for PCNA at 33.1 nM. **P1b-N1F, P1b-N3F** and **P1b-N4F** bound PCNA with K_D values of 101, 130 and 373 nM affinity, respectively. **P1b-N2F** exhibited a significant improvement in

affinity for PCNA from **P1**, indicating that appending the **N2F** tag to **P1** improved binding to PCNA.

P1b-N1F, P1b-N2F, P1b-N3F and **P1b-N4F** were investigated in a cell uptake assay to reveal **P1b-N1F** and **P1b-N3F** showing modest cell permeability and **P1b-N2F** displaying cell and nuclear permeability. These results indicate that only the $SV40_{126-132}$ NLS peptide is able to facilitate transport of the **P1** segment into the cell and nucleus.

A series of macrocyclic NLS tagged **P1** peptides were synthesised which contained a bimane cyclised p21-peptide, **P1bim**. **P1bim** was functionalised with a maleimide at the *N*-terminus and conjugated separately to each of the four NLS tags to give **P1c-N1F**, **P1c-N2F**, **P1c-N3F** and **P1c-N4F**.

P1c-N1F, **P1c-N3F**, **P1c-N4F** and **P1c-N2** all lacked nuclear permeability with **P1c-N4F** displaying toxicity to the cells. Importantly, **P1c-N2F** and **P1bimF** exhibited nuclear permeability. This indicates that the fluorescein tagged $SV40_{126-132}$ NLS is able to deliver **P1bim** to the nucleus. Interestingly, the untagged counterparts, **P1c-N2** and **P1bim**, only displayed cell permeability but not nuclear permeability, indicating the fluorescein tag alters the nuclear entry of **P1c-N2F** and **P1bimF**.

The K_D values for **P1bim**, **P1bimF**, **P1c-N2** and **P1c-N2F** binding to PCNA were 570 nM, 25 µM, 33.1 nM and 535 nM, respectively. We found fluorescein labelling, in **P1bimF** and **P1c-N2F**, lowers the binding affinity for PCNA, in contrast to the untagged analogues. Importantly, appending an $SV40_{126-132}$ NLS to **P1bim**, in **P1c-N2**, improved PCNA binding. This demonstrates that high affinity binding for PCNA can still be maintained with a short bimane cyclised p21-peptide with an appended NLS peptide. It also highlights $SV40_{126-132}$ as the ideal NLS of the four selected for high affinity PCNA binding.

This chapter highlights **P1bimF** and **P1c-N2F** as interesting leads to develop a viable pre-clinical cancer therapeutics. **P1bimF** and **P1c-N2F** are short p21-derived cyclic peptide-based candidates that are cell and nuclear permeable. Future study with these lead compounds should include investigation of their *in vitro* stability and activity against cancer cells. Sequence modifications in the p21-peptide scaffold should also be investigated for to improve PCNA binding affinity.

Chapter 4 Fluorescent sensor peptides

4.1 Design of sensor peptides

This chapter describes studies on developing a p21-derived peptide capable of exhibiting a fluorescence response in the presence of PCNA, thus allowing its detection. Such a PCNA fluorescent sensor could be used to investigate diseases associated with PCNA upregulation and cell proliferation, such as cancer.

Chapter One outlined the idea of introducing a solvatochromic amino acid into a peptide that selectively binds PCNA with an associated fluorescence response. Peptides derived from the p21$^{WAF/CIP1}$ protein, such as **p21$_{139\text{-}160}$**, are known to bind PCNA selectively, and were chosen as a template for the proposed peptide-based PCNA fluorescent sensor. The position of the solvatochromic amino acid within the peptide sequence, however, must be carefully chosen such that upon binding PCNA, the residue is embedded on the hydrophobic protein surface and shielded from the surrounding polar aqueous environment, giving rise to an increase in fluorescence. **p21$_{139\text{-}160}$** binds PCNA and inserts three residues (Met147, Phe150 and Tyr151) into a hydrophobic pocket on the protein surface, and thus present as ideal sites for solvatochromic amino acid insertion.

Three biosensor peptides were prepared in our laboratory by Aimee Horsfall [66] and these were based on the p21-derived sequence p21$_{141-155}$ (^{141}KRRQTSMTDFYHSKR^{155}) (**P5**), a truncated derivative of **p21$_{139-160}$**. The solvatochromic amino acid 4-*N*,*N*-dimethylamino-1,8-naphthalimidoalanine (4-DMNA) was inserted at positions 147, 150 or 151 (see Figure 43). The binding affinity of the subsequent peptides for PCNA was analysed by SPR to give K_D values of >25 μM, 20 μM and 921 nM, respectively. The PCNA binding of the 147-and 150-substituted peptides displayed a large non-specific component, which suggests that the 4-*N*,*N*-dimethylamino-1,8-naphthalimide (4-DMN) fluorophore is too large to be accommodated in the PCNA hydrophobic pocket.

In this study, the 4-dimethylaminophthalimide (4-DMAP) fluorophore was used in order to investigate whether using a smaller solvatochromic fluorophore improves PCNA binding and subsequent PCNA sensing (see Figure 43). 4-DMAP, a fluorophore of the same family as 4-DMN, was incorporated into a peptide sequence via its solvatochromic amino acid derivative, 4-dimethylaminothalimidoalanine (4-DAPA) (see Figure 43). 4-DAPA was introduced separately into the same p21-scaffold (**P5**), at positions 147, 150 and 151, to give three peptides, B1, B2 and B3 (see Table 16).

Fluorophores of the dimethylamino-phthalimide and -naphthalimide family are highly sensitive and exhibit exceptionally low fluorescence in water relative to those in apolar solvents, creating a greater signal-to-noise ratio. [56]
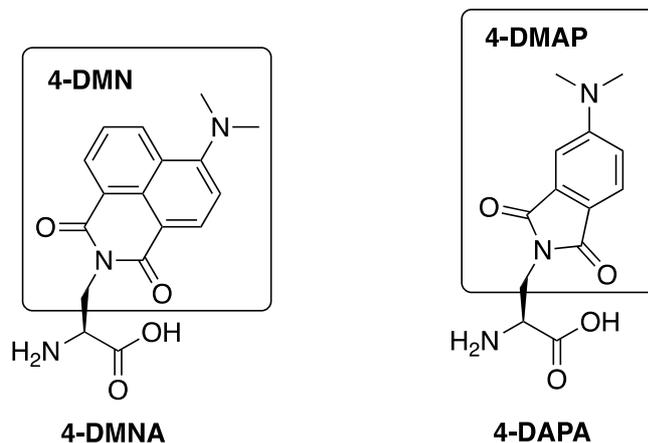
Figure 43: Structures of solvatochromic amino acids, 4-DMNA and 4-DAPA and their corresponding fluorophores, 4-DMN and 4-4-DMAP. Figure adapted.[56]

Table 16: p21-derived fluorescent peptide sensors for PCNA

Peptide Code	Sequence	Fluorophore
P5	H-KRRQTSMTDFYHSKR-NH$_2$	N/A
B1	H-KRRQTS**B$_1$**TDFYHSKR-NH$_2$	
B2	H-KRRQTSMTD **B$_1$**YHSKR-NH$_2$	
B3	H-KRRQTSMTDF**B$_1$**HSKR-NH$_2$	**B$_{1=}$**

4.2 Synthesis of sensor peptides

4-DAPA was incorporated into the **P5** scaffold via an on-resin derivatisation strategy (as described in Experimental). The peptide was assembled on-resin where an orthogonally protected diamino-propionic acid, Fmoc-Dap(Alloc)-OH, was coupled in the desired position of the peptide sequence to allow subsequent incorporation of fluorophore. After completion of the peptide sequence, the Allyloxy carbonyl (Alloc)-protecting group was selectively deprotected and *N,N'*-dimethyl-4-aminophthalic anhydride (**5**) was then coupled to the free amine.

Firstly, **5** was synthesised in two steps from commercially available 4-aminophthalic acid. Specifically, treatment of 4-aminophthalic acid with Pd/C in formaldehyde, under H_2 atmosphere, gave the corresponding diacid. Dehydration of this diacid on treatment with acetic acid gave **5** in 68% yield (see Scheme 9).
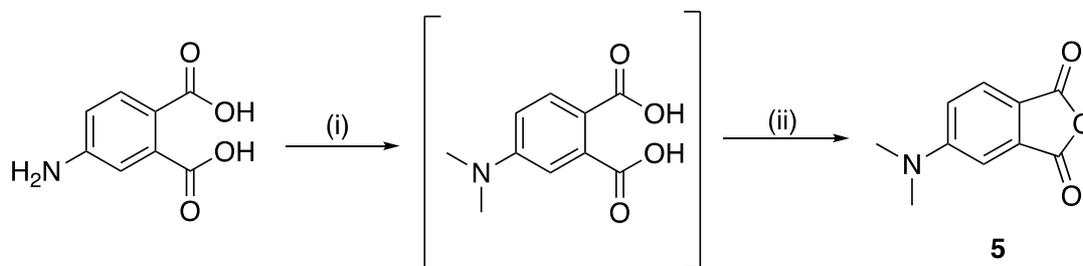
Scheme 9: Reagents and conditions: i) Pd/C, formaldehyde, H_2, 3 h, ii) AcOH, 50°C, 18 h (68%).

B1, **B2** and **B3** were assembled on-resin by Fmoc/tBu-SPPS (as described in Experimental). In brief, amino-acids were coupled onto a resin-bound growing peptide chain by addition of a solution of the Fmoc-protected amino acids (5 equiv), HATU (5 equiv) and DIPEA (10 equiv) in DMF, to the resin for 1 h. On-resin Fmoc-deprotections were achieved on treatment with 20% piperidine and 0.5 M HOBt. Successive amino acid couplings and Fmoc-deprotections then gave the desired peptide sequence. An orthogonally allyloxycarbonyl (Alloc)-protected diamino-propionic acid, Fmoc-Dap(Alloc)-OH was introduced at the 147, 150 and 151 positions, for subsequent derivatisation into 4-DAPA. The final amino-acid was protected with a Nα-Boc group where Boc-Lys(Boc)-OH was coupled to the peptide chain. Following assembly of the desired peptide sequence, the resin was suspended in dry DCM, and then reacted with Pd(PPh$_3$)$_4$ (0.8 equiv) and phenylsilane (25 equiv) under nitrogen for 15 min, to selectively remove the Alloc-protecting group. This treatment was repeated twice to ensure deprotection. The resin was reacted with a solution of **5** (3 equiv) and DIPEA (4 equiv) in NMP, under nitrogen, for 24 h. The resulting solution was drained and a solution of HBTU (6 equiv), HOBt (6 equiv) and DIPEA (12 equiv), in NMP, was added to resin and stirred for 2 h, and repeated twice to ring-close the phthalimide group. The peptides were simultaneously

cleaved from resin and globally deprotected by treatment with a solution of TFA (92.5%), DODT (2.5%), TIPS (2.5%) and water (2.5%) for 2 h to give **B1**, **B2** and **B3**. The resulting peptides were purified by semi-preparative RP-HPLC and identity confirmed by HRMS.
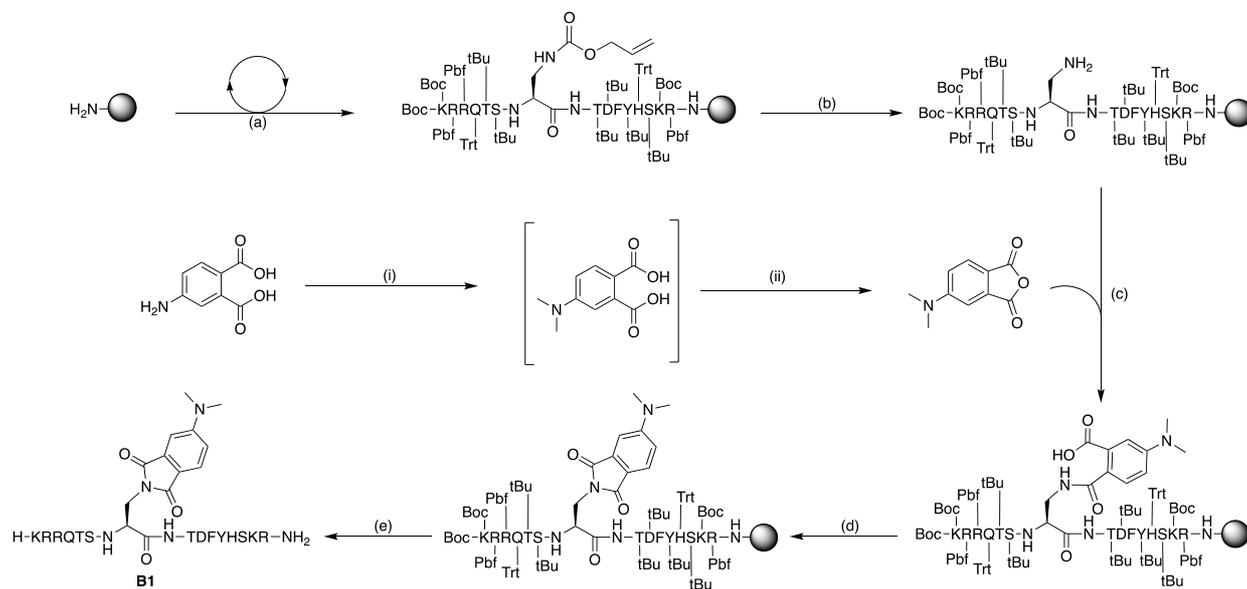
Scheme 10: general scheme for the synthesis of 4-DMAP-containing peptides (a-e) where **B1** is provided as an example, and the synthesis of 4-(*N,N*-dimethylamino)phthalic anhydride (i-ii). *Reagents and conditions:* a) repeated steps of coupling (Fmoc-protected amino-acid, HATU, DIPEA, DMF, 1 h) and deprotection (20% piperidine,0.5 M HOBt, DMF, 15 min).; b) Pd(PPh$_3$), phenylsilane, N$_2$, 15 min, repeated twice; c) 4-(*N,N*-dimethylamino)phthalic anhydride, DIPEA, NMP, N$_2$, o/n; d) HBTU/HOBt, DIPEA, NMP, 2 h; e) 92.5% TFA, 2.5% DODT, 2.5% TIPS, 2.5% H$_2$O, 2 h; i) Formaldehyde, Pd/C, H$_2$, MeOH,3 h; ii) Acetic acid, 50°C, N$_2$, 12 h.

4.3 Fluorescent characterisation of B1, B2 and B3

4.3.1 Fluorescence response in polar vs hydrophobic solvent (solvatochromism)

The solvatochromic properties of **B1**, **B2** and **B3** were first studied in a polar (Tris buffer) and hydrophobic (dioxane + 5 mM 18-crown-6). These conditions simulate the polar cell cytosol and the hydrophobic protein pocket environment, respectively. A sample of each peptide was prepared in buffer and dioxane/crown and the six resulting peptide samples were plated in triplicate in a 96-well plate.

A fluorescence emission spectrum (420-700 nm) was collected at excitation wavelength 390 nm for each peptide in each solvent system, followed by an excitation spectrum (300-500 nm)

with fixed emission wavelength 520 nm (see Figure 44). The emission spectra of **B1**, **B2** and **B3** in buffer revealed fluorescence maximum at 585, 580 and 585 nm, respectively. **B1**, **B2** and **B3** exhibited a fluorescence maximum at 515, 515 and 520 nm, respectively, in the dioxane/crown system. The maximum emission wavelength of all three peptides was thus blue shifted in going from polar (buffer) to hydrophobic (dioxane/crown) systems. These results are in good agreement with literature on 4-DMAP.[56]

As expected,[56] **B1**, **B2** and **B3** displayed much lower fluorescence intensity in the buffer system, relative to the dioxane/crown. The change in fluorescence intensity was calculated as the difference between the maximum fluorescence response of each peptide in buffer and dioxane/crown (see Table 17). **B1** produced the greatest change in fluorescence intensity between the two solvent systems, with a 300-fold increase in dioxane/crown. **B2** and **B3** displayed reduced 106-and 199-fold increases, respectively. These large changes in fluorescence intensity are consistent with literature that states that 4-DMAP exhibits exceptionally low fluorescence quantum yields in polar solvents, compared to hydrophobic. [65] The different fluorescence responses exhibited by **B1**, **B2** and **B3** indicate that changing the position of 4-DMAP, within the same peptide sequence, has an effect on the fluorescence of the resultant peptide. One explanation for this may be that the position of 4-DMAP changes peptide folding and thus how exposed the 4-DMAP fluorophore is to the solvent microenvironment, affecting its fluorescent properties.
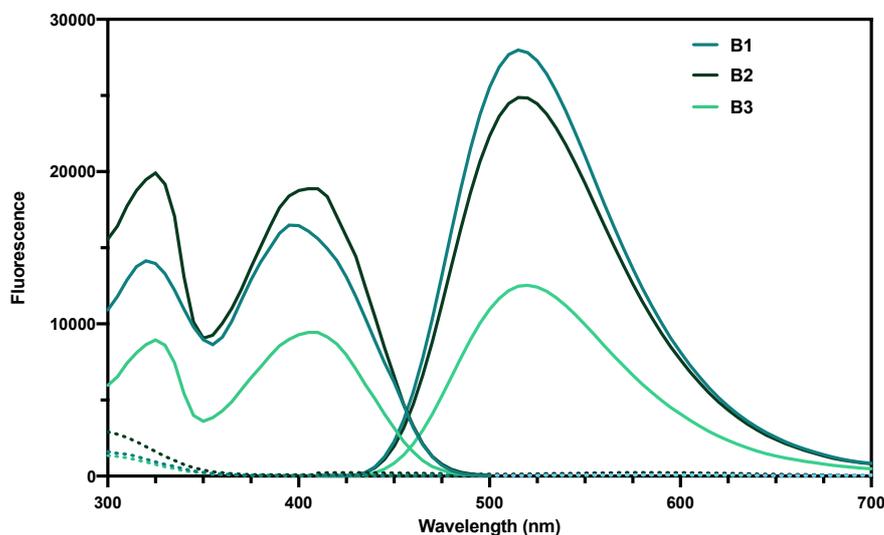
Figure 44: Excitation and emission fluorescence spectra of peptides **B1**, **B2** and **B3** in Tris buffer (dashed) and dioxane + 5 mM 18-crown-6 (solid). Emission spectrum (410-700 nm, ex. 390 nm). Excitation spectrum (300-500 nm, fixed em. 520 nm). All curves were smoothed in GraphPad Prism 9 with 10 neighbours.

Table 17: Changes in fluorescence intensity of **B1**, **B2** and **B3** in a buffer vs diox/crown solvent system.

Peptide	Max. Fluorescence response in buffer (arb. units)	Max. Fluorescence response in diox/crown (arb. units)	Fold change
B1	93	37987	300
B2	234	24866	109
B3	63	12529	199

4.3.2 Fluorescence response in the presence of PCNA

The fluorescence response of **B1**, **B2** and **B3** in the presence of PCNA was then examined in order to determine whether a fluorescence increase occurs in the presence of PCNA. Samples of each peptide (400 nM) with PCNA (800 nM) were prepared in Tris buffer (pH 7.4) and plated in triplicate in a 96-well plate. Solutions of each peptide (400 nM) in buffer, PCNA (800 nM) in buffer and the Tris buffer alone were also prepared and plated in triplicate, as controls. Fluorescence emission spectra were obtained for each peptide (ex. 390 nm, 410-700 nm) (see Figure 45). All three peptides, **B1**, **B2** and **B3** gave significantly enhanced fluorescence in the presence of PCNA. The change in fluorescence intensity was measured by

comparing the maximum fluorescence response of each peptide alone and with PCNA (see Table 18).

B3 produced the largest change in fluorescence intensity with a 7.91-fold increase upon exposure to PCNA. **B1** and **B2** displayed 5.50-and 3.91-fold increases, respectively. This indicates that 4-DMAP at position 151, as in **B3**, gave rise to a peptide which produced the largest 'turn on' fluorescence signal for PCNA, relative to peptide **B1** and **B2**. These results suggest **B3** interacts with PCNA resulting in the fluorophore being shielded from the surrounding polar solvent, resulting in higher fluorescence. The increase in fluorescence was accompanied by a blue shift in the emission maximum wavelength of **B3,** 560 nm in buffer and 525 nm in buffer with PCNA. This is consistent with a change from a polar to hydrophobic environment. Interestingly, only **B3** exhibited this blue shift, that was previously seen in the buffer vs dioxane/crown experiment. This suggests that only **B3** interacts sufficiently with PCNA, such that the 4-DMAP fluorophore experienced a change in solvent microenvironment (buffer vs PCNA hydrophobic pocket), causing an increase in fluorescence intensity and a blue shift in emission wavelength.
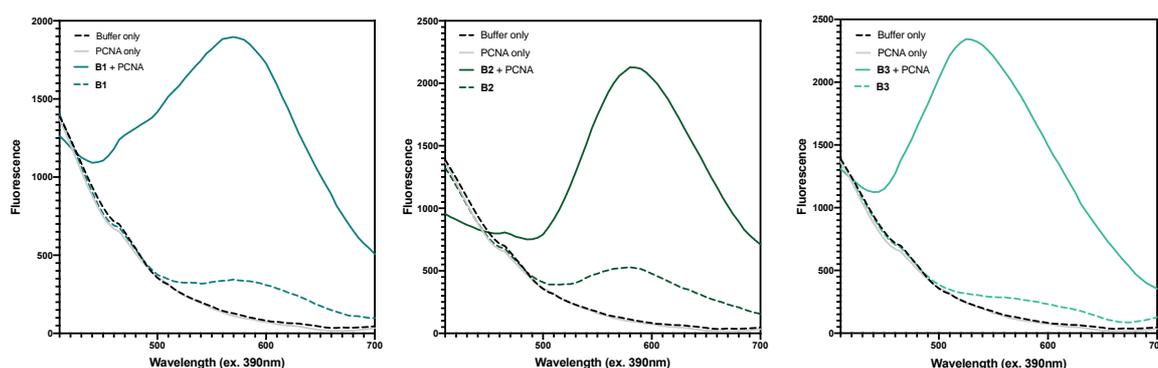
Figure 45: Fluorescence emission spectra (410-700 nm, ex. 390 nm) of **B1**, **B2** and **B3** (left to right) in the presence of PCNA (solid, coloured) and absence of PCNA (dashed coloured). All solutions were in Tris buffer. All curves were smoothed in GraphPad Prism 9 with 10 neighbours.

Table 18: Fluorescence response of **B1**, **B2** and **B3** in the presence or absence of PCNA.

Peptide	Max. Fluorescence Response (peptide only) (arb. units)	Max. Fluorescence Response (+PCNA) (arb. units)	Fold change
B1	349	1919	5.50
B2	550	2153	3.91
B3	299	2365	7.91

4.4   Binding affinity of sensor peptides

The sensor properties of **P1**, **P2** and **P3** were evaluated by first measuring the binding affinities for PCNA. To sense PCNA levels through fluorescence, the solvatochromic fluorophore-containing p21-peptides must bind to PCNA such that the fluorophore is shielded from surrounding solvent. The binding affinities of **B1**, **B2** and **B3** for PCNA were determined by SPR as previously described in Section 2.3, in order to determine how the position of introduced fluorophore affects PCNA binding (see Experimental for details). PCNA binding affinities of each peptide can also provide clarity on why **B3** displayed the largest fluorescence response in the presence of PCNA compared to **B1** and **B2**. The PCNA binding affinity of **P5** was also evaluated by SPR as a control. In brief, PCNA was first immobilised onto a gold sensor chip by EDC/NHS mediated coupling. Each peptide was serially diluted (1:2) eight times using buffer and flowed over the PCNA-loaded sensor chip from lowest to highest concentration. Changes in refractive index of the sensor chip, as a result of binding events, are converted to a response. The response versus peptide concentration was plotted and a line fitted to provide the equilibrium dissociation constant K_D.

P5 bound PCNA with a K_D of 12.3 nM, which is consistent with previous results. [89] **B1** displayed low binding affinity for PCNA at >25 μM. The maximum response of **B1**, indicated by the corresponding sensorgram, exceeded the theoretical maximum response expected (RU_{max}= 103) as calculated using the amount of PCNA loaded onto the sensor chip (see

Figure 46). [89] This suggests that interaction of **B1** with PCNA was largely non-specific with **B1** simply binding to the reference cell. This suggests the 4-DMAP fluorophore was too large for sufficient packing within the hydrophobic pockets of PCNA when inserted at position 147. A K_D of 4.61 µM was calculated for **B2**, however, the sensorgram curve also exceeded the theoretical maximum response, suggesting a non-specific binding interaction with PCNA (see Figure 46). **B3**, in contrast, interacted specifically with PCNA to give a K_D of 1.28 μM (see Figure 46). This suggests 4-DMAP insertion was most favourable at position 151, likely because the hydrophobic cleft of PCNA is large enough to accommodate a fluorophore at that position, such that the binding between PCNA and the peptide is not disturbed.

Interestingly, of the three previously synthesised DMNA-substituted peptides, the 151-DMNA substituted peptide bound PCNA with the highest affinity at 921 nM, relative to the 147-and 150-DMNA substituted peptides. [89]**B3** bound PCNA with a similar affinity to its 151-DMNA substituted counterpart at 1.28 μM affinity. This suggests both the 4-DMN and 4-DMAP fluorophores are well tolerated by the hydrophobic PCNA pocket when incorporation at the 151 position of **P5**. Hence, the 151-substituted peptides were chosen for further study as PCNA fluorescent sensors.

Table 19: Binding affinities of biosensor peptides and **P5** for PCNA. *The interaction of these peptides with PCNA was largely non-specific.

Peptide	Affinity K_D (μM)	K_D SE (nM)	χ^2
P5	12.3 nM	0.598	0.196
B1*	> 25.0	-	-
B2*	4.61	570	3.02
B3	1.28	260	0.0298
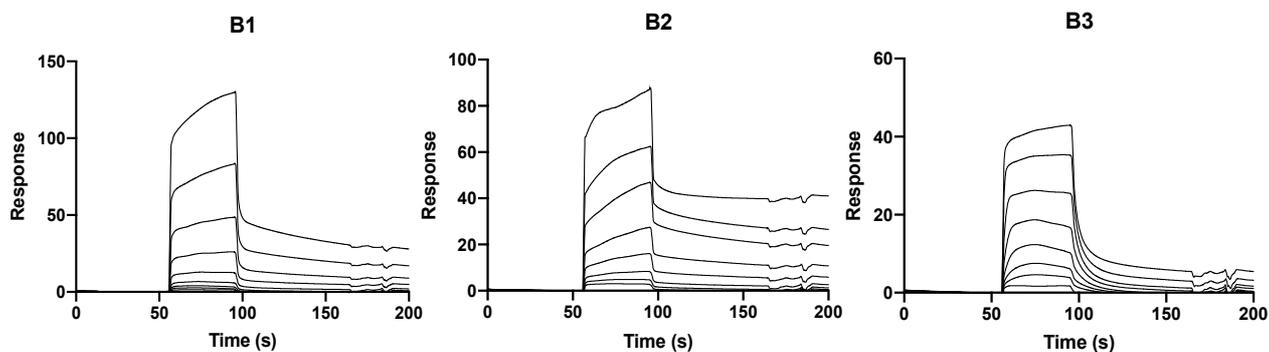
Figure 46: SPR sensorgrams of **B1**, **B2** and **B3**. The top curve of each sensorgram marks the highest concentration for each peptide, which is listed in Table S1. Each subsequent curve below represents a 1 in 2 serial dilution from the top concentration. All curves were smoothed in GraphPad Prism 9 with 10 neighbours.

4.5 PCNA titration

B3 was selected for evaluation in a dose response experiment with fluorescence monitored with increasing concentrations of PCNA. This was performed in order to determine the binding ratio of **B3**: PCNA needed to elicit the maximum fluorescence response. A sample of **B3** (3 μM), in HEPES buffer, was made up and plated in triplicate in a 96-well plate. PCNA, in HEPES, was added to the wells, to give a **B3**: PCNA ratio ranging from 12: 1 to 1: 15. Wells containing **B3**, PCNA and buffer only were also prepared as controls. Each well was excited at 390 nm and a fluorescence spectrum (460-700 nm) collected (see Figure 47) to reveal a maximum 3.5-fold increase in fluorescence in the presence of PCNA at a 1:2.5 ratio of peptide: PCNA.
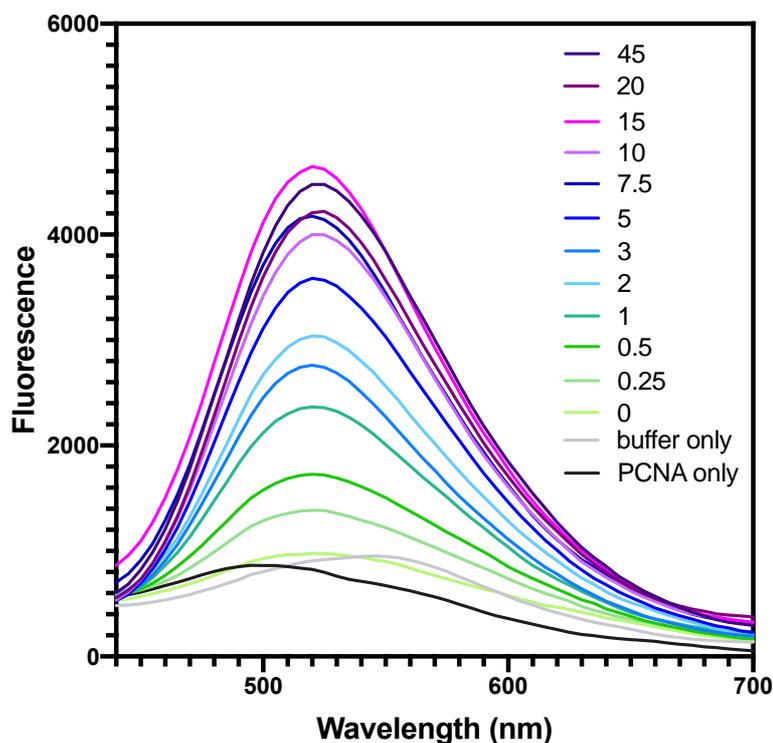
Figure 47: Fluorescence emission spectra (440-700 nm, ex. 390 nm) of **B3** combined with increasing concentrations of PCNA, to give final PCNA concentrations of 0, 0.25, 0.5, 1, 2, 3, 5, 7.5, 10, 15, 20 and 45 μM. All curves were smoothed in GraphPad Prism 9 with 10 neighbours.

4.6   Design of second-generation fluorescent sensor peptides

A functional PCNA sensor for live cell imaging requires that the peptide permeates the cell and nucleus where fluorescence can 'switch on' upon binding PCNA. A second generation of p21 sensor peptides was designed based on this requirement. Earlier studies revealed that the **p21$_{139-160}$** sequence is cell permeable (see Chapter 2), and was thus chosen as the scaffold for these peptides. As 151-substitued peptides were shown to maintain sufficient PCNA binding, two peptides were prepared, **B4** and **B5**, which have 4-DAPA and 4-DMNA substituted at position 151, respectively (see Table 20). The **p21$_{139-160}$** sequence was further modified with a Met147Ile substitution in order to avoid methionine oxidation issues in its synthesis as was evident in the preparation of 150-DMNA and 151-DMNA substituted peptides. [89]A Met147Ile substitution has been found to marginally increase the peptide affinity to PCNA. Methionine and isoleucine are similar in size and hydrophobicity, thus a substitution in this peptide design is unlikely to have detrimental effects on peptide binding and interaction with

PCNA. [90] An Asp149Glu modification was also introduced as earlier work (see Chapter 2) showed that p21-peptides undergo significant aspartimide formation, thus aspartic acid was substituted with glutamic acid here. An *N*-terminal cysteine residue was also introduced for subsequent conjugation to an NLS, in order to facilitate nuclear entry of the PCNA fluorescent sensor peptides. Furthermore, *N*-terminal protection was achieved with an acetyl cap. These modifications gave rise to the 151-DAPA-and 151-DMNA-substituted peptides **B4** and **B5**, respectively (see Table 20).

Table 20: peptide sequence of second generation of sensor peptides.

Peptide Code	Sequence	Fluorophore
p21$_{139-160}$	H-GRKRRQTSMTDFYHSKRRLIFS-NH$_2$	N/A
B4	Ac-CGRKRRQTSITEF**B$_1$**HSKRRLIFS-NH$_2$	**B$_{1=}$**
B5	Ac-CGRKRRQTSITEF**B$_2$**HSKRRLIFS-NH$_2$	**B$_{2=}$**

4.7 Synthesis of second-generation sensor peptides

4.7.1 Synthesis of B4

B4 was synthesised as per **B1-3** (See Methods and see Section 4.2). For **B4**, the Nα of the final amino-acid was acylated rather than Boc-protected as was seen with **B1-3**. After the final *N*-terminal Fmoc-deprotection, the peptide was acetylated by treating the resin-bound peptide with a solution of acetic anhydride (470 μL) and DIPEA (870 μL) in DMF (4 mL) for 15 min (see Scheme 11). **B4** was purified by semi-preparative RP-HPLC and its identity confirmed by HRMS.
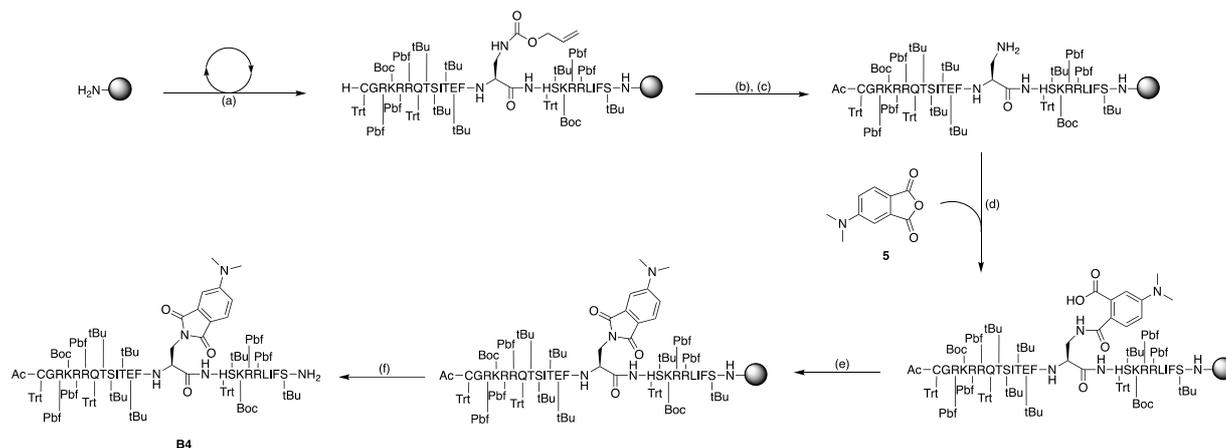
Scheme 11: Synthetic scheme for **B4** via an on-resin derivatisation strategy. *Reagents and Conditions*: a) repeated steps of coupling (Fmoc-protected amino-acid, HATU, DIPEA, DMF, 1 h) and deprotection (20% piperidine,0.5 M HOBt, DMF, 15 min).; b) Acetic anhydride, DIPEA, DMF, 15 min; c) Pd(PPh$_3$), phenylsilane, N$_2$, 15 min, repeated twice; d) **5**, DIPEA, NMP, N$_2$, o/n; e) HBTU/HOBt, DIPEA, NMP, 2 h; f) 92.5% TFA, 2.5% DODT, 2.5% TIPS, 2.5% H$_2$O, 2 h

4.7.2 Synthesis of B5

B5 was assembled via an Fmoc-SPPS based building block strategy, with the 4-DMN fluorophore introduced during the SPPS process. The first step involved synthesis of the Fmoc-protected amino-acid containing the 4-DMN fluorophore, *N*-α-Fmoc-(4-*N*,*N*-dimethylamino-1,8-naphthalimido)-alanine (Fmoc-DMNA) (**8**) in three steps from 4-*N*,*N*-dimethylamino naphthalic anhydride (**6**). Treatment of **6** with (S)-3-amino-2-(Boc-amino)-propionic acid in dioxane, gave the Boc-protected amino acid, *N*-α-Boc-(4-*N*,*N*-dimethylamino-1,8-naphthalimido)-alanine (Boc-DMNA) (**7**). This was then Boc-deprotected on treatment with TFA, to give a free primary amine, which was Fmoc-protected on reaction with Fmoc-OSu in dioxane, to give **8** in 20% yield over three steps (see Scheme 12).
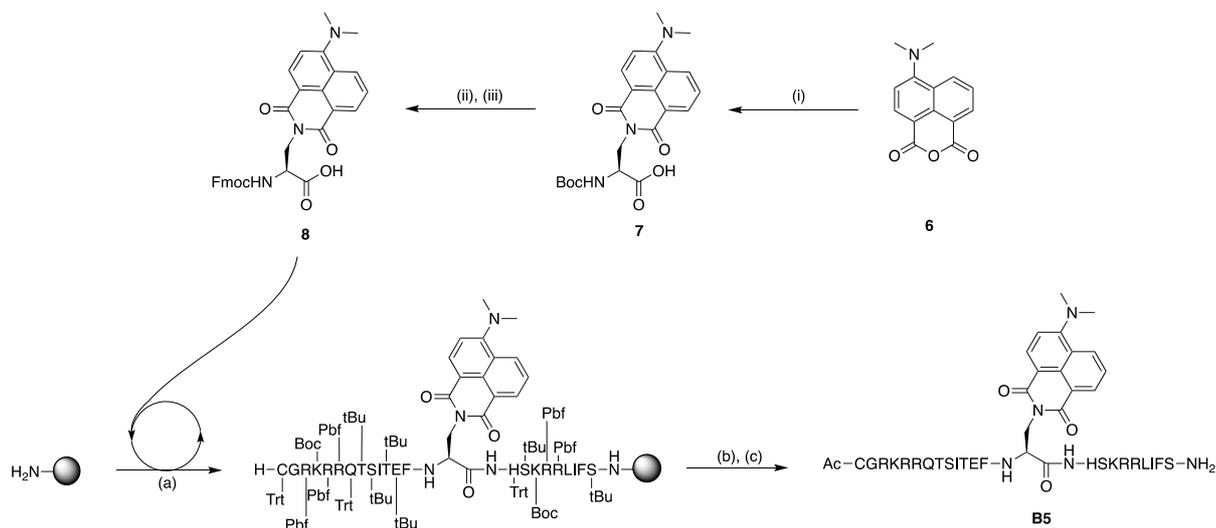
Scheme 12: Synthetic scheme for **B5** via an Fmoc-SPP building block approach. *Reagents and Conditions*: i) Boc-Dap-OH, NaHCO$_3$, dioxane/H$_2$O, N$_2$, reflux, 30 min, ii) TFA/DCM (1:1), rt, 3 h, iii) Fmoc-OSu, NaHCO$_3$, dioxane/H$_2$O (5:1), rt, 18 h (20% over 2 steps), a) repeated steps of coupling (Fmoc-protected amino-acid, HATU, DIPEA, DMF, 1 h) and deprotection (20% piperidine, 0.5 M HOBt, DMF, 15 min), b) Acetic anhydride, DIPEA, DMF, 15 min, c) 92.5% TFA, 2.5% DODT, 2.5% TIPS, 2.5% H$_2$O, 2 h.

The peptide was assembled on-resin by Fmoc/tBu-SPPS (described in Experimental and see Scheme 12). In brief, amino-acids were coupled by addition of a solution of the Fmoc-protected amino-acids (5 equiv), HATU (5 equiv) and DIPEA (10 equiv) in DMF, to the resin for 1 h. On-resin Fmoc-deprotections were achieved on treatment with 20% piperidine and 0.5 M HOBt. Successive amino acid couplings and Fmoc-deprotections then gave the desired peptide sequence. 4-DMNA was incorporated at position 151 by adding a solution of **8** (3 equiv) and DIPEA (6 equiv) in DMF (5 mL) to the resin for 3 h. Following the final *N*-terminal Fmoc-deprotection, the peptide was acetylated by treating the resin-bound peptide with a solution of acetic anhydride (470 μL) and DIPEA (870 μL) in DMF (4 mL) for 15 min. The peptide was cleaved from resin and globally deprotected on treatment with a solution of TFA (92.5%), DODT (2.5%), TIPS (2.5%) and water (2.5%) for 2 h. **B5** was purified by semi-preparative RP-HPLC and identity confirmed by HRMS.

4.8   Binding affinity of second generation sensor peptides

The binding affinities of **B4** and **B5** for PCNA were determined by SPR as before in Section 4.4 (see Experimental for details). None of these peptides gave rise to PCNA binding and

simply bound largely to the reference cell as opposed to the PCNA-loaded cell and thus K_D values could not be determined (see Figure 48). This was surprising given **p21139-160** reportedly binds PCNA with 2.5 nM affinity.[27] Additionally, p21-peptides with D149E and M1471 mutations have both been shown to bind PCNA with nanomolar affinity. [90] This suggests modifications made in **B4** and **B5**, relative to **p21139-160**, such as fluorophore insertion, likely reduced the binding to PCNA. As selective and high affinity binding to PCNA is imperative for a PCNA-sensor, further study was not pursued with **B4** and **B5**.
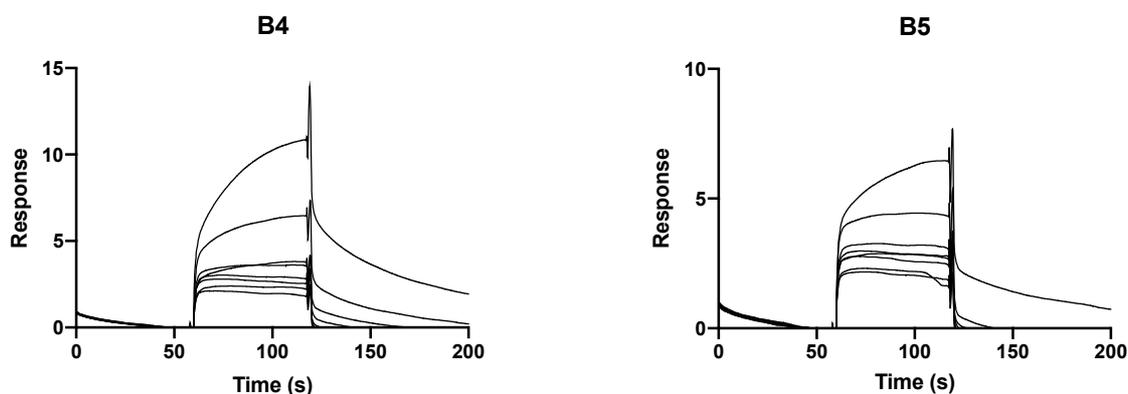
Figure 48: SPR sensorgrams of **B4** and **B5**. The top curve of each sensorgram marks the highest concentration for each peptide, which is listen in Table S1. Each subsequent curve below represents a 1 in 2 serial dilution from the top concentration. All curves were smoothed in GraphPad Prism 9 with 10 neighbours.

4.9 Chapter conclusions

Three p21-derived peptides were successfully synthesised, **B1**, **B2** and **B3**, with 4-DAPA introduced at positions 147, 150 and 151 of **P5**, respectively. **B1** exhibited the greatest change in fluorescence on changing from a polar solvent to a hydrophobic solvent, with a 300-fold change. **B1**, **B2** and **B3** all exhibited an increase in fluorescence in the presence of PCNA as expected with fold changes of 5.50, 2.91 and 7.91, respectively. **B3** displayed the largest change in fluorescence intensity, highlighting fluorophore insertion at 151 results in a larger 'turn on' fluorescence signal for PCNA. This was expected as **B3** displayed the highest affinity PCNA binding with a K_D of 1.28 μM, relative to **B1** and **B2**. This suggests that **B3** binds PCNA such that the 4-DMAP occupies the hydrophobic cleft of PCNA, becoming

shielded from the polar solvent. This results in a 7.91-fold increase in fluorescence on exposure of B3 to PCNA. **B1** and **B2** interacted with PCNA non-specifically, suggesting insertion of 4-DMAP at 147 or 150 disrupts PCNA binding, which likely explains the smaller 5.50-and 2.91-fold changes of **B1** and **B2**, respectively.

The three analogous DMNA-containing peptides produced a similar fluorescence response when exposed to PCNA, with the 150-substituted peptide giving the largest fold increase of 14.5. The 149-and 151-DMNA peptides produced a 6.5-and 7.0-fold increase in fluorescence when exposed to PCNA, respectively. As per the DAPA-peptides, only the 151-substituted of the DMNA-peptides displayed specific binding to PCNA, with a K_D of 921 nM, suggesting inclusion of 4-DMNA at positions 147 and 150 also perturbs PCNA binding. [89]

The fluorescence response of **B3** reached saturation at 2.5 equivalents of PCNA to peptide and upon saturation, exhibited a 3.5-fold fluorescence increase relative to peptide only. The fluorescence response of the analogous 151-DMNA substituted peptide was also saturated at 2.5 equivalents of PCNA. This gave a 10-fold increase in fluorescence relative to peptide only. Overall, in comparing the DMNA substituted peptides to the DAPA substituted peptides, both gave similar fluorescence and binding responses to PCNA, indicating that the change in fluorophore did not largely affect the peptide properties.

Lastly, in an effort to develop a cell and nuclear permeable PCNA sensor, 4-DMAP and 4-DMNA were inserted at position 151 of a modified **p21$_{139-160}$** scaffold, to give **B4** and **B5** respectively. These peptides bound largely to the reference cell of the sensor chip and as such K_{DS} could not be determined and further study was not pursued. This suggests the modifications made in **B4** and **B5**, from the native **p21$_{139-160}$** sequence reduced PCNA binding. We suggest that incorporating a large solvatochromic amino acid, relative to a natural amino acid, such as 4-DMNA or 4-DAPA, into different p21-derived sequences, alters

the folding and secondary structure of the modified peptide in unexpected ways, affecting the PCNA binding.

This work establishes position 151 of the **P5** sequence as the most favourable site to incorporate a solvatochromic amino acid, 4-DAPA, as in **B3**. This peptide was found to bind specifically to PCNA with a fluorescence response in the presence of PCNA. We found that substituting a larger solvatochromic amino acid, 4-DMNA, at the same position, gave rise to a peptide with comparable binding affinity to PCNA and fluorescence response when exposed to PCNA. Both 151-substituted peptides provide an interesting lead for developing a PCNA fluorescent sensor to detect cell proliferation levels. Future work would involve optimising both peptides for cell and nuclear permeability to enable in-cell experiments.

Chapter 5   Thesis conclusions and future directions

A viable peptide based therapeutic must be cell and nuclear permeable in order to target and inhibit PCNA for application as anticancer treatment. To enact its inhibitory effect, the peptide must also display high binding affinity to PCNA. This work centred on using the **p21$_{139\text{-}160}$** sequence, derived from the native parent protein p21$^{WAF1/CIP1}$, as a template to define the minimum scaffold required to permit cell permeability. Previously, **p21$_{139\text{-}160}$** truncation experiments identified a 12mer peptide (**P1**) as a high affinity PCNA binder, however, a fluorescein-tagged analogue, **P1F**, was not cell permeable. Chapter Two presents the synthesis of three new peptides (**P2**, **P3** and **P4**) where the **P1** sequence was extended by reintroducing positively charged arginine or lysine residues from the longer **p21$_{139\text{-}160}$** sequence, to determine if these residues are able to restore cell permeability. The **P3** sequence was found to be unusually prone to aspartimide formation during its synthesis. Efforts to optimise the deprotection conditions during **P3** synthesis found an increased concentration of HOBt in the deprotection solution was most effective in minimising aspartimide formation. Fluorescein-tagged derivatives, **P2F**, **P3F** and **P4F** were assayed in a cell permeability experiment with **P4F** displaying the most significant cellular entry, demonstrating extending the **P1** sequence at both termini and reintroducing charged residues (arginine and lysine) enhanced cell entry, relative to **P1**. **p21$_{139\text{-}160}$F**, however, showed significantly more cell permeability than **P4F**, suggesting cell entry is facilitated by the ^{157}LIFS^{160} region of the **p21$_{139\text{-}160}$** sequence. This region of hydrophobic residues increases the amphipathic nature of the peptide, a feature common amongst CPPs which likely aided the cell permeability of **p21$_{139\text{-}160}$F**. We found that simply lengthening the **P1** sequence to **P2**, **P3** and **P4** was not enough to facilitate substantial cell permeability, thus further modification of the **P1** sequence was investigated.

Chapter Three describes studies on synthesising two series of NLS tagged **P1** conjugate peptides; linear and macrocyclic. The linear **P1** peptide was conjugated to four NLS peptides (Tat$_{48-57}$, cMyc$_{320-328}$, SV40$_{126-132}$ and R6W3) to determine if these tags are able to facilitate cell and nuclear permeability. This was achieved by a thiol conjugation approach via a maleimide group. The resulting conjugates, peptides, **P1b-N1F**, **P1b-N2F**, **P1b-N3F** and **P1b-N4F** were administered to MDA-MB-468 breast cancer cells, to determine if the NLS peptides facilitated cell and nuclear uptake, which revealed only **P1b-N2F** as cell and nuclear permeable. Next a macrocyclic derivative of **P1**, **P1bim**, with a bimane-containing linker designed to constrain the backbone into a helix that is known to be adopted on binding to PCNA was conjugated to the four NLS peptides. The resulting peptide conjugates, **P1c-N1F**, **P1c-N2F**, **P1c-N3F** and **P1c-N4F** were separately administered to MDA-MB-468 breast cancer cells with the SV40$_{126-132}$ tagged **P1c-N2F**, displaying nuclear accumulation. Additionally, the control peptide, **P1bimF**, which contains a bimane linker and fluorescein tag but no NLS peptide, was also nuclear permeable. In contrast, analogues of **P1c-N2F** and **P1bimF** without the fluorescein tag (**P1c-N2** and **P1bim**) were only cell permeable. This highlights that in this instance, the fluorescein tag was altering the nuclear uptake. The binding affinity of **P1bimF** and **P1c-N2F**, and the untagged analogues (**P1bim** and **P1c-N2**) was determined by SPR. **P1c-N2** displayed the highest affinity for PCNA with a K$_D$ of 176 nM. Interestingly, the SV40$_{126-132}$ tagged linear peptide conjugate, **P1b-N2F**, displayed the highest affinity for PCNA of the linear peptide conjugates with a K$_D$ of 33.1 nM. This indicates that the SV40$_{126-132}$ NLS is the most favourable NLS of the four selected for facilitating high affinity PCNA binding. Additionally, these results show that potent nanomolar binding for PCNA is maintained with a short p21-peptide with an appended NLS peptide.

The work described in Chapters Two and Three defined the scaffold required for a p21-derived peptide to be cell and nuclear permeable, in order to target PCNA as potential cancer

treatment. The introduction of a bimane linker into a short p21-peptide facilitates cell entry, likely by constraining the helical secondary structure of p21-peptides. Conjugation of this short p21-peptide to a fluorescein-tagged $SV40_{126-132}$ NLS or simply a fluorescein tag, was found to confer nuclear permeability. Additionally, the linear derivative of the short p21-peptide was found to also accumulate within cell nuclei when attached to fluorescein-tagged $SV40_{126-132}$.

The two nuclear permeable macrocyclic bimane peptides, **P1bimF** and **P1c-N2F**, offer potential benefits as drug candidates over the linear **P1b-N2F**. In contrast to **P1b-N2F** and other documented nuclear permeable NLS tagged p21-peptides, [23, 29, 87] **P1bimF** and **P1c-N2F** are macrocyclic which is associated with improved proteolytic, *in vitro* stability and structural stability, relative to linear peptides. [84, 88] **P1bim** and **P1c-N2F** are both promising candidates for developing a viable pre-clinical peptide-based cancer therapeutic that targets PCNA.

Future work would involve *in vitro* assay of nuclear permeable **P1b-N2F**, **P1bimF** and **P1c-N2F** against cancer cells to determine if they inhibit cell proliferation by disrupting DNA-replication. This activity assay would also provide insight into whether the p21-peptide is effectively being transported into the nucleus to affect its function on PCNA. The proteolytic stability of the lead peptides would also be determined in order to define the half-life and potential biostability of the peptide. Future work would also involve investigating sequence mutations as a means to enhance the affinity of these short p21-peptides for PCNA. Additionally, we would introduce other linkers into the short p21-peptide in order to determine if other linkers also facilitate cell permeability or the bimane linker specifically imparts cell entry.

This thesis also presented studies on using the p21-sequence as a template to develop three potential peptide-based PCNA sensors. A viable peptide sensor for PCNA must display high

binding affinity for PCNA and upon binding, display a fluorescence response. A short p21-peptide (**P5**), known to bind PCNA with high affinity, was chosen as the starting template. p21-peptides insert Met147, Phe150 and Tyr151 into hydrophobic pockets on PCNA's surface upon binding. A solvatochromic fluorophore, 4-DMAP, was incorporated at each of these positions (147, 150 and 151) to afford **B1**, **B2** and **B3**, respectively. **B1**, **B2** and **B3** showed enhanced fluorescence in a hydrophobic (dioxane with 18-crown-5) environment compared to a polar solvent environment (Tris buffer) with 300-, 109-and 199-fold increases, respectively. These differences in fluorescence responses between **B1**, **B2** and **B3** indicates that changing the position of 4-DMAP in the same peptide sequence has an effect on the fluorescent properties of the resultant peptide. The position of 4-DMAP likely affects how the resultant peptide folds and thus how exposed the 4-DMAP fluorophore is to the solvent microenvironment which affects its fluorescent properties.

The fluorescence response of the peptides to PCNA was then evaluated in order to determine which peptide gave the largest 'turn on' signal for PCNA. **B3** exhibited the largest increase in fluorescence in the PCNA dosed buffer, with a 7.91-fold change. **B1** and **B2** displayed fold changes of 5.50 and 3.91, respectively. These results suggest **B3** binds PCNA and inserts the 4-DMAP fluorophore at position 151 into the hydrophobic cleft of PCNA, resulting in a large fluorescence increase. The binding affinity of these peptides for PCNA was determined by SPR, with only **B3** interacting specifically with PCNA to give a K_D of 1.28 μM. This demonstrated 151-substitution of 4-DMAP into the **P5** sequence allows interaction with PCNA and elicits a resultant 'turn on' fluorescence response. Fluorophore insertion was most favourable at position 151, likely because the hydrophobic cleft of PCNA is large enough to accommodate a fluorophore at that position. The poor binding interaction and fluorescence responses of **B1** and **B2** indicate fluorophore insertion at positions 147 and 150 disrupt PCNA binding. This suggests the hydrophobic cleft of PCNA is too small to accommodate 4-DMAP

at both these positions. Overall, the three analogous DAPA-containing p21-peptides produced a similar fluorescence response when exposed to PCNA to the previously studied DMNA-containing p21-peptides. This indicates that the smaller size of the 4-DMAP fluorophore did not drastically impact the properties of the corresponding p21-sensor peptides. These results suggest that the 4-DMAP and 4-DMNA fluorophores can both be successfully inserted into the p21-sequence.

A second generation of 151-substituted sensor peptides was prepared in an effort to obtain a viable cell and nuclear permeable PCNA sensor. 4-DMAP and 4-DMNA were each substituted into position 151 of a modified **p21$_{139\text{-}160}$** scaffold, giving **B4** and **B5**, respectively, where the **p21$_{139\text{-}160}$** sequence was used for its cell permeability capabilities. None of these peptides gave rise to PCNA binding and simply bound largely to the reference cell and thus K_D values could not be determined and further studies were not pursued.

This work demonstrated that the p21-sequence can be successfully exploited to incorporate a solvatochromic fluorophore as a sensing component for PCNA. For a p21$_{141\text{-}155}$ sequence, position 151 is the most favourable position to incorporate such a sensing component. Fluorescence studies highlighted that the 151-substituted p21-peptide was able to effectively bind PCNA and elicit a 'turn on' fluorescence signal to indicate PCNA binding. Future work would involve expanding the capability of the lead fluorescent sensor into quantitative sensing by attaching a second nonsolvatochromic fluorophore with a different emission wavelength, to allow ratiometric fluorescence measurements. Further study is required to optimise the scaffold developed to facilitate cell and nuclear permeability which would allow for live-cell PCNA sensing. For example, a cell permeable, macrocyclic bimane p21-peptide could be modified with a solvatochromic fluorophore and an SV40$_{126\text{-}132}$ NLS appended to it to confer nuclear permeability.

Overall, this thesis describes the design and synthesis of a cell and nuclear permeable p21-derived peptide that maintain high binding affinity for PCNA, as a first step toward a viable pre-clinical peptide that targets PCNA for application in broad spectrum cancer treatment. The p21-sequence was also used to synthesise a promising lead towards a 'turn on' fluorescent PCNA sensor, for monitoring cell proliferation and to study the role of PCNA in diseases such as cancer.

Chapter 6   Experimental

6.1   Materials

All reagents and solvents were purchased from Sigma Aldrich unless stated otherwise. All Fmoc-protected amino-acids and coupling agents [Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate (HATU) used were obtained from Chem-Impex International. The solvents *N,N*-dimethylformamide (DMF), dichloromethane (DCM) and piperidine were obtained from Merck & Co. Inc.; *N,N'*-Diisopropylethylamine (DIPEA) was obtained from Alfa Aesar.

The following standard Fmoc-protected amino-acids with orthogonal protecting groups were used unless stated otherwise: Fmoc-β-Ala-OH, Fmoc-L-Ala-OH, Fmoc-L-Asp(tBu)-OH, Fmoc-L-Thr(tBu)-OH, Fmoc-L-Glu(tBu)-OH, Fmoc-L-Pro-OH, Fmoc-L-Gly-OH, Fmoc-L-Cys(Trt)-OH, Fmoc-L-Cys(Mmt)-OH, Fmoc-L-Val-OH, Fmoc-L-Met-OH, Fmoc-L-Ile-OH, Fmoc-L-Leu-OH, Fmoc-L-Tyr(tBu)-OH, Fmoc-L-Phe-OH, Fmoc-L-His(Trt)-OH, Fmoc-L-Lys(Boc)-OH, Fmoc-L-Arg(Pbf)-OH, Fmoc-L-Trp(Boc)-OH, Fmoc-L-Gln(Trt)-OH, Fmoc-L-Asn(Trt)-OH.

6.2   Methods

Stock Solutions

Solution A (Deprotection: 20% piperidine with 0.5 M HOBt in DMF)

Piperidine (10 mL) and 3.4 g HOBt in DMF (10 mL) were combined with DMF (30 mL).

Solution B (Coupling: 0.5 M HATU in DMF)

HATU (1.9 g, 0.5 M) was dissolved in DMF (10 mL).

Solution C1 (High-acid TFA cleavage)

TFA (9.25 mL) was combined with DODT (250 μL), TIPS (250 μL) and H$_2$O (250 μL).

Solution C2 (High-acid TFA cleavage without DODT)

TFA (9.5 mL) was combined with TIPS (250 μL) and H$_2$O (250 μL).

Solution D1 (TNBS test)

DIPEA (50 μL) was added to DMF (950 μL).

Solution D2 (TNBS test)

5% aq. TNBS (100 μL) was added to DMF (900 μL).

Solution E (Acetylation)

Ac$_2$O (870 μL) and DIPEA (470 μL) were combined with DMF (10 mL).

General Method 1: Solid-Phase Peptide Synthesis

Rink Amide PL resin (644 mg, 0.2 mmol/g, Agilent) was swollen in 1:1 DMF/DCM (10 mL) for 15 min. The resin was treated with solution A (5 mL) for 15 min to remove the Fmoc-protecting group. The resin was washed with DMF (3 x 5 mL). Amino-acids were coupled with addition of the Fmoc-protected amino-acids (5 equiv), solution B (5 equiv, 2 mL) and DIPEA (10 equiv, 348 μL) in DMF (5 mL), to the resin and stirred intermittently for 1 h. The resin was washed with DMF (3 x 5 mL). Successive couplings and Fmoc-deprotections were repeated to achieve the desired peptide sequence. A TNBS test was used to verify couplings with steps repeated as necessary. A small amount (micro spatula-full) of resin was transferred and tested for the presence of a free amine by addition of solution D1 (5 μL) followed by solution D2 (5 μL). Clear/yellow resin beads indicate no free amine and red/orange beads indicate presence of free amine. For fluorescently tagged peptides, a solution of fluorescein-5-isothiocyanate (5 equiv) and DIPEA (10 equiv) in DMF (5 mL) was added to the resin and

stirred intermittently for 2 h. After the final Fmoc-deprotection or the fluorophore attachment, the resin was washed with DMF (5 x 5 mL), DCM (3 x 5 mL) and DMF (5 x 5 mL).

General Method 1a: Maleimide Functionalisation

A solution of 4-malemidobutyric acid (3 equiv), solution B (3 equiv, 600 μL) and DIPEA (10 equiv, 348 μL) in DMF (5 mL) was added to the *N*-terminal deprotection peptide on resin and stirred intermittently for 2 h. The solution was removed and the resin washed with DMF (5 x 5 mL), DCM (3 x 5 mL) and DMF (5 x 5 mL). A small amount (micro spatula-full) of resin was transferred and tested for the presence of a free amine by addition of solution D1 (5 μL) followed by solution D2 (5 μL).

General Method 1b: Acetylation

Solution E (10 mL) was added to the *N*-terminal deprotected peptide on resin and stirred intermittently for 15 min. The solution was removed and the resin successively washed with DMF (5 x 10 mL), DCM (2 x 10 mL) and DMF (5 x 10 mL).

General Method 2: Cleavage

The resin with the desired peptide sequence was dried by washing with DMF (5 x 5 mL), DCM (3 x 5 mL), DMF (5 x 5 mL) and diethyl ether (3 x 5 mL) with remaining ether allowed to evaporate overnight. The resin was added to solution C1 (10 mL) and rocked for 2 h*. The TFA solution was pipetted from the resin and concentrated to 0.5-1 mL under a nitrogen stream, then peptide precipitated with diethyl ether (10 mL) and the mixture cooled to -10°C overnight. The precipitate was pelleted by centrifugation (7600 rpm, 10 min), dried under a nitrogen stream, and dissolved in 1:1 ACN/H$_2$O, before syringe filtering and lyophilization to give the crude linear peptide as a fluffy powder.

104

* For arginine-rich, charged peptides **N1F**, **N2F**, **N3F and N4F** the cleavage time was extended to 5 h to ensure complete deprotection and scavenging of protecting groups

General Method 2a: Cleavage

The resin with completed peptide was dried by washing with DMF (5 x 5 mL), DCM (3 x 5 mL), DMF (5 x 5 mL) and diethyl ether (3 x 5 mL) with remaining ether allowed to evaporate overnight. The resin was added to solution C2 (10 mL) and rocked for 5 h. The TFA solution was pipetted from the resin and concentrated to 0.5-1 mL under a nitrogen stream, then peptide precipitated with diethyl ether (10 mL) and the mixture cooled to -10°C overnight. The precipitate was pelleted by centrifugation (7600 rpm, 10 min), dried under a nitrogen stream, and dissolved in 1:1 ACN/H_2O, before syringe filtering and lyophilization to give the crude linear peptide as a fluffy powder.

General Method 3: Automated Purification

Purification was performed by semi-preparative RP-HPLC using a Gilson GX-Prep HPLC on a Phenomenex Luna C18(2) column (250 x 10 mm) unless stated otherwise. The solvent gradient used is as stated in each protocol using solvent A = 0.1% TFA in water and solvent B = 0.1% TFA in acetonitrile. Appropriate fractions were combined and lyophilised on a Christ freeze dryer to give the purified product as a powder.

General Method 4: Thioether Conjugation Reaction

The cysteine-containing NLS peptide (15 mg) and maleimide-containing p21 peptide (1 equiv) were each dissolved in water (3 mL) and the solutions combined. The solution was shaken on an orbital rocker for 24 h, and the reaction progress monitored by RP-HPLC and MS. The solution was then lyophilised to give the crude peptide conjugate.

General Method 5: Bimane Cyclisation Method

Following linear peptide assembly, the Mmt protecting groups were selectively removed from the cysteine side chains by treatment of resin with 2% TFA in DCM (5 mL) for 1 min. The resin was washed with DCM (3 x 5 mL) and the TFA treatments were repeated until the solution was colourless upon addition to the resin (200-250 mL total volume). The resin was washed with DCM (5 x 5 mL) and DMF (5 x 5 mL). The resin was treated with a solution of dibromobimane (2 equiv) and DIPEA (4 equiv) in DMF (6 mL) for 3 h, stirring intermittently. The solution was then drained and the resin washed with DMF (5×5 mL) and DCM (5×5 mL).

General Method 6: On-resin derivatisation method:

The peptide sequence was assembled per '*General Method 1: Solid-Phase Peptide Synthesis*', where the *N*-terminus was protected with a Boc group (Boc-Lys(Boc)-OH). Fmoc-Dap(Alloc)-OH was incorporated at the desired position as per the standard coupling conditions. The resin was washed with DCM (5 mL). The resin was suspended in dry DCM (5 mL) in a sealed flask and purged N_2. Alloc-protecting group was then selectively removed by addition of Pd(PPh$_3$)$_4$ (0.8 equiv) and phenylsilane (25 equiv) to the resin and bubbled with N_2 for 15 min. The mixture was transferred to a fritted syringe and the solution drained and resin washed with DCM (3 x 5 mL). The resin was resuspended in dry DCM (5 mL) and transferred to a sealed flask and the Pd treatment repeated as before, twice more. The resin was then suspended in a solution of *N,N'-dimethyl-4-aminophthalic anhydride* (**5**) (2 equiv) and DIPEA (4 equiv) in NMP (5 mL) in a sealed flask purged with N_2, and bubbled with nitrogen overnight. The resin was then transferred to a fritted syringe, the solution drained and the resin washed with DCM (3 x 5 mL) and DMF (3 x 5 mL). Fluorophore ring closure was achieved by treatment of the resin with a solution of HBTU (6 equiv), HOBt (6 equiv) and

DIPEA (12 equiv) in NMP (12.5 mL) for 2 h, with intermittent stirring. The solution was drained and the resin washed with DMF (3 x 5 mL) and DCM (3 x 5 mL). The HBTU/HOBt treatment was repeated as before, twice more. The solution was drained and the resin washed with DMF (3 x 5 mL), DCM (3 x 5 mL) and cleaved from the resin as described by '*General Method 2: Cleavage*'.

6.3   Analysis

6.3.1 Analytical Methods

Product purity was confirmed by RP-HPLC on a Phenomenex Luna C18(2) column. Product identity was confirmed via High Resolution mass spectrometry on an Agilent 6230 ESI-TOF LCMS.

6.3.2 NMR spectroscopy

NMR spectra were acquired on an Agilent 500 MHz spectrometer or Oxford 600 MHz Spectrometer as specified in the appropriate compound experimental detailed as follows.

6.3.3 Plate reader experiments

All fluorescence values were obtained on a H4 Synergy Plate Reader, with Xenon light source and a slit width of 9.0. Gain was varied dependent on the concentration of the samples.

Solvent Comparison

A Tris buffer containing 20 mM Tris.HCl pH 7.4 with 100 mM NaCl, 1 mM EDTA, 2 mM DTT & 0.05% Tween20 was made. The peptides **B1**, **B2** and **B3** (5 μM) were dissolved in buffer, or dioxane with 5 mM 18-crown-6. The six solutions were then plated in triplicate (60 μL per well) into a black, clear-bottomed 96-well plate (CoStar 3905). Fluorescence spectra

were collected immediately with measurements obtained from the top of the plate, and a 5 nm step was used. An absorbance spectrum was first collected (300-700 nm). An emission spectrum for the peptides was collected (420-700 nm) at an excitation of 390 nm, and an excitation spectrum was collected (300-500 nm) with emission 520 nm.

Fluorescence response in the presence of PCNA

A Tris buffer containing 20 mM Tris.HCl pH 7.4 with 100 mM NaCl, 1 mM EDTA, 2 mM DTT & 0.05% Tween20 was made. The peptides **B1**-**B3** (1.6 µM) were prepared in buffer and combined with PCNA (1.6 µM) in buffer, in a 1:2 ratio. The three solutions were then plated in triplicate (60 µL per well) into a black 96-well plate (CoStar 3915). As controls, a sample of each peptide in buffer (800 nM), PCNA in buffer (600 nM) and buffer alone were also plated in triplicate. The plate was incubated at 4°C overnighted before fluorescence measurements were collected from the top of the plate and with a 10 nm step. An emission spectrum for was collected (410-700 nm) with excitation 390 nm was collected for **B1-B3**, and the controls.

PCNA titration

A HEPES buffer containing 10 mM HEPES, 150 mM NaCl, 3 mM EDTA, 0.05% Tween20 and adjusted to pH 7.4 with 2 N NaOH was made. B3 (6 µM) was dissolved in buffer (600 µL). PCNA was prepared in buffer (120 µL) at concentration: 0.5, 1, 2, 4, 6, 10, 15, 20, 30, 40 and 90 µM. The **B3** sample in buffer (50 µL) was then combined with PCNA at each concentration (50 µL) to give a final **B3** concentration of 3 µM and final PCNA concentrations as 0.25, 0.5, 1, 2, 3, 5, 7.5, 10, 15, 20 and 45 µM, respectively. The resulting 11 solutions were incubated for 15 min at rt. Each solution was then plated in triplicate (30 µL per well) in a black, clear bottom 384-well plate (Grenier Bio-One, 781209). As controls, PCNA (20 µM), B3 (3 µM) and buffer alone were plated in triplicate (30 µL per well). A

fluorescence spectrum (5 nm step) was collected for B3 with ex. 390 nm (em. 410-700 nm), ex. 420 nm (em. 440-700 nm), ex. 440 nm (em. 460-700 nm) and ex. 408 nm (em. 430-700 nm). A fluorescence point read was also collected with ex. 390, em. 567 nm; ex. 420 nm, em. 567 nm; ex. 440 nm, em. 550 nm; ex. 408 nm, em. 550 nm. All measurements were taken from the top of the plate with a 9.0 slit width used.

6.3.4 SPR experiments

A running buffer of 10 mM HEPES buffer with 150 mM NaCl, 3 mM EDTA and 0.05% Tween20 was made and adjusted to pH 7.4 with 10 M NaOH. A sensor chip (GE CM5, series S) was primed with the running buffer and preconditioned per the manufacturer's protocol with sequential injections (2 x 50 s, 30 µL/min) of 50 mM NaOH, 10 mM HCl, 0.1% SDS, 0.85% H_3PO_4 and glycine (pH 9.5). The sensor chip surface was activated with an injection of 0.2 M EDC and 50 mM NHS (600 s, 10 µL/min). The running buffer (245 µL) was used to dilute PCNA (5 µL, 12 mg/mL). Once preactivation step was complete, PCNA was further diluted in 10 mM NaOAc (~ pH 4.6), to a final concentration of 25 µL/mL by adding PCNA/HEPES (50 µL) to a solution of 100 mM NaOAc (50 µL) and water (400 µL). The final PCNA solution was injected over a flow cell at 10 µL/min until a stable ~1500 RU was reached. Both flow cells were then blocked with 1.0 M ethanolamine (pH 8.5) (600 s, 10 µL/min) before the chip was left for two hours to stabilise. After stabilisation, a final protein level of 1390 RU was achieved.

Peptides (~2 mg) were dissolved in milliQ H_2O (40 µL) and centrigued (7800 rpm, 10 min) to remove particulates. UV absorbance was used to determine peptide stock concentration, where 2 µL of the stock was diluted in water (40-fold) and triplicate measurements were taken with a Nanodrop2000, baselined to 750 nm absorbance.

The ε_λ and λ used for each peptide is specified in Table. The concentration of the stock solution was then calculated using $c = (A_\lambda/\varepsilon_\lambda l) \, DF$, where c is in molar, A_λ is the average absorbance across the triplicates, λ is the appropriate wavelength for each peptide, l is the pathlength (0.1 cm for Nanodrop), ε_λ is the molar absorptivity and DF is the dilution factor. Peptides were diluted into running buffer before further dilution as required.

Steady state affinity experiments were run at a 30 µL/min flow rate, with a 40 s starting contact time and a 60 s dissociation which was extended if a steady state was not reached. Each peptide was serially diluted 1 in 2, eight times and injected sequentially from lowest to highest concentration following a buffer only injection. Each injection was followed by a regeneration with 2 M NaCl (2 x 30 s, 30 µL/min). The data was analysed with the GE Biosystems software Biacore S200 Evaluation software. Data is summarised in Table S1.

Table S1: SPR data for peptides. λ = wavelength. ε_λ = extinction coefficient at given wavelength. Top Conc.= top concentration of peptide used in steady state affinity experiments by SPR. K_D = binding affinity of peptide for PCNA. SE= standard error. χ^2= chi^2 (gives a measure of accuracy of fitting). Ass/Diss time (s) = association/dissociation time of peptide binding to PCNA during steady state affinity experiments.

Peptide	λ (nm)	ε_λ	Top Conc. (nM)	Affinity K_D (nM)	K_D SE (nM)	χ^2	Ass/Diss time (s)
P1	205	56820	500	102	5.30	0.0701	
P2	205	67867	500	41.8	5.80	1.01	40/60
P3	205	65080	1000	45.9	4.30	0.301	40/60
P4	205	76120	500	37.9	10.0	7.06	60/90
P2F	490	86983	1000	202	18.0	0.556	40/60
P3F	490	86983	2000	190	8.70	0.100	40/60
P4F	490	86983	500	61.0	5.60	1.09	40/60
P1b-N1F	490	86983	1000	100	9.50	3.60	40/60
P1b-N2F	490	86983	250	33.1	4.20	2.62	40/60
P1b-N3F	490	86983	2000	130	15.0	0.770	40/60
P1b-N4F	490	86983	1000	372	30.0	3.70	40/60
P1bim	380	4694	10000	887	15.4	0.880	40/60
P1c-N2	380	4694	1000	175	27.0	5.58	40/40
P1bimF	380	4694	30000	25.2 μM	1.90 μM	0.0547	40/60
P1c-N2F	380	4694	3000	535	73.0	4.92	40/60
P5	205	67860	500	16.3	0.516	0.0855	40/60
B1*	420[56]	6500	20000	> 2.50 μM	-	-	40/60
B2*	420[56]	6500	10000	461	570	3.02	40/60
B3	420[56]	6500	2500	128	260	0.0298	40/60
B4*	420[56]	6500	100	25.6 μM	2.97 μM	0.298	60/90
B5*	450[56]	8800	100	121 μM	82.5 μM	0.267	60/90
*The interaction of these peptides with PCNA was deemed non-specific							

6.4 Cell permeability assay

Intracellular visualisation of peptides inT47D cells using fluorescence microscopy

T47D breast cancer cells were cultured at 37°C and 5% CO_2 in a 6-well plate, on glass coverslips in 2 mL of media per well consisting of RPMI (Sigma, R0883) supplemented with 2 mM L-Glutamine (Sigma (G7513) and 10% FBS. The cells were leftto reach 70% confluency before 4, 24 and 48 hour treatment with 5 uM of peptide. During and after treatment, samples were protected from light as much as possible. Following treatment, cells were washed with ice cold PBS (Gibco, 14190144) two times for 5 min. The cells were fixed

with 4% PFA (10% Neutral Nuffered Formalin, ChemSupply, #1258) for 10 min at rt. Samples were then washed twice for 5 min with PBS at rt and then permeabilised in 5% TritonX solution for 1 h at rt. The samples were washed twice for 5 min with PBS at rt before being stained with Phalloidin (Alexa Fluor 568 Phalloidin, Invitrogen A12380; 1/40 in 5% BSA in PBS) for 20 minutes. The samples were washed twice for 5 min with PBS at rt before the nuclei in the samples were stained with DAPI (4',6-Diamidino-2-Phenylindole, Dihydrochloride, Invitrogen D1306; 1/2000 in PBS) for 1 minute. The sample was washed twice with PBS and the coverslips mounted onto slides using DAKO fluorescent mounting medium (S302380-2) and sealed with clear nail polish (Sally Hansen). The slides were left to dry overnight, and the samples were imaged the following day with Olympus IX73 Inverted Fluorescence Microscope (Adelaide Microscopy).

The cell nuclei were visualised using a standard UV long-pass filter. For fluorescein-tagged peptides, a standard Intermediate-blue long-pass filter was used. For visualisation of cellular F-actin (phalloidin staining) a standard Cy5 long-pass filter was used. The images were taken with a 40x and 100x (oil) objectives.

Intracellular visualisation of peptides in MDA-MB-468 cells using confocal microscopy

MDA-MB-468 mKate is a breast cancer cell line which is modified to express nuclear fluorescent mKate protein (ex. 588 nm, em. 635 nm). MDA-MB-468 mKate cells were cultured on glass coverslips at 37°C and 5% CO_2 in a 6-well plate in media (2 mL) containing DMEM (Sigma, D5671) supplemented with 2 mL L-Glutamine (Sigma, G7513 and 1 mM sodium pyruvate (Sigma, S8636). The cells were left to reach 70% confluency before 24 h treatment with 10 μM peptide. Once the cells were treated, samples were protected from light as much as possible. Following treatment, the cells were washed with ice cold PBS (Gibco, 14190144) (2 x 5 min). The cells were fixed with 4% PFA (10% Neutral Buffered Formalin,

ChemSupply, #1258) for 10 min at rt and then washed with PBS (2 x 5 min). The coverslips were mounted onto slides with DAKO fluorescent mounting medium (S302380-2) and sealed with clear nail polish (Sally Hansen) before being left to dry overnight. Samples were imaged with Confocal Olympus FV3000 microscope (Adelaide Microscopy).

The nuclei of cells were visualised using a 594nm laser at a detection of 600-700 nm. For imaging the FITC fluorophore, a 488 nm laser was used at a detection of 490-534 nm. For imaging of the bimane fluorophore, a 405 nm laser was used at a detection of 410-485 nm. The images were taken using a 30x silicon oil objection with a 2x zoom setting (Olympus, Cell Sens), bringing the total magnification to 60x.

6.5   Syntheses

Note: All peptides presented are C-terminally amidated.
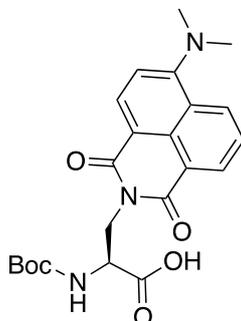
Boc-DMNA   **(7)**,   *N-α-Boc-(4-N,N'-Dimethylamino-1,8-nathpthalamido)alanine.*   4-Dimethylamino-1,8-nathphthalic anhydride (0.50 g, 2.07 mmol) was dissolved in 1,4-dioxane (25 mL) brought to reflux under an N_2 atmosphere. A suspension of 3-amino-2-(Boc-amino)-propionic acid (Boc-Dap-OH, 0.42 g, 1 equiv, 2.07 mmol) and sodium bicarbonate (0.87 g, 5 equiv, 10.4 mmol) in water (6.25 mL) was then added slowly to the naphthalic anhydride solution over 15 min, such that reflux was maintained. The reaction proceeded at reflux for a further 1h before cooling to rt and concentrated *in vacuo* to ~20 mL. The residue was taken up in water (50 mL) and washed with ether (3 x 100 mL). The aqueous layer was then acidified to pH 1 with 1 M HCl and extracted with DCM (6 x 100 mL). The combined organic layers were dried with MgSO$_4$, filtered and the solvent removed *in vacuo* to afford N-α-Boc-(4-N,N'-dimethylamino-1,8-nathpthalamido)alanine as an orange residue. ^{1}H NMR (500 MHz, CDCl$_3$) δ 8.62 – 8.55 (m, 1H), 8.49 (dd, *J* = 8.1, 2.6 Hz, 1H), 8.46 (d, *J* = 8.9 Hz, 1H), 7.68 (q, *J* = 7.6 Hz, 1H), 7.12 (dd, *J* = 8.3, 1.6 Hz, 1H), 4.76 – 4.52 (m, 2H), 3.70 (d, *J* = 1.0 Hz, 10H), 3.24 – 3.06 (m, 6H).
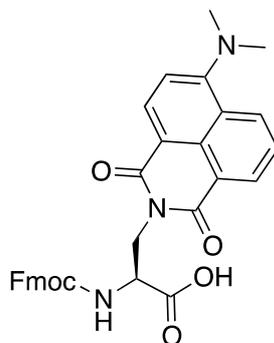
Fmoc-DMNA (8), *N-α-Fmoc-(4-N,N'-Dimethylamino-1,8-nathpthalamido)alanine*. N-α-Boc-(4-N,N'-Dimethylamino-1,8-nathpthalamido)alanine (0.85 g, 1.99 mmol) was dissolved in a 1:1 solution of DCM/TFA (40 mL) and stirred at rt for 2 h. The reaction mixture was concentrated *in vacuo*, azeotroped with chloroform (3 x 50 mL) and then redissolved in water (12 mL) with NaHCO$_s$ (0.84 g, 5 equiv, 9.95 mmol). A solution of Fmoc-Osu (0.74 g, 1.1 equiv, 2.19 mmol) was in dioxane (25 mL) was added to the amino acid solution over 5 min. The reaction mixture was stirred at rt for 3 h and then concentration to ~20 mL *in vacuo*. The residue was dissolved in water (50 mL) and washed with ether (5 x 50 mL). The aqueous layer was acidified to pH 1 with 1 M HCl and extracted with DCM (6 x 100 mL). The combined organic layers were dried with MgSO$_4$ and concentrated *in vacuo*. The crude product was loaded onto celite and purified by column chromatography on silica using 3:1 ethyl acetate/hexane with 1% AcOH. The pure fractions were combined and the solvent removed *in vacuo*. The residue was azeotroped with toluene to give the Fmoc-DMNA as an orange solid (0.22 g, 20% yield over two steps). NMR was consistent with literature: [56] ^{1}H NMR (500 MHz, CDCl$_3$) δ 8.58 (d, *J* = 7.3 Hz, 1H), 8.46 (d, *J* = 8.3 Hz, 1H), 8.40 (d, *J* = 8.4 Hz, 1H), 7.68 (t, *J* = 7.4 Hz, 2H), 7.62 (t, *J* = 8.0 Hz, 1H), 7.50 (d, *J* = 6.7 Hz, 1H), 7.46 (d, *J* = 7.5 Hz, 1H), 7.32 (t, *J* = 7.7 Hz, 2H), 7.25 – 7.20 (m, 1H), 7.20 – 7.13 (m, 1H), 7.03 (d, *J* = 8.3 Hz, 1H), 6.11 (d, *J* = 7.4 Hz, 1H), 4.88 (s, 1H), 4.71 (dd, *J* = 20.3, 11.1 Hz, 2H), 4.22 (d, *J* = 8.6 Hz, 1H), 4.16 (t, *J* = 9.3 Hz, 1H), 4.00 (s, 1H), 3.07 (s, 6H) ppm; HRMS (ESI+) Expected [M+H]$^+$ for C$_{32}$H$_{27}$N$_3$O$_6$ (549.1900): 500.1978, observed: [M+H]$^+$ 500.1965.
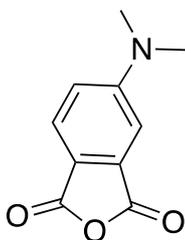
4-DMAP anhydride (5), *N,N'-dimethyl-4-laminophthalic anhydride..* A suspension of 4-aminophthalic acid (0.50 g, 2.76 mmol) in methanol (150 mL) was placed under a N_2 atmosphere before gradual addition of formaldehyde (37% wt., 15 mL), followed by Pd/C (200 mg, 1.88 mmol). The reaction was allowed to proceed under a H_2 atmosphere for 3 h. The solution was then filtered through Celite and the celite washed with methanol (2 x 50 mL), the filtrate was dried under vacuum to give the corresponding diacid. The diacid was suspended in acetic anhydride (50 mL) and stirred at 50 °C under a N_2 atmosphere for 12 h. The rection mixture was concentrated *in vacuo* before azeotroping with toluene (2 x 100 mL). The crude product was purified by column chromatography on silica using DCM to give the final product as yellow crystals (0.35 g, 68% over two steps). NMR was consistent with literature:[55] ^{1}H NMR (600 MHz, CDCl$_3$) δ 7.75 (d, *J* = 8.6 Hz, 1H), 7.08 (d, *J* = 2.4 Hz, 1H), 6.94 (dd, *J* = 8.7, 2.4 Hz, 1H), 3.15 (s, 6H) ppm.

P2, *RKRRQTSMTDFYHSK.* The peptide was synthesised as per *General Method 1*, and then cleaved from resin by *General Method 2*. The crude peptide was purified per *General Method 3*, to afford a white powder. Analytical RP-HPLC: 0-100% over 15 min, R_t (C18) = 7.5 min; HRMS (ESI+) Expected [M+2H]$^{2+}$ for [C$_{82}$H$_{134}$N$_{30}$O$_{23}$S]: 970.5057, observed: [M+2H]$^{2+}$ 970.5058.

P2F, FITC-Aβ*RKRRQTSMTDFYHSK*. The peptide was synthesised as per *General Method 1*, and then cleaved from resin by *General Method 2*. The crude peptide was purified per *General Method 3*, to afford a yellow powder. Analytical RP-HPLC: 0-50% over 15 min, R_t (C18) = 11.2 min; HRMS (ESI+) Expected $[M+4H]^{4+}$ for $[C_{106}H_{152}N_{32}O_{29}S_2]$: 600.7750, observed: $[M+4H]^{4+}$ 600.7727.

P3, *RQTSMTDFYHSKRR*. The peptide was synthesised as per *General Method 1*, and then cleaved from resin by *General Method 2*. The crude peptide was purified per *General Method 3*, to afford a white powder. Analytical RP-HPLC: 0-50% over 15 min, R_t (C18) = 8.9 min; HRMS (ESI+) Expected $[M+4H]^{4+}$ for $[C_{76}H_{122}N_{28}O_{22}S]$: 453.7330, observed: $[M+4H]^{4+}$ 453.7330.

P3F, FITC-Aβ*RQTSMTDFYHSKRR*. The peptide was synthesised as per *General Method 1*, and then cleaved from resin by *General Method 2*. The crude peptide was purified per *General Method 3*, to afford a yellow powder. Analytical RP-HPLC: 0-50% over 15 min, R_t (C18) = 11 min; HRMS (ESI+) Expected $[M+6H]^{6+}$ for $[C_{100}H_{140}N_{30}O_{28}S_2]$: 379.5040, observed: $[M+6H]^{6+}$ 379.5032.

P4, *RKRRQTSMTDFYHSKRR*. The peptide was synthesised as per *General Method 1*, and then cleaved from resin by *General Method 2*. The crude peptide was purified per *General Method 3*, to afford a white powder. Analytical RP-HPLC: 0-100% over 15 min, R_t (C18) = 7.5min; HRMS (ESI+) Expected $[M+4H]^{4+}$ for $[C_{94}H_{158}N_{36}O_{25}S_1]$: 563.8073, observed: $[M+4H]^{4+}$ 563.8068.

P4F, FITC-Aβ*RKRRQTSMTDFYHSKRR*. The peptide was synthesised as per *General Method 1*, and then cleaved from resin by *General Method 2*. The crude peptide was purified per *General Method 3*, to afford a yellow powder. Analytical RP-HPLC: 0-50% over 15 min, R_t (C18) = 10 min; HRMS (ESI+) Expected $[M+3H]^{3+}$ for $[C_{118}H_{174}N_{40}O_{31}S_2]$: 904.7683, observed: $[M+3H]^{3+}$ 904.7681.

N1F, *FITC-AβGRKKRRQRRRC*. The peptide was synthesised as per *General Method 1*, and then cleaved from resin by *General Method 2*. The crude peptide was purified per *General Method 3*, to afford a yellow powder. Analytical RP-HPLC: 0-50% over 15 min, R_t (C18) = 9.5 min; HRMS (ESI+) Expected $[M+3H]^{3+}$ for $[C_{82}H_{131}N_{35}O_{18}S_2]$: 653.6696, observed: $[M+3H]^{3+}$ 653.6609.

N2F, *FITC-AβPKKKRKVC*. The peptide was synthesised as per *General Method 1*, and then cleaved from resin by *General Method 2*. The crude peptide was purified per *General Method 3*, to afford a yellow powder. Analytical RP-HPLC: 0-50% over 15 min, R_t (C18) = 11.2 min; HRMS (ESI+) Expected $[M+4H]^{4+}$ for $[C_{67}H_{100}N_{18}O_{14}S_2]$: 362.1855, observed: $[M+4H]^{4+}$ 362.1915.

N3F, *FITC-AβPAAKRVKLDC*. The peptide was synthesised as per *General Method 1*, and then cleaved from resin by *General Method 2*. The crude peptide was purified per *General Method 3*, to afford a yellow powder. Analytical RP-HPLC: 0-50% over 15 min, R_t (C18) = 12.5 min; HRMS (ESI+) Expected $[M+4H]^{4+}$ for $[C_{71}H_{102}N_{18}O_{18}S_2]$: 390.6843, observed: $[M+4H]^{4+}$ 390.6866.

N4F, *FITC-AβRRWWRRWRRC*. The peptide was synthesised as per *General Method 1*, and then cleaved from resin by *General Method 2*. The crude peptide was purified per *General Method 3*, to afford a yellow powder. Analytical RP-HPLC: 0-50% over 15 min, R_t (C18) = 13 min; HRMS (ESI+) Expected $[M+3H]^{3+}$ for $[C_{96}H_{126}N_{34}O_{68}S_2]$: 692.6589, observed: $[M+3H]^{3+}$ 692.6472.

N2, *PKKKRKVC*. The peptide was synthesised as per *General Method 1*, and then cleaved from resin by *General Method 2*. The crude peptide was purified per *General Method 3*, to afford a white powder. Analytical RP-HPLC: 0-50% over 15 min, R_t (C18) = 5.2 min; HRMS (ESI+) Expected $[M+2H]^{2+}$ for $[C_{43}H_{84}N_{16}O_8S]$: 493.3267, observed: $[M+2H]^{2+}$ 493.3270.

P1b, *mal-RQTSMTDFYHSK*. The peptide was synthesised as per *General Method 1*, and then cleaved from resin by *General Method 2a*. The crude peptide was purified per *General Method 3*, to afford a white powder. Analytical RP-HPLC: 0-50% over 15 min, R_t (C18) = 5.2 min; HRMS (ESI+) Expected $[M+4H]^{4+}$ for $[C_{72}H_{105}N_{21}O_{23}S]$: 416.9430, observed: $[M+4H]^{4+}$ 416.9403.
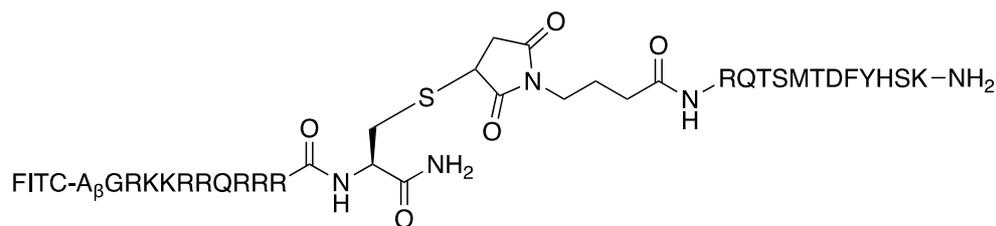

P1b-N1F, *FITC-AβGRKKRRQRRRC[suc-RQTSMTDFYHSK]*. Peptide conjugation of purified **P1b** and **N1F** was carried out by *General Method 4: Thioether Conjugation Reaction* and the peptide was purified per *General Method 3* to afford a yellow powder. Analytical RP-HPLC: 0-50% over 15 min, R_t (C18) = 9.5 min; HRMS (ESI+) Expected $[M+6H]^{6+}$ for $[C_{154}H_{236}N_{56}O_{41}S_3]$: 604.6289, observed: $[M+6H]^{6+}$ 604.6287.
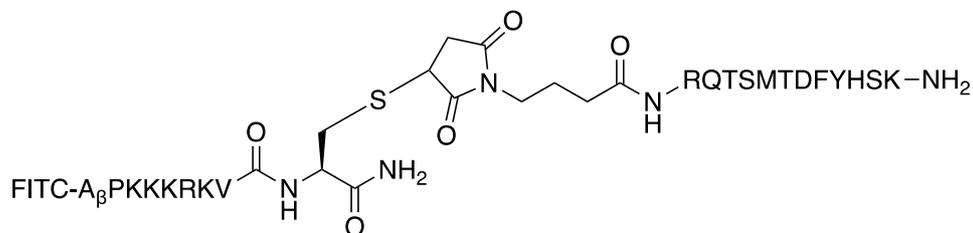
P1b-N2F, *FITC-AβPKKKRKVC[suc-RQTSMTDFYHSK].* Peptide conjugation of purified

P1b and **N3F** was carried out by *General Method 4: Thioether Conjugation Reaction* and the

peptide was purified per *General Method 3* to afford a yellow powder. Analytical RP-HPLC:

0-50% over 15 min, R_t (C18) = 11.2 min; HRMS (ESI+) Expected $[M+5H]^{5+}$ for

$[C_{139}H_{205}N_{39}O_{37}S_3]$: 622.6982, observed: $[M+5H]^{5+}$ 622.6968.
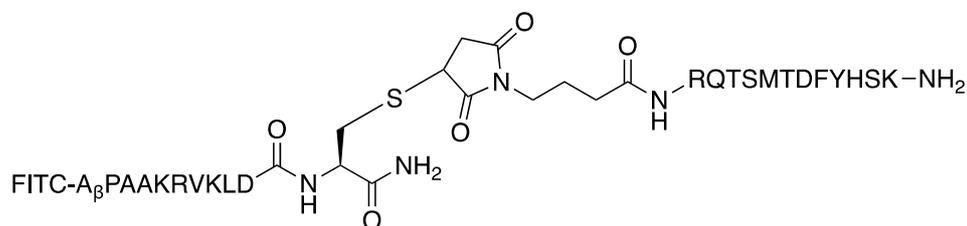
P1b-N3F, *FITC-AβPAAKRVKLDC[suc-RQTSMTDFYHSK].* Peptide conjugation of purified

P1b and **N3F** was carried out by *General Method 4: Thioether Conjugation Reaction* and the

peptide was purified per *General Method 3* to afford a yellow powder. Analytical RP-HPLC:

0-50% over 15 min, R_t (C18) = 12.5 min; HRMS (ESI+) Expected $[M+5H]^{5+}$ for

$[C_{143}H_{207}N_{39}O_{41}S_3]$: 645.4973, observed: $[M+5H]^{5+}$ 645.4969.
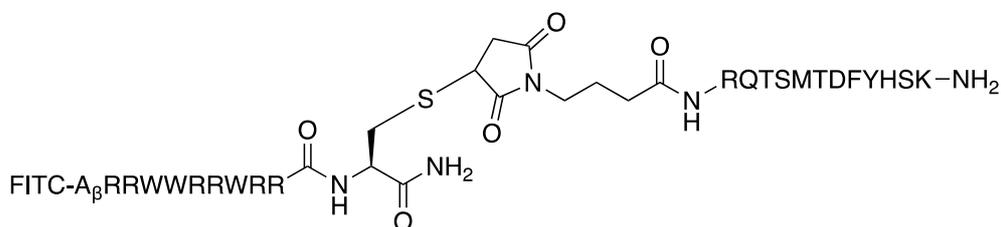
P1b-N4F, *FITC-AβRRWWRRWRRC[suc-RQTSMTDFYHSK].* Peptide conjugation of purified

P1b and **N4F** was carried out by *General Method 4: Thioether Conjugation Reaction* and the

peptide was purified per *General Method 3* to afford a yellow powder. Analytical RP-HPLC:

0-50% over 15 min, R_t (C18) = 13 min; HRMS (ESI+) Expected $[M+6H]^{6+}$ for

$[C_{168}H_{231}N_{55}O_{39}S_3]$: 624.1236, observed: $[M+6H]^{6+}$ 624.1239.
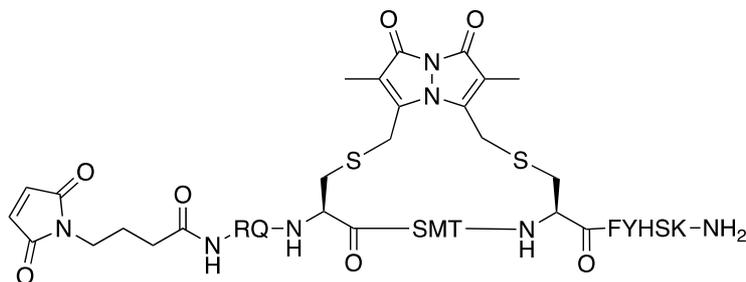
P1c, *mal-RQC(-)SMTC(Bim-)FYHSK*. The linear peptide was synthesised per *General Method 1*. The peptide was then cyclised by the *General Method 5: Bimane Cyclisation Method* and then cleaved from resin by *General Method 2b*. The crude peptide was purified per *General Method 3*, to afford a white powder. Analytical RP-HPLC: 0-50% over 15 min, R_t (C18) = 10.6 min; HRMS (ESI+) Expected $[M+3H]^{3+}$ for $[C_{80}H_{111}N_{23}O_{22}S_3]$: 614.9223, observed: $[M+3H]^{3+}$ 614.9252.
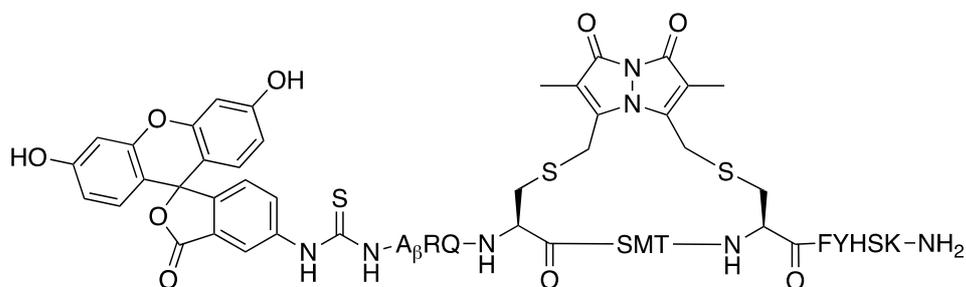
P1bimF, *FITC-AβRQC(-)SMTC(Bim-)FYHSK-NH₂*. The linear peptide was synthesised per *General Method 1*. The peptide was cyclised by the *General Method 5: Bimane Cyclisation Method* prior to β-Alanine coupling and FITC attachment. The cyclised and fluorescently tagged peptide was then cleaved from resin by *General Method 2a*. The crude peptide was purified per *General Method 3*, to afford a yellow powder. Analytical RP-HPLC: 0-50% over

15 min, R_t (C18) = 12 min; HRMS (ESI+) Expected $[M+3H]^{3+}$ for $[C_{97}H_{122}N_{24}O_{24}S_4]$: 712.6059, observed: $[M+3H]^{3+}$ 712.5945.
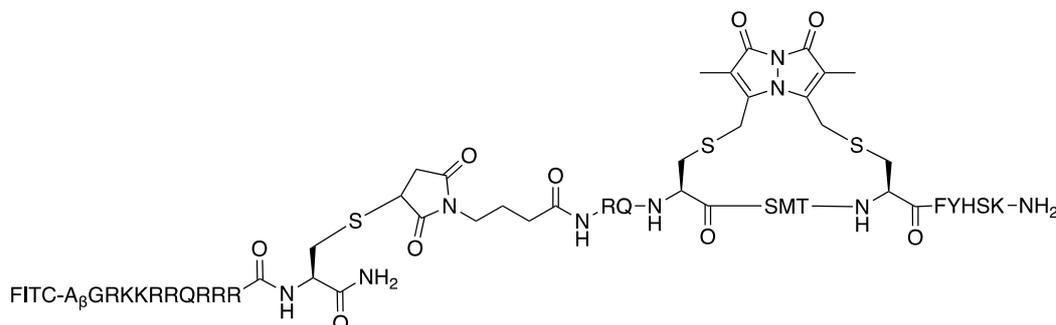
P1c-N1F, *FITC-A$_\beta$GRKKRRQRRRC[suc-RQC(-)SMTC(Bim)FYHSK].* Peptide conjugation of purified **P1c** and **N1F** was carried out by *General Method 4: Thioether Conjugation Reaction* and the peptide was purified per *General Method 3* to afford a yellow powder. Analytical RP-HPLC: 0-50% over 15 min, R_t (C18) =11 min; HRMS (ESI+) Expected $[M+5H]^{5+}$ for $[C_{162}H_{242}N_{58}O_{40}S_5]$: 760.9536, observed: $[M+5H]^{5+}$ 760.9534.
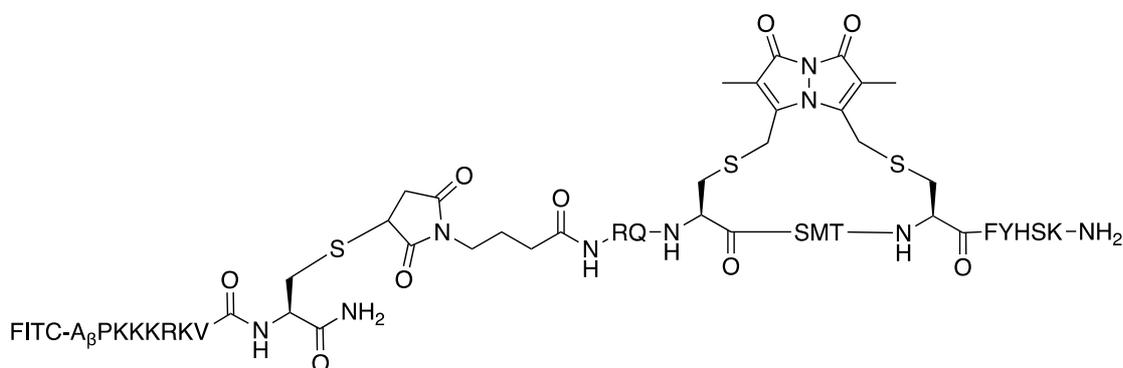
P1c-N2F, *FITC-A$_\beta$PKKKRKVC[suc-RQC(-)SMTC(Bim)FYHSK-NH2].* Peptide conjugation of purified **P1c** and **N2F** was carried out by *General Method 4: Thioether Conjugation Reaction* and the peptide was purified per *General Method 3* to afford a yellow powder. Analytical RP-HPLC: 0-50% over 15 min, R_t (C18) =11.2 min; HRMS (ESI+) Expected $[M+6H]^{6+}$ for $[C_{147}H_{211}N_{41}O_{36}S_5]$: 548.7502, observed: $[M+6H]^{6+}$ 548.7506.
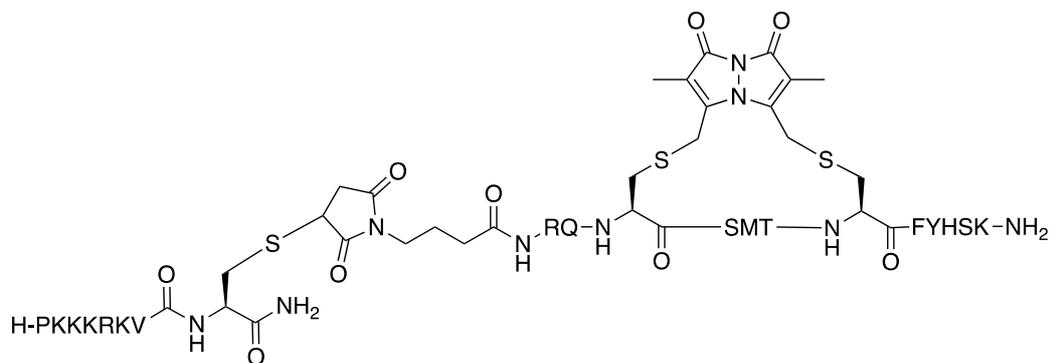
P1c-N2, *PKKKRKVC[suc-RQC(-)SMTC(Bim)FYHSK].* Peptide conjugation of purified **P1c** and **N2** was carried out by *General Method 4: Thioether Conjugation Reaction* and the peptide was purified per *General Method 3* to afford a white powder. Analytical RP-HPLC: 0-50% over 15 min, R_t (C18) = 9.5 min; HRMS (ESI+) Expected $[M+4H]^{4+}$ for $[C_{123}H_{195}N_{39}O_{30}S_4]$: 707.6032, observed: $[M+4H]^{4+}$ 707.6026.
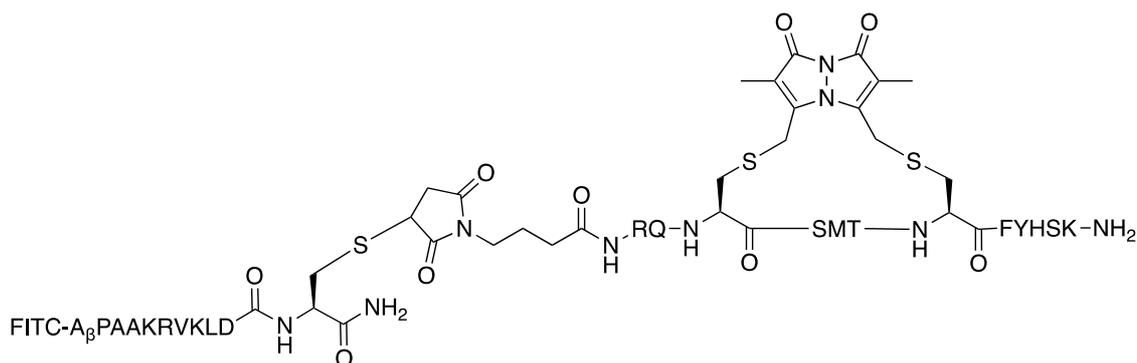
P1c-N3F, *FITC-A_βPAAKRVKLDC[suc-RQC(-)SMTC(Bim)FYHSK].* Peptide conjugation of purified **P1c** and **N3F** was carried out by *General Method 4: Thioether Conjugation Reaction* and the peptide was purified per *General Method 3* to afford a yellow powder. Analytical RP-HPLC: 0-50% over 15 min, R_t (C18) = 12 min; HRMS (ESI+) Expected $[M+6H]^{6+}$ for $[C_{151}H_{213}N_{41}O_{40}S_5]$: 567.7494, observed: $[M+6H]^{6+}$ 567.7497.
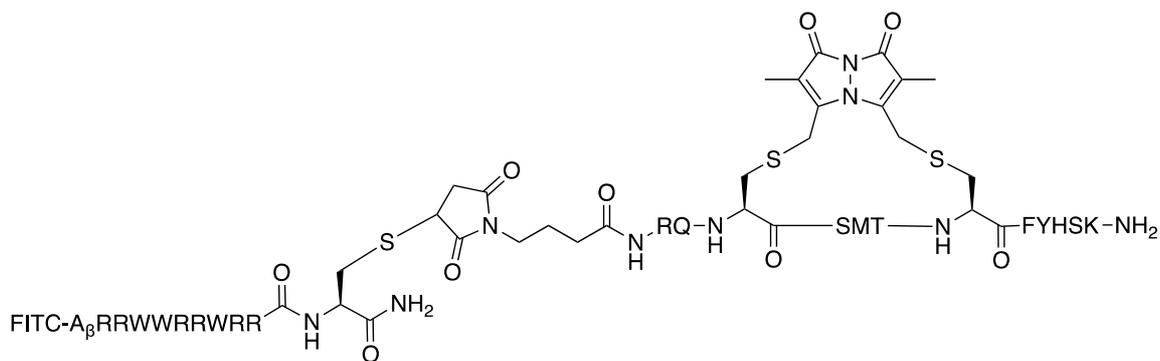
P1c-N4F, *FITC-A$_\beta$RRWWRRWRRC[suc-RQC(-)SMTC(Bim)FYHSK]*. Peptide conjugation of purified **P1c** and **N4F** was carried out by *General Method 4: Thioether Conjugation Reaction* and the peptide was purified per *General Method 3* to afford a yellow powder. Analytical RP-HPLC: 0-50% over 15 min, R_t (C18) = 12.6 min; HRMS (ESI+) Expected [M+6H]$^{6+}$ for [C$_{176}$H$_{237}$N$_{57}$O$_{38}$S$_5$]: 653.7906, observed: [M+6H]$^{6+}$ 653.7907.
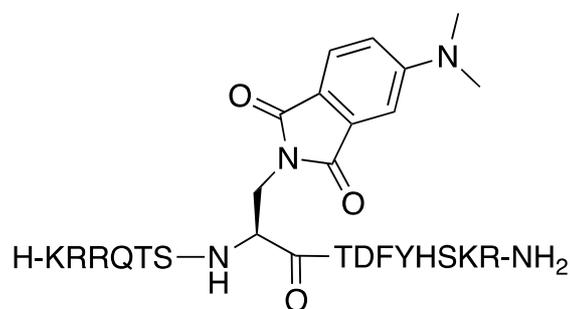
B1, *KRRQTS(4-DAPA)TDFYHSKR*. The peptide was assembled per *General Method 1* using commercial Fmoc-L-amino acid building blocks, where the *N*-terminal was protected with a Boc group (Boc-Lys(Boc)-OH). 4-DMAP was incorporated using the *General Method 6: On-resin Derivatisation Method*. The peptide was cleaved from the resin per *General Method 2*. The peptide was then purified by *General Method 3* to afford a yellow powder. Analytical

RP-HPLC: 0-100% over 15 min, R_t (C18) = 8 min; HRMS (ESI+) Expected [M+4H]$^{4+}$ for [$C_{90}H_{138}N_{32}O_{25}$]: 517.7706, observed: [M+4H]$^{4+}$ 517.7707.
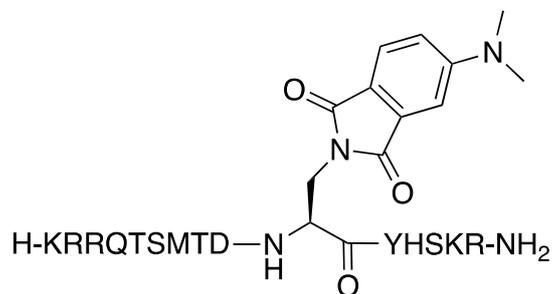
B2, *KRRQTSMTD(4-DAPA)YHSKR.* The peptide was assembled per *General Method 1* using commercial Fmoc-L-amino acid building blocks, where the *N*-terminal was protected with a Boc group (Boc-Lys(Boc)-OH). 4-DMAP was incorporated using the *General Method 6: On-resin Derivatisation Method*. The peptide was cleaved from the resin per *General Method 2*. The peptide was then purified by *General Method 3* to afford a yellow powder. Analytical RP-HPLC: 0-50% over 15 min, R_t (C18) = 9.5 min; HRMS (ESI+) Expected [M+4H]$^{4+}$ for [$C_{86}H_{138}N_{32}O_{25}S$]: 513.7636, observed: [M+4H]$^{4+}$ 513.7640.
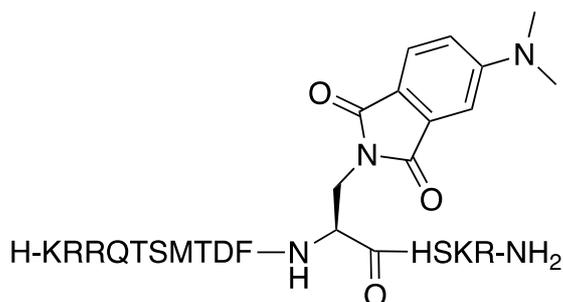
B2, *KRRQTSMTDF(4-DAPA)HSKR.* The peptide was assembled per *General Method 1* using commercial Fmoc-L-amino acid building blocks, where the *N*-terminal was protected with a Boc group (Boc-Lys(Boc)-OH). 4-DMAP was incorporated using the *General Method 6: On-resin Derivatisation Method*. The peptide was cleaved from the resin per *General Method 2*. The peptide was then purified by *General Method 3* to afford a yellow powder.

Analytical RP-HPLC: 0-50% over 15 min, R_t (C18) = 9.5 min; HRMS (ESI+) Expected $[M+4H]^{4+}$ for $[C_{86}H_{138}N_{32}O_{24}S]$: 509.7649, observed: $[M+4H]^{4+}$ 509.7654.
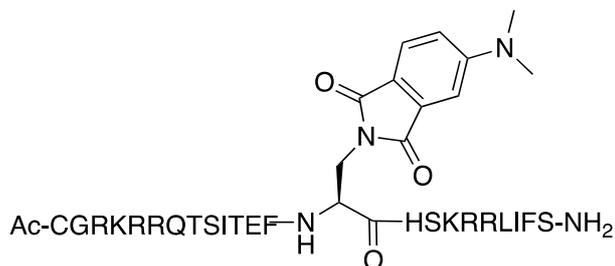
B4, *Ac-CGRKRRQTSITEF(4-DAPA)HSKRRLIFS*. The peptide was assembled per *General Method 1* using commercial Fmoc-L-amino acid building blocks. Following the final *N*-terminal Fmoc deprotection, the peptide was capped by acetylation per *General Method 1b: Acetylation*. 4-DMAP was incorporated using the *General Method 6: On-resin Derivatisation Method*. The peptide was cleaved from the resin per *General Method 2*. The peptide was then purified by *General Method 3* to afford a yellow powder. Analytical RP-HPLC: 0-50% over 15 min, R_t (C18) = 11.9 min; HRMS (ESI+) Expected $[M+4H]^{4+}$ for $[C_{131}H_{212}N_{46}O_{34}S]$: 752.4077, observed: $[M+4H]^{4+}$ 752.4068
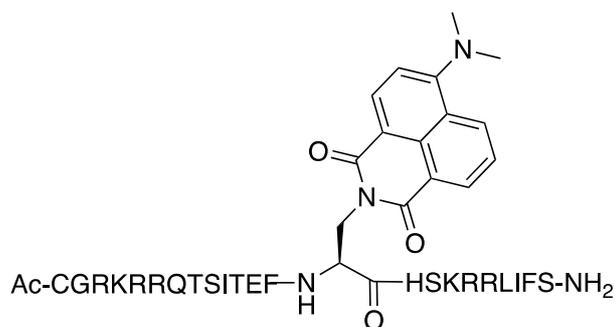
B5, *Ac-CGRKRRQTSITEF(4-DMNA)HSKRRLIFS*.. The peptide was assembled per *General Method 1* and 4-DMNA was incorporated by reacting the resin-bound peptide with a solution of Fmoc-DMNA (**8**) (synthesised in-house) (3 equiv) and DIPEA (6 equiv) in DMF (5 mL) for 3h. Following the final *N*-terminal Fmoc deprotection, the peptide was capped by acetylation per *General Method 1b: Acetylation.* The peptide was cleaved from the resin per

General Method 2 and then purified by *General Method 3* to afford a yellow powder.

Analytical RP-HPLC: 0-50% over 15 min, R_t (C18) = 12 min; HRMS (ESI+) Expected

$[M+4H]^{4+}$ for $[C_{135}H_{214}N_{46}O_{34}S]$: 764.9116, observed: $[M+4H]^{4+}$ 764.9104.

Appendix

Appendix 1: MS and RP-HPLC for Chapter 2
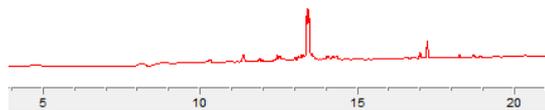
Figure S1: Analytical RP-HPLC of crude peptide prepared by D2 conditions, 0-50% aq. ACN over 15 min (from 5-20 min), visualised at 220 nm.
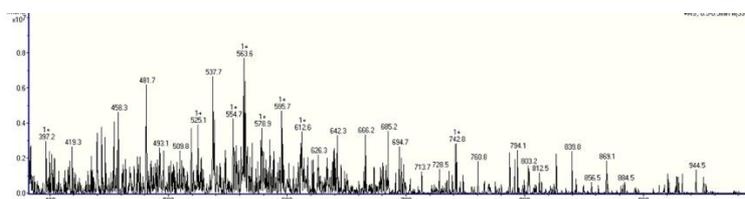
Figure S2: MS of crude peptide prepared by D2 conditions. Target peptide (1812 au): 363 [M+5H]$^{5+}$, 454 [M+4H]$^{4+}$, 605 [M+3H]$^{3+}$ *m/z*.
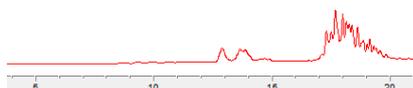
Figure S3: Analytical RP-HPLC of crude peptide prepared by D3 conditions, 0-50% aq. ACN over 15 min (from 5-20 min), visualised at 220 nm.
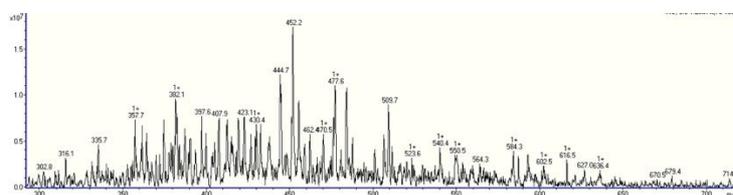
Figure S4: MS of crude peptide prepared by D3 conditions. Target peptide (1812 au): 363 [M+5H]$^{5+}$, 454 [M+4H]$^{4+}$, 605 [M+3H]$^{3+}$ *m/z*.
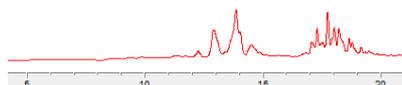
Figure S5: Analytical RP-HPLC of crude peptide prepared by D4 conditions, 0-50% aq. ACN over 15 min (from 5-20 min), visualised at 220 nm.
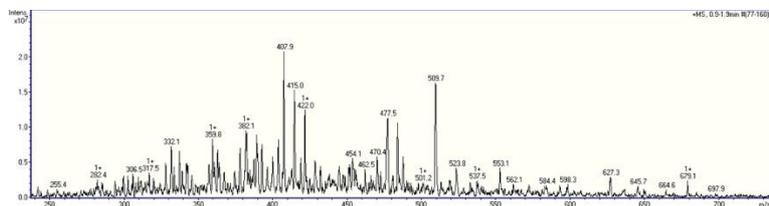
Figure S6: MS of crude peptide prepared by D4 conditions. Target peptide (1812 au): 363 [M+5H]$^{5+}$, 454 [M+4H]$^{4+}$, 605 [M+3H]$^{3+}$ *m/z*.

Figure S7: Analytical RP-HPLC of crude peptide prepared by D5 conditions, 0-50% aq. ACN over 15 min (from 5-20 min), visualised at 220 nm.
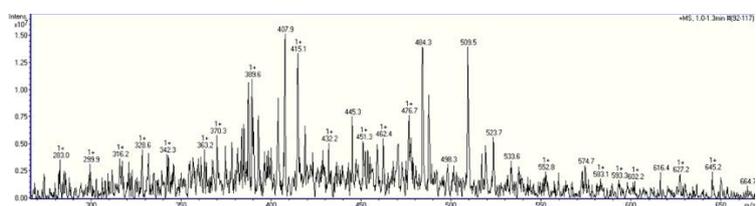
Figure S8: MS of crude peptide prepared by D5 conditions. Target peptide (1812 au): 363 [M+5H]$^{5+}$, 454 [M+4H]$^{4+}$, 605 [M+3H]$^{3+}$ *m/z*.
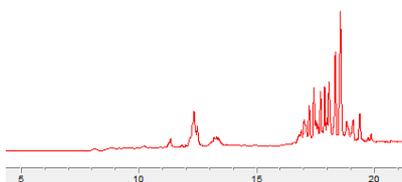
Figure S9: Analytical RP-HPLC of crude peptide prepared by D6 conditions, 0-50% aq. ACN over 15 min (from 5-20 min), visualised at 220 nm.
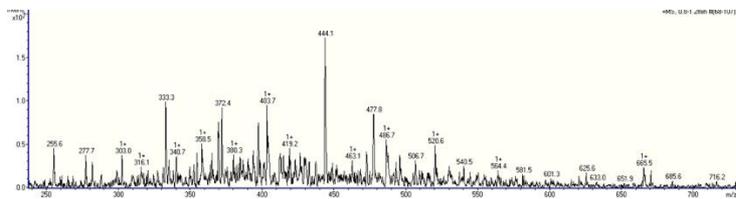
Figure S10: MS of crude peptide prepared by D6 conditions. Target peptide (1812 au): 363 [M+5H]$^{5+}$, 454 [M+4H]$^{4+}$, 605 [M+3H]$^{3+}$ *m/z*.

Appendix 2: MS and RP-HPLC for Chapter 3
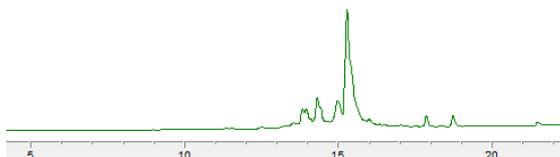
Figure S11: Analytical RP-HPLC of crude **P1a-N1F**, 0-50% aq. ACN over 15 min (from 5-20 min), visualised at 254 nm.

**Figure S12**: Analytical RP-HPLC of crude **P1a-N3F**, 0-50% aq. ACN over 15 min (from 5-20 min), visualised at 254 nm.

Figure S13: Analytical RP-HPLC of crude **P1a-N2F**, 0-50% aq. ACN over 15 min (from 5-20 min), visualised at 254 nm.
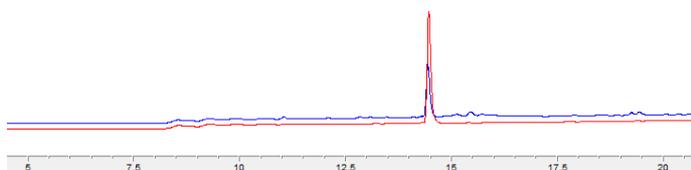
Figure S14: Analytical RP-HPLC overlay of crude **P3a-N1F** (blue) and **P3a** (red), 0-50% aq. ACN over 15 min (from 5-20 min), visualised at 220 nm.
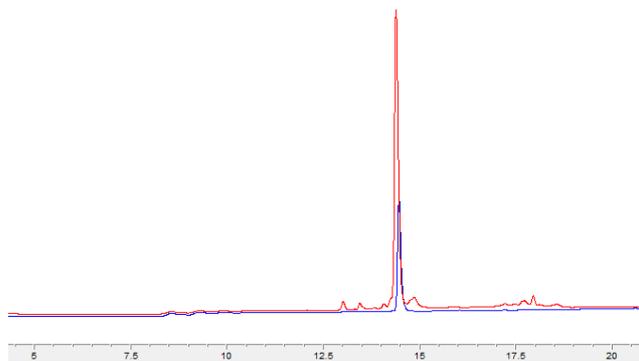
Figure S15: Analytical RP-HPLC overlay of crude **P3a-N4F** (blue) and **P3a** (red), 0-50% aq. ACN over 15 min (from 5-20 min), visualised at 220 nm.

Appendix 3: Characterisation

NMR spectra
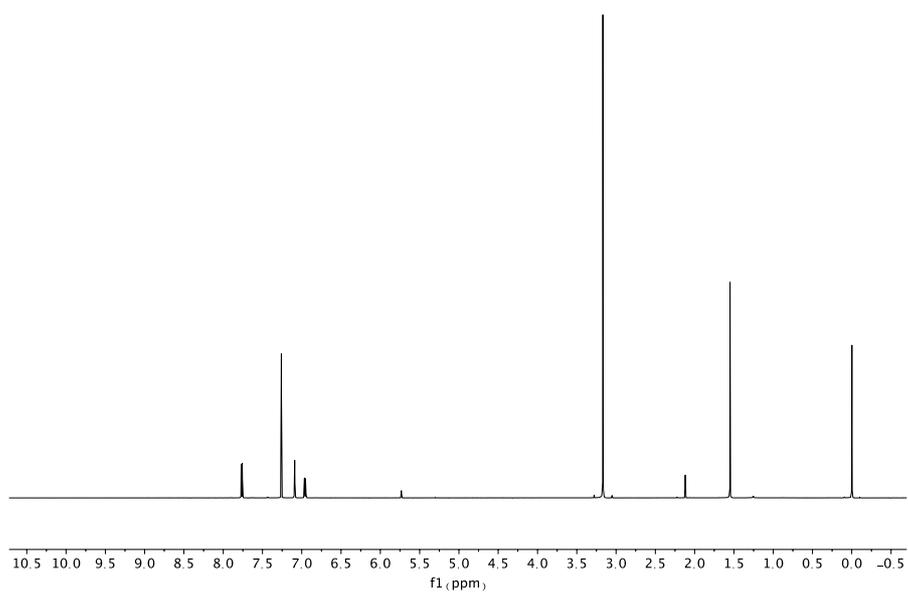
Figure S16: 600 MHz spectrum of **5** in CDCl₃
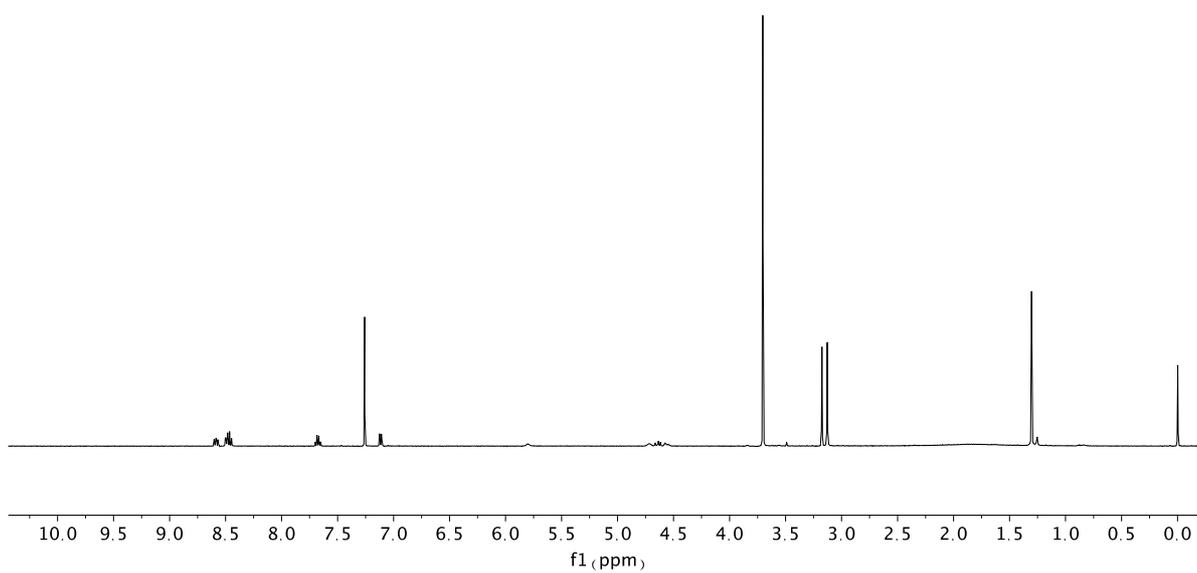
Figure S17: 500 MHz of **7** in CDCl₃
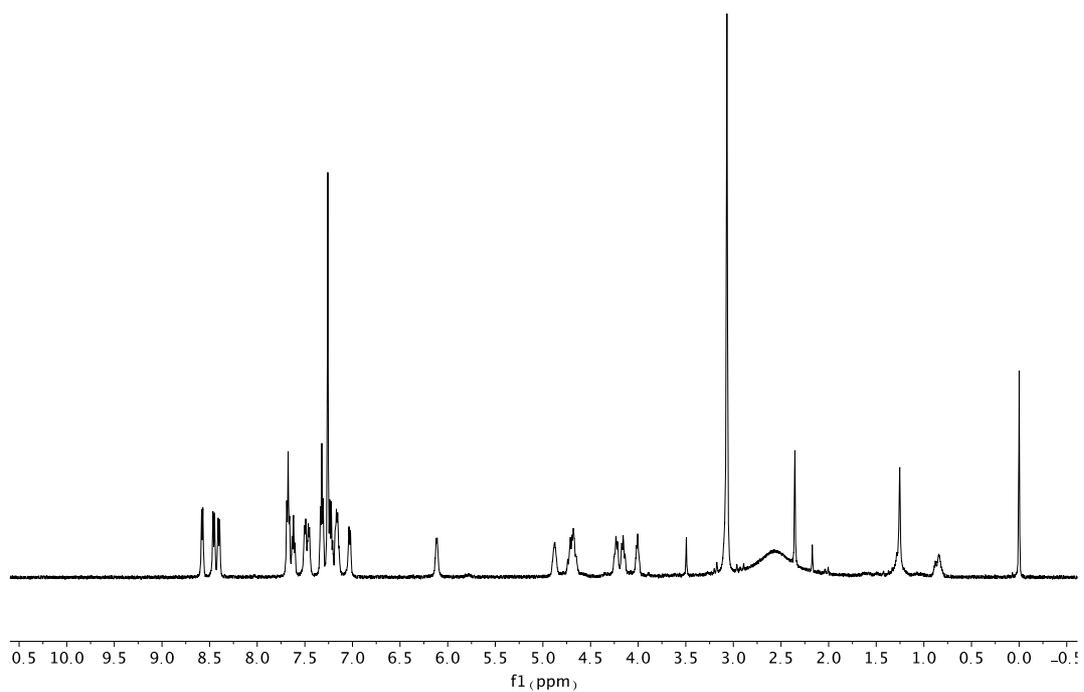
Figure S18: 500 MHz spectrum of **8** in CDCl$_3$.

HPLC spectra
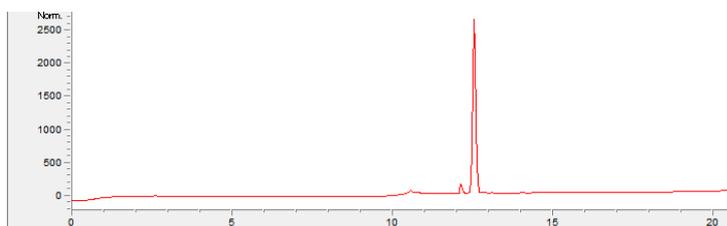
Figure S19: Analytical RP-HPLC C18 spectrum for **P2**, 0-100% aq. ACN over 15 min (from 5-20 min) visualised at 220 nm.
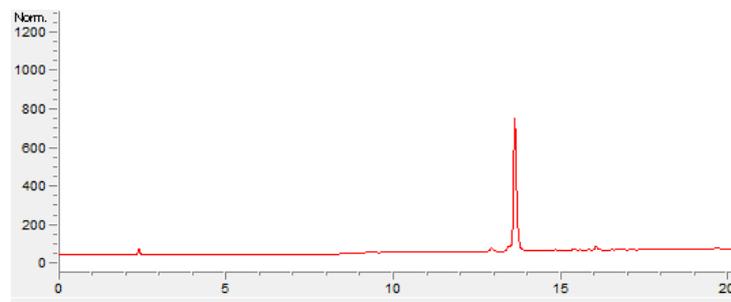
**Figure S20:** Analytical RP-HPLC C18 spectrum for **P3**, 0-50% aq. ACN over 15 min (from 5-20 min) visualised at 220 nm.
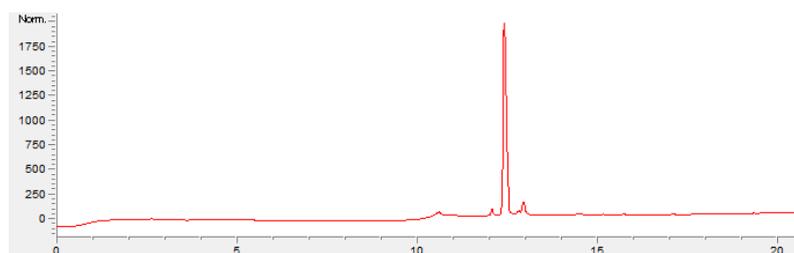

Figure S21: Analytical RP-HPLC C18 spectrum for **P4**, 0-100% aq. ACN over 15 min (from 5-20 min) visualised at 220 nm.
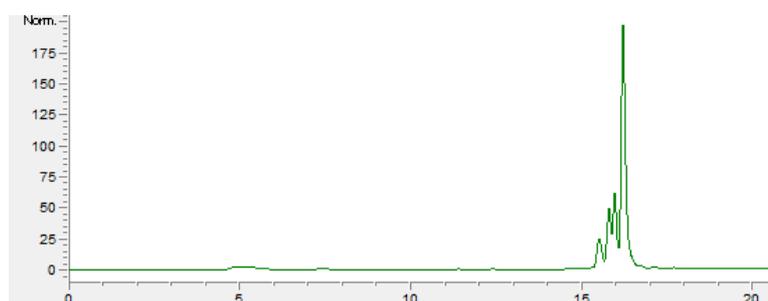
Figure S22: Analytical RP-HPLC C18 spectrum for **P2F** 0-50% aq. ACN over 15 min (from 5-20 min) visualised at 254 nm.
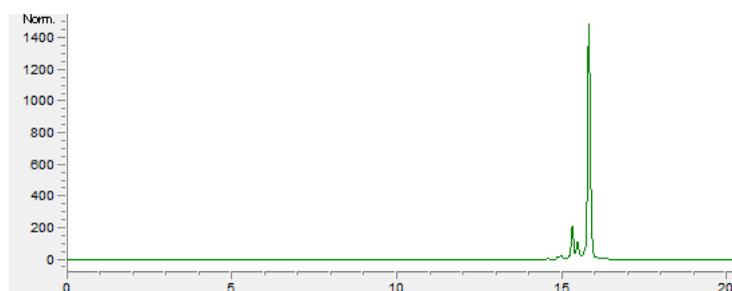

**Figure S23:** Analytical RP-HPLC C18 spectrum for **P3F**, 0-50% aq. ACN over 15 min (from 5-20 min) visualised at 254 nm.
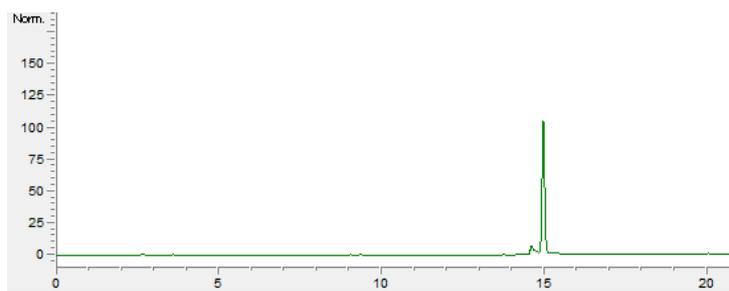
Figure S24: Analytical RP-HPLC C18 spectrum for **P4F**, 0-50% aq. ACN over 15 min (from 5-20 min) visualised at 254 nm.
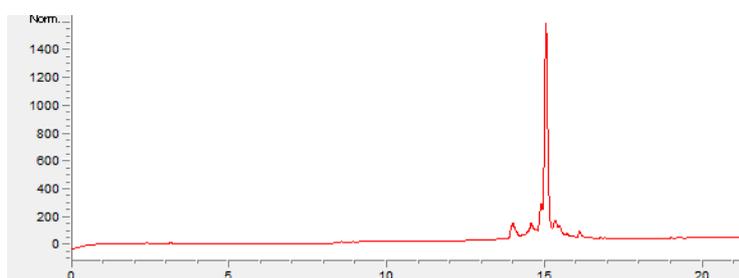
Figure S25: Analytical RP-HPLC C18 spectrum for **P1b** 0-50% aq. ACN over 15 min (from 5-20 min) visualised at 220 nm.
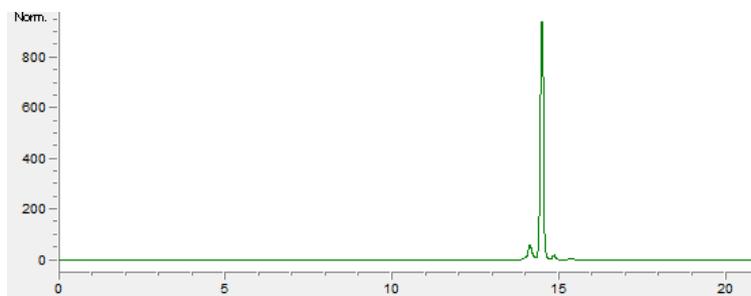
Figure S26: Analytical RP-HPLC C18 spectrum for **N1F** 0-50% aq. ACN over 15 min (from 5-20 min) visualised at 254 nm.
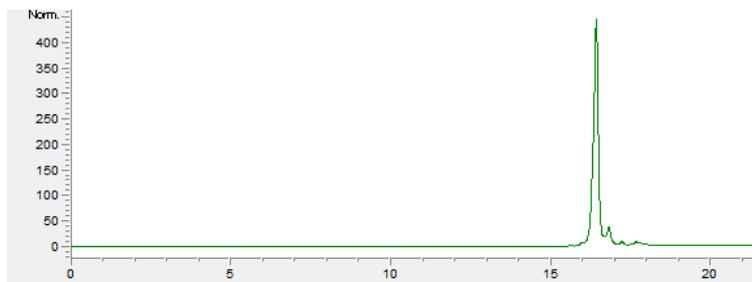
Figure S27: Analytical RP-HPLC C18 spectrum for **N2F** 0-50% aq. ACN over 15 min (from 5-20 min) visualised at 254 nm.
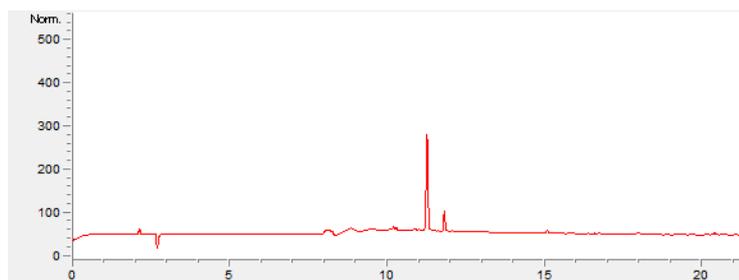

Figure S28: Analytical RP-HPLC C18 spectrum for **N2** 0-50% aq. ACN over 15 min (from 5-20 min) visualised at 220 nm.
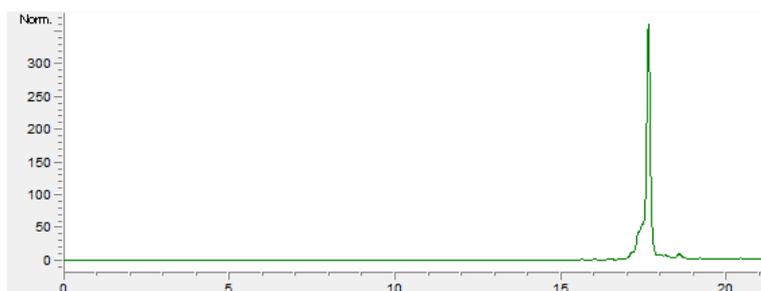
Figure S29: Analytical RP-HPLC C18 spectrum for **N3F** 0-50% aq. ACN over 15 min (from 5-20 min) visualised at 254 nm.
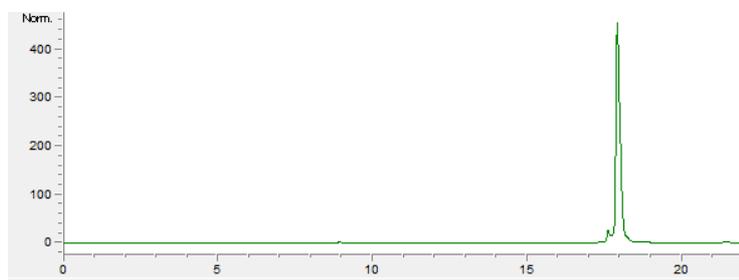
Figure S30: Analytical RP-HPLC C18 spectrum for **N4F** 0-50% aq. ACN over 15 min (from 5-20 min) visualised at 254 nm.
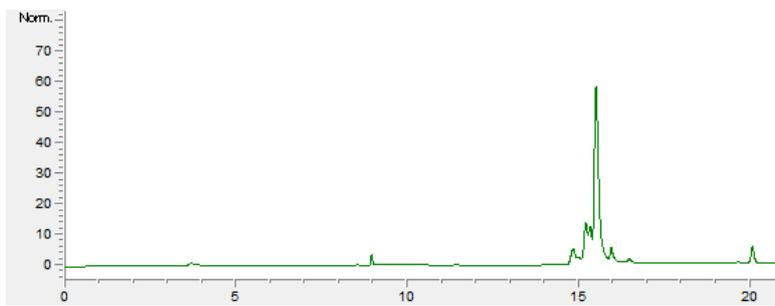
Figure S31: Analytical RP-HPLC C18 spectrum for **P1b-N1F** 0-50% aq. ACN over 15 min (from 5-20 min) visualised at 254 nm.
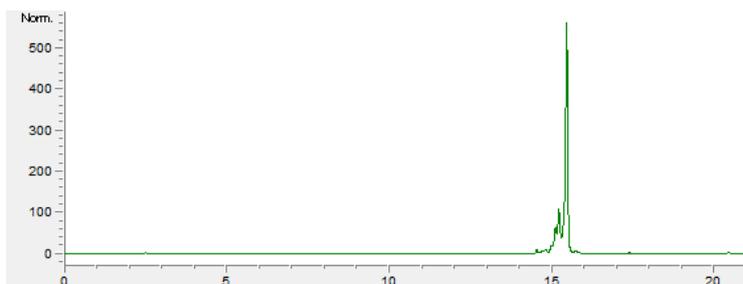

Figure S32: Analytical RP-HPLC C18 spectrum for **P1b-N2F** 0-50% aq. ACN over 15 min (from 5-20 min) visualised at 254 nm.
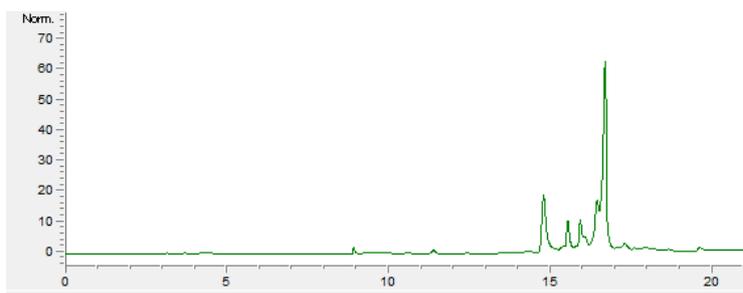

Figure S33: Analytical RP-HPLC C18 spectrum for **P1b-N3F** 0-50% aq. ACN over 15 min (from 5-20 min) visualised at 254 nm.
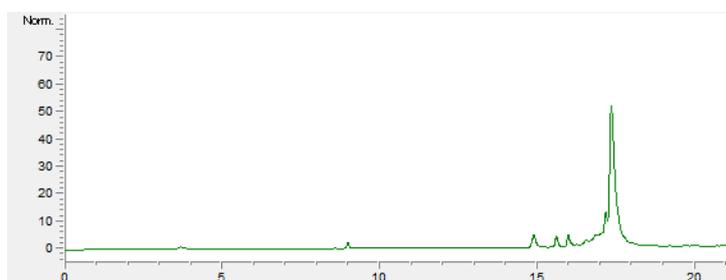
Figure S34: Analytical RP-HPLC C18 spectrum for **P1b-N4F** 0-50% aq. ACN over 15 min (from 5-20 min) visualised at 254 nm.
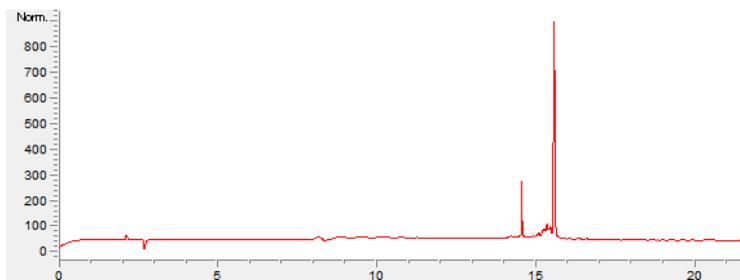
Figure S35: Analytical RP-HPLC C18 spectrum for **P1c** 0-50% aq. ACN over 15 min (from 5-20 min) visualised at 220 nm.
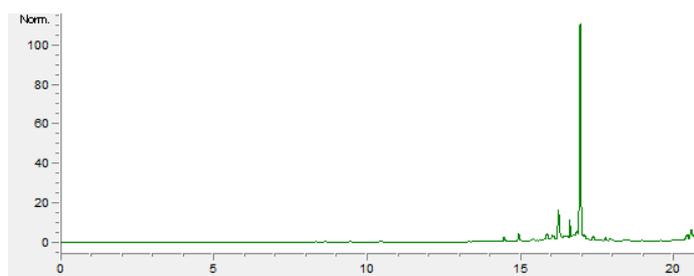
Figure S36: Analytical RP-HPLC C18 spectrum for **P1bimF** 0-50% aq. ACN over 15 min (from 5-20 min) visualised at 254 nm.
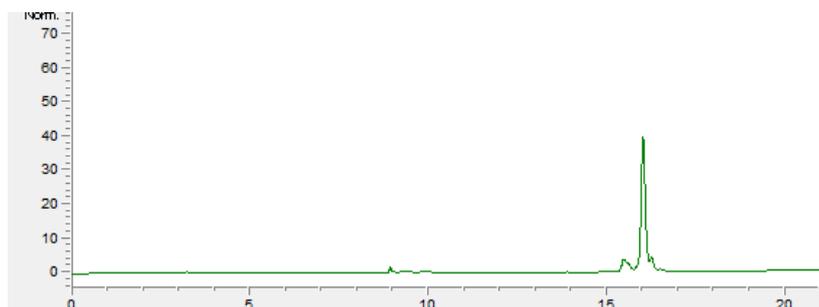
Figure S37: Analytical RP-HPLC C18 spectrum for **P1c-N1F** 0-50% aq. ACN over 15 min (from 5-20 min) visualised at 254 nm.
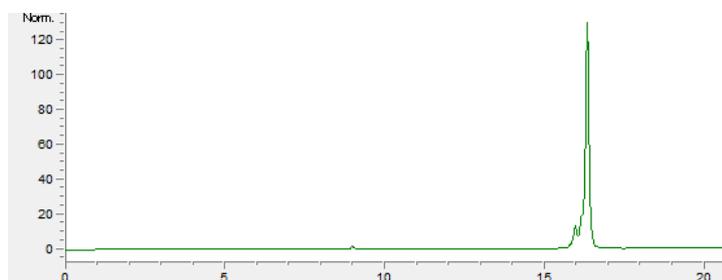

Figure S38: Analytical RP-HPLC C18 spectrum for **P1c-N2F** 0-50% aq. ACN over 15 min (from 5-20 min) visualised at 254 nm.
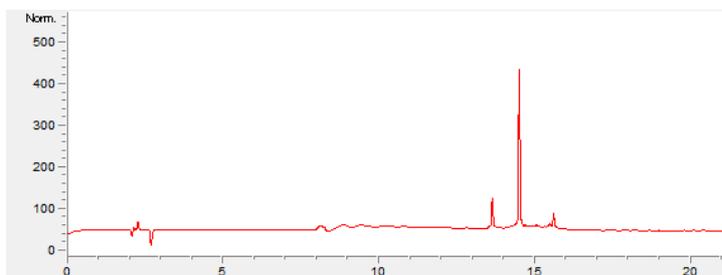
Figure S39: Analytical RP-HPLC C18 spectrum for **P1c-N2** 0-50% aq. ACN over 15 min (from 5-20 min) visualised at 220 nm.
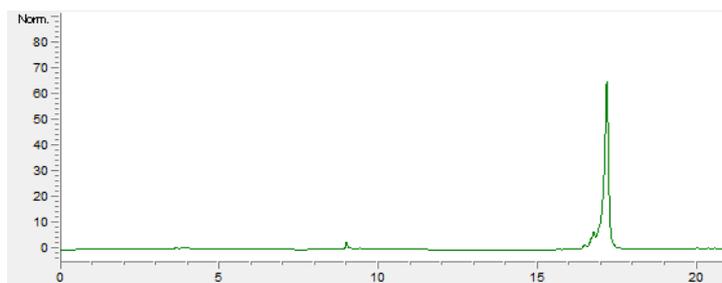
Figure S40: Analytical RP-HPLC C18 spectrum for **P1c-N3F** 0-50% aq. ACN over 15 min (from 5-20 min) visualised at 254 nm.
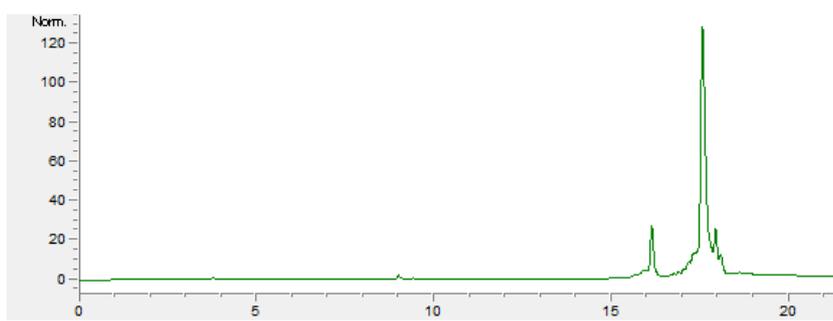

Figure S41: Analytical RP-HPLC C18 spectrum for **P1c-N4F** 0-50% aq. ACN over 15 min (from 5-20 min) visualised at 254 nm.
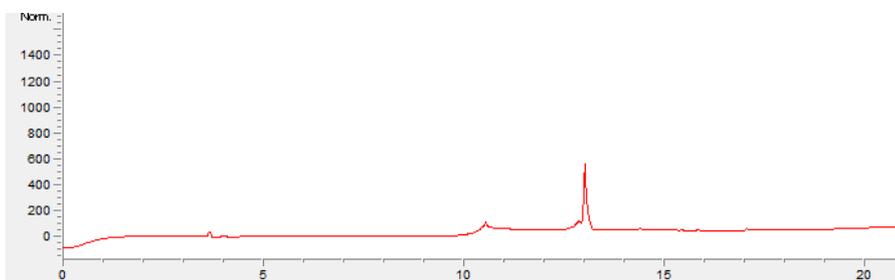

Figure S42: Analytical RP-HPLC C18 spectrum for **B1** 0-100% aq. ACN over 15 min (from 5-20 min) visualised at 220 nm.
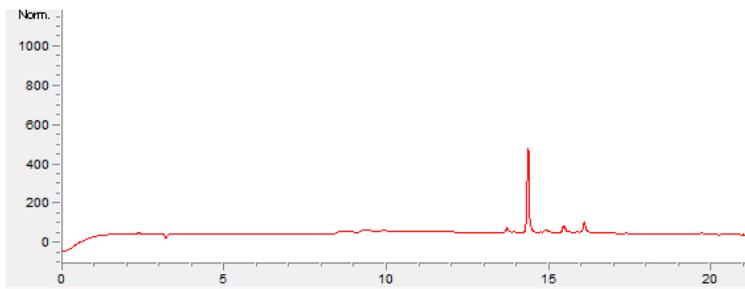
Figure S43: Analytical RP-HPLC C18 spectrum for **B2** 0-50% aq. ACN over 15 min (from 5-20 min) visualised at 220 nm.
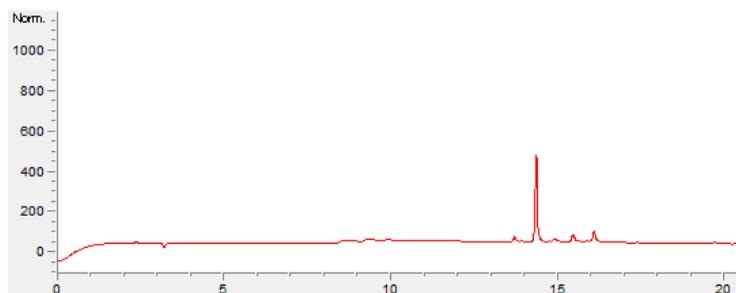
Figure S44: Analytical RP-HPLC C18 spectrum for **B3** 0-50% aq. ACN over 15 min (from 5-20 min) visualised at 220 nm.

**Figure S45**: Analytical RP-HPLC C18 spectrum for **B4** 0-50% aq. ACN over 15 min (from 5-20 min) visualised at 220 nm. Peaks spanning over 15 and 16.5 min are artefacts of the C18 column.
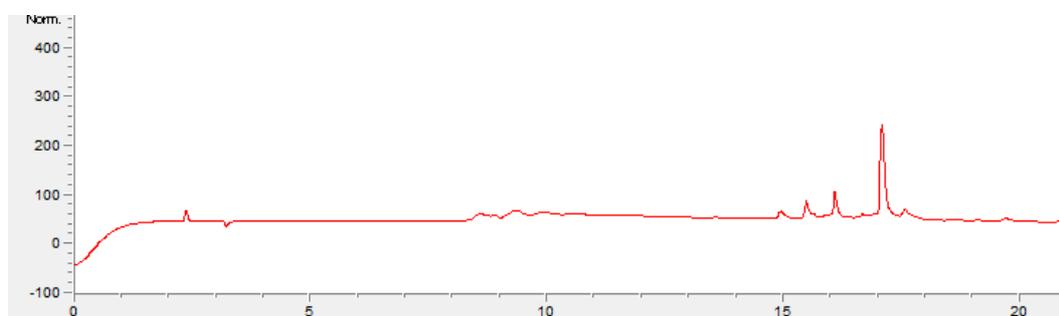
**Figure S46:** Analytical RP-HPLC C18 spectrum for **B5** 0-50% aq. ACN over 15 min (from 5-20 min) visualised at 220 nm. Peaks spanning over 15 and 16.5 min are artefacts of the C18 column.

Appendix 4: Publications

Publications from this work

A. J. Horsfall, **T. Chav,** Z. Kikhtyak, J. L. Pederick, W. Kowalczyk, D. B. Scanlon, W. D. Tilley, T. E. Hickey, A. D. Abell and J. B. Bruning, Nature Chemical Biology 2021

Manuscript in preparation

ABSTRACT

Human Proliferating Cell Nuclear Antigen (PCNA) mediates DNA replications and repair, and thus inhibiting PCNA interactions can shut down DNA-replication which may find application in developing novel cancer treatments. Here, a short p21-derived peptide system is optimised to provide a modular nuclear permeable peptidomimetic to target PCNA in breast cancer cells. We demonstrate that a p21-fluorescent macrocycle provides the smallest optimal scaffold to permit cellular entry, and determine an optimal NLS sequence derived from SV40 to confer nuclear uptake. Lastly, we show that attaching a fluorescein tag dramatically alters the cellular distribution of the p21 peptidomimetics. This study has identified an inherently fluorescent p21-based peptidomimetic scaffold which provides a significant advance toward a peptide-based therapeutic to inhibit PCNA.

References

1.   Mabonga, L.; Kappo, A. P., Peptidomimetics: A Synthetic Tool for Inhibiting Protein–Protein Interactions in Cancer. *Int J Pept Res Ther* **2019,** *26* (1), 225-241.

2.   Watkins, A. M.; Arora, P. S., Structure-based inhibition of protein-protein interactions. *Eur J Med Chem* **2015,** *94*, 480-8.

3.   Tsomaia, N., Peptide therapeutics: targeting the undruggable space. *Eur J Med Chem* **2015,** *94*, 459-70.

4.   Lu, H.; Zhou, Q.; He, J.; Jiang, Z.; Peng, C.; Tong, R.; Shi, J., Recent advances in the development of protein-protein interactions modulators: mechanisms and clinical trials. *Signal Transduct Target Ther* **2020,** *5* (1), 213.

5.   Mabonga, L.; Kappo, A. P., Protein-protein interaction modulators: advances, successes and remaining challenges. *Biophys Rev* **2019,** *11* (4), 559-581.

6.   Ran, X.; Gestwicki, J. E., Inhibitors of protein-protein interactions (PPIs): an analysis of scaffold choices and buried surface area. *Curr Opin Chem Biol* **2018,** *44*, 75-86.

7.   Smith, M. C.; Gestwicki, J. E., Features of protein-protein interactions that translate into potent inhibitors: topology, surface area and affinity. *Expert Rev Mol Med* **2012,** *14*, e16.

8.   Lee, A. C.; Harris, J. L.; Khanna, K. K.; Hong, J. H., A Comprehensive Review on Current Advances in Peptide Drug Development and Design. *Int J Mol Sci* **2019,** *20* (10).

9.   Behrendt, R.; White, P.; Offer, J., Advances in Fmoc solid-phase peptide synthesis. *J Pept Sci* **2016,** *22* (1), 4-27.

10.   Sampson, W. R.; Patsiouras, H.; Ede, N. J., The Synthesis of Difficult Peptides Using 2-Hydroxy-4-Methoxybenzyl or Pseudoproline Amino Acid Building Blocks: a Comparative Study. *J Pept Sci* **1999,** *5*, 403-409.

11.   Cardona, V.; Eberle, I.; Barthélémy, S.; Beythien, J.; Doerner, B.; Schneeberger, P.; Keyte, J.; White, P. D., Application of Dmb-Dipeptides in the Fmoc SPPS of Difficult and Aspartimide-Prone Sequences. *Int J Pept Res Ther* **2008,** *14* (4), 285-292.

12.   White, P.; Keyte, J. W.; Bailey, K.; Bloomberg, G., Expediting the Fmoc solid phase synthesis of long peptides through the application of dimethyloxazolidine dipeptides. *J Pept Sci* **2004,** *10* (1), 18-26.

13.   Arora, P. S., 187 Plucking the high hanging fruit: a systematic approach for targeting protein interfaces. *J Biomol Struct Dyn* **2013,** *31* (sup1), 120-121.

14.   Muttenthaler, M.; King, G. F.; Adams, D. J.; Alewood, P. F., Trends in peptide drug discovery. *Nat Rev Drug Discov* **2021,** *20* (4), 309-325.

15.   Lau, J. L.; Dunn, M. K., Therapeutic peptides: Historical perspectives, current development trends, and future directions. *Bioorg Med Chem* **2018,** *26* (10), 2700-2707.

16.   Behrendt, R.; Huber, S.; White, P., Preventing aspartimide formation in Fmoc SPPS of Asp-Gly containing peptides--practical aspects of new trialkylcarbinol based protecting groups. *J Pept Sci* **2016,** *22* (2), 92-7.

17.   Neumann, K.; Farnung, J.; Baldauf, S.; Bode, J. W., Prevention of aspartimide formation during peptide synthesis using cyanosulfurylides as carboxylic acid-protecting groups. *Nat Commun* **2020,** *11* (1), 982.

18.   Quibell, M.; Owen, D.; Packman, L. C.; Johnson, T., Suppression of Piperidine-mediated Side Product Formation for Asp(OBut)-containing Peptides by the Use of *N*-(2-hydroxy-4-methoxybenzyl) (Hmb) Backbone Amide Protection. *J. Chem. Soc. Chem. Commun.* **1994**, 2343-2344.

19.   Michels, T.; Dölling, R.; Haberkorn, U.; Mier, W., Acid-mediated prevention of aspartimide formation in solid phase peptide synthesis. *Org Lett* **2012,** *14* (20), 5218.

20.   Haack, T.; Mutter, M., Serine derived oxazolidines as secondary structure disrupting, solubilizing building blocks in peptide synthesis. *Tetrahedron Lett* **1992,** *33* (12), 1589-1992.

21.   Johnson, T.; Quibell, M.; Owen, D.; Sheppard, R. C., A Reversible Protecting Group for the Amide Bond in Peptides. Use in the Synthesis of 'Difficult Sequences' *J. Chem. Soc. Chem. Commun.* **1993**, 369-372.

22.   Miyachi, K.; Fritzler, M. J.; Tan, E. M., Autoantibody to a nuclear antigen in proliferating cells. *J Immun Balt* **1978,** *121* (6), 2228.

23.   Baker, R. D.; Howl, J.; Nicholl, I. D., A sychnological cell penetrating peptide mimic of p21 WAF1/CIP1 is pro-apoptogenic. *Peptides* **2007,** *28* (4), 731-740.

24.   Kroker, A. J.; Bruning, J. B., p21 Exploits Residue Tyr151 as a Tether for High-Affinity PCNA Binding. *Biochemistry* **2015,** *54* (22), 3483.

25.   Horsfall, A. J.; Abell, A. D.; Bruning, J., Targeting PCNA with peptide mimetics for therapeutic purposes. *ChemBioChem* **2020,** *21*, 442-450.

26.   Dieckman, L. M.; Freudenthal, B. D.; Washington, M. T., PCNA structure and function: insights from structures of PCNA complexes and post-translationally modified PCNA. *Subcell Biochem* **2012,** *62*, 281-99.

27.   Gulbis, J. M.; Kelman, Z.; Hurwitz, J.; Donnell, M.; Kuriyan, J., Structure of the C-Terminal Region of p21WAF1/CIP1Complexed with Human PCNA. *Cell* **1996,** *87* (2), 297-306.

28.   Roberts, G. C. K., *Encyclopedia of Biophysics*. 2013.

29.   Yan, L.; Jerard, H.; Joan, M., Cell-cycle inhibition by independent CDK and PCNA binding domains in p21Cip1. *Nature* **1995,** *375* (6527), 159.

30.   Wegener, K. L.; McGrath, A. E.; Dixon, N. E.; Oakley, A. J.; Scanlon, D. B.; Abell, A. D.; Bruning, J. B., Rational Design of a 310 -Helical PIP-Box Mimetic Targeting PCNA, the Human Sliding Clamp. *Chemistry* **2018,** *24* (44), 11325-11331.

31.   Derakhshankhah, H.; Jafari, S., Cell penetrating peptides: A concise review with emphasis on biomedical applications. *Biomed Pharmacother* **2018,** *108*, 1090-1096.

32.   Yang, N. J.; Hinner, M. J., Getting across the cell membrane: an overview for small molecules, peptides, and proteins. *Methods Mol Biol* **2015,** *1266*, 29-53.

33.   Torchilin, V. P., Recent Approaches to Intracellular Delivery of Drugs and DNA and Organelle Targeting. *Annu. Rev. Biomed. Eng.* **2006,** *8*, 343-375.

34.   Shinoda, W., Permeability across lipid membranes. *Biochim Biophys Acta* **2016,** *1858* (10), 2254-2265.

35.   Zhang, R.; Qin, X.; Kong, F.; Chen, P.; Pan, G., Improving cellular uptake of therapeutic entities through interaction with components of cell membrane. *Drug Deliv* **2019,** *26* (1), 328-342.

36.   Campbell, J. E.; Cohall, D., Pharmacodynamics—A Pharmacognosy Perspective. In *Pharmacognosy*, 2017; pp 513-525.

37.   Tashima, T., Intelligent substance delivery into cells using cell-penetrating peptides. *Bioorganic Med Chem Lett* **2017,** *27* (2), 121-130.

38.   Guo, Z.; Peng, H.; Kang, J.; Sun, D., Cell-penetrating peptides: Possible transduction mechanisms and therapeutic applications. *Biomed Rep* **2016,** *4* (5), 528-534.

39.   Habault, J.; Poyet, J. L., Recent Advances in Cell Penetrating Peptide-Based Anticancer Therapies. *Molecules* **2019,** *24* (5).

40.   Delaroche, D.; Aussedat, B.; Aubry, S.; Chassaing, G.; Burlina, F.; Clodic, G.; Bolbach, G.; Lavielle, S.; Sagan, S., Tracking a new cell-penetrating (W/R) nonapeptide, through an enzyme-stable mass spectrometry reporter tag. *Anal Chem* **2007,** *79* (5), 1932-8.

41.   Gagat, M.; Zielinska, W.; Grzanka, A., Cell-penetrating peptides and their utility in genome function modifications (Review). *Int J Mol Med* **2017,** *40* (6), 1615-1623.

42.   Deepthi, A.; Raju, S.; Kalyani, A.; Udaya Kiran, M.; Vanaja, A., Targeted Drug Delivery to the Nucleus and its Potential Role in Cancer Chemotherapy *J Pharm Sci* **2013,** *5* (2), 48-56.

43.   Lange, A.; Mills, R. E.; Lange, C. J.; Stewart, M.; Devine, S. E.; Corbett, A. H., Classical nuclear localization signals: definition, function, and interaction with importin alpha. *J Biol Chem* **2007,** *282* (8), 5101-5.

44.   Hodel, M. R.; Corbett, A. H.; Hodel, A. E., Dissection of a nuclear localization signal. *J Biol Chem* **2001,** *276* (2), 1317-25.

45.   Ragin, A. D.; Morgan, R. A.; Chmielewski, J., Cellular Import Mediated by Nuclear Localization Signal Peptide Sequences. *Chem Biol* **2002,** *8*, 943-948.

46.   Kim, Y. H.; Han, M. E.; Oh, S. O., The molecular mechanism for nuclear transport and its application. *Anat Cell Biol* **2017,** *50* (2), 77-85.

47.   Dang, C. V.; Lee, W. M. F., Identification of the Human c-myc Protein Nuclear Translocation Signal. *Mol Cell Biol* **1988,** *8* (10), 4048-4054.

48.   Opanasopit, P.; Rojanarata, T.; Apirakaramwong, A.; Ngawhirunpat, T.; Ruktanonchai, U., Nuclear localization signal peptides enhance transfection efficiency of chitosan/DNA complexes. *Int J Pharm* **2009,** *382* (1-2), 291-5.

49.   Derossi, D.; Chassaing, G.; Prochiantz, A., Trojan peptides- the penetratin system for intracellular delivery. *Trends Cell Biol* **1998,** *8*, 84-87.

50.   Durzynska, J.; Przysiecka, L.; Nawrot, R.; Barylski, J.; Nowicki, G.; Warowicka, A.; Musidlak, O.; Gozdzicka-Jozefiak, A., Viral and other cell-penetrating peptides as vectors of therapeutic agents in medicine. *J Pharmacol Exp Ther* **2015,** *354* (1), 32-42.

51.   Borrelli, A.; Tornesello, A. L.; Tornesello, M. L.; Buonaguro, F. M., Cell Penetrating Peptides as Molecular Carriers for Anti-Cancer Agents. *Molecules* **2018,** *23* (2).

52.   Fu, X.; Liang, C.; Li, F.; Wang, L.; Wu, X.; Lu, A.; Xiao, G.; Zhang, G., The Rules and Functions of Nucleocytoplasmic Shuttling Proteins. *Int J Mol Sci* **2018,** *19* (5).

53.   Xiao, Z.; Prieto, D.; Conrads, T. P.; Veenstra, T. D.; Issaq, H. J., Proteomic patterns: their potential for disease diagnosis. *Mol Cell Endocrinol* **2005,** *230* (1-2), 95-106.

54.   Steffen, P.; Kwiatkowski, M.; Robertson, W. D.; Zarrine-Afsar, A.; Deterra, D.; Richter, V.; Schluter, H., Protein species as diagnostic markers. *J Proteomics* **2016,** *134*, 5-18.

55.   Matthieu, S.; Barbara, I., Synthesis of anhydride precursors of the environment-sensitive fluorophores 4-DMAP and 6-DMN. *Nat Protoc* **2007,** *2* (12), 3219.

56.   Loving, G.; Imperiali, B., A Versatile Amino Acid Analogue of the Solvatochromic Fluorophore 4-N,N-Dimethylamino-1,8-naphthalimide- A Powerful Tool for the Study of Dynamic Protein Interactions. *J. Am. Chem. Soc* **2008,** *130*, 13630-13638.

57.   Pazos, E.; Vazquez, O.; Mascarenas, J. L.; Vazquez, M. E., Peptide-based fluorescent biosensors. *Chem Soc Rev* **2009,** *38* (12), 3348-59.

58.   Wegener, K. L.; McGrath, A. E.; Dixon, N. E.; Oakley, A. J.; Scanlon, D. B.; Abell, A. D.; Bruning, J. B., Rational Design of a 310-Helical PIP-Box Mimetic Targeting PCNA, the Human Sliding Clamp. *Chem Eur J* **2018,** *24* (44), 11238-11238.

59.   Kameyama, K.; Takami, H.; Unemura, S.; Osamura, Y. R.; Wada, N.; Sugino, K.; Mimura, T.; Ito, K., PCNA and Ki-67 as Prognostic Markers in Human Parathyroid Carcinomas. *Ann Surg Oncol* **2000,** *7* (4), 301-305.

60.   Thom, M.; Palmer, A.; Cattell, V.; Cook, T., Proliferating cell nuclear antigen (PCNA) as a diagnostic marker of acute cellular rejection in routinely processed biopsies of renal allografts. *Nephrol Dial Transplant* **1994,** *9* (2), 153.

61.   Raibaut, L.; Vasseur, W.; Shimberg, G. D.; Saint-Pierre, C.; Ravanat, J. L.; Michel, S. L. J.; Seneque, O., Design of a synthetic luminescent probe from a biomolecule binding domain: selective detection of AU-rich mRNA sequences. *Chem Sci* **2017,** *8* (2), 1658-1664.

62.   Loving, G. S.;   Sainlos, M.; Imperiali, B., Monitoring protein interactions and dynamics with solvatochromic fluorophores. *Trends Biotechnol* **2010,** *28* (2), 73-83.

63.   Klymchenko, A. S., Solvatochromic and Fluorogenic Dyes as Environment-Sensitive Probes: Design and Biological Applications. *Acc Chem Res* **2017,** *50* (2), 366-375.

64. Fakhari, M. A.; Rokita, S. E., A new solvatochromic fluorophore for exploring nonpolar environments created by biopolymers. *Chem Commun (Camb)* **2011,** *47* (14), 4222-4.

65. Matthieu, S.; Barbara, I., Tools for investigating peptide–protein interactions: peptide incorporation of environment-sensitive fluorophores via on-resin derivatization. *Nat Protoc* **2007,** *2* (12), 3201.

66. Horsfall, A. J., PhD, University of Adelaide.

67. Jullian, M.; Hernandez, A.; Maurras, A.; Puget, K.; Amblard, M.; Martinez, J.; Subra, G., N-terminus FITC labeling of peptides on solid support: the truth behind the spacer. *Tetrahedron Lett* **2009,** *50* (3), 260-263.

68. Pedersen, S. L.; Tofteng, A. P.; Malik, L.; Jensen, K. J., Microwave heating in solid-phase peptide synthesis. *Chem Soc Rev* **2012,** *41* (5), 1826-44.

69. Palasek, S. A.; Cox, Z. J.; Collins, J. M., Limiting racemization and aspartimide formation in microwave-enhanced Fmoc solid phase peptide synthesis. *J Pept Sci* **2006,** *13* (3), 143-148.

70. Samson, D.; Rentsch, D.; Minuth, M.; Meier, T.; Loidl, G., The aspartimide problem persists: Fluorenylmethyloxycarbonyl-solid-phase peptide synthesis (Fmoc-SPPS) chain termination due to formation of N-terminal piperazine-2,5-diones. *J Pept Sci* **2019,** *25* (7), e3193.

71.	Ralhan, K.; KrishnaKumar, V. G.; Gupta, S., Piperazine and DBU: a safer alternative for rapid and efficient Fmoc deprotection in solid phase peptide synthesis. *RSC Advances* **2015,** *5* (126), 104417-104425.

72.	Kikhtyak, Z., RA, University of Adelaide.

73.	Wade, J. D.; Mathieu, M. N.; Macris, M.; Tregear, G. W., Base-induced side reactions in Fmoc-solid phase peptide synthesis: Minimization by use of piperaizine as Nα-deprotection reagent. *Lett Pept Sci* **2000,** *7*.

74.	Frankel, A. D.; Pabo, C. O., Cellular uptake of the tat protein from human immunodeficiency virus. *Cell* **1988,** *55*, 1189-1193.

75.	Green, M.; Loewenstein, P. M., Autonomous functional domains of chemically synthesised human immunodeficiency virus tat trans-activator protein. *Cell* **1988,** *55*, 1179-1188.

76.	Kalderon, D.; Roberts, B. L.; Richardson, W. D.; Smith, A. E., A Short Amino Acid Sequence Able to Specify Nuclear Location. *Cell* **1984,** *39*, 499-508.

77.	Gatos, D.; Patrianakou, S.; Hatzi, O.; Barlos, K., Comparison of the stepwise and convergent approaches in the solid-phase synthesis of rat Tyr-atriopeptin II. *Lett Pept Sci* **1997,** *4*, 177-184.

78.	Koniev, O.; Wagner, A., Developments and recent advancements in the field of endogenous amino acid selective bond forming reactions for bioconjugation. *Chem Soc Rev* **2015,** *44* (15), 5495-551.

79.   Robey, F. A.; Fields, R. L., Automated Synthesis of N-Bromoacetyl-Modified Peptides for the Preparation of Synthetic Peptide Polymers, Peptide-Protein Conjugates, and Cyclic Peptides. *Anal Biochem* **1989,** *177*, 373-377.

80.   Monso, M.; Kowalczyk, W.; Andreu, D.; de la Torre, B. G., Reverse thioether ligation route to multimeric peptide antigens. *Org Biomol Chem* **2012,** *10* (15), 3116-21.

81.   Martinez-Jothar, L.; Doulkeridou, S.; Schiffelers, R. M.; Sastre Torano, J.; Oliveira, S.; van Nostrum, C. F.; Hennink, W. E., Insights into maleimide-thiol conjugation chemistry: Conditions for efficient surface functionalization of nanoparticles for receptor targeting. *J Control Release* **2018,** *282*, 101-109.

82.   Manikwar, P.; Zimmerman, T.; Blanco, F. J.; Williams, T. D.; Siahaan, T. J., Rapid identification of fluorochrome modification sites in proteins by LC ESI-Q-TOF mass spectrometry. *Bioconjug Chem* **2011,** *22* (7), 1330-6.

83.   Schnaible, V.; Przybylski, M., Identification of Fluorescein-5′-Isothiocyanate-Modification Sites in Proteins by Electrospray-Ionization Mass Spectrometry. *Bioconjug Chem.* **1999,** *10*, 861-866.

84.   Horsfall, A. J.; Dunning, K. R.; Keeling, K. L.; Scanlon, D. B.; Wegener, K. L.; Abell, A. D., A Bimane-Based Peptide Staple for Combined Helical Induction and Fluorescent Imaging. *Chembiochem* **2020,** *21* (23), 3423-3432.

85.   Ravasco, J.; Faustino, H.; Trindade, A.; Gois, P. M. P., Bioconjugation with Maleimides: A Useful Tool for Chemical Biology. *Chemistry* **2019,** *25* (1), 43-59.

86.   Fontaine, S. D.; Reid, R.; Robinson, L.; Ashley, G. W.; Santi, D. V., Long-term stabilization of maleimide-thiol conjugates. *Bioconjug Chem* **2015,** *26* (1), 145-52.

87.   Cayrol, C.; MKnibiehler, M.; Ducommun, B., p21 binding to PCNA causes G1 and G2 cell cycle arrest in p53-deficient cells. *Oncogene* **1998,** *16*, 311-320.

88.   Shibata, K.; Suzawa, T.; Soga, S.; Mizukami, T.; Yamada, K.; Hanai, N.; Yamasaki, M., Improvement of biological activity and proteolytic stability of peptides by coupling with a cyclic peptide. *Bioorganic Med Chem Lett* **2003,** *13* (15), 2583-2586.

89.   Horsfall, A. J.; Chav, T.; Bruning, J. B.; Abell, A. D., A turn-on fluorescent PCNA sensor. *Bioorg Med Chem Lett* **2021,** *41*, 128031.

90.   Horsfall, A. J.; Vandborg, B. A.; Kowalczyk, W.; Chav, T.; Scanlon, D. B.; Abell, A. D.; Bruning, J. B., Unlocking the PIP-box: A peptide library reveals interactions that drive high-affinity binding to human PCNA. *J Biol Chem* **2021,** *296*, 100773.