Approaches to study protein interactions

Submitted in fulfilment of the degree

Doctor of Philosophy (Chemistry)

Aimee Jade Horsfall

B. Sc. (M&DD), MPhil (Chem. Sci.)

The University of Adelaide

School of Physical Sciences

Department of Chemistry

Supervisors:

Prof. Andrew D. Abell & Dr. John B. Burning

April 2021

To my family, and especially the numerous fearless, strong and intelligent women who are my role models every day.

TABLE OF CONTENTS

Table of contents iii

Abstract xii

Declaration xv

Abbreviations xvii

Acknowledgments xviii

Publications from this work xx

Accepted for publication xx

Manuscripts in preparation xx

Additional publications xxi

Manuscripts in preparation xxi

Presentations of this work xxii

Chapter 1. Introduction 1

1.1 Protein-protein interactions in disease 3

Proliferating Cell Nuclear Antigen 3

1.1.1 Inhibitors of protein interactions 4

PCNA-p21 interaction 5

1.2 Peptides and peptidomimetics to target protein interactions 6

1.2.1 Protein structure 6

α-helix 7

3_{10}-helix 8

β-strand/sheet 8

1.2.2 Defining peptide secondary structure 9

Side-chain constraints 10

1.2.3 Determining secondary structure of peptidomimetics 11

Nuclear magnetic resonance spectroscopy 12

Circular dichroism 12

X-ray crystallography 13

1.3 Assessing biological function of peptides and peptidomimetics 14

1.3.1 Binding assays 14

1.3.2 Cell permeability 15

Cell uptake imaging assays 16

Considerations of introducing auxiliary fluorescent tags 16

1.3.3 Nuclear permeability 17

1.3.4 Proteolytic stability 17

1.4 Peptides as protein sensors 17

1.4.1 Solvatochromic amino-acids 18

1.5 Solid-phase peptide synthesis 19

1.5.1 Common side-reactions 19

1.6 Aims for this thesis and thesis outline 21

References 23

Chapter 2. A bimane-based peptide staple for combined helical induction and fluorescent imaging ... 29

Abstract ... 31

Statement of authorship ... 32

2.1 Introduction ... 35

2.2 Results and discussion ... 35

2.2.1 Synthesis ... 35

2.2.2 Structural studies ... 37

i-i+4 and i-i+7 constrained peptides ... 38

i-i+3, i-i+5 and i-i+6 constrained peptides ... 38

i-i+2 constrained peptides ... 40

2.2.3 Fluorescence characterisation ... 42

2.2.4 Cellular imaging ... 42

2.3 Conclusion ... 43

2.4 Experimental ... 43

2.4.1 General information ... 43

2.4.2 Synthesis and characterisation ... 44

2.4.3 Fluorescence assays ... 47

2.4.4 Bimane cyclisation optimisation experiments ... 47

2.4.5 NMR structure calculations ... 47

2.4.6 Circular dichroism experiments ... 48

2.4.7 Cell imaging ... 48

2.5 Acknowledgements ... 48

References ... 49

Supplemental Data. ... 53

S2.1 Characterisation data ... 55

S2.2 %Helicity calculations ... 65

S2.3 NMR structure assignment ... 65

S2.4 NMR structure calculations ... 67

References ... 75

Chapter 3. Approaches to introduce helical structure in cysteine-containing peptides with a bimane group ... 77

Abstract ... 79

Statement of authorship ... 80

3.1 Introduction ... 82

3.2 Results and discussion: ... 83

Design and synthesis ... 83

Structural analysis: ... 86

Alpha-helical ERα binding conformation ... 88

Fluorescence characterisation ... 89

3.3 Conclusions ... 91

3.4 Experimental section ... 91

3.4.1 Synthesis and characterisation ... 91

3.4.2 NMR analysis ... 94

3.4.3 Bimane attachment of unprotected peptide in solution and Ellman reagent test ... 94

3.4.4 Computational modelling of bimane-peptide conformation by energy minimisation 95
3.4.5 Fluorescence characterisation 95
3.4.6 Circular dichroism 95
3.5 Acknowledgments 95
3.6 Conflicts of interest 95
References 96
Supplemental Data. 99
S3.1 Additional characterisation data 101
S3.2 Ellman reagent test absorbance spectra 110
S3.3 CD spectra 111
%Helicity calculations 111
S3.4 NMR chemical shift analysis tables 112
S3.5 Supplementary computational modelling images 113
S3.6 Fluorescence/absorbance characterisation 115
References 116
Chapter 4. Targeting PCNA with peptide mimetics for therapeutic purposes 119
Abstract 121
Statement of authorship 123
4.1 Introduction 125
4.2 PCNA structure and function 125
4.3 The PCNA-interacting protein (PIP) box and p21 126
4.4 PCNA-binding partners display structural variation 127
4.5 The non-canonical PIP-box and sequence variation 128
4.6 PIP-box divergence 128
4.7 Chemical modification of the PIP-box and therapeutic design 133
4.8 Summary and outlook 135
4.9 Acknowledgements 135
References 136
Chapter 5. Unlocking the PIP-box: A peptide library reveals interactions that drive high-affinity binding to human PCNA 139
Abstract 141
Statement of authorship 142
5.1 Introduction 144
5.2 Results 146
5.3 Discussion 150
5.3.1 Conserved glutamine position 1: Glutamine enhances PCNA affinity. 151
5.3.2 Conserved hydrophobic position 4: Not too big and not too small 152
5.3.3 Conserved aromatic positions 7/8: Hydrogen bonds create tighter interactions between peptide and PCNA. 153
5.3.4 Non-conserved position 2/3: A positively charged residue at P3 increases affinity 154
5.3.5 Non-conserved position 5/6: Side-chains with hydrogen bond donor and acceptor character increase affinity. 155
5.3.6 Individual PIP-box modifications must be considered in the context of the entire PIP-box sequence. 156

5.3.7 Rationally designed PIP-box sequences: Designing the highest affinity PCNA partner to date…… *158*
5.4 Conclusions and outlook 160
5.5 Methods 161
5.5.1 Peptide synthesis *161*
5.5.2 Protein expression and purification *161*
5.5.3 SPR assays *162*
5.5.4 Protein-peptide co-crystallisation experiments *163*
5.5.5 Computational modelling *163*
5.6 Data availability 163
5.7 Acknowledgements 163
5.8 Funding 164
5.9 Conflicts of interest 164
References 165
Supplemental Data. 167
S5.1 Synthesis and characterisation of peptides 169
S5.1.1 Peptide synthesis *169*
S5.2 SPR 172
S5.3 Co-crystal experiments 173
S5.4 Computational modelling of PCNA monomers bound to $p21_{\mu}$ peptides 177
S5.4.1 Analysis of structures *189*
S5.5 Comparison of structures to native sequences 190
References 191
Chapter 6. A cell permeable bimane-constrained PCNA-interacting peptide 193
Abstract 195
Statement of authorship 196
6.1 Introduction 199
6.2 Results and discussion 200
6.2.1 Design and synthesis *200*
6.2.2 Binding affinity *201*
6.2.3 Structural analysis *202*
6.2.4 Cell imaging *207*
6.3 Conclusions 209
6.4 Acknowledgements 210
6.5 Conflicts of interest 210
References 211
Supplemental Data. ……...... 213
S6.1 Experimental 215
S6.1.1 General information *215*
S6.1.2 Synthesis and characterisation *215*
S6.1.3 General procedure for cleavage and isolation: *215*
S6.1.4 SPR protocol: *219*
S6.1.5 NMR analysis *220*
S6.1.6 Protein expression and purification *220*
S6.1.7 Protein-peptide co-crystallisation experiments *220*

S6.1.8 Computational modelling ... *221*
S6.1.9 Cell imaging ... *221*
S6.2 Supplementary information ... 222
S6.2.1 HPLC spectra ... *222*
S6.2.2 NMR characterisation ... *225*
S6.2.3 SPR information ... *227*
S6.2.4 Crystallographic data statistics ... *228*
S6.2.5 PCNA-peptide structures supplementary data ... *228*
S6.2.6 Main-chain assignments and secondary shift analysis ... *235*
References ... **238**

Chapter 7. A nuclear permeable peptidomimetic to target PCNA ... **239**
Abstract ... **241**
Statement of authorship ... **242**
7.1 Introduction ... 247
7.2 Results ... 248
7.2.1 Identifying the shortest sequence that maintains high affinity PCNA binding ... *249*
7.2.2 The influence of positive charge on cell uptake of p21 peptides ... *249*
7.2.3 Introducing an NLS to increase cell uptake ... *251*
7.2.4 Macrocyclic p21 peptides appended to NLS sequences ... *251*
7.3 Discussion ... 253
7.4 Acknowledgements ... 254
7.5 Competing interests ... 254
References ... **255**
Supplemental Data. ... **257**
S7.1 Methods ... 259
S7.1.1 Synthesis and characterisation of peptides ... *259*
S7.1.2 General SPR protocol: ... *261*
S7.1.3 Cell uptake and imaging experiments ... *262*
S7.2 Peptide synthesis supplementary data ... 263
S7.2.1 Supplementary cell images ... *266*
References ... **267**

Chapter 8. A turn-on fluorescent PCNA sensor ... **269**
Abstract ... **271**
Statement of authorship ... **272**
8.1 Introduction ... 275
8.2 Results and discussion ... 277
8.3 Conclusions ... 281
8.4 Acknowledgments ... 282
8.5 Conflicts of interest ... 282
References ... **283**
Supplemental Data. ... **285**
S8.1 Experimental ... 287
S8.1.1 General information ... *287*
S8.1.2 Fluorophore synthesis ... *287*

S8.1.3 Peptide synthesis *288*
S8.1.4 Plate reader experiments *291*
S8.1.5 SPR protocol *292*
S8.2 Supplementary Information 293
S8.2.1 Fluorescence Data *293*
S8.2.2 SPR data *294*
References 295

Chapter 9. Summary, future directions and outlook 297

Supplemental Data. 307
S9.1 Experimental 309
S9.1.1 General information *309*
S9.1.2 Synthesis and characterisation *309*
Lactam peptides *310*
Bimane peptide *312*
Thioether peptides *313*
All-hydrocarbon peptides *315*
Azide + Alkyne peptides *316*
S9.1.3 SPR Protocol *318*
S9.2 Supplementary information 319
S9.2.1 SPR *319*
References 319

Appendix 1. An inherently fluorescent peptide constraint to define secondary structure: Moving away from auxiliary tags 321

Statement of authorship 323
A1.1 Background 325
A1.2 Method and Applications 325
On-resin attachment of bimane *326*
Solution-phase bimane attachment: *326*
Funding *326*
Conflict of Interest *326*
References 327

Appendix 2. Other contributions 329

A silk-based functionalization architecture for single fiber imaging and sensing 331
A2.1 Background 331
Contribution: *331*
Carbon-binding peptides to functionalise nanodiamond 332
A2.2 Background 332
Contribution *332*
A2.3 Synthesis and characterisation of peptides 332
References *333*
Protein detection enabled using functionalised silk-binding peptides on a silk-coated optical fibre 334
A2.4 Abstract 334
Contribution *334*

Defining high affinity inverse agonism in PPARγ ... 335

A2.5 Abstract ... 335

Contribution ... 335

General method for peptide synthesis and characterisation ... 335

Identification and characterisation of a novel L-amino-acid ligase from a pathogenic bacterium ... 337

A2.6 Background: ... 337

Contribution: ... 337

A2.7 Methods ... 337

A2.8 Compound characterisation ... 338

A2.9 Results ... 338

Synthesis of modifiable photocleavable MALDI mass tags using copper click chemistry ... 342

A2.10 Background ... 342

Contribution ... 342

A2.11 Synthesis and characterisation ... 342

Design and synthesis of bone-targeting prodrugs for the selective treatment of breast cancer bone metastasis ... 344

A2.12 Background ... 344

Contribution ... 345

A2.13 Synthesis and characterisation ... 345

ABSTRACT

This thesis presents studies on series of modified peptides to study protein-protein interactions. Specifically, the use of new fluorescent peptide modifications to influence secondary structure and secondly, development of a peptidomimetic scaffold to target Proliferating Cell Nuclear Antigen (PCNA). **Chapter 1** introduces the importance of protein-protein interactions, their role in disease, and the difficulty in targeting these large-surface-area interactions. Peptides present as an ideal compromise between the large size of proteins and the drug-like properties of small molecules, to study such interactions. The importance of secondary structure and methods to stabilise a conformation favourable to binding specific proteins are outlined in order to address limitations associated with the use of peptides as therapeutics. Methods used to prepare such peptides, and determine the resulting conformations are outlined, along descriptions of techniques to characterise biological activity of peptides and their interaction with proteins.

Chapter 2 details studies on the synthesis of peptide macrocycles that are constrained by a bimane group that covalently links two cysteine residues. The methodology allows preparation of nine macrocycles that range in size from 16 to 31 atoms. These peptides are cell permeable, where blue fluorescence corresponding to the bimane is present in the cell cytosol. CD and NMR secondary shift analysis revealed that the *i*-*i*+4 bimane-constrained pentapeptide is α-helical. **Chapter 3** extends the investigation presented in Chapter 2 and demonstrates the *i*-*i*+4 bimane constraint introduces 3_{10}-helical structure into a 12 amino-acid sequence known to target Estrogen Receptor alpha (ERα). This same peptide adopts an α-helical geometry in the presence of 20% TFE, and when the peptide is bound to ERα as shown by computational modelling. This demonstrates sufficient flexibility in the bimane-constrained macrocycles to adopt α-helical conformation. Interestingly, helical structure is also adopted in acyclic peptides, where a bimane-modified cysteine is six amino-acids away from a tryptophan or phenylalanine residue. The fluorescence properties of the bimane-modified peptides are also presented. The methodology to introduce the bimane constraint into peptides and define helical structure is summarised in a mini-review-style Focus article in Appendix 1.

Chapter 4 reviews the structure-activity relationship of peptides and molecules that bind PCNA, and summarises the current knowledge of features that allow specific and high affinity interaction with PCNA. **Chapter 5** presents a series of 51 modified p21-derived peptides (141-155) with sequence modifications to the PCNA-interacting motif (known as the PIP-box). SPR analysis identified seven peptides that bind PCNA with higher affinity than the native p21 sequence (12.3 nM). The PCNA-binding affinity was correlated to the binding conformation of these peptides bound to PCNA, as studied by X-ray crystallography and computational modelling, and highlights a series of important hydrogen bonding networks that modulate PCNA binding affinity. Collectively, these data elucidate the rational design of a new PCNA-binding peptide that binds with the highest affinity reported to date (1.21 nM). A series of five macrocyclic p21 peptides is presented in **Chapter 6**, where a range of covalent constraints were installed by dithiol *bis*-alkylation in a p21 peptide (143-154) containing two cysteines at positions 146 and 149. The

binding affinity of the resulting *i*-*i*+4 constrained macrocycles for PCNA was determined by SPR to reveal the bimane-constrained p21 peptide as the most potent at 570 nM. The secondary structure adopted by the macrocycles bound to PCNA was studied by X-ray crystallography and computational modelling, and indicated that the bimane-constrained peptides are the only macrocycle to adopt the key 3_{10}-helical binding conformation. Additionally, the bimane peptide was cell permeable, in comparison to a linear fluorescein-tagged p21 peptide of the same length, where confocal microscopy revealed blue fluorescence corresponding to the bimane in the cell cytosol. **Chapter 7** examines the role of the peptide sequence flanking PCNA-binding motif of the p21 peptide, and reveals that a short p21 peptide of 12 amino-acids (143-154) can bind PCNA with 102 nM affinity. These short p21 peptides, however, are not cell permeable and a series of nuclear locating sequence (NLS)-tags were appended to a p21 peptide in an endeavour to improve cell and nuclear uptake, but were unsuccessful. Albeit, nuclear permeability was achieved when the NLS-tags were instead appended to the macrocyclic p21 bimane-constrained peptide presented in Chapter 6. Only the bimane-constrained p21 macrocycle appended to a SV40 NLS-tag via a thiol-maleimide linkage was nuclear permeable. Interestingly, nuclear entry was only permitted when a fluorescein was coupled to the *N*-terminus of the SV40-tag.

Chapter 8 explores the use of solvatochromic amino-acids in a p21 sequence (141-155), at the positions known to insert into the hydrophobic cleft on the PCNA surface (147, 150 and 151), to determine whether a peptide could be utilised as a selective turn-on fluorescent PCNA sensor. Two different solvatochromic amino-acids were utilised, 4-DMNA and 4-DMAP, and the fluorescence properties and affinity of the six resulting peptides for PCNA was characterised. This revealed that only the peptides with 4-DMNA or 4-DMAP at position 150 maintained high affinity binding to PCNA. A 10-fold increase in fluorescence intensity, relative to the peptide alone, was achieved in the presence of 2.5 equivalents of PCNA for the 151-substituted 4-DMNA peptide. The 151-substituted 4-DMAP peptide only produced a 3.5-fold change in fluorescence under the same conditions.

Chapter 9 provides an overall summary of this thesis, including the methodology to introduce a bimane modification into short peptides and the optimal configurations to introduce helical structure. It also details key mutations presented in Chapter 5 to improve binding affinity of peptides for PCNA, along with the importance of stabilising a 3_{10}-helical binding conformation (Chapter 6). Chapter 9 presents a new series of nine *i*-*i*+4 constrained macrocyclic peptidomimetics that include the high affinity sequence modifications, in order to consolidate these distinct achievements. The binding affinity of these macrocycles for PCNA, and corresponding linear analogues, was determined by SPR and revealed that in all except for one case, the binding affinity of the macrocycle was greater than the linear analogue. The position of a polar group (e.g. amide or acyl thioether) in asymmetric constraints alters PCNA binding affinity, where higher affinity for PCNA was observed when the polar group is nearer the *C*-terminal end of the constraint. The most potent peptide of this series is a Lys/Glu lactam macrocycle which binds PCNA with 8.16 nM affinity, the highest affinity peptidomimetic that binds PCNA to date. Finally, this chapter outlines future directions to build on this work and develop a peptidomimetic that targets PCNA for application as a cancer therapeutic.

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 acknowledge that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I also 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.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

Aimee Jade Horsfall

26/04/2021

ABBREVIATIONS

1D, one-dimensional; **2D**, two-dimensional; **3D**, three-dimensional; **a.u.**, atomic units; **aa**, amino-acid/s; **ACN**, acetonitrile; **ADP**, adenosine diphosphate; **AF-2**, activation function 2; **Ala (A)**, alanine; **ANFF**, Australian national fabrication facility; **APIM**, AlkB homologue 2 PCNA-interacting motif; **ARC**, Australian Research Council; **Arg (R)**, arginine; **Asn (N)**, asparagine; **Asp (D)**, aspartic acid; **ATP**, adenosine triphosphate; **Bim**, *syn*-9,10-(CH_3,CH_3)bimane; **Boc**, tert-butoxycarbonyl; **BRo5**, beyond rule of 5; **BSA**, Bovine Serum Albumin; **%BSA**, percentage buried surface area; **c.f.**, compare with; **CD**, circular dichroism; **CNBP**, Centre for Nanoscale Biophotonics; **CO**, carbonyl; **CSY**, cyanosulfurylide; **Cys (C)**, cysteine; **Cα**, alpha-carbon; **DAPA**, 4-*N,N*-dimethylaminophthalimidoalanine; **DAPI**, 4′,6-diamidino-2-phenylindole; **dBB**, dibromobimane; **DCM**, dichloromethane; **DEAE**, diethylethanolamine; **DIC,** *N,N'*-diisopropylcarbodiimide; **DIPEA**, *N,N'*-diisopropylethylamine; **DMAP**, 4-(*N,N*-dimethylamino)phthalimide; **Dmb**, dimethoxylbenzyl; **DMEM**, Dubelco's modified eagle media; **DMF**, *N,N'*-dimethylformamide; **DMN**, 4-*N,N*-dimethylamino-1,8-naphthalimide; **DMNA**, 4-*N,N*-dimethylamino-1,8-naphthalimidoalanine; **DNA**, deoxyribonucleic acid; **DODT**, 2,2′-(ethylenedioxy)diethanethiol; **dsDNA**, double stranded deoxyribonucleic acid; **DTNB**, 5,5'-dithio-*bis*-2-nitrobenzoic acid; **DTT,** dithiothreitol; **E. Coli,** *Escherichia Coli;* **e.g.**, for example; **EDC,** 1-ethyl-3-(3-diethylaminopropyl)carbodiimide; **EDTA**, ethylenediaminetetraacetic acid; **equiv**, equivalent/s; **ERα**, Estrogen Receptor alpha; **ESI+**, positive-mode electrospray ionisation; **ESI-TOF**, electrospray ionisation time of flight; **FACS**, fluorescence assisted cell sorting; **FBS**, fetal bovine serum; **FEN-1**, Flap endonuclease 1; **FITC**, fluorescein isothiocyanate; **Fmoc**, 9-fluorenylmethoxycarbonyl; **FP**, fluorescence polarisation; **FRET**, Förster resonance energy transfer; **Gln (Q),** glutamine; **Glu (E),** glutamic acid; **Gly (G)**, glycine; **h**, hour/s; **HATU**, (1-[bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxide hexafluorophosphate; **Hcy**, homocysteine; **HEPES,** 4-(2-hydroyethyl)-1-piperazineethanesulfonic acid; **His (H),** histidine; **Hmb**, 2-hydroxy-4-methoxybenzyl; **HMBC**, heteronuclear multiple-bond correlation; **HOBt,** 1-hydroxybenzotriazole hydrate; **HPLC**, high performance liquid chromatography; **HRMS,** high resolution mass spectrometry; **HSQC**, heteronuclear single quantum coherence; **Hα**, alpha-proton; **i.e.**, that is; **IDCL**, interdomain connecting loop; **Ile (I),** isoleucine; **IPTG,** isopropyl β-D-1-thiogalactopyranoside; **ITC**, isothermal calorimetry; **LB,** Lennox broth; **LBD**, ligand binding domain; **LCMS**, liquid chromatography mass spectrometry; **Leu (L)**, leucine; **Lys (K)**, lysine; **mBB**, monobromobimane; **MeOH**, methanol; **Met (M)**, methionine; **MF**, molecular formula; **min**, minute/s; **Mmt**, 4-monomethoxytrityl; **Mpe**, 1-(4-methoxyphenyl)ethyl ester; **MRE**, molar residue ellipticity; **MS**, mass spectrometry; **MW**, microwave; **Mw**, molecular weight; **NH**, amide proton; **NHS**, *N*-hydroxysuccinimide; **NIH**, national institute of health (USA); **NLS**, nuclear locating sequence; **NMR**, nuclear magnetic resonance; **NOE**, nuclear Overhauser effect; **NOESY**, nuclear Overhauser effect spectroscopy; **PAF**, platelet-activating factor; **PARG**, poly(ADP-ribose)glycohydrolase; **Pbf**, 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl; **PBS,** phosphate buffered saline; **PCNA**, Proliferating Cell Nuclear Antigen; **PEG**, poly-ethylene glycol; **PFA**, paraformaldehyde; **PG**, protecting group; **Phe (F)**, phenylalanine; **PIP**, PCNA-interacting protein/peptide; **pol**, polymerase; **PPI**, protein-protein interaction; **ppm**, parts per million; **Pro (P)**, proline; **PyBOP**, (benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate; **RFC**, replication factor C; **RMSD**, root mean square deviation; **ROESY**, rotating-frame nuclear Overhauser effect correlation spectroscopy; **RP-HPLC**, reverse-phase high performance liquid chromatography; **rt**, room temperature; **RTPS**, research training program stipend; **RU**, response units; **SDS**, sodium dodecyl sulfate; **Ser (S),** serine; **SI**, supplementary information; **SPE**, solid-phase extraction; **SPPS**, solid-phase peptide synthesis; **SPR**, surface plasmon resonance; **SRC2**, Steroid Receptor Coactivator 2; **SV40**, simian virus 40; **TAT**, transactivator of transcription; **tBu**, *tert*-butyl; **TCEP**, tris(2-carboxyethyl)phosphine; **TFA**, trifluoroacetic acid; **TFE**, 2,2,2-trifluoroethanol; **Thr (T)**, threonine; **TIPS**, triisopropylsilane; **TNB**, 5-thio-2-nitrobenzoate; **TNBS**, 2,4,6-trinitrobenzenesulfonic acid (picrylsulfonic acid); **TOCSY**, total correlation spectroscopy; **Tris,** trisaminomethane; **Trp (W)**, tryptophan; **Trt**, trityl; **Tyr (Y)**, tyrosine; **UV**, ultra-violet; **UV-Vis**, ultra-violet to visible; **v/v**, volume per volume; **Val (V)**, valine; **VDW**, van der Waal; **w/v**, weight per volume; **WRN**, Werner; **Xaa**, any amino-acid; **XPG**, Xeroderma pigmentosum group G.

ACKNOWLEDGMENTS

They say it takes a village to raise a child, and if my thesis has been the child I've been tending to for the last four years, that saying definitely holds true. I could not have completed this feat without the support of a large number of people, a small number of whom I will list here, but to the many more have shaped my PhD journey, both directly and indirectly – thank you.

Firstly, to my supervisors. Andrew, you have been my postgraduate supervisor for over six years and have shaped me into the researcher I am today. Thank you for your unwavering support and encouraging me to challenge myself, "back myself", and pursue every opportunity. John, thank you for all your encouragement and sharing your passion for the science, especially on the PCNA project – it has made me excited to see where the science goes.

To Denis, I always look forward to our chats and updating you on our latest results, or picking your brain on my latest problem. Thank you for sharing your friendship, mentorship and enthusiasm for peptide chemistry with me, as well as introducing me to so many great researchers in the field, in particular to Wioleta. Thank you for letting me into your lab Wioleta, and all the hints, tricks and friendship you've shared along the way.

I have been fortunate to have worked with so many great people and made a number of friendships throughout my PhD, especially in the Abell and Bruning labs. Thank you all for the laughs, it has been a pleasure to work alongside and collaborate with you, and I hope we cross paths again soon. Most especially, to Kathryn, Dion and Rouven – I cannot thank you enough for all your friendship and support, in and out of the lab – the many tears (both happy and otherwise), laughs and beers shared, I know we have a bond that won't fade and I cannot wait to celebrate with you all as you finish up. Thank you also to Theresa and Beth, who have been instrumental to the progress of the PCNA project, and worked so tirelessly – I couldn't ask for more diligent collaborators to work with and am so very proud of what we have achieved in the last couple years, you girls are awesome.

Further abroad, I would also like to thank the CNBP for all the opportunities I have been afforded during my PhD, and providing such a supportive and nurturing environment in which to train as a researcher. I can only hope I can help cultivate some of what we have in my new workplaces. Thank you to all within the Centre for your support, especially to Mark, Kathy & Mel for leading the way. Similarly, thanks to all of those within the chemistry department at Adelaide, and most especially to the technical staff, Matt, Phil, Peter and Gino, that have always kept things running smoothly. To our extended Chemistry family at Adelaide, I will miss the comraderie and consistency of Friday knock-offs.

To Drs. Tash and Oli, you have both now been such a long-standing, consistent source of support in my chemistry life, throughout undergraduate and now research – here's to many more wines to come. And to my oldest friends, Rach, Klay and Cohen – I don't think I could've made it this far without you, starting with the leap to University. I know there are many more dinners to be shared as we all start yet another Chapter.

Closer to home, to my family – the whole lot of you. Thank you for your interest (feigned or not), letting me attempt to explain what I do year after year, providing different perspectives, and challenging me to find better ways to talk about science.

Mum & Dad – thank you. For everything. You have always gone above and beyond to provide us with opportunities, and pushed us to improve. You both encourage us to question the status quo, to ask why, and problem solve; but more than that, you have taught us to be dependable, to lend a hand and the meaning of hard work – I do truly believe these qualities are the reason I am any good at what I do. Lachlan, thank you for keeping me grounded. You have always been a role model to me, you're work ethic and sense of humour are second to none, and I always look forward to seeing you, whether that's to catch up, joke, share music, or talk about nothing in particular at all.

Lastly, to Pat. I think we make a pretty good team. Thank you for keeping me sane, being my sounding board when things go sideways, driving me to the lab late at night and refilling my wine glass. This journey would not have been the same without you, nor do I think I could have achieved half as much without your support. It will be strange not working together all day now, but I'm excited for what the next chapter brings.

Happy reading, or just pawing through the pretty pictures.

PUBLICATIONS FROM THIS WORK

MiniReview:

A. J. Horsfall, A. D. Abell and J. B. Bruning, Targeting PCNA with Peptide Mimetics for Therapeutic Purposes, *ChemBioChem*, 2019, **21**(4) 442-450, doi.org/10.1002/cbic.201900275.

Research Article:

A. J. Horsfall, K. R. Dunning, K. L. Keeling, D. B. Scanlon, K. L. Wegener and A. D. Abell, A bimane-based peptide staple for combined helical induction and fluorescent imaging, *ChemBioChem*, 2020, **21**(23) 3423-3432, doi.org/10.1002/cbic.202000485.

Research Article:

A. J. Horsfall, T. Chav, J. B Bruning and A. D. Abell, A turn-on fluorescent sensor for PCNA, *Journal of Bioorganic and Medicinal Chemistry Letters*, 2021, **41** 128931, doi.org/10.1016/j.bmcl.2021.128031.

Research Article:

A. J. Horsfall, B. A. Vandborg, W. Kowalczyk, T. Chav, D. B. Scanlon, A. D. Abell and J. B. Bruning, Unlocking the PIP-box: A peptide library reveals interactions that drive high-affinity binding to human PCNA, *Journal of Biological Chemistry,* 2021, **296** 100773, doi.org/10.1016/j.jbc.2021.100773.

Research Article:

A. J. Horsfall, D. P. McDougal, D. B. Scanlon, John B. Bruning, and A. D. Abell, Approaches to introduce helical structure in cysteine-containing peptides with a bimane group, *ChemBioChem*, 2021, **22**(17) 2711-2720, doi.org/10.1002/cbic.202100241.

Accepted for publication

Research Article:

A. J. Horsfall, B. A. Vandborg, Z. Kikhtyak, D. B. Scanlon, T. E. Hickey, W. D. Tilley, J. B. Bruning and A. D. Abell, A cell permeable, fluorescent peptidomimetic inhibitor of PCNA, *RSC Chemical Biology*, 2021, In press, doi.org/10.1039/D1CB00113B.

Focus Article:

A. J. Horsfall, and A. D. Abell, A bifunctional fluorescent peptide modification to influence secondary structure: Moving away from auxiliary tags, *Australian Journal of Chemistry,* 2021, In press.

Manuscripts in preparation

Research Article:

A. J. Horsfall, T. Chav, Z. Kikhtyak, J. L. Pederick, W. Kowalczyk, D. B. Scanlon, W. D. Tilley, T. E. Hickey, D. J. Peet, A. D. Abell and J. B. Bruning, A nuclear permeable peptidomimetic to inhibit PCNA, *Manuscript in preparation.*

ADDITIONAL PUBLICATIONS

Research Article:

P. K. Capon, J. Li, **A. J. Horsfall**, S. Yagoub, E. P. Schartner, A. Khalid, R. W. Kirk, M. S. Purdey, K. R. Dunning, R. A. McLaughlin* and A. D. Abell*, A silk-based functionalization architecture for single fiber imaging and sensing, *Advanced Functional Materials*, 2021, 2010713 doi.org/10.1002/adfm.202010713.

Research Article:

P. K. Capon, **A. J. Horsfall**, J. Li, E. P. Schartner, A. Khalid, M. S. Purdey, K. R. Dunning, R. A. McLaughlin and A. D. Abell*, Protein detection enabled using functionalised silk-binding peptides on a silk-coated optical fibre, *RSC Advances*, 2021, **11** 22334-22342, doi.org/10.1039/D1RA03584C.

Manuscripts in preparation

Research Article:

P. K. Capon, P. Reineck, **A. J. Horsfall**, W. Kowalczyk, M. D. Torelli, O. A. Shandarova, M. S. Purdey and A. D. Abell*, Carbon-binding peptides to functionalise nanodiamond, *Manuscript in preparation.*

Research Article:

R. L. Frkic, B. Jovcevski, E. Bruning, M. R. Chang, **A. J. Horsfall**, W. Kowalczyk, B. Pascal, T. M. Kamenecka, J. Harmer, P. R. Griffin and J. B. Bruning*, Defining high affinity inverse agonism in PPARγ, *Manuscript in preparation.*

Research Article:

J. L. Pederick, N. Hao, **A. J. Horsfall**, A. D. Abell, K. E. Shearwin and J. B. Bruning*, Identification and characterisation of a novel L-amino-acid ligase from a pathogenic bacterium, *Manuscript in preparation.*

Research Article:

R. Becker, S. Fritter, **A. J. Horsfall**, A. D. Abell* and A. C. W. Zannettino*, Design and synthesis of bone-targeting prodrugs for the selective treatment of breast cancer bone metathesis, *Manuscript in preparation.*

PRESENTATIONS OF THIS WORK

University of Adelaide Chemistry Department – Final PhD seminar – February 2021

“PCNA and Peptides: Development of a peptide-based PCNA inhibitor”

RACI Victoria Peptide User Group Online Symposium Oral Presentation – October 2020

“An inherently fluorescent 3_{10}-helical peptide constraint”

This talk was awarded the ‘Best ECR Presentation’ award.

36th European Peptide Symposium Poster Presentations - September 2020 – *Postponed due to COVID-19*

“An inherently fluorescent helical peptide constraint”

“The PIP-box unlocked: A high affinity peptide-based PCNA inhibitor for therapeutic application”

This abstract was awarded a ‘Registration Grant for Young Scientists’.

IPAS Seminar Series Oral Presentation – August 2020

“Going big: Targeting protein-protein interactions”

CNBP Virtual Conference Oral Presentation – July 2020

“Going big: Targeting protein-protein interactions”

This talk was awarded the ‘Best Presentation’ award in the section ‘Chemistry into Biology’

CNBP Virtual Conference Poster Presentation – July 2020

“Turning peptides into drugs: How to go from biology, to chemistry and back into biology”

This talk was awarded the ‘Best Poster’ award in the section ‘Chemistry into Biology’

BioPhotonics Careers Symposium Poster Presentation – November 2019

“Why you should use peptides”

This presentation was awarded the ‘Best Poster’ award.

CNBP Conference Oral Presentation – October 2019

“Peptides: Why you should use them and fibre functionalisation”

CNBP Conference Poster Presentation – October 2019

“Peptides as protein biosensors”

This presentation was awarded the ‘Best Poster Presentation Award’

BLiSS Adelaide Conference Poster Presentation – September 2019

“Designing better protein sensors: Peptide puzzle pieces”

CNBP Seminar Series Oral Presentation – September 2019

“Stop Growing! Peptide-based cell proliferation biosensors and inhibitors of PCNA”

13th Australian Peptide Conference Poster Presentation – September 2019

“Exploiting the p21 template: Development of a broad-spectrum cancer diagnostics and peptidomimetic therapeutic”

This abstract was awarded a travel grant.

7th Solid-Phase Peptide Symposium Poster Presentation – September 2019

“Exploiting the p21 template: Development of a broad-spectrum cancer diagnostics and peptidomimetic therapeutic”

This presentation was awarded the ‘Best Poster Presentation Award’. This abstract was awarded a travel grant.

Invited Oral Presentation at CSIRO, Clayton – May 2019

“Toward a therapeutic PCNA-interaction inhibitor”

IPAS Student Presentations – February 2019

“The PCNA-p21 Jigsaw Puzzle: Design of therapeutics and biosensors”

This presentation was awarded the ‘Tanya Monro Presentation Award for best talk’.

RACI Victoria Peptide User Group Meeting Oral Presentation – November 2018

“Toward a therapeutic PCNA-interaction inhibitor: Modification of the p21 template”

Adelaide Protein Group Student Award Night Poster Presentation – May 2018

“A fluorescent constraint to induce secondary structure”

CNBP Conference Poster Presentation – November 2018

“Modification of the p21 template: Toward a therapeutic PCNA-interaction inhibitor”

12th Australian Peptide Conference Poster Presentation – October 2017

“A fluorescent constraint to induce secondary structure”

This abstract was awarded a travel grant.

6th Solid-Phase Peptide Symposium Poster Presentation – October 2017

“Introduction of a new fluorescent constraint on-resin”

This presentation was awarded the ‘Best Student Poster’ award. This abstract was awarded a travel grant.

CNBP Conference Poster Presentation – November 2017

“A fluorescent constraint to induce secondary structure”

University of Adelaide Chemistry Department PhD Introductory Seminar – November 2017

“Peptides, protein-protein interactions and their secondary structure”

CNBP Mid-year Workshop Oral Presentation – July 2017

“A bifunctional fluorescent peptide constraint to introduce secondary structure”

Nothing in life is to be feared, it is only to be understood.
Now is the time to understand more, so that we may fear less.

Marie Curie.

Chapter 1.

INTRODUCTION

1.1 Protein-protein interactions in disease

Protein-protein interactions are key to almost all biochemical processes and are consequently an ideal point to intervene, in order to prevent or treat disease. Protein-protein interaction interfaces occur between large, generally featureless, surfaces that are difficult to inhibit with small molecules which are readily displaced from the interaction site by proteins.[1-3] Thus, protein-protein interactions have historically been considered ‘undruggable’. Despite these being generally featureless interfaces, protein-protein interactions occur with a high degree of specificity, conferred by the three-dimensional shape of the associated proteins. Even intrinsically disordered proteins, which do not have a fixed three-dimensional structure, participate in protein interactions and adopt a well-defined structure at the interaction interface, upon binding. The specificity and unique structural interface of each interaction enables tight regulation of biochemical processes and selective binding of one protein over another. Judicious targetting of such interfaces theoretically allows access to therapeutics with minimal to no side-effects, where only the aberrant interaction of interest is inhibited. For example, the protein Steroid Receptor Coactivator 2 (SRC2) possesses an α-helical domain which positions three conserved leucines on the same face of the helical barrel, and allows these side-chains to selectively embed in a hydrophobic groove on Estrogen Receptor alpha (ERα).[4-6] SRC2 and other estrogen receptor coactivators and corepressors share this conserved leucine-rich motif, however unique sequence changes interspersed around this motif, that differ between the various coregulators, allows tight regulation over the biological functions of ERα. An approach to stabilise the binding geometry of a peptidomimetic to mimic the interaction of SCR2 with ERα is discussed in Chapter 3. Another example of how small sequence modifications regulate the selectivity of protein interactions, is the interactions of cell cycle regulator protein Proliferating Cell Nuclear Antigen (PCNA), which associates with over 200 different protein partners through the same binding site.[7-11] A major focus of this thesis is to understand the structure activity relationship of these PCNA-protein interactions.

Proliferating Cell Nuclear Antigen

PCNA is a toroidal, homotrimeric protein with six-fold pseudosymmetry, comprised of three 27 kDa monomers.[11-13] Each monomer consists of two, nearly structurally-identical, domains which are connected by a flexible loop, termed the interdomain connecting loop (IDCL). PCNA is loaded over primer-template junctions by Replication Factor C (RFC),[14-17] and encircles dsDNA to act as a docking platform and key mediator of DNA-replication and -repair processes. PCNA slides uni-directionally with the progressing replication fork orchestrating events at the DNA, and thereby increases the processivity of a large number of enzymes required for DNA-replication and -repair, such as polymerase delta (pol δ), despite having no enzymatic function itself.[11, 18, 19] More than 200 proteins are known to bind PCNA and interact with the protein dynamically throughout the DNA-replication and -repair processes (Figure 1).[7, 10, 11, 20-23] This central role, accompanied by the fact PCNA is upregulated in many cancers, highlights it as a key inhibition target to develop broad-spectrum cancer therapeutics.[8, 9, 23-26]

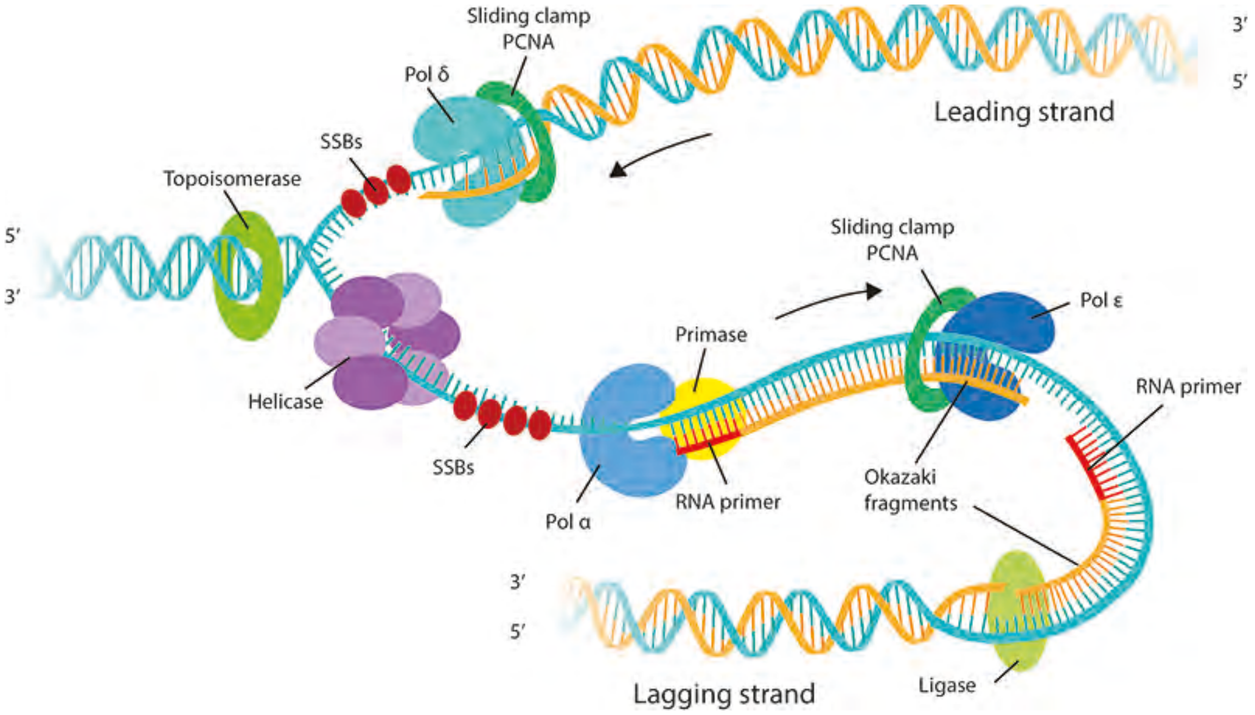


Figure 1: Cartoon representation of the process of DNA-replication, for which PCNA plays a central role. Image credit NUS.[27]

The large majority of PCNA partners bind at one of three degenerate binding sites, nestled between the two domains of a monomer, under the IDCL. Almost all partners adopt a 3_{10}-helical turn that inserts three hydrophobic residues onto PCNA's surface. This turn is generally flanked by two short β-strands which make electrostatic interactions with the PCNA *C*-terminal tail and the IDCL.[10, 22, 28, 29] Many PCNA partners bind through a consensus sequence defined as the PCNA-interacting protein motif, or PIP-box motif, defined as Qxxφxxψψ where Q is glutamine, 'x' is any amino-acid, φ is a hydrophobic residue and ψ an aromatic residue (commonly phenylalanine or tyrosine).[30]

1.1.1 Inhibitors of protein interactions

Small molecules are often not effective inhibitors of protein-protein interactions, owed to their comparatively small size relative to the surface area involved in protein interactions. Furthermore, targeting one protein specifically over another with small molecules is difficult, due to relatively featureless protein surfaces that only possess shallow grooves. In comparison, proteins are able to make highly specific interactions with large surface areas. However, the use of proteins as therapeutics is generally limited to extracellular targets owed to poor cell permeability and bioavailability of many proteins. Peptides provide an effective compromise between the specificity of proteins and cell permeability of small drug-like molecules. Peptides, short strings of (arbitrarily < 50) amino-acids, are able to participate in much larger interactions compared to small molecules, and are therefore more effective inhibitors of protein-protein interactions.[2, 31-33] In addition, a native protein interaction can inspire design of a peptide sequence that mimics the native interface, where a peptide sequence can be derived directly from a protein known to target the protein of interest (Figure 2). For example, the p21 protein is known to bind PCNA, through a conserved motif located between residues 144-151. A p21-derived peptide comprising this binding motif, residues 139-160 of p21,

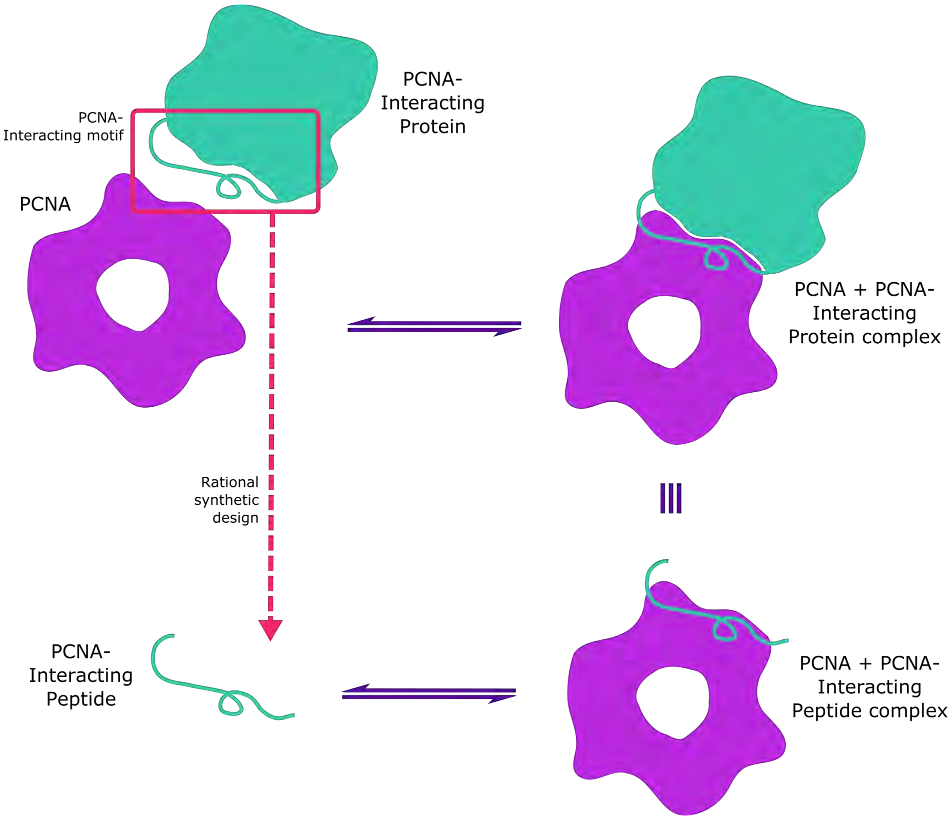


Figure 2: Design of peptides to mimic protein-protein interactions, specifically PCNA, inspired from native sequences.

is also able to bind PCNA and mimics the native PCNA-p21 interaction. Unusual, modified or non-proteinogenic amino-acids can be readily incorporated into such peptides to enhance the interaction and fortify or disguise the structures from degradation by native proteases. Additionally, peptides are able to adopt the same three-dimensional conformation as the parent peptide upon binding the protein target, that gives rise to the high specificity of the native interaction. However, a peptide has many degrees of freedom and is conformationally flexible, which consequently can incur a large entropy penalty upon binding a target and adopting a more rigid binding conformation. Additioanlly, this flexibility enables rapid degradation of and native peptide sequences by proteases which may limit the use of native linear peptides as therapeutics. These limitations must be addressed in order to reap the benefits of a peptide scaffold to target protein-protein interactions.

PCNA-p21 interaction

An example of a peptide that mimics a native protein interaction is a p21-derived peptide that binds PCNA. The cell-cycle regulator protein, p21$^{CIP1/WAF1}$ (referred to herein as p21), is the highest affinity binding partner of PCNA reported, and upon binding PCNA shuts down DNA-replication. PCNA-bound p21 competitively inhibits all other PCNA partners, required for DNA-replication and repair, from interacting with the PCNA surface.[34-36] p21 binds at the same location on the PCNA surface, the PIP-box binding site, as the large majority of other proteins required for DNA-replication and -repair.[12, 37, 38] Although p21 is an intrinsically disordered protein, the structure of the interaction interface with PCNA is well defined. The PCNA-interacting sequence of p21 forms a short β-sheet with the PCNA *C*-terminal tail, followed by a single 3_{10}-helical turn and an electrostatically stabilised β-strand with the IDCL.[12] The 3_{10}-helical turn is anchored onto the PCNA surface by three hydrophobic residues of p21's classical PIP-box, ^{144}**Q**TS**M**TD**FY**^{151}, which are inserted into a deep hydrophobic cleft. The glutamine residue is inserted into the so-called Q-pocket where it makes three main-chain hydrogen bonds with Ala205, Ala252 and Pro253. This glutamine residue has been shown to be critical to p21's ability to bind PCNA, along with the three hydrophobic residues (Met, Phe, Tyr).[37] A short peptide derived from p21, residues 139-160, has been shown to bind PCNA with comparable affinity to the native protein, where binding affinities are reported in the low nanomolar range (2.5-90 nM, technique dependent). Furthermore, this 22mer peptide can itself effectively inhibit PCNA and has been shown to outcompete the major processive polymerase, pol δ, and shut down SV40 DNA replication in an *in vitro* activity assay.[37] This peptide therefore provides a useful template to begin design of a PCNA inhibitor for application as a broad-spectrum cancer therapeutic. However, the subtleties in both structure and sequence, to which p21 owes its high PCNA binding affinity, are not well understood.

1.2 Peptides and peptidomimetics to target protein interactions

Peptides, though effective protein interaction inhibitors, are generally not successful therapeutics. Peptides comprised entirely of native L-amino-acids are readily degraded by endogenous proteases, thereby limiting peptide half-life and the potential therapeutic action. Incorporation of D-amino-acids or non-proteinogenic amino-acids generally increases *in vivo* half-life, however such modifications require strategic design to maintain binding of the target protein. Furthermore, short peptides are inherently flexible in an aqueous environment, and generally do not display a defined secondary structure. Consequently, a large entropic penalty is incurred upon a short peptide binding the target protein and adopting a defined binding conformation. This limitation, in conjunction with the short half-life of native peptides, can be addressed by introducing a chemical modification to preorganise the peptide into a conformation favourable to binding.[39-46] This in turn enhances therapeutic action. These so-called peptidomimetics have reduced peptide character, and therefore decreased protease recognition and improved *in vivo* half-life.[39, 47, 48] Many peptidomimetics also show increased cell permeability relative to native peptide counterparts.[47-52] Judicious introduction of such a chemical modification can give rise to a peptidomimetic with a defined structure that mimics the native protein binding conformation.[40, 43, 53, 54]

1.2.1 Protein structure

Protein structure is defined by four categories: Primary, secondary, tertiary and quaternary (Figure 3).[55] Primary structure is described by the sequence of amino-acids linked by amide bonds. Secondary structure is defined by the hydrogen bond network that forms between main-chain amides of a peptide sequence and gives rise to discrete structural motifs such as α-helices and β-sheets. These motifs interact through additional secondary interaction networks of the amino-acid side-chains to form domains, which is termed the tertiary structure. Lastly, multiple domains, from one or many peptide strands can associate to form a functional protein with a quaternary structure. For example, PCNA is formed from a 261 amino-acid sequence (primary structure) that folds into a series of α-helices and β-sheets (secondary structure), that further associate into two almost symmetrical domains where the α-helices line one face of the protein monomer (tertiary structure). Three identical monomers then associate to form the functional ring-shaped

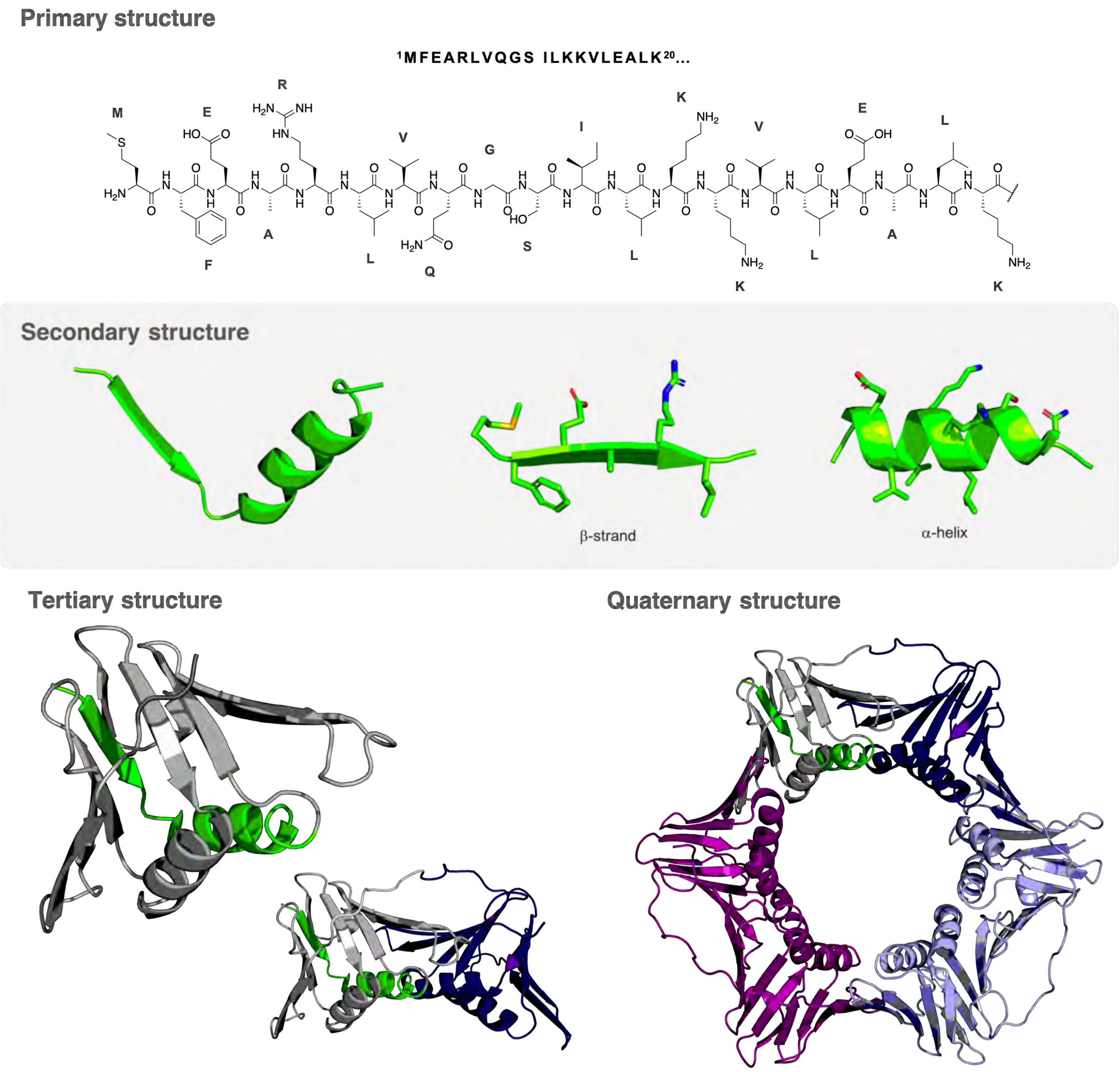

Primary structure
1MFEARLVQGS ILKKVLEALK20...
M
E
R
V
G
I
K
V
E
L
F
A
L
Q
S
L
K
L
A
K
Secondary structure
β-strand
α-helix
Tertiary structure
Quaternary structure

Figure 3: The four levels of protein structure. Primary structure, linear amino-acid sequence; secondary structure, main-chain hydrogen bonding to form discrete motifs such as β-sheets and α-helices; tertiary structure, side-chain interactions to form domains of a protein; quaternary structure, association of multiple domain or protein units to form a functional protein. Crystal structure of PCNA (PDB ID: 1VYM).

protein (quaternary structure). The secondary structure of short peptides, or peptidomimetics, plays an integral role in correctly orientating the amino-acid side-chains to effectively interact with a protein surface. The defining features of key secondary structural elements are described below.

α-helix

An α-helix is the most prevalent secondary structure motif in proteins and is defined as a barrel shaped coil of amino-acids (Figure 4), where each turn is a rise of 5.4 Å. Systematically, an α-helix is also described as a 3.6_{13}-helix, as there are 3.6 amino-acids per helical turn and 13 main-chain atoms. Every main-chain amide participates in a hydrogen bond, where the amide NH donates to a carbonyl four residues earlier in the sequence (i.e. *i*-*i*+4 hydrogen bonding) and results in a backbone dihedral angles of -60/-45° (φ/ψ,

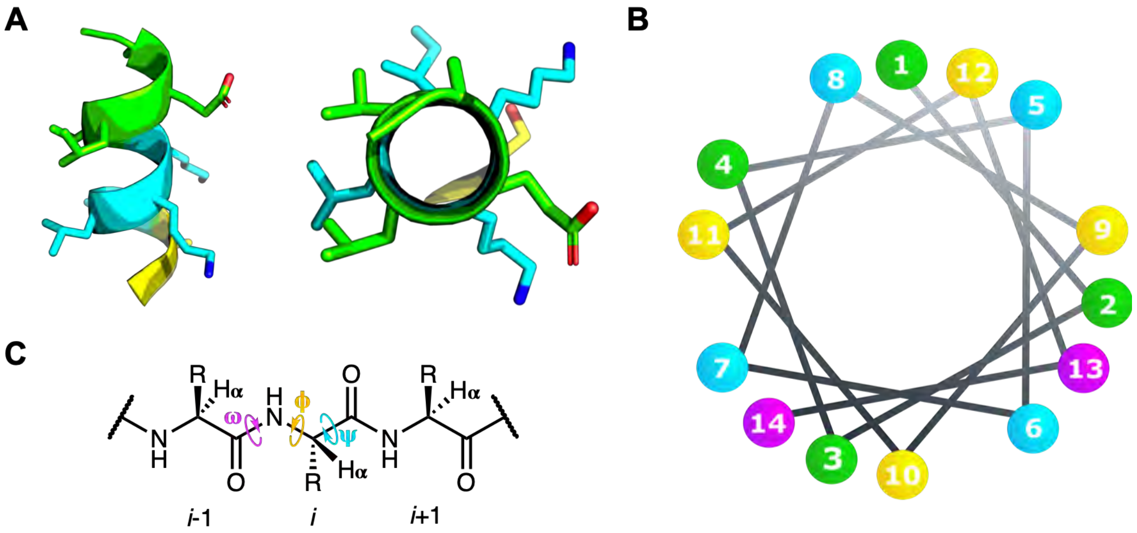


Figure 4: Alpha-helical peptide secondary structure. **A** A cartoon helix with side-chains represented as sticks; each turn of the helical is represented in a different colour, green, blue or yellow. **B** The 1st (i), 5th (*i*+4), and 8th (*i*+7) residue side-chain lie on the same side of the helical barrel, shown by the barrel diagram. **C** Dihedral torsion angles, phi (φ), psi (ψ) and omega (ω) define the type of peptide secondary structure present.

Figure 4C) or sum of 105°. Consequently, every fourth, seventh, eleventh (and so on) amino-acid side-chain is positioned on the same side of the helical barrel, but slightly off-set from one another.

3_{10}-helix

A 3_{10}-helix, is also a barrel shaped structure, where there is only three amino-acids per turn comprising of ten main-chain atoms. This results in a tighter coil, with a rise of 6.0 Å per helical turn and backbone dihedral angles on -49/-26° (φ/ψ) or sum of -75°. The amide NH of one amino-acid hydrogen bonds to a main-chain carbonyl of a residue three amino-acids earlier (*i*-*i*+3), where every third amino-acid side-chain is aligned along a face of the helical barrel. The 3_{10}-helix is less common than an α-helix and these motifs are often observed the termini of α-helical motifs. However, this motif is of particular interest to this thesis as a 3_{10}-helical turn is adopted upon peptides binding PCNA.

β-strand/sheet

A β-sheet is the second most common secondary structure motif, after α-helices. β-strands are an extended peptide arrangement, where the backbone amides are planar, and then associate with one another to form β-sheets. These strands interact through hydrogen bonding of the main-chain amides of another strand. β-Strands are directional and many strands can associate to form large β-sheets. Parallel β-strands have dihedral angles of -120/115° and occur where the *N*-terminus of each peptide strand is in the same direction, which requires the two segments to be relative distance from one other in the protein sequence, or from strands from different proteins. The amino-acid side chains of opposing strands in the resulting structure point toward one another, and there is a slight offset between the carbonyl and amide that hydrogen bond together. Consequently, a parallel sheet is less stable than an anti-parallel β-strand.

Antiparallel β-strands are defined by strands that run in opposite direction (*N*- to *C*-terminus, and *C*- to *N*-terminus). This results in the main-chain carbonyl and amides of opposing strands aligning, and the side-chains point away from one another, giving rise to a more stable structure. The dihedrals of an anti-parallel β-strand are -140/135° and commonly occur with sections of amino-acid sequence that are continuous or close together in a protein sequence, commonly connected by a β-hairpin.

1.2.2 Defining peptide secondary structure

Protein interactions, or protein-peptide interactions, occur with a defined secondary structure. This secondary structure precisely orientates the amino-acid side-chains, in a predictable manner, to interact optimally with the binding interface. For instance, an amino-acid (*i*) side-chain is positioned on the same side of the barrel as the side-chain of an amino-acid four (*i*+4), seven (*i*+7) or eleven (*i*+11) residues away. This information can be leveraged to rationally design chemical modifications for peptides in order to preorganise the peptide backbone into a conformation favourable to binding.

Many methodologies have been developed to chemically modify a peptide and stabilise a bioactive binding conformation.[2, 3, 40, 56, 57] One of the earliest and still most prevalent methods is to covalently link two, appropriately separated, amino-acid side-chains. Only a linker with optimal length, flexibility and geometry will effectively stabilise a particular secondary structure. The linker should be positioned such that it does not interfere adversely with the binding of the peptidomimetic to a protein target, and the chemical composition of a linker may be optimised to enhance any interactions it may have with the binding interface. Additionally, the position and character of the linker has been shown to influence cell permeability, or act as a useful handle to impart additional functionality.

Structural information about the protein interaction of interest is invaluable to effectively design a peptidomimetic that successfully mimics the structure of the native protein interaction. For example, it is known that SCR2 interacts with ERα through an α-helical motif that inserts three leucine residues into a hydrophobic cleft of the protein surface. Co-crystalstructures of short coactivator peptides bound to ERα have further highlighted that only one face of the helical barrel interacts with the protein.[58] This information together can inform design of a constrained peptidomimetic, with regard to where a linker can be positioned to stabilise the α-helical geometry.[4-6] That is, it must not remove the key leucine residues, and should not be positioned on the same side of the helical barrel as those leucine residues, which may hinder binding to ERα. If the protein obscured a second face of the peptide helix upon binding, this would also need to be considered when choosing the location of a linker in a constrained peptidomimetic. Covalently linking an amino-acid (*i*) that points away from the interacting face, with a residue either 4 (*i*+4) or 7 (*i*+7) or 11 (*i*+11) residues away is known to stabilise and α-helix (Figure 5). It has also been shown that an *i*-*i*+4 linker positioned at different locations in a peptide sequence, along the same face of the helix, can drastically change the helicity and cell permeability of the resulting peptidomimetic.[59, 60] Additionally, multiple linkers could be introduced to achieve a more stable helical structure.[39, 48]

Other secondary structure conformations can also be mimicked by constrained peptidomimetics. In the case of a 3_{10}-helix, generally *i* and *i*+3 separated residue side-chains are linked, or *i* and *i*+2 for short β-strands. X-ray crystal structures of peptides bound to a target protein, particularly those with non-classical

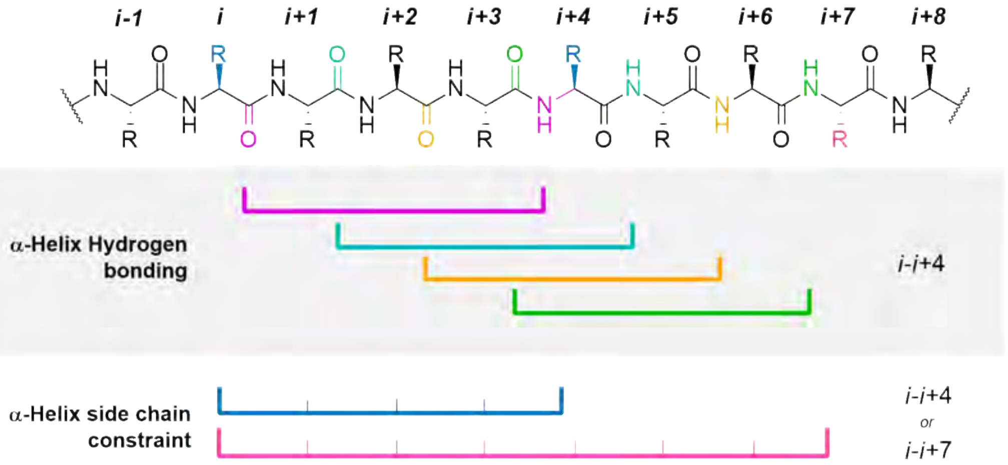


Figure 5: α-helical hydrogen bonding network and side-chain constraint spacing to introduce α-helical structure.

conformations, can better inform where to install a linker and used to help predict the optimal length of the linker. For example, in the case of a $p21_{139-160}$ peptide bound to PCNA, it might be expected that the 3_{10}-helical turn is best stabilised by a *i*-*i*+3 linker, however the side-chains of 146 (*i*) and 149 (*i*+4) residues are directed toward one another and were deemed better candidates to introduce a lactam linker.[61] Indeed, the less traditional linker position did stabilise the 3_{10}-helical conformation both in solution and bound to PCNA.

However, despite the best efforts to rationally design such peptidomimetics, there are often seemingly unpredictable outcomes. In particular, the location of a covalent linker can impact not only the structure but also the proteolytic stability, cell permeability, or may indirectly influence the binding affinity for the protein. This can often be rationalised by how the linker folds against the peptide, or interacts with the protein when bound to the target. Such factors are often difficult to anticipate and much trial and error is generally required to uncover a viable lead compound.

Side-chain constraints

The disulphide bond is nature's method most common method to constrain two amino-acid side-chains and stabilise peptide structure and is observed in a number of cysteine rich peptides such as insulin and cyclotides. Appropriately positioned disulphide bonds have also been introduced into short peptides in order to stabilise helical geometry in peptides.[62] However, such 'linkers' are vulnerable to reducing conditions *in vivo* and as such are not ideal for therapeutic application where a stable secondary structure is imperative. Another approach, inspired by nature, is the lactam bridge where a lysine and aspartic or glutamic acid side-chain are connected via an amide bond (Figure 6).[63-66] This approach is particularly effective at stabilising α-helical geometries, but the directionality of the lactam (i.e. Lys/Asp or Asp/Lys) has been shown to impact on the resulting peptidomimetic's properties.

A variety of non-natural amino-acids have been developed to introduce other linker types which are generally orthogonal to routine solid-phase peptide synthesis conditions (Figure 6). 1,2,3-triazole linkers are introduced by Huisgen 1,3-cycloaddition of propargylglycine and azido-lysine, or other amino-acid analogues such as propargylserine and azidoornithine.[67-71] All hydrocarbon linkers can be installed with the

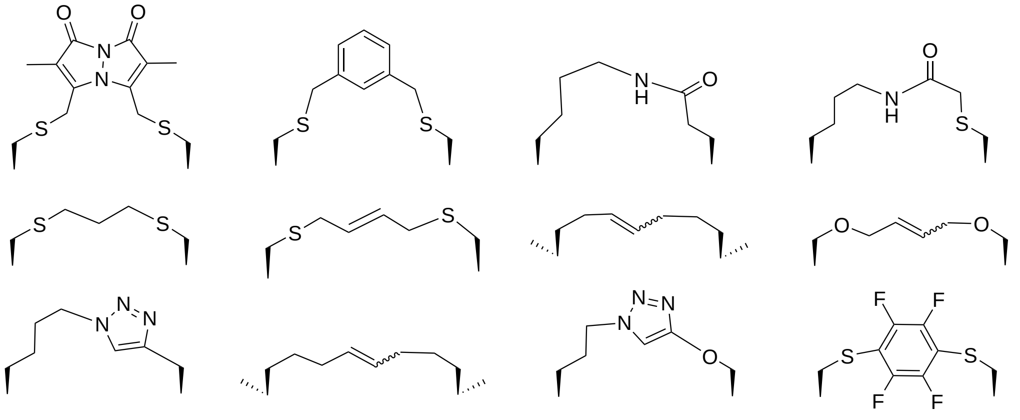


Figure 6: Representative examples of different covalent linkers to stabilise secondary structure in peptides.

now commercially available Fmoc-protected 5-pentenylalanine building blocks and linked under metathesis conditions.[6, 39, 43, 44, 48, 72] The resulting linker can be one of three variants - the unsaturated E or Z isomer, or further reduced to the fully saturated hydrocarbon. Each of these linkers will have a different impact on the properties on the peptidomimetic, in terms of structure, permeability and target binding affinity. Such all-hydrocarbon constrained peptidomimetics are reported to be particularly cell permeable, owing to the hydrophobic patch introduced by the linker, and include a number of analogues that are currently in clinical trials[32, 57] such as MDMX/MDM2 inhibitor, ALRN-6924.[73, 74] A thioether linker is similarly hydrophobic but can be oxidised to modulate the peptidomimetic's properties. All of the methodologies discussed thus far can all be classed as 'one component' constraint systems, where both linker halves are inherently part of the peptide, often comprised of unusual amino-acids, and in many cases result in asymmetric linkers. Two component constraint systems, in contrast, more commonly give rise to a symmetric linker. One methodology, that does not demand unnatural amino-acids or orthogonal protecting groups is to utilise dithiol *bis*-alkylation chemistry, where two appropriately positioned cysteines are reacted with a di-halo species to install a peptide constraint.[75, 76] Many reagents have been utilised to introduced dithiol *bis*-alkylation linkers including dibromotetrafluorobenzene, dibromoxylene, dichloroacetone and diiodoacetamide azobenzene. This approach has the added benefit that many macrocyclic peptidomimetics can be generated from the one parent peptide. Here, a new peptide linker will be introduced that is incorporated into cysteine-containing peptides by dithiol *bis*-alkylation. Dibromobimane, upon reaction with a peptide containing two cysteines, gives rise to an inherently fluorescent macrocycle.

1.2.3 Determining secondary structure of peptidomimetics

Structure plays a critical role in dictating protein interactions and many techniques have been developed to elucidate the secondary structure of peptides and proteins. Solution-based methods generally indicate the conformational average of the population, where a flexible peptide may sample a variety of conformers. The dynamic nature of these conformers can sometimes be deconvoluted from such data sets. Many peptides undergo a conformational change upon binding the protein target, and as such it is important to study the conformation of the peptide alone, and bound to the protein. Techniques such as X-ray crystallography provide the conformation of a peptide, or peptide-bound to a protein, in the solid-state. This provides a snap-shot of one, generally stable, conformation of the peptide or peptidomimetic. Consequently, there may be discrepancies between the solution-state determined structure and solid-state structure, but this may highlight important conformational transitions the peptide or peptidomimetic undergoes. Techniques that are utilised in this thesis to define the secondary structure of peptides and peptidomimetics are described, in brief, below.

Nuclear magnetic resonance spectroscopy

Nuclear magnetic resonance (NMR) provides a number of avenues to determine structure of peptides and proteins. Different secondary structures give rise to distinct dihedral angles between the α-proton (Hα) and amide-proton (NH) of an amino-acid, and unique hydrogen bonding patterns. This in turn gives rise to a distinct chemical environment for these protons. This can be indicated by long range NOEs that arise from distant parts of the peptide sequence and variable temperature NMR can indicate the presence of hydrogen bonds present in a peptide, where a shift of <-0.4 ppb/K is considered indicative of the amide participating in a hydrogen bond. The $J_{\mathrm{H\alpha NH}}$ coupling constant and chemical shift of the main-chain resonances can also be diagnostic for a particular secondary structure. For example, an α-helical structure results in the amide- and α-proton of neighbouring amino-acids to be positioned close in space, and consequently gives rise to a low $J_{\mathrm{H\alpha NH}}$ (< 6 Hz). In contrast, an extended peptide structure, such as in a β-strand, results in a larger dihedral angle and a high $J_{\mathrm{H\alpha NH}}$ (> 8 Hz). This is described by the Karplus equation, shown below:

$$^{3}J_{\mathrm{H\alpha NH}} = 6.4\cos^{2}\theta - 1.4\cos\theta + 1.9$$

where $\theta = |\phi - 60°|$, *and* ϕ *is the HαNH dihedral angle*

A list of proton and carbon chemical shifts for a type of resonance (Hα, NH etc.), that belong to a particular amino-acid in a random coil or unstructured environment has been developed.[77-80] It has been shown, that relative to these 'random coil' chemical shift values, the Hα and NH resonances will be shifted upfield in an α-helical conformation due to the deshielding environment of the structure, or downfield for an extended β-strand like conformation. Conversely, the alpha-carbon (Cα) and carbonyl-carbon (CO) resonances will be shifted downfield in an α-helical structure and upfield in a β-strand motif. The difference between the expected and experimentally observed values can be calculated to give the 'secondary shift' for each resonance and the pattern of these secondary shifts for a peptide sequence can be assessed to reveal the secondary structure of the peptide. A secondary shift threshold has been set to deem whether a secondary shift is significant (0.1 ppm for proton or 0.7 ppm for carbon resonances); three consecutive secondary shifts that surpass this threshold are necessary to confidently conclude the presence of a secondary structure motif. If the secondary shifts are close to 0, then it is deemed that there is no defined secondary structure.

Circular dichroism

The far UV region of a circular dichroism (CD) spectrum is often used to determine the secondary structure components of protein, and peptides.[81, 82] Additional electronic transitions are possible in peptides with a defined secondary structure, relative to a highly flexible peptide. The characteristic hydrogen bonding patterns of particular secondary structural motifs, and the corresponding electronic transitions, give rise to distinct CD spectra. For example, α-helical motifs have a CD maximum at 190 nm, and two minima and 208 and 222 nm with a minima ratio of 0.9 (Figure 7, purple). 3_{10}-helical structures have the same pattern of peaks however the ratio between the 208 and 222 nm minima is 0.6, where the 208 nm minimum is deeper.[83] These values were determined from proteins of known structure, and it has since been shown that in short peptides the higher wavelength minimum often exists closer to 215 nm than 222 nm.[83]

Extended β-sheets display one distinct minimum at 230 nm. The CD spectrum of flexible peptides or proteins, with no defined secondary structure, displays a single deep minimum below 190 nm (Figure 7, yellow). A CD spectrum represents the average conformation of the species present. Hence, for peptides that are somewhat flexible and sample a variety of conformers, or contain multiple different secondary structure elements, the CD spectrum reflects this and for example, may combine the elements of both a random coil and α-helical structure (Figure 7, green).

X-ray crystallography

co-crystal structures, of a peptide bound to a protein, determined using X-ray crystallography provides invaluable information about the atomic interactions between the peptide and protein, or intramolecular interactions that stabilise the binding conformation of the peptide. In contrast to the other methods described here which are all solution-based structure information, X-ray crystallography captures a single solid-state conformation of the peptide upon interacting with its partner. Co-crystal structures are often difficult to obtain and requires high purity of both components to provide invaluable atomic resolution of interactions between the peptide and protein. The resolution of the resulting electron density diagram can, in some cases, provide information about stabilised vs. flexible regions of the peptide bound to the protein. X-ray crystal structures can be analysed in detail through a variety of servers that determine buried surface area peptide amino-acids on a protein surface, hydrophobic and hydrogen bond networks,[84, 85] and

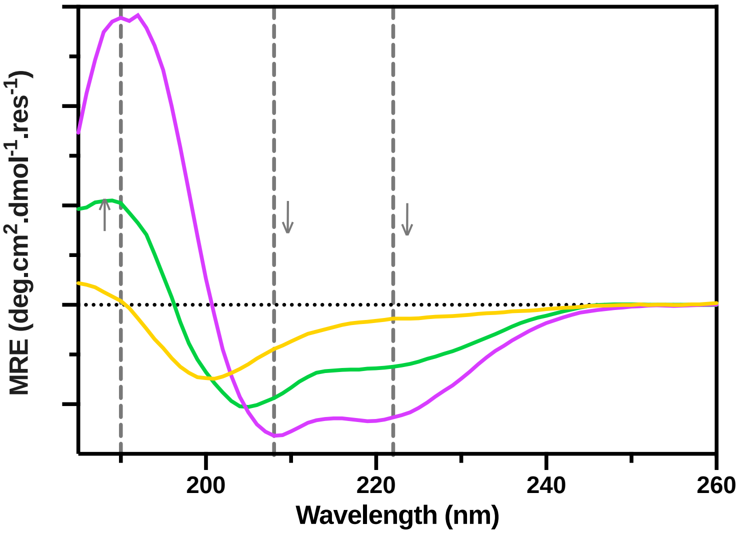


Figure 7: General features of an α-helical circular dichroism spectrum (purple), random coil (yellow) and a combination of α-helical and random coil conformers (green). Arrows represent the minima and maxima expected for a classic α-helical spectrum. MRE, molar residue ellipticity.

secondary structure characterisation.[86-88] Such structures can also be visualised[89, 90] assessed to determine the optimal position to incorporate chemical modifications on protein-bound peptides, such as covalent constraints to stabilise the bioactive binding conformation. Co-crystal structures also provide an important starting point to develop computational models, to investigate the binding conformation of similar analogues and predict the goodness of interaction, or understand differences in binding affinity where co-crystal structures of related analogues cannot be obtained.

1.3 Assessing biological function of peptides and peptidomimetics

The success of a peptide or peptidomimetic, to modulate a protein interaction, is dependent on a number of factors. The ability of such a molecule to interact with the protein of interest is characterised by a binding affinity, that is influenced by the structure-based information discussed above. However, the ability of a peptide or peptidomimetic to affect a biological function also requires it to localise at the appropriate site in a cell or tissue, which may involve transitioning through the cell membrane, and in the cases discussed here, requires further translocation into the nucleus. The half-life of the peptidomimetic *in vivo* should be long enough for the peptide to persist and affect its desired function. Additional considerations include testing the activity of the peptidomimetic *in vitro* and *in vivo* to ensure that the appropriate target is inhibited and that this inhibition gives rise to the hypothesised therapeutic activity; as well as toxicity of the peptidomimetic itself and its metabolites, followed by an assessment for appropriate dose.

1.3.1 Binding assays

A number of techniques have been developed to determine the strength of a protein interaction which include fluorescence polarisation (FP), isothermal titration calorimetry (ITC), NMR titration, microscale thermophoresis (MST) and surface plasmon resonance (SPR) assays. Here, SPR is used as the primary binding assay technique used due to the ability to screen ligands in high-throughput and minimal consumption of protein.

SPR requires immobilisation of a protein onto a gold sensor chip coated with dextran, where 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide/*N*-hydroxy succinimide (EDC/NHS) mediated coupling is a common method to covalently link protein lysines to the carboxylated dextran coating. A microfluidics system then flows samples with increasing concentration of the ligand (here, peptide) over the protein surface, termed the contact or association phase. Buffer is passed over the chip between each ligand sample, to observe dissociation of the ligand from the protein, termed the dissociation phase (Figure 8B). Binding is detected by a change in the reflected angle of a laser beam, directed at the bottom of the gold chip, caused by binding events that alter the mass and electron corona of the biological analytes on the chip surface (Figure 8A). This change in angle of the reflected beam is converted to a ‘response’ of the protein. Equilibrium binding is reached during the contact phase, when the association and dissociation of the ligand are equal, which results in a plateau of the response that is termed the steady-state response. The steady state response at each concentration of ligand is then plotted against concentration of the sample. This curve is then fitted to generate a binding affinity (K_D) value. A ligand that specifically binds a

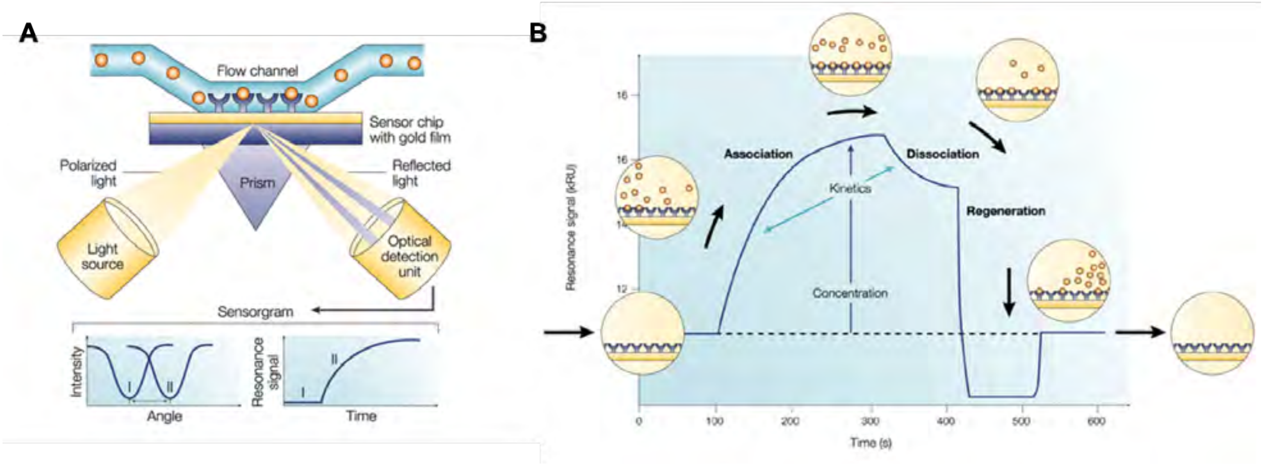

A
Flow channel
Sensor chip with gold film
Polarized light
Reflected light
Prism
Light source
Optical detection unit
Sensorgram
Intensity
I
II
Angle
Resonance signal
Time
B
Resonance signal (kRU)
18
16
14
12
Association
Dissociation
Kinetics
Concentration
Regeneration
0
100
200
300
400
500
600
Time (s)

Figure 8: Surface Plasmon Resonance. **A** Cartoon schematic of the chip and detection method of protein binding. **B** A generalised SPR sensorgram and the corresponding binding or dissociation phase. Adapted from Cooper 2002.[91]

protein will, upon increasing concentration of the ligand, saturate all binding sites and reach a maximum response related to the amount of protein on the chip. Non-specific binding, that is, binding to an undesired or random site on the protein, is often indicated when steady state binding cannot be reached, even with an increase in contact time. Further, the response only increases (exponentially) with increasing concentration of the ligand. This can be further confirmed when the maximum response, for a particular amount of protein on a chip, is exceeded. The maximum binding response can be determined by the following equation:

$$R_{max} = \frac{R_{protein} * Valency_{protein} * M_{peptide}}{M_{protein}}$$

where, R is the response in response units (RU) and M is the molecular weight in a.u.

Lastly, with high quality data, kinetic information about the binding pharmacokinetics can be derived. That is, the association rate (k_a), and dissociation rate (k_d), of a ligand to the protein can be calculated. These values indicate the speed a ligand binds or unbinds from a protein, which reflects the residency time of a ligand on a protein. This is particularly important for designing therapeutic drugs where a slow dissociation constant is generally considered desirable to affect function. Additionally, these rates can highlight key differences in the behaviour of ligands, independent of the protein binding affinity.

1.3.2 Cell permeability

Peptides and peptidomimetics generally do not conform to Lipinski's rule of five for oral bioavailability and often fall in the 'beyond rule of 5' (BRo5) class of compounds.[92] Factors that influence the passive permeability of peptides through the cell membrane are: overall charge, amphiphilicity, secondary structure and hydrogen bond donors/acceptors.[60, 92-98] Greater positive charge can often increase cell permeability, as is observed for poly-Arg or poly-Lys peptides and epitomised by cell permeable peptide TAT, where the positive-charge character is shown to destabilise the membrane to allow translocation to the cytoplasm.[99-101] A stabilised secondary structure, particularly helical structures, have been shown to enhance cell uptake. However, structures that are too rigid (e.g. high degree of helicity) have decreased cell uptake in some studies.[60] Cell uptake is further enhanced for scaffolds that are amphipathic in nature, such as a helix that possesses a positively-charged face and a hydrophobic face.[52, 60] This has been demonstrated by a number of studies on synthetic peptides, and is exemplified by the naturally derived α-helical peptide from Penetratin (residues 43-58), or its synthetic mimic - the R6W3 peptide. Lastly, reducing the number of hydrogen bond donors and in turn, increasing the number of hydrogen bond acceptors, has been shown to increase cell uptake.[95]

Cell uptake imaging assays

Investigation of a biologically relevant molecule either *in vitro* or *in vivo*, generally requires attachment of an auxiliary tag such as a fluorophore, or incorporation of labels such as radio-isotopes, to allow detection in a biological setting. Incorporation of a radio-label requires introducing a building-block containing a pre-installed radio-isotope into the synthesis, and thus may be an undesirable route as it requires resynthesis of an existing compound, or may not be synthetically viable if yields are poor due to the cost of some radio-isotopes. Additionally, detection of radio-isotopes requires specialist equipment which may not be readily accessible. Introduction of a fluorophore however, is a very common approach to enable tracking of a molecule through a biological system as a fluorophore can often be introduced post-synthetically with generally cheap reagents, and can be detected using routine experiments such as fluorescence microscopy. However, introducing a fluorophore to a scaffold, with an optimised binding mode and potency, can have dramatic and often unpredictable effects on the cell permeability, structure and activity, particularly for smaller molecular weight compounds.[102-105]

Considerations of introducing auxiliary fluorescent tags

Attachment of a fluorophore to a peptide, or peptidomimetic, can drastically alter cell uptake. Attaching a fluorophore, generally through an amine, will alter the overall peptide charge and hydrophobicity, and the number of hydrogen bond acceptors and donors. The fluorophore may change the conformation of the peptide and the solvent exposed groups, and simultaneously impact the secondary structure that has been optimised to adopt a conformation favourable to binding. The location of fluorophore attachment may also interfere with the previously optimised protein-binding interaction. The auxiliary fluorophore may dramatically alter the weight, particularly for small molecular weight compounds, and influence the overall biophysical properties of the compound of interest. Together these factors can, and have been shown to, change the cell uptake and localisation of the peptide or peptidomimetic – adversely or favourably.[102, 103] However, these facts are often overlooked, due to the simplicity of fluorophore introduction and analysis; but it is imperative to note that the actual cell uptake, localisation and intracellular activity of the therapeutically useful compound of interest, that may progress for further testing, is not accurately represented by the uptake and location of a fluorescently tagged analogue. Alternative approaches need to be developed that incorporate imaging tags from the start of compound development, or label-free imaging approaches, such that cellular assays can be conducted on the therapeutically useful molecule and *in vitro* assays more accurately reflect the behaviour compound of interest. One avenue of research is to stabilise peptide secondary structure with inherently fluorescent chemical moieties.

1.3.3 Nuclear permeability

Cell permeability is only the first hurdle that needs to be surmounted for some therapeutic molecules, such as those developed here. Many important protein targets are located within intracellular compartments, such as the nucleus. Various molecules accumulate at different locations within the cell, for example, highly positively-charged small molecules are known to gather in the (negatively-charged) mitochondria.[106, 107] However, targeting of more specialised compartments such as the nucleus requires greater design – especially for larger molecules which are not amenable to passive diffusion. Nuclear uptake generally requires recognition by active transporters such as importin-α. Many viruses contain proteins with sequences that have been identified to specifically target the nucleus. These nuclear locating sequences (NLS) have been characterised and used directly, or as inspiration to design new sequences with nuclear permeability. Such NLS-tags are tethered to a cargo molecule to permit uptake into the nucleus.[108] Only once the compound of interest has gained access to the cellular locale that houses the protein target can *ex vivo* studies, in cells or tissue, be carried out to assess function.

1.3.4 Proteolytic stability

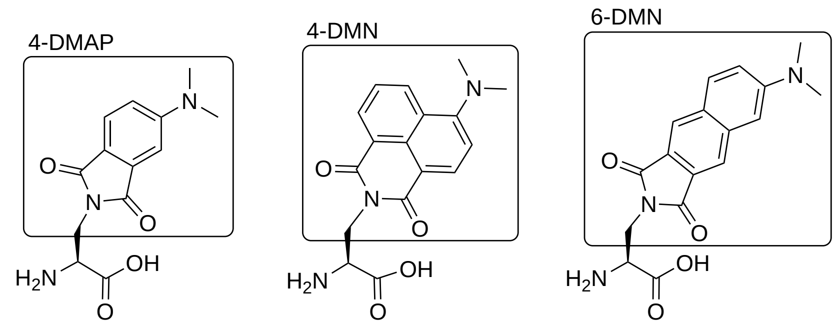
Therapeutically viable molecules must be sufficiently stable to persist within the body long enough to affect function. Peptides are generally regarded as poor drugs owing to their short half-lives *in vivo*, where native peptide sequences are readily digested by endogenous proteases. Peptidomimetics, or amino-acid sequences with reduced peptide character such as constrained peptides, often display increased half-lives. Peptidomimetics or constrained peptides owe the increased half-life to one or a combination of factors including: a defined secondary structure that may not be readily recognised by proteases (where β-strand is preferred); the covalent constraint may physically inhibit the peptidomimetic from entering a protease active site; or non-proteinogenic amino-acids that are not well recognised by proteases. Many synthetically produced peptidomimetics include *N*- and *C*-termini protecting groups, for example with an acetyl or amide respectively, to limit recognition by endopeptidases. Prior to *ex vivo* or *in vivo* assays, the compound half-life can be assessed *in vitro* against a battery of common and ubiquitous proteases which commonly include trypsin, chymotrypsin and the promiscuous proteinase K. Such assays require incubation of a peptidomimetic with a small amount of protease and monitoring the degradation over time by HPLC. The peptide fragments can be distinguished by MS and the protease cleavage sites can therefore be identified. A comparison between the linear precursor and cyclised peptidomimetic can reveal any protective effect provided by installing the covalent constraint.

1.4 Peptides as protein sensors

Peptides provide unique scaffolds to develop biosensors that selectively detect proteins, due to their small size (relative to antibodies) and selective protein surface binding (compared to small molecules). Such biosensors provide the added advantage of allowing real-time monitoring of cellular processes, compared to antibody techniques, such as histological staining, which only captures a point in time. Many methodologies have been developed to induce an increase or change in fluorescence upon a sensor interacting with a target protein, including FRET, metal-ion binding, aggregation and environmentally-sensitive fluorophores.[109, 110] These peptide-based biosensors are still subject to the aforementioned limitations of peptides as drugs, but also exhibit their benefits. In particular peptides provide a modifiable scaffold that targets proteins with a high degree of specificity. Here we investigate developing a sensor for PCNA by including a solvatochromic amino-acid in a p21 peptide. A sensor for PCNA, would allow the state of DNA-replication and cell cycle regulations, of which PCNA is central, to be monitored by real-time fluorescence microscopy.

1.4.1 Solvatochromic amino-acids

Solvatochromic amino-acids, are amino-acids with an environmentally sensitive fluorophore appended, generally to the side-chain.[111-113] Such fluorophores undergo a large change in quantum yield, upon being exposed to a particular environmental trigger which can include pH, temperature, presence of a molecule or metal ion, or change in solvent polarity.[113, 114] This change in quantum yield, when sufficiently large, provides an ‘on’ and ‘off’ state of the fluorophore. A wide range of solvatochromic fluorophores have been developed which vary in size, emission wavelength, environmental trigger and the change in quantum yield between the ‘on’ and ‘off’ state. Solvatochromic fluorophores that respond to a change in solvent polarity are particularly useful to develop protein sensors, where the fluorophore is not fluorescent in a polar solvent and highly fluorescent in a hydrophobic environment. Judicious introduction of such a fluorophore, into a peptide known to target a protein of interest, gives rises to a peptide-based sensor. Such a peptide is not fluorescent in the aqueous cell cytosol, but highly fluorescent upon binding a protein and embedding the fluorophore into a hydrophobic cavity on the protein surface, thereby ‘sensing’ the protein.[110, 112, 114, 115] A number of solvatochromic amino-acids have been developed and of particular note are those from the dimethylaminophthalimide family (Figure 9). It is imperative that inclusion of a solvatochromic amino-acidin a peptide known to target a protein of interest, does not perturb the peptide-protein binding. Additionally, the solvatochromic amino-acid must be positioned in the peptide sequence at a location the embeds the amino-acid into a hydrophobic cavity on the protein surface, in order to give fluorescence response. Structural information, such as X-ray crystal structures of the peptide bound to the protein, can significantly rise to the maximum



4-DAPA	**4-DMNA**	**6-DMNA**
ex. 390 nm	ex. 440 nm	ex. 390 nm
em (Tris) 580 nm	em (Tris) 554 nm	em (Tris) 625 nm
em (dioxane) 497 nm	em (dioxane) 512 nm	em (dioxane) 520 nm
$I_{dioxane}/I_{Tris}$ = 4500	$I_{dioxane}/I_{Tris}$ = 1200	$I_{dioxane}/I_{Tris}$ = 1400

Figure 9: Dimethylaminophthalimide family of solvatochromic amino-acids. Tris is representative of a polar solvent and dioxane representative of a hydrophobic environment. Image adapted from Loving 2008.[112]

aid the design of a peptide-based sensor. However some trial and error is often required to investigate the optimal position within a peptide sequence to include a solvatochromic amino-acid, to give a large fluorescence response upon binding the protein, and maintain high binding affinity. Additionally, a number of solvatochromic amino-acids of different sizes may need to be trialled.

1.5 Solid-phase peptide synthesis

This thesis presents a large number of peptides and modified peptides, which are all prepared by solid-phase peptide synthesis (SPPS). SPPS is the premier method for synthesising long peptide sequences, where a Fmoc/tBu protecting strategy is most commonly employed in modern SPPS.[116, 117] A solid support, generally polystyrene-based based beads, is functionalised with a linker containing a reactive group such as an amine, halogen or alcohol, to facilitate attachment of an amino-acid building block. Here, a Rink amide linker functionalised resin is used primarily, and contains a free amine that can be subsequently coupled to a carboxylic acid of an Fmoc-protected amino-acid. In Fmoc/tBu SPPS, the Nα-group is Fmoc-protected, and the amino-acid side-chain is generally protected with an acid-labile protecting group such as tBu, Boc, Trt or Pbf. Following coupling of the amino-acid, the Fmoc group is deprotected to liberate the *N*-terminal amine, and allow a subsequent amino-acid to be coupled with addition of an amide-coupling agent. This results in the peptide sequence assembly from the *C*-terminus, to the *N*-terminus. The linear peptide can then be cleaved from the resin by acidolysis, which simultaneously liberates the acid-labile side-chain protecting groups. This is commonly achieved by addition of TFA to the resin, in the presence of additives, referred to as scavengers, that ‘scavenge’ the liberated protecting group by-products. Judicious selection of low-acid or non-acid labile side-chain protecting groups, can allow for selective deprotection of key functional groups whilst the peptide remains attached to the (high-acid labile) solid-support. This enables further modification of the peptide, such as the introduction of a fluorophore or reactive handle (e.g. maleimide), macrocyclisation, or conjugation to a second peptide fragment.

This process can be automated with the use of automated peptide synthesisers; however, this is generally restricted to the assembly of linear and branched sequences, where further derivatisation must be carried out manually. Such automated synthesisers may also incorporate microwave technology that allows inclusion of a microwave heating step during both the coupling and deprotection phase. This significantly increases the reaction rate and allows assembly of a peptide (e.g. 15 amino-acids) within 2-3 hours, and often improves synthesis quality by improving reaction yield and limiting prolonged exposure to reagents that may induce side reactions (e.g. piperidine). Additionally, the microwave irradiation results in elongation or unfolding of the peptide chain to ensure the *N*-terminal group is readily accessible to reagents and thus increases the reaction efficiency. This is particularly pertinent in hydrophobic amino-acid sequences which are prone to folding back of themselves, and burying the *N*-terminal amine.

1.5.1 Common side-reactions

Judicious selection of reaction conditions is required to minimise side-reactions and improve synthesis quality and yield. These conditions include, but are not limited to, the choice of solid-support, amide coupling agent, deprotection additives, solvent, and amino-acid protecting groups. Further, these conditions must often be reoptimized for each peptide sequence. A large number of side-reactions can occur including, but not limited to, aspartimide formation, pyroglutamate formation, guanidinylation, racemisation, as well as various alkylation, oxidation or acyl transfer events.[118-120] For example, aspartic acid residues which are followed by a small amino-acid (often glycine) are often prone to aspartimide formation (Figure 10). This side-reaction is catalysed by bases such as piperidine, and as such prolonged exposure to piperidine should be minimised. Fmoc deprotection is generally carried out with 20% piperidine in DMF for 10 min. The basic nature of piperidine leads to deprotonation of main-chain amides. A deprotonated amide *i*+1 to the aspartic acid residue can undergo nucleophilic attack on the aspartic acid, or asparagine, side-chain carbonyl resulting in dehydration and formation of the cyclic aspartimide. This product can be detected in MS as a M-18 peak, where M is the mass of the desired product. However, the aspartimide can further react with either water or piperidine to ring open (Figure 10). Either an α- or β- amino-acid can be liberated upon attack of the carbonyl and ring opening, furthermore racemisation is possible as either face can be attacked. Where the aspartimide is ring-opened with water, the resulting product is of the same mass as the desired product and cannot be detected by MS, and the HPLC retention time may not be significantly altered. Work has shown that synthesis of p21 peptides (unpublished) can be prone to significant aspartimide formation, though the aspartic acid (residue 149) is followed by the relatively large phenylalanine residue (residue 150). Small amounts of aspartimide formation may be manageable if the resulting by-product is separable from the desired product, and does not significantly impact yield. However, for the synthesis of $p21_{139-160}$, aspartimide formation is significant, especially compared to shorter analogues such as $p21_{141-155}$. This is likely a combination of two main factors: less time exposed to piperidine, as there as less amino-acids to deprotect and couple following introduction of the aspartic acid residue; and the change in location of the aspartic acid relative to the solid-support thereby altering the backbone flexibility necessary for nucleophilic attack. A number of aspartic acid analogues with bulky protecting groups, such as β-3-methylpent-3-yl (Mpe), cyanosulfurylides (CSY), and backbone protecting groups 2-hydroxy-4-methoxy benzyl (Hmb), or 2,4-dimethoxybenzyl (Dmb) on the preceeding amine, that

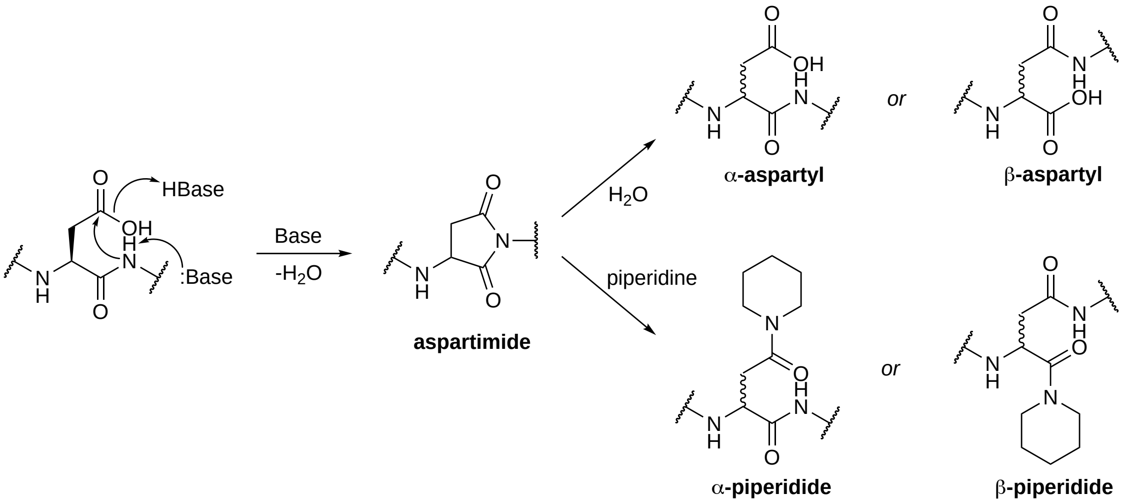


Figure 10: Generalised mechanism for aspartimide formation and subsequent racemisation or piperidide formation.

sterically inhibit aspartimide formation; or building blocks such as pseudoproline dipeptides, can be incorporated.[121] However, such building blocks are expensive which limits their use in large or numerous syntheses. Acidic modifiers, such as HOBt or formic acid, can be added to the deprotection mixture to decrease aspartimide formation; or deprotection times can be reduced to limit the exposure of the peptide to piperidine.[122] However, reducing deprotection times may lead to incomplete Fmoc deprotection, and consequently products with one or multiple amino-acid deletions may be present in the final peptide mixture. Microwave irradiation can accelerate the deprotection reaction, and therefore reduce the time necessary for the peptide to be exposed to piperidine and limit aspartimide formation while still achieving complete Fmoc deprotection.[120] Access to specialist equipment such as microwave peptide synthesisers is not always possible and we have used acidic modifiers, and sequences less prone to aspartimide formation, where possible.

1.6 Aims for this thesis and thesis outline

This thesis presents studies on using and developing peptides and peptidomimetics to study and inhibit protein-protein interactions, with a particular focus on defining secondary structure that favours binding. The thesis contains two broad aims: First, to develop methodology to introduce an inherently fluorescent modification, that comprises a bimane group and characterise the influence this group has on secondary structure. Second, to investigate the structure-activity relationship of peptides that bind the human sliding clamp protein, PCNA, that mediates DNA-replication and repair and is upregulated in cancers, in order to develop a peptidomimetic that targets PCNA (Chapters 5-10).

The first aim is addressed in Chapters 2-4. **Chapter 2** presents methodology to introduce a peptide side-chain that comprises a bimane group, by reaction of dibromobimane with short cysteine-containing peptides to generate inherently fluorescent macrocycles. The structure of the resulting macrocycles is studied by NMR and CD, to investigate the secondary structure in range of bimane-constrained macrocycle sizes. The bimane fluorescence is also used to directly image the macrocycles in cells. **Chapter 3** presents studies on bimane-modified peptides in a sequence known to target ERα. The influence of two types of bimane modification on peptide secondary structure are investigated. In the first type of modification, the bimane covalently connects two cysteine side-chains to form a macrocycle, and in the second modification the bimane is appended to a cysteine side-chain in an acyclic peptide and participates in a non-covalent interaction with an aromatic residue. The resulting secondary structure is studied by NMR and CD, and the fluorescence properties of the bimane modified peptides are characterised. Appendix 1 summarises this methodology and summarises benefits of using the bimane constraint to influence secondary structure in peptides, in the form of a focus article. The bimane peptide modification is also utilised in Chapters 6, 8 and 9, to create fluorescent macrocycles that target PCNA.

Chapters 4-9 present work to address the second thesis aim. The key features required to target PCNA are discussed in the **Chapter 4** minireview, which is built upon in **Chapter 5** with studies that probe the PCNA-binding interface with modified p21-derived peptides. The binding affinity of these peptides is determined by SPR and structure bound to PCNA studied by X-ray crystallography and computational modelling studies, in order to better understand the associated structure and hydrogen bonding networks that determine high affinity binding to PCNA. The ability of amino-acid side-chain constraints to stabilise the 3_{10}-helical binding conformation of p21 peptide bound to PCNA is then considered in **Chapter 6** with five macrocycles prepared from a p21 peptide with two cysteines by dithiol *bis*-alkylation. The peptides are studied by SPR, X-ray crystallography, and computationally modelling to reveal a correlation between the binding conformation on f the macrocycle and high affinity binding to PCNA. The most potent peptide is also studied by NMR, and imaged in cells. An extensive study in **Chapter 7** defines the minimum length of a p21 peptide necessary to maintain high affinity binding to PCNA, and investigates how the sequence that flanks the PIP-box contributes to cell permeability. Short p21 peptides are conjugated to NLS sequences in order to confer nuclear permeability. This chapter also presents a comparison of the intracellular distribution of peptides with or without an auxiliary fluorescein-tag. **Chapter 8** presents p21 peptides that are modified to include solvatochromic amino-acids, in order to develop a peptide capable of producing a fluorescent response on interacting with PCNA, thereby acting as a sensor. The fluorescent properties are characterised, and the interaction of the modified peptides with PCNA is determined by SPR. Lastly, **Chapter 9** summarises the findings of this thesis and highlights the next steps necessary to develop a peptidomimetic that targets PCNA, *en route* to developing a cancer therapeutic for preclinical trial. Additionally, Chapter 9 presents initial work toward combining the p21 modifications that were identified to increase affinity for PCNA (Chapter 5) into a peptidomimetic with a defined secondary structure (Chapter 6).

REFERENCES

1. D. E. Scott, M. T. Ehebauer, T. Pukala, M. Marsh, T. L. Blundell, A. R. Venkitaraman, C. Abell and M. Hyvonen, Using a fragment-based approach to target protein-protein interactions, *ChemBioChem*, 2013, **14**, 332-342.
2. N. Tsomaia, Peptide therapeutics: Targeting the undruggable space, *Eur. J. Med. Chem.*, 2015, **94**, 459-470.
3. A. M. Watkins and P. S. Arora, Structure-based inhibition of protein-protein interactions, *Eur. J. Med. Chem.*, 2015, **94**, 480-488.
4. A. M. Leduc, J. O. Trent, J. L. Wittliff, K. S. Bramlett, S. L. Briggs, N. Y. Chirgadze, Y. Wang, T. P. Burris and A. F. Spatola, Helix-stabilized cyclic peptides as selective inhibitors of steroid receptor-coactivator interactions, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 11273-11278.
5. K. W. Nettles, J. B. Bruning, G. Gil, J. Nowak, S. K. Sharma, J. B. Hahm, K. Kulp, R. B. Hochberg, H. Zhou, J. A. Katzenellenbogen, B. S. Katzenellenbogen, Y. Kim, A. Joachmiak and G. L. Greene, NFkappaB selectivity of estrogen receptor ligands revealed by comparative crystallographic analyses, *Nat. Chem. Biol.*, 2008, **4**, 241-247.
6. T. E. Speltz, S. W. Fanning, C. G. Mayne, C. Fowler, E. Tajkhorshid, G. L. Greene and T. W. Moore, Stapled Peptides with gamma-Methylated Hydrocarbon Chains for the Estrogen Receptor/Coactivator Interaction, *Angew. Chem. Int. Ed. (English)*, 2016, **55**, 4252-4255.
7. G. L. Moldovan, B. Pfander and S. Jentsch, PCNA, the maestro of the replication fork, *Cell*, 2007, **129**, 665-679.
8. I. Stoimenov and T. Helleday, PCNA on the crossroad of cancer, *Biochem. Soc. Trans.*, 2009, **37**, 605-613.
9. W. Strzalka and A. Ziemienowicz, Proliferating cell nuclear antigen (PCNA): a key factor in DNA replication and cell cycle regulation, *Annals of Botany*, 2011, **107**, 1127-1140.
10. A. De Biasio and F. J. Blanco, Proliferating cell nuclear antigen structure and interactions: too many partners for one dancer?, *Adv. Prot. Chem. Struct. Bio.*, 2013, **91**, 1-36.
11. E. M. Boehm, M. S. Gildenberg and M. T. Washington, The Many Roles of PCNA in Eukaryotic DNA Replication, *Enzymes*, 2016, **39**, 231-254.
12. J. M. Gulbis, Z. Kelman, J. Hurwitz, M. O'Donnell and J. Kuriyan, Structure of the C-Terminal Region of p21 WAF1/CIP1 Complexed with Human PCNA, *Cell*, 1996, **87**, 297-306.
13. L. M. Dieckman, B. D. Freudenthal and M. T. Washington, PCNA structure and function: insights from structures of PCNA complexes and post-translationally modified PCNA, *Subcell. Biochem.*, 2012, **62**, 281-299.
14. F. Uhlmann, J. Cai, E. Gibbs, M. O'Donnell and J. Hurwitz, Deletion analysis of the large subunit p140 in human replication factor C reveals regions required for complex formation and replication activities, *J. Biol. Chem.*, 1997, **272**, 10058-10064.
15. J. Majka and P. M. J. Burgers, The PCNA–RFC Families of DNA Clamps and Clamp Loaders, *Prog. Nucleic Acid Res. Mol. Biol.*, 2004, **78**, 227-260.
16. O. Chilkova, P. Stenlund, I. Isoz, C. M. Stith, P. Grabowski, E. B. Lundström, P. M. Burgers and E. Johansson, The eukaryotic leading and lagging strand DNA polymerases are loaded onto primer-ends via separate mechanisms but have comparable processivity in the presence of PCNA, *Nucleic Acids Res.*, 2007, **35**, 6588-6597.
17. O. Parnas, A. Zipin-Roitman, B. Pfander, B. Liefshitz, Y. Mazor, S. Ben-Aroya, S. Jentsch and M. Kupiec, Elg1, an alternative subunit of the RFC clamp loader, preferentially interacts with SUMOylated PCNA, *EMBO J.*, 2010, **29**, 2611-2622.
18. M. De March, N. Merino, S. Barrera-Vilarmau, R. Crehuet, S. Onesti, F. J. Blanco and A. De Biasio, Structural basis of human PCNA sliding on DNA, *Nat. Comm.*, 2017, **8**, 13935.
19. M. De March, S. Barrera-Vilarmau, E. Crespan, E. Mentegari, N. Merino, A. Gonzalez-Magana, M. Romano-Moreno, G. Maga, R. Crehuet, S. Onesti, F. J. Blanco and A. De Biasio, p15PAF binding to PCNA modulates the DNA sliding surface, *Nucleic Acids Res.*, 2018, **46**, 9816–9828.
20. G. Maga and U. Hubscher, Proliferating cell nuclear antigen (PCNA): A dancer with many partners, *J. Cell Sci.*, 2003, **116**, 3051-3060.
21. A. M. Pedley, M. A. Lill and V. J. Davisson, Flexibility of PCNA-protein interface accommodates differential binding partners, *PLoS One*, 2014, **9**, e102481.
22. H. Li, M. Sandhu, L. H. Malkas, R. J. Hickey and N. Vaidehi, How Does the Proliferating Cell Nuclear Antigen Modulate Binding Specificity to Multiple Partner Proteins?, *J. Chem. Inf. Model.*, 2017, **57**, 3011-3021.
23. A. Gonzalez-Magana and F. J. Blanco, Human PCNA Structure, Function and Interactions, *Biomolecules*, 2020, **10**, 570.
24. S. N. Naryzhny and H. Lee, Characterization of proliferating cell nuclear antigen (PCNA) isoforms in normal and cancer cells: There is no cancer-associated form of PCNA, *FEBS Lett.*, 2007, **581**, 4917-4920.
25. S. Y. Park, M. S. Jeong, C. W. Han, H. S. Yu and S. B. Jang, Structural and Functional Insight into Proliferating Cell Nuclear Antigen, *J. Microbiol. Biotechnol.*, 2016, **26**, 637-647.
26. L. Dieckman, Something's gotta give: How PCNA alters its structure in response to mutations and the implications on cellular processes, *Prog. Biophys. Mol. Biol.*, 2020, DOI: 10.1016/j.pbiomolbio.2020.10.008.
27. T. N. U. o. Singapore, How is DNA replicated?, https://www.mechanobio.info/genome-regulation/how-is-dna-replicated/).
28. J. B. Bruning and Y. Shamoo, Structural and thermodynamic analysis of human PCNA with peptides derived from DNA polymerase-delta p66 subunit and flap endonuclease-1, *Structure*, 2004, **12**, 2209-2219.

29. K. Hara, M. Uchida, R. Tagata, H. Yokoyama, Y. Ishikawa, A. Hishiki and H. Hashimoto, Structure of proliferating cell nuclear antigen (PCNA) bound to an APIM peptide reveals the universality of PCNA interaction, *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.*, 2018, **74**, 214-221.

30. E. Warbrick, PCNA binding through a conserved motif, *BioEssays*, 1998, **20**, 195-199.

31. K. Estieu-Gionnet and G. Guichard, Stabilized helical peptides: overview of the technologies and therapeutic promises, *Expert Opin. Drug Dicov.*, 2011, **6**, 937-936.

32. K. Fosgerau and T. Hoffmann, Peptide therapeutics: current status and future directions, *Drug Discov. Today*, 2015, **20**, 122-128.

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

34. S. Waga, G. J. Hannon, D. Breach and B. Stillman, The p21 inhibitor of cyclin-dependent kinases controls DNA replication by interaction with PCNA, *Nature*, 1994, **369**, 574-578.

35. Z.-Q. Pan, J. T. Reardon, L. Li, H. Flores-Rozas, R. Legerski, A. Sancar and J. Hurwitz, Inhibition of Nucleotide Excision Repair by the Cyclin-dependent Kinase Inhibitor p21, *J. Biol. Chem.*, 1995, **37**, 22008-22016.

36. Y. Luo, J. Hurwitz and J. Massague, Cell-cycle inhibition by independent CDK and PCNA binding domains in p21Cip1, *Nature*, 1995, **375**, 159-161.

37. F. Goubin and B. Ducommun, Identification of binding domains on the p21Cip1 cyclin-dependent kinase inhibitor, *Oncogene*, 1995, **10**, 2281-2287.

38. D. I. Zheleva, N. Z. Zhelev, P. M. Fischer, S. V. Duff, E. Warbrick, D. G. Blake and D. P. Lane, A Quantitative Study of the in Vitro Binding of the C-Terminal Domain of p21 to PCNA: Affinity, Stoichiometry, and Thermodynamics, *Biochemistry*, 2000, **39**, 7388-7397.

39. G. H. Bird, N. Madani, A. F. Perry, A. M. Princiotto, J. G. Supko, X. He, E. Gavathiotis, J. G. Sodroski and L. D. Walensky, Hydrocarbon double-stapling remedies the proteolytic instability of a lengthy peptide therapeutic, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 14093-14098.

40. T. A. Hill, N. E. Shepherd, F. Diness and D. P. Fairlie, Constraining cyclic peptides to mimic protein structure motifs, *Angew. Chem. Int. Ed. (English)*, 2014, **53**, 13020-13041.

41. S. Roy, P. Ghosh, I. Ahmed, M. Chakraborty, G. Naiya and B. Ghosh, Constrained alpha-Helical Peptides as Inhibitors of Protein-Protein and Protein-DNA Interactions, *Biomedicines*, 2018, **6**.

42. C. Morrison, Constrained peptides' time to shine?, *Nat. Rev. Drug Discovery*, 2018, **17**, 531-533.

43. L. K. Henchey, A. L. Jochim and P. S. Arora, Contemporary strategies for the stabilization of peptides in the alpha-helical conformation, *Curr. Opin. Chem. Biol.*, 2008, **12**, 692-697.

44. Y. W. Kim, T. N. Grossmann and G. L. Verdine, Synthesis of all-hydrocarbon stapled alpha-helical peptides by ring-closing olefin metathesis, *Nat. Protoc.*, 2011, **6**, 761-771.

45. A. P. Higueruelo, H. Jubb and T. L. Blundell, Protein-protein interactions as druggable targets: recent technological advances, *Curr. Opin. Pharmacol.*, 2013, **13**, 791-796.

46. A. Jamieson and N. Robertson, Regulation of protein-protein interactions using stapled peptides, *Rep. Org. Chem*, 2015, **5**, 65-74.

47. A. Muppidi, K. Doi, S. Edwardraja, E. J. Drake, A. M. Gulick, H. G. Wang and Q. Lin, Rational design of proteolytically stable, cell-permeable peptide-based selective Mcl-1 inhibitors, *J. Am. Chem. Soc.*, 2012, **134**, 14734-14737.

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

49. A. Muppidi, Z. Wang, X. Li, J. Chen and Q. Lin, Achieving cell penetration with distance-matching cysteine cross-linkers: a facile route to cell-permeable peptide dual inhibitors of Mdm2/Mdmx, *Chem. Commun.*, 2011, **47**, 9396-9368.

50. M. M. Madden, A. Muppidi, Z. Li, X. Li, J. Chen and Q. Lin, Synthesis of cell-permeable stapled peptide dual inhibitors of the p53-Mdm2/Mdmx interactions via photoinduced cycloaddition, *Bioorg. Med. Chem. Lett.*, 2011, **21**, 1472-1475.

51. Y. Tian, Y. Jiang, J. Li, D. Wang, H. Zhao and Z. Li, Effect of Stapling Architecture on Physiochemical Properties and Cell Permeability of Stapled alpha-Helical Peptides: a Comparative Study, *ChemBioChem*, 2017, **18**, 2087-2093.

52. S. R. Perry, T. A. Hill, A. D. de Araujo, H. N. Hoang and D. P. Fairlie, Contiguous hydrophobic and charged surface patches in short helix-constrained peptides drive cell permeability, *Org. Biomol. Chem.*, 2018, **16**, 367-371.

53. A. D. Abell, N. A. Alexander, S. G. Aitken, H. Chen, J. M. Coxon, M. A. Jones, S. B. McNabb and A. Muscroft-Taylor, Synthesis of Macrocyclic Beta-Strand Templates by Ring Closing Metathesis, *J. Org. Chem.*, 2009, **74**, 4354-4356.

54. K. C. Chua, M. Pietsch, X. Zhang, S. Hautmann, H. Y. Chan, J. B. Bruning, M. Gutschow and A. D. Abell, Macrocyclic protease inhibitors with reduced peptide character, *Angew. Chem. Int. Ed. (English)*, 2014, **53**, 7828-7831.

55. L. Pauling, R. B. Corey and H. R. Branson, The Structure of Proteins: Two Hydrogen-bonded Helical Configurations of the Polypeptide Chain, *Proceedings of the National Academy of Sciences*, 1951, **37**, 205-211.

56. C. G. Cummings and A. D. Hamilton, Disrupting protein-protein interactions with non-peptidic, small molecule alpha-helix mimetics, *Curr. Opin. Chem. Biol.*, 2010, **14**, 341-346.

57. M. Klein, Stabilized helical peptides: overview of the technologies and its impact on drug discovery, *Expert Opin. Drug Dicov.*, 2017, **12**, 1117-1125.

58. P. Y. Maximov, B. Abderrahman, S. W. Fanning, S. Sengupta, P. Fan, R. F. Curpan, D. M. Q. Rincon, J. A. Greenland, S. S. Rajan, G. L. Greene and V. C. Jordan, Endoxifen, 4-Hydroxytamoxifen and an Estrogenic Derivative Modulate Estrogen Receptor Complex Mediated Apoptosis in Breast Cancer, *Mol. Pharmacol.*, 2018, **94**, 812-822.

59. K. L. Keeling, O. Cho, D. B. Scanlon, G. W. Booker, A. D. Abell and K. L. Wegener, The key position: influence of staple location on constrained peptide conformation and binding, *Org. Biomol. Chem.*, 2016, **14**, 9731-9735.

60. G. H. Bird, E. Mazzola, K. Opoku-Nsiah, M. A. Lammert, M. Godes, D. S. Neuberg and L. D. Walensky, Biophysical determinants for cellular uptake of hydrocarbon-stapled peptide helices, *Nat. Chem. Biol.*, 2016, **12**, 845-852.

61. K. L. Wegener, A. E. McGrath, N. E. Dixon, A. J. Oakley, D. B. Scanlon, A. D. Abell and J. Bruning, Rational design of a 310-helical PIP-box mimetic targeting PCNA - the human sliding clamp, *Chem. Eur. J.*, 2018, **24**, 11325-11331.

62. Maria Pellegrini, Miriam Royo, Michael Chorev and D. F. Mierke, Conformation consequnces for i, i+3 cystine linkages: nucleation for alpha-helicity?, *J. Pept. Res.*, 1997, **49**, 404-414.

63. P. Luo, D. T. Braddock, R. M. Subramanian, S. C. Meredith and D. G. Lynn, Structural and Thermodynamic Characterization of a Bioactive Peptide Model of Apolipoprotein E: Side-Chain Lactam Bridges To Constrain the Conformation, *Biochemistry*, 1994, **33**, 12367-12377.

64. M. E. H. Jr, C. L. Gannon, C. M. Kay and R. S. Hodges, Lactam Bride Stabilization of Alpha-Helical Peptides: Ring Size, Orientation and Positional Effects, *J. Pept. Sci.*, 1995, **1**, 274-282.

65. J. W. Taylor, The Synthesis and Study of Side-Chain Lactam-Bridged Peptides, *Pep. Sci.*, 2002, **66**, 49-75.

66. H. N. Hoang, R. W. Driver, R. L. Beyer, T. A. Hill, A. D. de Araujo, F. Plisson, R. S. Harrison, L. Goedecke, N. E. Shepherd and D. P. Fairlie, Helix Nucleation by the Smallest Known alpha-Helix in Water, *Angew. Chem. Int. Ed. (English)*, 2016, **55**, 8275-8279.

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

68. A. D. Pehere, M. Pietsch, M. Gutschow, P. M. Neilsen, D. S. Pedersen, S. Nguyen, O. Zvarec, M. J. Sykes, D. F. Callen and A. D. Abell, Synthesis and extended activity of triazole-containing macrocyclic protease inhibitors, *Chem. Eur. J.*, 2013, **19**, 7975-7981.

69. Y. H. Lau, Y. Wu, P. de Andrade, W. R. Galloway and D. R. Spring, A two-component 'double-click' approach to peptide stapling, *Nat. Protoc.*, 2015, **10**, 585-594.

70. B. B. Metaferia, M. Rittler, J. S. Gheeya, A. Lee, H. Hempel, A. Plaza, W. G. Stetler-Stevenson, C. A. Bewley and J. Khan, Synthesis of novel cyclic NGR/RGD peptide analogs via on resin click chemistry, *Bioorg. Med. Chem. Lett.*, 2010, **20**, 7337-7340.

71. L. Zhang, T. Navaratna and G. M. Thurber, A Helix-Stabilizing Linker Improves Subcutaneous Bioavailability of a Helical Peptide Independent of Linker Lipophilicity, *Bioconjug Chem*, 2016, **27**, 1663-1672.

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

73. G. Philippe, Y. H. Huang, O. Cheneval, N. Lawrence, Z. Zhang, D. P. Fairlie, D. J. Craik, A. D. de Araujo and S. T. Henriques, Development of cell-penetrating peptide-based drug leads to inhibit MDMX:p53 and MDM2:p53 interactions, *Biopolymers*, 2016, **106**, 853-863.

74. L. A. Carvajal, D. B. Neriah, A. Senecal, L. Benard, V. Thiruthuvanathan, T. Yatsenko, S. R. Narayanagari, J. C. Wheat, T. I. Todorova, K. Mitchell, C. Kenworthy, V. Guerlavais, D. A. Annis, B. Bartholdy, B. Will, J. D. Anampa, I. Mantzaris, M. Aivado, R. H. Singer, R. A. Coleman, A. Verma and U. Steidl, Dual inhibition of MDMX and MDM2 as a therapeutic strategy in leukemia, *Sci Transl Med*, 2018, **10**.

75. D. P. Fairlie and A. D. de Araujo, Stapling peptides using cysteine crosslinking, *Biopolymers*, 2016, **106**, 843-852.

76. L. Peraro, T. R. Siegert and J. A. Kritzer, Conformational Restriction of Peptides Using Dithiol Bis-Alkylation, *Methods Enzymol.*, 2016, **580**, 303-332.

77. M. P. Williamson, Secondary-Structure Dependent Chemical Shifts in Proteins, *Biopolymers*, 1990, **29**, 1423-1431.

78. D. S. Wishart and B. D. Sykes, Chemical Shifts as a Tool for Structure Determination, *Methods Enzymol.*, 1994, **239**, 363-392.

79. T. Asakura, K. Taoka, M. Demura and M. P. Williamson, The relationship between amide proton chemical shifts and secondary structure in proteins, *J. Biomol. NMR*, 1995, **6**, 227-236.

80. D. S. Wishart, Interpreting protein chemical shift data, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2011, **58**, 62-87.

81. S. M. Kelly, T. J. Jess and N. C. Price, How to study proteins by circular dichroism, *Biochim. Biophys. Acta*, 2005, **1751**, 119-139.

82. R. W. Woody, The development and current state of protein circular dichroism, *Biomed. Spectrosc. Imaging*, 2015, **4**, 5-34.

83. N. E. Shepherd, H. N. Hoang, G. Abbenante and D. P. Fairlie, Single Turn Peptide Alpha Helices with Exceptional Stability in Water, *J. Am. Chem. Soc.*, 2005, **127**, 2974-2983.

84. D. Piovesan, G. Minervini and S. C. Tosatto, The RING 2.0 web server for high quality residue interaction networks, *Nucleic Acids Res.*, 2016, **44**, W367-374.

85. K. Stierand, P. C. Maass and M. Rarey, Molecular complexes at a glance: automated generation of two-dimensional complex diagrams, *Bioinformatics*, 2006, **22**, 1710-1716.

86. A. Bhattacharya, R. Tejero and G. T. Montelione, Evaluating protein structures determined by structural genomics consortia, *Proteins: Struct., Funct., Genet.*, 2007, **66**, 778-795.

87. W. Kabsch and C. Sander, Dictionary of Protein Secondary Structure: Pattern Recognition of Hydrogen-Bonded and Geometrical Features, *Biopolymers*, 1983, **22**, 2577-2637.

88. M. Heinig and D. Frishman, STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins, *Nucleic Acids Res.*, 2004, **32**, W500-502.

89. Schrodinger, LLC, unpublished work.

90. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, UCSF Chimera - a visualization system for exploratory research and analysis, *J. Comput. Chem.*, 2004, **25**, 1605-1612.

91. M. A. Cooper, Optical biosensors in drug discovery, *Nat Rev Drug Discov*, 2002, **1**, 515-528.

92. P. Matsson, B. C. Doak, B. Over and J. Kihlberg, Cell permeability beyond the rule of 5, *Adv. Drug Delivery Rev.*, 2016, **101**, 42-61.

93. A. Prochiantz, Protein and peptide transduction, twenty years later a happy birthday, *Adv Drug Deliv Rev*, 2008, **60**, 448-451.

94. X. Antoniou and T. Borsello, Cell Permeable Peptides: A Promising Tool to Deliver Neuroprotective Agents in the Brain, *Pharmaceuticals (Basel)*, 2010, **3**, 379-392.

95. N. Barlow, D. K. Chalmers, B. J. Williams-Noonan, P. E. Thompson and R. S. Norton, Improving Membrane Permeation in the Beyond Rule-of-Five Space by Using Prodrugs to Mask Hydrogen Bond Donors, *ACS Chem. Biol.*, 2020, **15**, 2070-2078.

96. G. B. Santos, A. Ganesan and F. S. Emery, Oral Administration of Peptide-Based Drugs: Beyond Lipinski's Rule, *ChemMedChem*, 2016, **11**, 2245-2251.

97. A. T. Jones and E. J. Sayers, Cell entry of cell penetrating peptides: tales of tails wagging dogs, *J Control Release*, 2012, **161**, 582-591.

98. C. Bechara and S. Sagan, Cell-penetrating peptides: 20 years later, where do we stand?, *FEBS Lett.*, 2013, **587**, 1693-1702.

99. H. D. Herce, A. E. Garcia, J. Litt, R. S. Kane, P. Martin, N. Enrique, A. Rebolledo and V. Milesi, Arginine-rich peptides destabilize the plasma membrane, consistent with a pore formation translocation mechanism of cell-penetrating peptides, *Biophys. J.*, 2009, **97**, 1917-1925.

100. T. B. Potocky, A. K. Menon and S. H. Gellman, Cytoplasmic and nuclear delivery of a TAT-derived peptide and a beta-peptide after endocytic uptake into HeLa cells, *J. Biol. Chem.*, 2003, **278**, 50188-50194.

101. A. Mishra, G. H. Lai, N. W. Schmidt, V. Z. Sun, A. R. Rodriguez, R. Tong, L. Tang, J. Cheng, T. J. Deming, D. T. Kamei and G. C. Wong, Translocation of HIV TAT peptide and analogues induced by multiplexed membrane and cytoskeletal interactions, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 16883-16888.

102. H. H. Szeto, P. W. Schiller, K. Zhao and G. Luo, Fluorescent dyes alter intracellular targeting and function of cell-penetrating tetrapeptides, *The FASEB Journal*, 2005, **19**, 118-120.

103. D. Birch, M. V. Christensen, D. Staerk, H. Franzyk and H. M. Nielsen, Fluorophore labeling of a cell-penetrating peptide induces differential effects on its cellular distribution and affects cell viability, *Biochim. Biophys. Acta. Biomembr.*, 2017, **1859**, 2483-2494.

104. M. P. Luitz, A. Barth, A. H. Crevenna, R. Bomblies, D. C. Lamb and M. Zacharias, Covalent dye attachment influences the dynamics and conformational properties of flexible peptides, *PLoS One*, 2017, **12**, e0177139.

105. S. F. Hedegaard, M. S. Derbas, T. K. Lind, M. R. Kasimova, M. V. Christensen, M. H. Michaelsen, R. A. Campbell, L. Jorgensen, H. Franzyk, M. Cárdenas and H. M. Nielsen, Fluorophore labeling of a cell-penetrating peptide significantly alters the mode and degree of biomembrane interaction, *Sci. Rep.*, 2018, **8**, 6327.

106. K.-Y. Tan, C.-Y. Li, Y.-F. Li, J. Fei, B. Yang, Y.-J. Fu and F. Li, Real-Time Monitoring ATP in Mitochondrion of Living Cells: A Specific Fluorescent Probe for ATP by Dual Recognition Sites, *Anal. Chem.*, 2017, **89**, 1749-1756.

107. D. de la Fuente-Herreruela, V. Gonzalez-Charro, V. G. Almendro-Vedia, M. Moran, M. A. Martin, M. P. Lillo, P. Natale and I. Lopez-Montero, Rhodamine-based sensor for real-time imaging of mitochondrial ATP in living fibroblasts, *Biochim. Biophys. Acta, Bioenerg.*, 2017, **1858**, 999-1006.

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

109. E. Pazos, O. Vazquez, J. L. Mascarenas and M. E. Vazquez, Peptide-based fluorescent biosensors, *Chem. Soc. Rev.*, 2009, **38**, 3348-3359.

110. Q. Liu, J. Wang and B. J. Boyd, Peptide-based biosensors, *Talanta*, 2015, **136**, 114-127.

111. M. Sainlos and B. Imperiali, Synthesis of anhydride precursors of the environment-sensitive fluorophores 4-DMAP and 6-DMN, *Nat. Protoc.*, 2007, **2**, 3219-3225.

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

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

114. L. Choulier and K. Enander, Environmentally sensitive fluorescent sensors based on synthetic peptides, *Sensors*, 2010, **10**, 3126-3144.

115. K. E. Sapsford, J. B. Blanco-Canosa, P. E. Dawson and I. L. Medintz, Detection of HIV-1 Specific Monoclonal Antibodies Using Enhancement of Dye-Labeled Antigenic Peptides, *Bioconjugate Chem.*, 2010, **21**, 393-398.

116. G. B. Fields and R. L. Noble, Solid phase peptide synthesis utilizing 9-fluorenylmethoxycarbonyl amino acids, *Int. J. Pept. Protein Res.*, 1990, **35**, 161-214.

117. G. S. Vanier, Microwave-assisted solid-phase peptide synthesis based on the Fmoc protecting group strategy (CEM), *Methods Mol. Biol.*, 2013, **1047**, 235-249.

118. A. A. Mazurov, S. A. Andronati, T. I. Korotenko, V. Y. Gorbatyuk and Y. E. Shapiro, Formation of pyroglutamylglutamine (or asparagine) diketopiperazine in ‘non-classical’ conditions: a side reaction in peptide synthesis, *Int. J. Pept. Protein Res.*, 1993, **42**, 14-19.

119. J. D. Wade, M. N. Mathieu, M. Macris and G. W. Tragear, Base-induced side reactions in Fmoc-solid phase peptide synthesis Minimization by use of piperazine as Na-deprotection reagent, *Lett. Pept. Sci.*, 2000, **7**, 107-112.

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

121. A.-B. M. Abdel-Aal, G. Papageorgiou, R. Raz, M. Quibell, F. Burlina and J. Offer, A backbone amide protecting group for overcoming difficult sequences and suppressing aspartimide formation, *J. Pept. Sci.*, 2016, **22**, 360-367.

122. T. Michels, R. Dolling, U. Haberkorn and W. Mier, Acid-Mediated Prevention of Aspartimide Formation in Solid Phase Peptide Synthesis, *Org. Lett.*, 2012, **14**, 5218-5221.

Chapter 2.

A BIMANE-BASED PEPTIDE STAPLE FOR COMBINED HELICAL INDUCTION AND FLUORESCENT IMAGING

Research Article: ChemBioChem **2020**, 21(23), 3423-3432

A bimane-based peptide staple for combined helical induction and fluorescent imaging

Aimee J. Horsfall,[1-3] Kylie R. Dunning,[2-4] Kelly L. Keeling,[1,3] Denis B. Scanlon,[1,3] Kate L. Wegener[3,5*] and Andrew D. Abell[1-3*]

[1] The Department of Chemistry, School of Physical Sciences, The University of Adelaide, North Terrace, Adelaide, SA 5005, Australia.

[2] The ARC Centre of Excellence for Nanoscale BioPhotonics (CNBP), The University of Adelaide, North Terrace, Adelaide, SA 5005, Australia.

[3] Institute for Photonics and Advanced Sensing (IPAS), The University of Adelaide, North Terrace, Adelaide, SA 5005, Australia.

[4] Robinson Research Institute, Adelaide Medical School, The University of Adelaide, North Terrace, Adelaide, SA 5005, Australia.

[5] School of Biological Sciences, The University of Adelaide, North Terrace, Adelaide, SA 5005, Australia

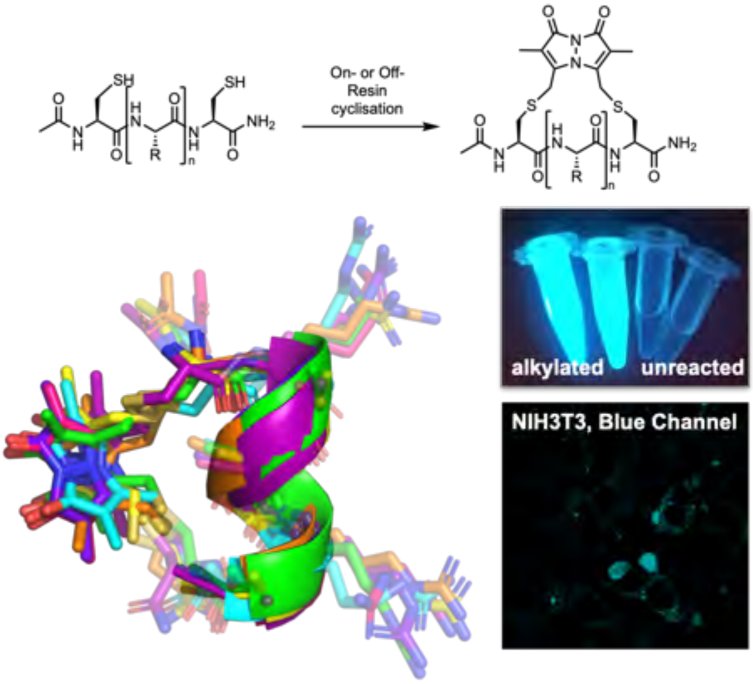
*corresponding author

First published online: 23rd July 2020

© 2020 ChemBioChem, Wiley

ABSTRACT

The thiol-selective fluorescent imaging agent, dibromobimane, is repurposed to cross-link cysteine and homocysteine containing peptides, with the resulting bimane-linker acting as both a structural constraint and a fluorescent tag. Macrocyclisation was conducted on nine short peptides containing two cysteines and/or homocysteines, both on-resin and in buffered aqueous solution, to give macrocycles ranging in size from 16 (*i*-*i*+2) to 31 (*i*-*i*+7) atoms. The structures were defined by CD, NMR structure calculations using Xplor-NIH, NMR secondary shift and $J_{H\alpha NH}$ analyses, to reveal helical structure in the *i*-*i*+4 (**1**, **2**), and *i*-*i*+3 (**5**) constrained peptides. Cellular uptake studies were conducted with three of the macrocycles. Subsequent confocal imaging revealed punctate fluorescence within the cytosol indicative of peptides trapped in endocytic vesicles. These studies demonstrate that dibromobimane is an effective tool to define secondary structure within short peptides, whilst simultaneously introducing a fluorescent tag suitable for common cell-based experiments.



A repurposed fluorescent thiol labelling agent, dibromobimane, is used to form a peptide constraint to introduce helical structure into short peptides. The cyclisation reaction can be conducted on-resin following SPPS or in-solution. The resultant inherently fluorescent peptides are compatible with cell-uptake imaging experiments and mitigate the need for introduction of a secondary tag.

STATEMENT OF AUTHORSHIP

Title of Paper	A bimane-based peptide staple for combined helical induction and fluorescent imaging
Publication Status	■ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	**Research Article:** A. J. Horsfall, K. R. Dunning, K. L. Keeling, D. B. Scanlon, K. L. Wegener and A. D. Abell, *ChemBioChem*, 2020, **21**(23) 3423-3432.

Principal Author

Name of Principal Author (Candidate)	Aimee J Horsfall		
Contribution to the Paper	Designed & carried out all experiments, peptide synthesis, analysed and discussed results, wrote and edited manuscript		
Overall percentage (%)	60%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	15/02/2021

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	K R Dunning		
Contribution to the Paper	Performed cell uptake assays, confocal imaging, discussed results, and edited manuscript		
Signature		Date	15/2/2021

Name of Co-Author	Kelly L Keeling		
Contribution to the Paper	Conceived initial idea, peptide synthesis and edited manuscript		
Signature		Date	28/2/2021

Name of Co-Author	Denis B Scanlon		
Contribution to the Paper	Peptide synthesis, discussed results, and edited manuscript		
Signature		Date	01/03/2021

Name of Co-Author	Kate L Wegener		
Contribution to the Paper	Conceived initial idea, NMR assignment and structure calculations, discussed results, and edited manuscript		
Signature		Date	02/03/2021

Name of Co-Author	Andrew D Abell		
Contribution to the Paper	Conceived initial idea, supervised AJH & KLK, discussed results, wrote and edited manuscript.		
Signature		Date	15/2/2021

2.1 Introduction

Peptide-based probes and therapeutics can be constrained into a biologically active conformation by the judicious introduction of a covalent constraint that typically links two amino-acid side-chains. This then reduces the entropic penalty on binding to a target enzyme or receptor,[1, 2] while also enhancing proteolytic stability and often membrane permeability.[2-5] For example, a constraint linking the *i* and *i*+4 amino-acid side-chains of a peptide is commonly used to pre-form an α-helix.[6-8] A 3_{10} helix is stabilised by linking *i* and *i*+3 amino-acid side-chains,[9-13] while an *i*-*i*+2 constraint can establish a β-strand to provide a basis for protease inhibitors.[14-17] A number of synthetic approaches have been developed to introduce such constraints, including lactam bridges,[18-20] all-hydrocarbon staples,[16, 17, 21-23] 1,2,3-triazoles[24-26] and perfluoro aryl-based linkers.[27] Each system has inherent benefits and limitations.[5, 6, 22, 28-33]

A variety of dihalo-reagents has also been used to define peptide backbone geometry by cross-linking cysteine side chains. These alkylating agents include dibromo-*m*-xylene,[34, 35] hexafluorobenzene/dibromotetrafluorobenzene,[27, 28] 1,3-diiodo-propane and 1,4-diiodo-butane.[36] Cysteine *bis*-alkylation has emerged as a method for rapid derivatisation of a peptide sequence to screen a multitude of linker chemistries.[34, 37] This work has been recently extended to seleno-cysteines.[38].

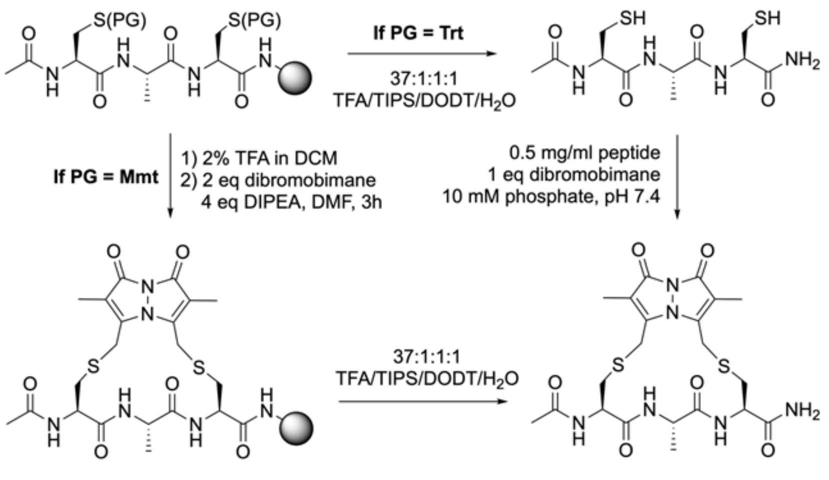
An auxiliary fluorescent tag can be attached to these peptidomimetics and fluorescence monitored by confocal microscopy to allow imaging and tracking in a biological environment. However, introduction of a tag can adversely influence conformation,[39] cell permeability,[40-42] and the biological interaction under investigation.[6,43-45] A significant advance would come with the development of a linker able to both conformationally constrain a peptide backbone, while also fluorescing to allow imaging and tracking. Here we explore the a bimane-based linker as a thiol-specific fluorescent label and structure constraint.[46-48]

Dibromobimane, and the closely related monochloro- or monobromobimane have been used extensively in the literature as fluorescent labelling agents for the detection of thiol species.[49-52] Monohalobimanes have also been routinely used for distance mapping of protein residues by fluorescence quenching of bimane fluorescence in proximity to tryptophan or tyrosine residues.[53-56] A range of cysteine and homocysteine (Hcy) containing peptides was produced (*i*-*i*+2 through to *i*-*i*+7 Cys/Hcy), and reacted with dibromobimane, to produce a range of macrocycles (16 to 31 atoms, see **1**-**9** in Table. 1). The effect of the bimane constraint on backbone geometry was assessed by NMR and CD spectroscopy. Fluorescence was shown to be independent of macrocycle size or pH (between 6 and 8). Finally, we demonstrate that the bimane-linked peptides fluorescence allows imaging by confocal microscopy.

2.2 Results and discussion

2.2.1 Synthesis

The linear peptides used for the preparation of **1**-**9** were synthesised by standard solid-phase peptide synthesis (SPPS) procedures as described in the Experimental section. The reaction of dibromobimane with the precursor to peptide **1** was initially optimised under solution-phase conditions, by varying peptide concentration, dibromobimane equivalents, and pH using a 96-well plate platform. Formation of **1** was



8

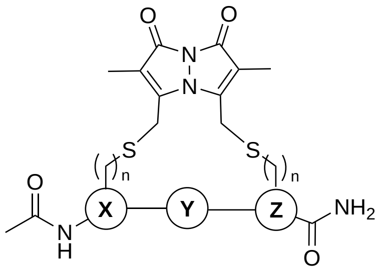
Scheme 1. General synthetic scheme, with exemplar peptide **8**. All linear precursor peptides were assembled by SPPS on Rink Amide resin and then reacted with dibromobimane to give **1**-**9**, or monobromobimane to give **10** by one or both of these methods, as specified below. Trityl cysteine protecting groups (PG = Trt) were used for solution-phase cyclisation, where the peptide was globally deprotected and cleaved from the resin before reaction with dibromobimane. Peptides **1**-**10** were prepared in this way. 4-Monomethoxytrityl cysteine protecting groups (PG = Mmt) were used for on-resin cyclisation, and selectively removed with repetitive treatments with 2% TFA in DCM to unmask the thiol group for reaction with dibromobimane. Peptides **1**, **3** and **8** were synthesised in this manner.

monitored by measuring an increase in fluorescence intensity (ex. 385 nm, em. 477 nm, see Figures S1 and S2), where bromobimane has negligible fluorescence.[50]

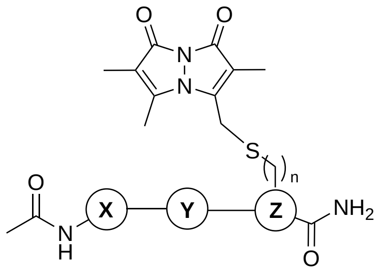
The optimum reaction conditions were determined to be 0.5 mg/mL peptide, and one equivalent of dibromobimane in 10 mM phosphate buffer at pH 7.8 (Figure S1). Product **1** formation was deemed complete after 5 min as no further increase in bimane fluorescence was observed (Figure S2). This was supported by reverse phase HPLC and mass spectrometry (MS), without evidence of dimer or other by-product formation. The use of phosphate buffer at pH 7.8 in this reaction appears to be important, where the use of ammonium acetate, with its limited buffering capacity at neutral pH,[57] slowed the reaction significantly as seen by a smaller change in fluorescence over time (Figure S2). It is worth noting that the linear precursor to **1** showed limited solubility above 1 mg/mL, which likely explains the reduced fluorescence observed under these conditions. Higher peptide concentrations may be tolerated by other amino-acid sequences.

The optimised conditions were then used to prepare peptides **2**-**9** using a conservative reaction time of 30 min to ensure complete formation of the macrocycles (Scheme 1, where PG=Trt). Complete conversion from linear to cyclised peptide was apparent in all cases by MS, irrespective of macrocycle size. This is of note since the formation of other constraints (e.g. an all-hydrocarbon staple) are reported to be strongly size dependent.[22, 58] 2,2,2-Trifluoroethanol (10% of the final reaction volume) was added in preparations of **1**-**4** and **6**-**7**, to promote anticipated helical folding.[59] The acyclic analogue **10** was prepared as a control for circular dichroism studies by reaction with monobromobimane.

Table 1. Structures of the bimane cyclised peptides (**1**-**9**) and the acyclic bimane peptide **10**. X, Y and Z represent the linear amino-acid sequence, where C is L-cysteine, A is L-alanine, R is L-arginine, D-Cys is D-cysteine and Hcy is L-homocysteine.



Peptide	n	X	Y	Z	Constraint
1	1	C	AAA	C	*i*-*i*+4
2	1	D-Cys	AAA	C	*i*-*i*+4
3	1	C	RAAARA	C	*i*-*i*+7
4	1	D-Cys	RAAARA	C	*i*-*i*+7
5	1	C	AA	C	*i*-*i*+3
6	1	C	ARAA	C	*i*-*i*+5
7	1	C	ARAAA	C	*i*-*i*+6
8	1	C	A	C	*i*-*i*+2
9	2	Hcy	A	Hcy	*i*-*i*+2



Peptide	n	X	Y	Z	Constraint
10	1	A	A	C	-

A solid-phase synthesis protocol was also investigated for the synthesis of bimane-constrained peptides. A 4-monomethoxytrityl (Mmt) thiol protecting group was used as it allows selective on-resin deprotection and subsequent cyclisation. Reaction of the resin bound linear peptides with two equivalents of dibromobimane and four equivalents of DIPEA in DMF, for three hours, gave the *i*-*i*+4, *i*-*i*+7 and *i*-*i*+2 constrained peptides **1**, **3** and, **8** respectively (Scheme 1, where PG=Mmt). This provides some advantage over the earlier solution-based methodology as it requires fewer isolation steps.

2.2.2 Structural studies

The structures of peptides **1**-**10** were defined using a combination of circular dichroism (CD), NMR secondary shift analysis,[60, 61] $J_{H\alpha NH}$ values and NMR structure calculations. NMR and CD spectra were collected under standard literature conditions.[18, 28, 33] Structure calculations were carried out with the Xplor-NIH program,[62-64] employing NOE distance restraints and chemical shift based dihedral angle restraints (DANGLE[65] from within CCPNMR Analysis[66]). The resulting data are summarised in Figure 1, 2 and Figure S3-5 and discussed in more detail below.

i-i+4 and i-i+7 constrained peptides

Peptide 1: The NMR-derived structure of the *i-i*+4 constrained pentapeptide **1** revealed a single turn helix, with φ (C-N-Cα-C) and ψ (N-Cα-C-N) dihedral angles of -61° and -40° for the central alanine residues (Table S13). These values are close to those anticipated for an ideal α-helix (φ: -57°, ψ: -47° or sum of -105°) and are consistent with other *i-i*+4 constrained peptides.[67] The Hα, and 2 of 3 NH resonances, for **1** are shifted upfield relative to random coil values (i.e. negative secondary shifts), while Cα resonances are shifted downfield (positive secondary shifts), consistent with an α-helix (Figure 1A, blue). The carbonyl carbon (CO) secondary shifts were positive as anticipated for helical structure (Figure 1A, blue). $J_{\text{HαNH}}$ values < 6 Hz for the central alanine residues and an ellipticity minimum between 215 and 220 nm in the CD spectrum of **1** are also consistent with α-helical structure (Figure 1C, blue).[30] The percentage helicity was calculated[13] using the MRE intensity at 215 nm to be 53% (Figure 1C, Table S1). It should be noted that the bimane does not contribute to the CD spectrum, as has been observed for perfluoro aryl constraints,[9] since acyclic peptide **10** (bimane attached to a single cysteine residue) gave an essentially identical spectrum to its unmodified precursor peptide (Figure S3).

Peptide 2: Peptide **2** with its *N*-terminal D-Cys residue in place of L-Cys of peptide **1**, showed a deeper ellipticity minimum at 215-220 nm in its CD spectrum (Figure 1C, purple). Peptide **2** with its larger NMR secondary shifts (negative/upfield for Hα and positive/downfield for carbon shifts) and smaller $J_{\text{HαNH}}$ values for Ala2 and Ala3 (Figure 1A, purple) suggests that it forms a somewhat more stable α-helix than **1**. The percentage helicity for peptide **2** was calculated to be 84% (Figure 1C, Table S1). The ratio between the 215 and 207 nm MRE intensities is 0.62, suggests a 3_{10}-helicity.[13]

The NMR-calculated structure of **2** is, however, consistent with ideal α-helical structure for residues Ala2 and Ala3 (φ/ψ angles of -61°/-40° and -60°/-42° respectively), and only slight deviation from this ideal at residue Ala4 (Ala4 φ/ψ angles of -70°/-32°) (Table S13).

Peptide 3: The *i-i*+7 constrained peptide **3** was determined to have some α-helical content (15%) that is disrupted at residue 4 as evidenced by φ/ψ angles of approximately -60°/-40° for residues 2, 3 and 5-7, compared to -85°/-0.7° at Ala4 (Table S13). In support, the Hα and Cα resonances are shifted upfield and downfield respectively (relative to the random coil literature shifts) and a minimum was apparent in the CD spectrum between 210 and 220 nm (Figure 1, green).

Peptide 4: Peptides **3** and **4**, with an *N*-terminal L-Cys and D-Cys respectively, gave very similar NMR and CD data suggesting common structures (Figure 1, yellow). This is supported by NMR structure calculations with **4** forming an α-helix, disrupted at residue Ala4 (φ/ψ angles ~-60°/-40° for residues 2, 3 and 5-7, and -81°/- 9.1° for Ala4) (Table S13). It thus appears that the introduction of an *N*-terminal D-Cys residue does not significantly impact overall structure, with a percentage helicity of 22% compared to 15% for **3**.

i-i+3, i-i+5 and i-i+6 constrained peptides

The structures of the *i-i*+3, *i-i*+5, and *i-i*+6 bimane cyclised peptides **5**, **6** and **7** were also determined (Figure 2) in order to explore the effect of the linker at a range of positions.

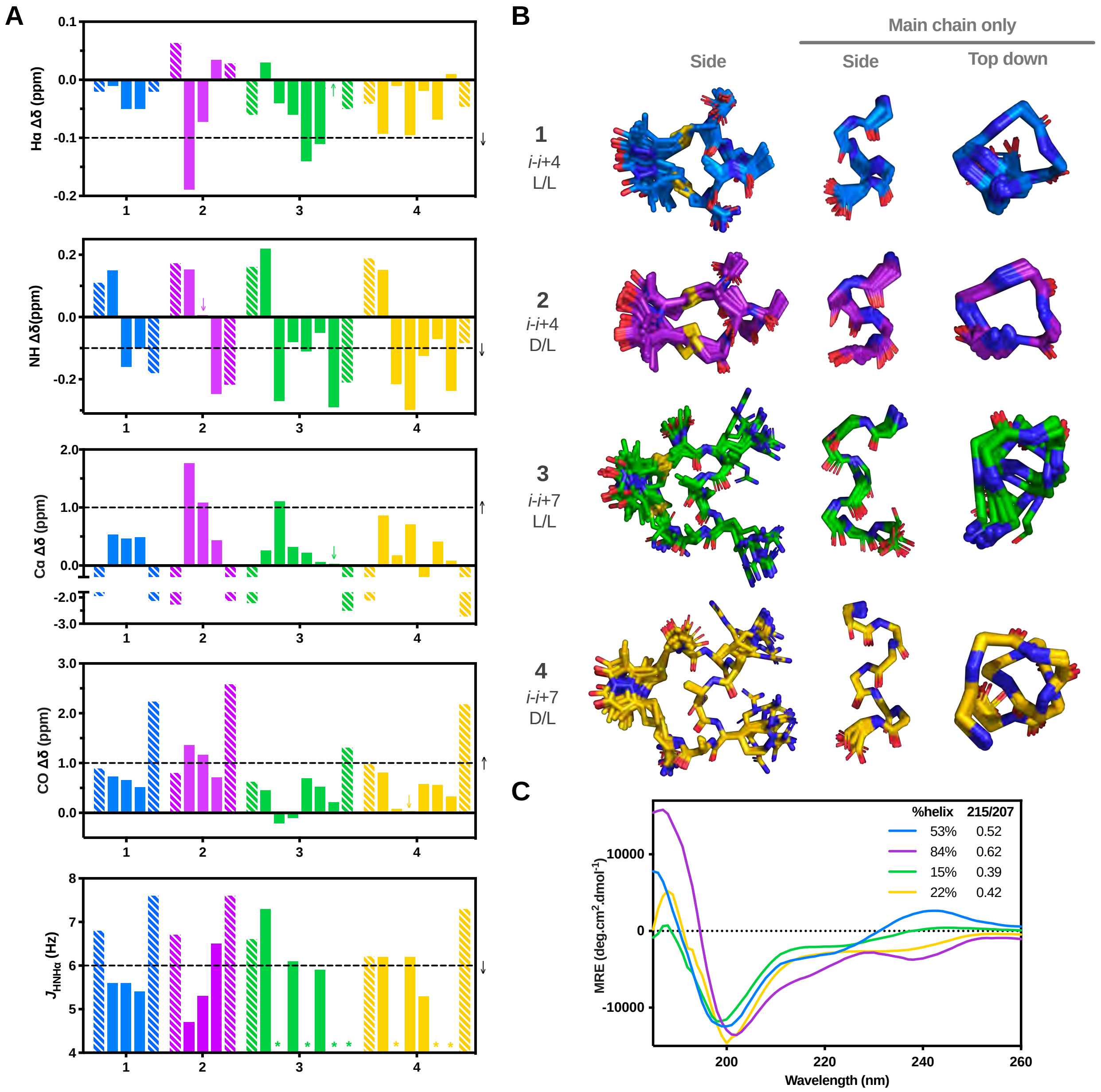


Figure 1. Structural data for peptides **1** (blue), **2** (purple), **3** (green) and **4** (yellow). **A)** NMR structural data (10% aq. D_2O, pH ~5): Hα, Cα and carbonyl (CO) carbon secondary shifts are calculated relative to the random coil literature values reported by Wishart (2011) and nearest neighbour corrections applied,[60] each column represents a consecutive residue. Dashed columns represent bimane-modified cysteine residues, whose secondary shifts were calculated relative to random coil cysteine (reduced) data, as more appropriate reference values were not available. Stars (*) represent data that could not be reliably extracted from the spectrum. The horizontal dashed lines and arrow indicate the generally accepted threshold (and direction) to indicate helical structure. Three consecutive residues surpassing the horizontal dashed lines strongly indicates the presence of helical structure. **B)** 20 lowest energy NMR structures calculated using Xplor-NIH.[62-64] All structures are orientated with the *N*-terminus toward the top of the page. Carbon atoms are colour coded to the peptide, nitrogen (dark blue), oxygen (red) and sulfur (yellow), hydrogens are not shown for clarity. **C)** CD spectra collected in 10 mM phosphate, pH 7.2, at 298K The %helix content was calculated using the MRE intensity at 215 nm using the method reported in Shepherd (2005)[13] for peptides of five amino-acids length or greater, and the ratio of the 215/207 nm MRE intensity which is indicates of 3_{10}-helical (0.6) or α-helical (0.9) structure.

Peptide 5: The CD spectrum of peptide **5** (*i*-*i*+3) shows a deep minimum at 220 nm, shallow minimum at 203 nm and strong maximum at 190 nm (Figure 2C, dark red), indicating helical structure. The two central alanine residues gave φ and ψ angles of approximately -60°/-40°, similar values to those observed for **1** (Table S13). The $J_{H\alpha NH}$ values for Ala2 and Ala3 were relatively small (5.4 and 6.1 Hz) which further supports helical-like structure. However, the Ala2 and Ala3 secondary shifts are opposite in magnitude, indicating the structure is not uniform with a complete turn requiring a longer sequence.[30]

Peptide 6: The *i*-*i*+5 25-membered constrained peptide **6** shows helical structure calculated to be 31% (Figure 2C, Table S1), which is consistent with the large positive secondary shifts for all Cα and CO resonances and negative Hα secondary shifts for the central Ala3 and Ala4 residues (Figure 2A, light blue). However, NMR structure calculations indicate that this helical structure is somewhat “unwound” and distorted from ideal α-helicity, with larger φ/ψ angles (see Table S13). All residues were located in the helical region of Ramachandran space, but only 56% of these were found in the ‘most favoured’ region. The remainder are classified as ‘allowed’ with the identity of the favoured/allowed residues varying between conformers, consistent with a degree of strain.[68-70] This conclusion is supported by large standard deviation in the φ/ψ angles of the 20 lowest energy conformers. Closer analysis of the NMR structure calculations reveal three similar but distinct conformations for **6** (Tables S14-S16).

The limited flexibility and nature of the bimane constraint likely does not allow 1.5 helical turns expected for a 6-mer with ideal α-helical configuration, resulting in a distorted helix. While peptide **6** with its *i*-*i*+5 bimane linker does not form an ideal α-helix, this constraint may be of interest in the study of non-canonical helical structures such as that of the *i*-*i*+5 constrained oxytocin hormone.[71-73]

Peptide 7: Some helical character (25%, Figure 2C and Table S1) was also apparent for the *i*-*i*+6 constrained peptide **7**, with a CD minimum at 215-222 nm and negative Hα and NH secondary shifts, positive Cα and CO secondary shifts, and three $J_{H\alpha NH}$ values (for Ala2, 5 and 6) below 5 Hz (Figure 2A, dark blue). Some perturbation of helicity is again apparent around Arg3 and Ala4 in the NMR calculated structure (Table S13 and Figure 2B, dark blue), with notably larger $J_{H\alpha NH}$ values of 7.6 and 6.4 Hz respectively (Figure 2A, dark blue, c.f. 4.4-4.9 Hz for the other residues). This suggests that an *i*-*i*+6 linker results in a distorted structure.

i-i+2 constrained peptides

An *i* to *i*+2 side-chain constraint is known to effectively stabilise a β-strand geometry particularly in short peptides,[3] with the constraint length being important.[14, 17, 26] With this in mind the *i*-*i+2* cysteine and homocysteine bimane constrained peptides **8** and **9** were prepared.

Peptides 8 and 9: NMR secondary shifts and CD spectra for peptides **8** and **9** were collected. However how meaningful these data are for such short peptides is unclear, especially in the absence of long-range NOEs. A $J_{H\alpha NH}$ nearing 8 Hz for peptide **9**, and large negative Hα and NH secondary shifts, and a positive Cα secondary shift, suggest the larger macrocycle provides a more extended (β-strand-like) structure than **8** (Figure S4, orange).

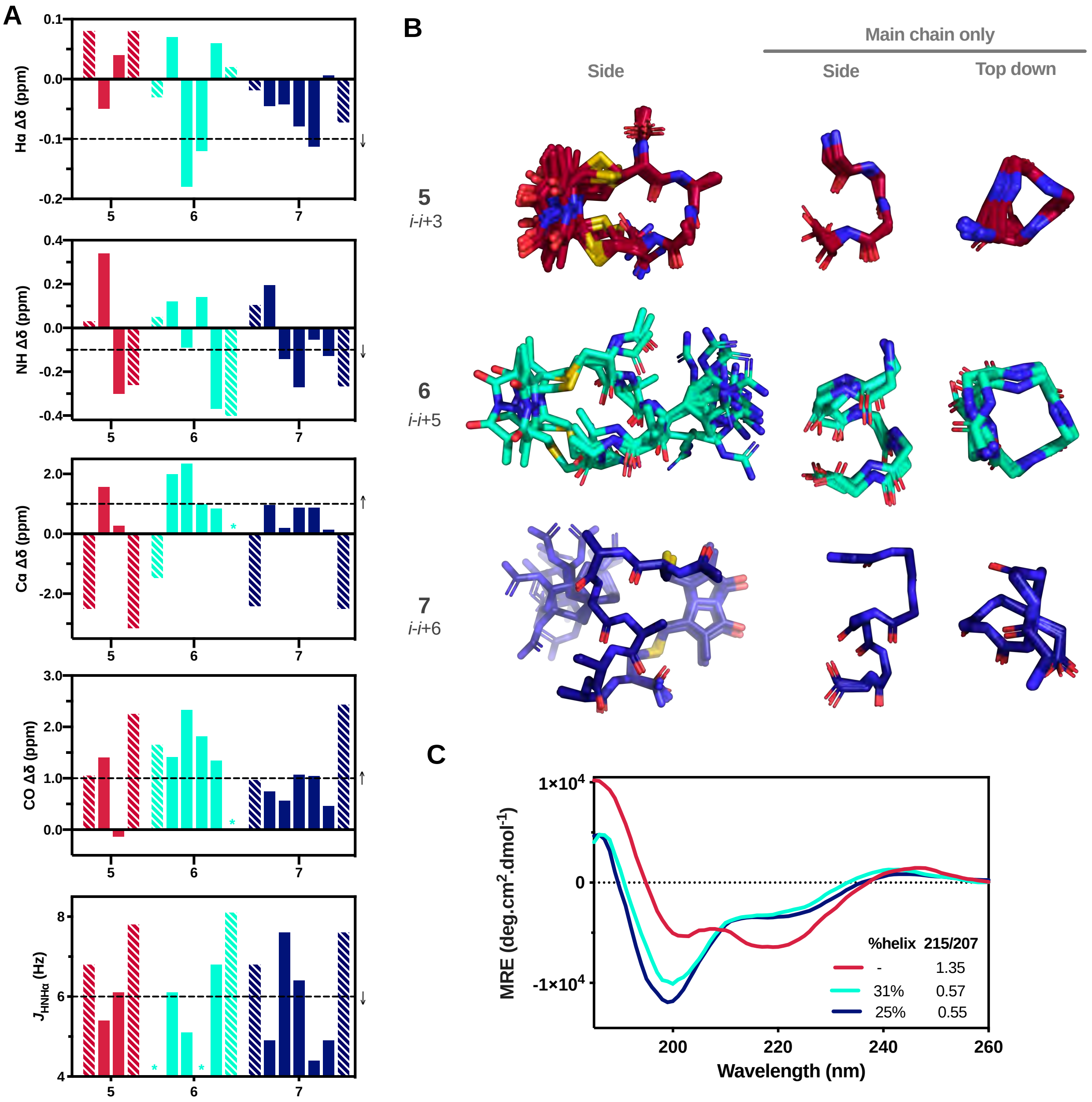


Figure 2. Structural data for **5** (dark red), **6** (light blue), and **7** (navy blue). **A)** NMR structural data (10% aq. D_2O, pH ~5): Hα, Cα and CO carbon secondary shifts are calculated relative to the random coil literature values reported by Wishart (2011) and nearest neighbour corrections applied,[60] each column represents a consecutive residue. Dashed columns represent cysteine residues, whose secondary shifts were calculated relative to random coil cysteine data. Stars (*) represent data that could not be reliably extracted from the spectra. The horizontal dashed lines and arrow, define the generally accepted threshold to indicate helical structure. Three consecutive residues surpassing the horizontal dashed lines strongly indicate the presence of helical structure. The final panel shows CD spectra collected in 10 mM phosphate at pH 7.2, 298K **B)** 20 lowest energy NMR structures calculated using Xplor-NIH.[62-64] All structures are orientated with the *N*-terminus toward the top of the page. Carbon atoms are colour coded to the peptide, nitrogen (dark blue), oxygen (red) and sulfur (yellow), the hydrogens are omitted for clarity. **C)** The CD spectra was collected in 10 mM phosphate at pH 7.2, 298K. The %helix content was calculated using the MRE intensity at 215 nm using the method reported in Shepherd (2005)[13] for peptides of five amino-acids length or greater, and the ratio of the 215/207 nm MRE intensity which is indicates of 3_{10}-helical (0.6) or α-helical (0.9) structure.

2.2.3 Fluorescence characterisation

The fluorescence spectra of peptides **1**, **3** and **5**-**10** revealed that all gave an excitation maxima at 385 nm and emission at 477-480 nm (Figure S6), consistent with literature.[47, 50] Peptide **2** and **4**, with an *N*-terminal D-cysteine, gave and emission maximum at 480 nm, with the excitation maximum shifted to 420 nm, and a shoulder in the excitation spectrum at 385 nm. Importantly, all peptides displayed good photostability as required for imaging, with fluorescence intensity decreasing by less than 8% after more than 1900 irradiations over six hours (high energy Xenon Flash, Figure S7, Table S18). Peptide **1** shows an approximate limit of detection of 25 nM (Figure S8) with essentially identical fluorescence intensity over a pH range of 6-8 (Figure S9).

2.2.4 Cellular imaging

The bimane-constrained peptides **7**, **9** and **10** were chosen as representative examples to demonstrate that bimane fluorescence is clearly resolved from background cellular autofluorescence. The peptides (5 µM) were incubated with NIH/3T3 cells for 4 h, the cells were then washed with medium, incubated for a further 16 h, and then imaged by confocal microscopy using simultaneous brightfield and fluorescent images (Figure 3). Transitioning through the z-plane it was confirmed that the image taken was indeed within the cell and not on the cell membrane. The brightfield image allowed us to confirm we were always within the cell. Further, imaging parameters were set using cells that were not exposed to fluorescent peptides such that no fluorescence signal was captured due to endogenous fluorescent molecules found within cells (autofluorescence).

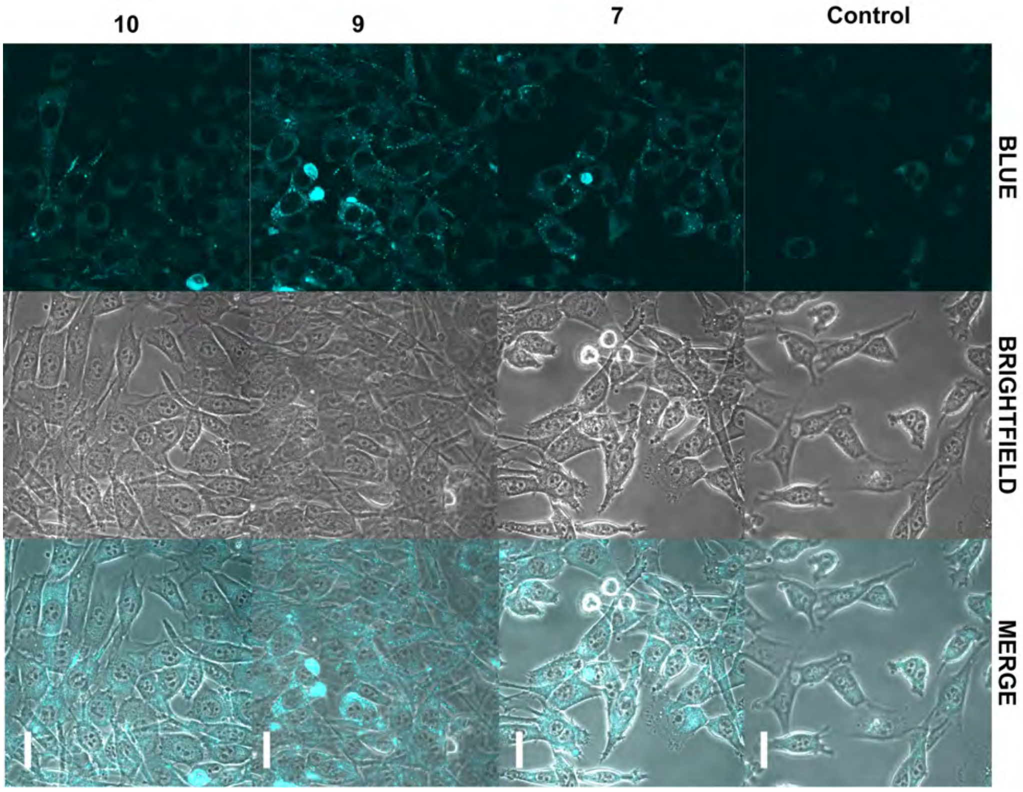


Figure 3. Confocal microscopy images of peptides **7**, **9** and **10** at 5 µM incubated for 4 h with NIH/3T3 cells, then washed and imaged at ex. 405 nm, em. 420-520 nm (blue). For the negative control, cells were incubated in medium with no peptide (*Control*) as described in the Experimental section. Scale bar = 50 µm.

All three peptides were clearly resolved, with fluorescence distinguished from cellular autofluorescence observed in the negative control (Figure 3, 'Control'). The bimane-peptides are thus suitable for cell-based imaging experiments. High intensity punctate fluorescence was observed in the cells for all three peptides, consistent with uptake into endocytic vesicles.[74, 75] Penetration of cell membranes is somewhat surprising given that these peptides lack sequences known to promote cellular uptake.[35] However, both mono- and dibromobimane, on their own, are able to penetrate cell membranes.[19a] Whether the bimane linker promotes the uptake of attached cargo sequences awaits further study, but it is clear from the current experiments that it is fully compatible with fluorescent imaging of peptide cellular uptake.

2.3 Conclusion

In summary, nine peptides containing two cysteine and/or homocysteine residues, were synthesised and cyclised by reaction with dibromobimane. This cyclisation was carried out both on-resin – immediately follow SPPS assembly (2 equiv dibromobimane, 4 equiv DIPEA, DMF, 3h), or rapidly in phosphate buffered solution (0.5 mg/mL peptide, 1 equiv dibromobimane, 10 mM phosphate at pH 7.4, 30 min). These peptides, with terminal cysteines, varied in length from three to eight amino-acids and upon cyclisation gave rise to macrocycles of 15 to 31 atoms in size. The secondary structure of the resultant macrocycles was defined by CD, NMR structure calculations using Xplor-NIH, NMR secondary shift and $J_{H\alpha NH}$ analyses, to reveal α-helical structure in the *i*-*i*+4 (**1**, **2**), and *i*-*i*+3 (**5**) constrained peptides. The substitution of a D-Cys, for L-Cys, at the *N*-terminal position in peptides **1** and **3**, to give peptides **2** and **4**, did not significantly impact the secondary structure. Further studies are required to investigate the ability of the bimane linker to stabilise, or indeed propagate, helical structure in longer peptide sequences.

The bimane constrained peptides are highly fluorescent (ex. 385 nm, em. 477 nm). Three bimane constrained peptides were assayed in a cell uptake experiment with NIH/3T3 cells and then imaged by confocal microscopy. The peptides were clearly distinguished from background autofluorescence, as highly fluorescent and punctate within the cells. This demonstrates the bimane fluorescence is sufficient and further modification – such as incorporation of an auxiliary tag – is not necessary to study bimane-constrained peptides in a common cell-based experiment.

2.4 Experimental

2.4.1 General information

Unless otherwise indicated, all starting materials were purchased from commercial sources and used without further purification. ^1H and ^{13}C NMR spectra were recorded on an Agilent 500 MHz or a Varian Inova 600 MHz instrument in 10% aq. D_2O at 298K, pH ~5, and referenced to DSS at 0 ppm. ES suppression sequences were used. ROESY and zTOCSY were obtained for all compounds with 300 ms mixing times. gHMBC and HSQC experiments were also collected with extended relaxation time of 2 s. Chemical shifts are reported in ppm (δ). Signals are reported as s (singlet), br. s (broad singlet), d (doublet), t (triplet) or m (multiplet). Full ^1H assignments are included in the SI (Tables S2-11). High-resolution mass spectra were collected using an Agilent 6230 ESI-TOF LCMS. RP-HPLC solvents were (A) H_2O with 0.1% TFA and (B) ACN with 0.1% TFA. Purity of all compounds was confirmed by analytical RP-HPLC on an

Agilent 1260 HPLC equipped with a Phenomenex Luna C18(2) column (250 x 4.6mm) over a gradient of 0-100% B (15 min). Purification was carried out by semi-preparative HPLC on a Gilson GX-Prep RP-HPLC system, or using Supelco SPE cartridges as specified in the individual compound sections. All graphs were generated using GraphPad Prism 8 software. NMR generated structures were visualised using PyMOL for the final Figures. Fluorescence spectra were collected on a BioTeK H4 Synergy Plate Reader.

2.4.2 Synthesis and characterisation

All peptides were synthesised by the Fmoc solid-phase peptide synthesis protocol detailed below, with L-amino-acids unless otherwise specified, and *N*-terminally acetylated. Peptides were then cleaved (and simultaneously globally deprotected) from the resin and cyclised in solution. Alternatively, following acetylation, the cysteine residues' side-chain were selectively deprotected and the peptide cyclised on resin prior to cleavage from the resin. See Scheme 1 and the following general procedures.

General procedure for solid-phase peptide synthesis:

Rink Amide PL resin (250 mg, 0.4 mmol/g, Agilent) was swollen in 1:1 DMF/DCM (10 mL) for 15 min. The Fmoc-protecting group was removed by treatment of the resin with a solution of 20% piperidine and 0.1 M HOBt in DMF (5 mL) for 15 min. The solution was drained and the resin washed with DMF (3 x 5 mL). Amino-acid couplings were achieved by addition of a solution of Fmoc-protected amino-acid (5 equiv), HATU (5 equiv) and DIPEA (10 equiv) in DMF (5 mL), to the resin and stirred intermittently for 1 h. The solution was drained and the resin washed with DMF (5 x 5 mL). The *N*-terminal Fmoc-protecting group was removed by treatment of the resin with a solution of 20% piperidine and 0.1 M HOBt in DMF (5 mL) for 15 min, the solution was drained and the resin washed with DMF (5 x 5 mL). A TNBS test* was used to verify each coupling (negative/colourless) and deprotection (positive/red) step, with steps repeated as necessary. Successive couplings and Fmoc-deprotections were repeated to achieve the desired sequence.

*TNBS Test:[76] A small spatula of swollen resin taken out and 1 drop each of TNBS (100 µL 5% w/v picrylsulfonic/trinitrobenzenesulfonic acid in H_2O added to 900 µL of DMF) and DIPEA solutions (100 µL in 900 µL of DMF) added and allowed to develop for 1 min. Clear/yellow beads indicated no free amine (negative), while red/orange beads showed free amine was present (positive).

General procedure for acetylation:

After the final Fmoc-deprotection, the *N*-terminus was protected with an acetyl functionality by reaction with acetic anhydride (470 µL) and DIPEA (870 µL) in DMF (10 mL) for 15 min. The resin was washed with DMF (3 x 5 mL), and DCM (5 x 5 mL) then dried with diethyl ether (3 x 5 mL). Remaining ether was allowed to evaporate overnight before the peptide was cleaved.

General procedure for cleavage and isolation:

The peptide was cleaved from the resin by addition of 92.5:2.5:2.5:2.5 TFA/TIPS/DODT/H_2O (10 mL) and rocked for 1 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. 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 white fluffy powder.

Bimane attachment of unprotected peptide in solution:

Linear peptide (20 mg, 0.5 mg/mL final) was dissolved in 1:1 ACN/H_2O (4 mL), and added to 100 mM phosphate, pH 7-8 (4 mL), [TFE (4 mL) for peptides **1**-**4**, **6** and **7**] and made up to 36 mL with water before dibromobimane (peptides **1**-**9**, 1 equiv) – or monobromobimane in the case of **10** – dissolved in methanol (4 mL) was added and the solution rocked for 30 min. The solution was then further diluted with water and lyophilized to give the crude bimane-peptide as a pale yellow fluffy powder. Peptides were purified by RP chromatography over a gradient of ACN in water as specified below. Pure fractions were combined and lyophilized to give the final purified product as a pale yellow fluffy powder.

Bimane attachment on protected resin-anchored peptide:

The bimane was alternately incorporated whilst the peptide was attached to the resin in the case of peptides **1**, **3** and **8** by incorporation of Fmoc-Cys(Mmt)-OH in place of Fmoc-Cys(Trt)-OH. After the *N*-terminal acetylation, the Mmt groups were removed by repetitive treatment of the resin with 2% TFA in DCM (5 mL) for 1 min, followed by washing with DCM (3 x 5 mL). Treatments were repeated until the solution no longer turned yellow on addition to the resin (~ 150-200 mL total). The resin was then further washed with DCM (5 x 5 mL) and DMF (5 x 5 mL), before a solution of dibromobimane (2 equiv) and DIPEA (4 equiv) in DMF (5 mL) was added to the resin and reacted for 3 h with intermittent stirring. The solution was then removed and the resin washed with DMF (5 x 5 mL) and DCM (5 x 5 mL), then dried with diethyl ether (3 x 5 mL). The bimane-peptide was then cleaved from the resin and isolated as above to give the products as a fluffy yellow powder. Peptides were purified by RP chromatography as specified below. Pure fractions were combined and lyophilized to give the final purified product as a pale yellow fluffy powder.

Characterisation

Peptide 1: Ac-CAAAC-Am (1-5 [*i*-*i*+4] bimane cyclised). Synthesis was carried out per the general procedure with Fmoc-Cys(Trt)-OH and Fmoc-Ala-OH. The crude product was purified by use of a SPE-C18 column, over a gradient of 6% ACN in water (20 mL), 8% (20 mL), 10% (40 mL), 11% (40 mL), 30% (20 mL), 100% (10 mL) and eluted over 10-11% ACN in water. ^{1}H NMR (10% aq. D_2O, 500 MHz): δ 8.59 (d, *J*=5.6 Hz, 1 H), 8.44 (d, *J*=6.8 Hz, 1 H), 8.12 - 8.18 (m, 2 H), 8.09 (d, *J*=5.6 Hz, 1 H), 7.54 (br. s., 1 H), 7.28 (br. s., 1 H), 4.22 - 4.35 (m, 3 H), 3.87 - 4.11 (m, 4 H), 3.10 - 3.25 (m, 3 H), 2.91 - 3.06 (m, 1 H), 2.03 (s, 3 H), 1.89 (s, 3 H), 1.88 (s, 3 H), 1.31 - 1.49 (m, 9 H) ppm; ^{13}C NMR (10% aq. D_2O, 126 MHz): δ 177.8, 177.7, 177.6, 177.1, 176.8, 174.7, 151.7, 117.6, 117.4, 93.6, 56.3, 56.1, 53.0, 36.5, 36.4, 28.6, 24.4, 19.4, 19.1, 18.4, 9.1 ppm; HRMS (ESI+) Expected $[M+H]^+$ for $C_{27}H_{38}N_8O_8S_2$: 667.2332, observed: $[M+H]^+$ 667.2330.

Peptide 2: Ac-cAAAC-Am (1-5 [*i*-*i*+4] bimane cyclised). Synthesis was carried out per the general procedure with Fmoc-Ala-OH, Fmoc-L-Cys(Trt)-OH and Fmoc-D-Cys(Trt)-OH. The crude was purified by use of a SPE-C18 column, over a gradient of 10% ACN in water (40 mL), 11% (60 mL), 50% (10 mL) and eluted over 11% ACN in water. ^{1}H NMR (599 MHz, 10% aq. D_2O): δ 8.54 (d, *J*=4.7 Hz, 1 H), 8.49 (d, *J*=6.7 Hz, 1 H), 8.25 (d, *J*=5.3 Hz, 1 H), 8.10 (d, *J*=7.6 Hz, 1 H), 7.99 (d, *J*=6.5 Hz, 1 H), 7.36 - 7.54 (m, 1 H), 7.29 (s, 1 H), 4.59 - 4.63 (m, 1 H), 4.55 - 4.59 (m, 1 H), 4.32 - 4.39 (m, 1 H), 4.22 - 4.28 (m, 1 H), 4.10 - 4.16 (m, 1 H), 3.91 - 4.05 (m, 4 H), 3.12 - 3.26 (m, 3 H), 3.04 - 3.10 (m, 1 H), 2.01 (s, 3 H), 1.85 - 1.93 (m, 6 H), 1.40 (s, 9 H) ppm; ^{13}C NMR (151 MHz, 10% aq. D_2O): δ 178.0, 177.8, 177.0, 176.7, 174.4, 165.4, 165.3, 152.2, 151.9, 117.9, 117.3, 55.9, 55.8, 54.1, 53.4, 52.7, 36.4, 35.7, 28.4, 24.5, 18.9, 18.8, 9.1, 9.1 ppm; HRMS (ESI+) Expected $[M+H]^+$ for $C_{27}H_{38}N_8O_8S_2$: 667.2332, observed: $[M+H]^+$ 667.2298.

Peptide 3: Ac-CRAAARAC-Am (1-8 [*i*-*i*+7] bimane cyclised). Synthesis was carried out per the general procedure with Fmoc-Cys(Trt)-OH, Fmoc-Arg(Pbf)-OH and Fmoc-Ala-OH. The crude product was purified by semi-preparative HPLC on a Supelco C18 250x10 mm 5 µM column, over a gradient of 25-40% ACN in water over 15 minutes and eluted at 4.5 min. ^{1}H NMR (500 MHz, 10% aq. D_2O) δ 8.65 (d, *J* = 7.3 Hz, 1 H), 8.48 (d, *J* = 6.6 Hz, 1 H), 8.05 - 8.22 (m, 6 H), 7.43 (s, 1 H), 7.15 - 7.29 (m, 3 H), 4.14 - 4.58 (m, 8 H), 3.82 - 4.13 (m, 4 H), 3.11 - 3.34 (m, 8 H), 2.89 - 3.10 (m, 4 H), 1.53 - 2.16 (m, 17 H), 1.31 - 1.52 (m, 12 H) ppm; ^{13}C NMR (126 MHz, 10% aq. D_2O) δ 178.0, 177.5, 176.9, 176.8, 176.5, 176.1, 175.9, 175.8, 174.7, 165.5, 165.1, 159.7, 151.2, 150.8, 120.2, 117.9, 117.3, 56.3, 56.1, 56.0, 55.7, 53.6, 52.8, 52.7, 52.5, 43.3 (br), 30.3, 27.2, 27.0, 24.4, 19.4, 19.1, 19.0, 18.8, 9.2, 9.1 ppm; HRMS (ESI+) Expected $[M+3H]^{3+}$ for $C_{42}H_{67}N_{17}O_{11}S_2$: 350.8294, observed: $[M+H]^+$ 350.8298.

Peptide 4: Ac-cRAAARAC-Am (1-8 [*i*-*i*+7] bimane cyclised). Synthesis was carried out per the general procedure with Fmoc-Ala-OH, Fmoc-Arg(Pbf)-OH, Fmoc-L-Cys(Trt)-OH and Fmoc-D-Cys(Trt)-OH. The crude product was purified by semi-preparative HPLC on a Supelco C18 250x10 mm 5 µM column, over a gradient of 25-40% ACN in water over 15 minutes and eluted at 4.5 min. ^{1}H NMR (599 MHz, 10% aq. D_2O) δ 8.58 (d, *J* = 6.2 Hz, 1H), 8.51 (d, *J* = 6.2 Hz, 1 H), 8.24 (d, *J* = 7.3 Hz, 1 H), 8.14 - 8.19 (m, 3 H), 8.12 (d, *J* = 5.3 Hz, 1 H), 7.94 (d, *J* = 6.2 Hz, 1 H), 7.49 (s, 1H), 7.28 (s, 1 H), 7.17-7.24 (m, 2 H), 4.48 - 4.54 (m, 1 H), 4.16 - 4.37 (m, 7 H), 3.93 - 4.06 (m, 4 H), 2.92 - 3.28 (m, 8 H), 2.00 - 2.06 (m, 3 H), 1.54 - 1.96 (m, 12H), 1.33 - 1.48 (m, 12 H) ppm; ^{13}C NMR (151 MHz, 10% aq. D_2O) δ 177.9, 177.6, 177.1, 176.8, 176.1, 175.1, 165.6, 165.4, 165.0, 159.4, 151.1, 119.8, 117.9, 117.2, 116.3, 56.9, 56.4, 56.1, 55.5, 53.2, 52.7, 52.6, 52.2, 43.1, 30.3, 30.0, 27.6, 27.2, 26.9, 24.3, 19.0, 18.9, 18.8, 18.7, 9.0, 8.9, 8.4 ppm; HRMS (ESI+) Expected [M+2H]$^{2+}$ for $C_{42}H_{67}N_{17}O_{11}S_2$: 525.7402, observed [M+2H]$^{2+}$ 525.7394.

Peptide 5: Ac-CAAC-Am (1-4 [*i*-*i*+3] bimane cyclised). Synthesis was carried out per the general procedure with Fmoc-Cys(Trt)-OH and Fmoc-Ala-OH. The crude product was purified by use of a SPE-C18 column, over a gradient of 5% ACN in water (20 mL), 6% (20 mL), 7% (20 mL), 8% (20 mL), 9% (50 mL), 10% (20 mL), 30% (20 mL), 100% (10 mL) and eluted over 9-10% ACN in water. ^{1}H NMR (10% aq. D_2O, 500 MHz): δ 8.78 (d, *J*=5.4 Hz, 1 H), 8.35 (d, *J*=6.8 Hz, 1 H), 8.06 (d, *J*=7.8 Hz, 1 H), 7.94 (d, *J*=6.1 Hz, 1 H), 7.47 (br. s., 1 H), 7.25 (br. s., 1 H), 4.56 - 4.67 (m, 1 H), 4.34 - 4.40 (m, 1 H), 4.22 - 4.30 (m, 1 H), 3.96 - 4.07 (m, 3 H), 3.83 (d, *J*=14.9 Hz, 1 H), 3.21 - 3.27 (m, 1 H), 3.07 - 3.17 (m, 3 H), 2.01 (s, 3 H), 1.90 (s, 3 H), 1.88 (s, 3 H), 1.43 (d, *J*=7.3 Hz, 3 H), 1.35 (d, *J*=7.1 Hz, 3 H) ppm; ^{13}C NMR (10% aq. D_2O, 126 MHz): δ 178.1, 177.1, 176.7, 176.6, 174.8, 165.2, 152.1, 151.6, 117.5, 117.3, 55.6, 54.9, 53.9, 52.7, 36.8, 35.7, 28.7, 27.9, 24.4, 19.2, 18.7, 9.1, 8.8 ppm; HRMS (ESI+) Expected [M+H]$^+$ for $C_{24}H_{33}N_7O_7S_2$: 596.1961, observed: [M+H]$^+$ 596.1956.

Peptide 6: Ac-CARAAC-Am (1-6 [*i*-*i*+5] bimane cyclised). Synthesis was carried out per the general procedure with Fmoc-Cys(Trt)-OH and Fmoc-Ala-OH. The crude product was purified by semi-preparative HPLC on a Supelco C18 250x10 mm 5 μm column, over a gradient of 25-40% ACN in water over 15 minutes, eluting at 4.5 min.^{1}H NMR (10% aq. D_2O, 500 MHz): δ 8.57 (d, *J*=6.1 Hz, 1 H), 8.35 - 8.42 (m, 2 H), 8.15 (d, *J*=5.1 Hz, 1 H), 7.93 (d, *J*=8.1 Hz, 1 H), 7.88 (d, *J*=6.8 Hz, 1 H), 7.44 (br. s., 1 H), 7.25 (br. s., 1 H), 7.21 (t, *J*=5.1 Hz, 1 H), 3.83 - 4.66 (m, 12 H), 2.90 - 3.29 (m, 6 H), 2.03 (s, 3 H), 1.84 – 1.93 (m, 10 H), 1.56 – 1.83 (m, 3 H), 1.32 - 1.47 (m, 9 H) ppm; ^{13}C NMR (10% aq. D_2O,126 MHz): δ 177.5, 176.8, 176.7, 174.5, 165.2, 165.1, 151.6, 117.2, 57.3, 55.7, 52.5, 52.3, 43.1, 36.4, 28.5, 27.8, 27.1, 24.3, 18.8, 18.5, 9.0 ppm; HRMS (ESI+) Expected [M+H]$^+$ for $C_{33}H_{50}N_{12}O_9S_2$: 823.3343, observed: [M+H]$^+$ 823.3336.

Peptide 7: Ac-CARAAAC-Am (1-7 [*i*-*i*+6] bimane cyclised). Synthesis was carried out per the general procedure with Fmoc-Cys(Trt)-OH, Fmoc-Arg(Pbf)-OH and Fmoc-Ala-OH. The crude product was purified by semi-preparative HPLC on a Supelco C18 250x10 mm 5 μM column, over a gradient of 25-40% ACN in water over 25 minutes and eluted at 5 min. ^{1}H NMR (10% aq. D_2O, 500 MHz): δ 8.64 (d, *J*=4.9 Hz, 1 H), 8.43 (d, *J*=6.6 Hz, 1 H), 8.20 (d, *J*=4.9 Hz, 1 H), 8.08 - 8.15 (m, 3 H), 8.05 (d, *J*=7.6 Hz, 1 H), 7.42 (s, 1 H), 7.25 (br. s., 1 H), 7.22 (d, *J*=5.5 Hz, 1 H), 4.16 - 4.31 (m, 3 H), 3.92 - 4.08 (m, 4 H), 2.87 - 3.25 (m, 6 H), 2.03 (s, 3 H), 1.85 - 1.93 (m, 7 H), 1.52 - 1.81 (m, 3 H), 1.34 - 1.48 (m, 9 H) ppm; ^{13}C NMR (10% aq. D_2O,126 MHz): δ 178.0, 177.6, 176.8, 176.0, 172.5, 165.1, 163.6, 159.5, 152.3, 145.3, 117.3, 117.2, 114.0, 82.4, 53.3, 52.7, 52.5, 43.3, 35.5, 28.9, 28.1, 24.3, 21.3, 18.8, 18.7, 18.6, 9.0 ppm; HRMS (ESI+) Expected [M+H]$^+$ for $C_{36}H_{55}N_{13}O_{10}S_2$: 894.3714, observed: [M+H]$^+$ 894.3840.

Peptide 8: Ac-CAC-Am (1-3 [*i*-*i*+2] bimane cyclised). Synthesis was carried out per the general procedure with Fmoc-Cys(Trt)-OH and Fmoc-Ala-OH. The crude product was purified by use of a SPE-C18 column, over a step-wise gradient of 6% ACN in water (20 mL), 8% (20 mL), 10% (40 mL), 11% (40 mL), 30% (20 mL), 100% (10 mL), eluting over 10-11% ACN. ^{1}H NMR (10% aq. D_2O, 500 MHz): δ 8.77 (d, *J*=6.4 Hz, 1 H), 8.36 (d, *J*=6.8 Hz, 1 H), 7.99 (d, *J*=8.1 Hz, 1 H), 7.42 (br. s., 1 H), 7.25 (br. s., 1 H), 4.51 - 4.59 (m, 1 H), 4.37 - 4.46 (m, 1 H), 3.87 - 4.07 (m, 4 H), 3.27 (dd, *J*=13.8, 6.5 Hz, 1 H), 2.99 - 3.18 (m, 3 H), 2.01 (s, 3 H), 1.91 (s, 3 H), 1.88 (s, 3 H), 1.36 (d, *J*=7.1 Hz, 3 H) ppm; ^{13}C NMR (10% aq. D_2O, 126 MHz): δ 177.2, 177.0, 176.6, 174.7, 165.7, 152.9, 152.5, 118.5, 118.0, 55.6, 55.2, 52.8, 36.4, 35.9, 28.9, 28.3, 24.4, 18.7, 9.4, 9.1 ppm; HRMS (ESI+) Expected [M+H]$^+$ for $C_{21}H_{28}N_6O_6S_2$: 525.1590, observed: [M+H]$^+$ 525.1575.

Peptide 9: Ac-HcyAlaHcy-Am (1-3 [*i*-*i*+2] bimane cyclised). Synthesis was carried out per the general procedure with Fmoc-Ala-OH and Fmoc-Hcy(Trt)-OH, apart from coupling of Hcy(Trt), which was achieved by addition of

Fmoc-Hcy(Trt)-OH (2 equiv), HATU (2 equiv) and DIPEA (4 equiv) in DMF (5 mL) to the resin, and stirring intermittently for 16 h. The crude product was purified by use of a SPE-C18 column, over a gradient of 7% ACN in water (20 mL), 8% (20 mL), 9% (60 mL), 10% (20 mL), 30% (20 mL), 100% (10 mL) and eluted over 9-10% ACN. ^{1}H NMR (10% aq. D_2O, 500 MHz): δ 8.64 (d, *J*=7.8 Hz, 1 H), 8.47 (d, *J*=7.8 Hz, 1 H), 8.27 (d, *J*=6.6 Hz, 1 H), 7.64 (br. s., 1 H), 7.15 (br. s., 1 H), 3.84 - 4.16 (m, 4 H), 2.61 - 2.95 (m, 4 H), 2.05 - 2.28 (m, 4 H), 2.02 (s, 3 H), 1.86 (s, 6 H), 1.36 (d, *J*=7.1 Hz, 3 H) ppm; ^{13}C NMR (126 MHz, 10% aq. D_2O) δ 178.1, 177.0, 176.9, 175.3, 165.4, 150.8, 150.4, 117.0, 55.6, 54.9, 51.9, 34.5, 33.4, 33.2, 32.0, 30.6, 28.2, 27.4, 24.5, 19.1, 8.8, 8.7 ppm; HRMS (ESI+) Expected $[M+H]^+$ for $C_{23}H_{32}N_6O_6S_2$: 553.1903, observed: $[M+H]^+$ 553.1866.

Peptide 10: Ac-AAC-Am (Cys reacted bimane, acyclic). Synthesis was carried out per the general procedure with Fmoc-Ala-OH and Fmoc-Cys(Trt)-OH. The crude product was purified by use of a SPE-C18 column, over a gradient of 5% ACN in water (100 mL), 6% (100 mL), 7% (50 mL), 10% (10 mL), 100% (10 mL) and eluted over 5-7% ACN in water. ^{1}H NMR (10% aq. D_2O, 500 MHz): δ 8.36 (d, *J*=6.1 Hz, 1 H), 8.19 - 8.31 (m, 2 H), 7.51 (br. s., 1 H), 7.21 (br. s., 1 H), 4.43 - 4.50 (m, 1 H), 4.18 - 4.38 (m, 2 H), 3.85 - 3.98 (m, 2 H), 3.11 - 3.24 (m, 1 H), 2.86 - 3.05 (m, 1 H), 2.43 (s, 3 H), 2.01 (s, 3 H), 1.84 - 1.92 (m, 3 H), 1.78 (s, 3 H), 1.30 - 1.40 (m, 6 H) ppm; ^{13}C NMR (10% aq. D_2O, 126 MHz): δ 178.1, 177.7, 176.9, 176.8, 165.7, 165.0, 152.4, 150.8, 116.2, 114.2, 110.0, 55.3, 52.6, 52.5, 35.6, 27.5, 24.3, 19.1, 18.9, 13.8, 8.8, 8.3 ppm; HRMS (ESI+) Expected $[M+H]^+$ for $C_{21}H_{30}N_6O_6S$: 495.2026, observed: $[M+H]^+$ 495.2019.

2.4.3 Fluorescence assays

All fluorescence values were measured (from the bottom) in a black polystyrene, clear bottomed, tissue culture treated polystyrene 96-well plate (Corning Inc. costar® 3605) at an excitation wavelength of 385 nm and emission wavelength of 477 nm unless otherwise specified, on an H4 Synergy Plate Reader, with Xenon light source and a slit width of 9.0. Gain was varied depending on the concentration range of the samples analysed.

2.4.4 Bimane cyclisation optimisation experiments

A stock solution of the linear precursor peptide to **1** was prepared in 50% aq. ACN, and dibromobimane stock solutions in 100% methanol. Each reaction solution contained a final concentration of 10% methanol (from dibromobimane stock), 5% ACN (from peptide stock) and 10% TFE. A stock solution of 100 mM phosphate was added such that the final phosphate concentration of the reaction was 10 mM. Peptide and dibromobimane were added in the required ratios, as shown in the Figure S1. However dibromobimane was always the final reagent added to each well and the 96-well plate scanned by the plate reader immediately after its addition to the last well. The fluorescence was measured as described above with a gain of 60.

2.4.5 NMR structure calculations

CCPNMR Analysis[66] was used to assign the ROESY spectra (as collected above) and generate the distance restraints. Upper distance limits were based on the peak volume and all lower limits set to zero. Dihedral angle restraints were generated by the DANGLE functionality[65] within Analysis. Structure calculations were performed using Xplor-NIH v2.48[63, 64] using the general protocol from Bermejo and Schwieters[62] and NOE and dihedral angle restraints (Table S12). Parameters and topology for the bimane were generated first using PRODRG,[77] before manually adjusting distances and angles based on those from small molecule bimane analogues.[78] Initial structures were generated using full ambiguous restraints, which were refined manually by iterative analysis of the 20 lowest energy structures (of 100). Final structures contained no NOE violations > 0.5 Å, and no angle violations > 5 Å.

The 20 lowest energy structures were validated by analysis with the PSVS server.[68] It was shown that the RMSD for bond angles was no greater than 1.5° and bond length RMSD no greater than 0.007 Å for all structures. PROCHECK showed that for all peptides at least 99% of residues lay within the most favoured region of Ramachandran space, and none in generous or disallowed regions; the only exception was peptide **6** for which 56% were in the most favoured region and 44% in the allowed region. Φ and ψ angles for most peptides were analysed using the STRIDE program,[79] and the average values are displayed in Table S13, along with the secondary structure designations most commonly assigned for each residue. Heavy atom RMSD values for the 20 lowest energy conformers are shown in Table S17.

2.4.6 Circular dichroism experiments

Precise peptide concentrations were determined by NMR spectroscopy on an Agilent 500 MHz spectrometer. A wet1D ^{1}H spectrum with 256 scans and pulse angle of 90° was acquired for each sample, and integration of the aliphatic signals below 4 ppm used to determine the sample concentration by comparison to a caffeine standard using the 'determine concentration' function in vNMRj. The samples were then diluted into 10 mM phosphate (pH 7.2) to a final concentration of 250 µM. Data were collected on a Jasco J-810 spectropolarimeter in a 1 mm cuvette (UniSA Biophysical Characterisation Lab). Each spectrum is the average of 8 scans with a scan rate of 50 nm/min and pitch of 1 nm; the final spectra were smoothed with the in-built Savitzky-Golay function at a convolution width of 7.

2.4.7 Cell imaging

NIH/3T3 cells (ATCC CRL-1658) were maintained in Dulbecco's Modified Eagle Media (DMEM)/Glutamax supplemented with 10% fetal bovine serum (complete DMEM). Cells were incubated at 37°C with 6% CO_2 in a humidified incubator. For cellular uptake experiments NIH/3T3 cells were plated at 0.7 x 10^6 cells/2 mL of complete DMEM in 35 mm glass bottomed confocal dishes (Cell E&G, Houston TX, USA). Twenty-four hours post plating complete DMEM was removed and cells were washed thoroughly with pre-warmed, serum free DMEM and peptides added at 5 µM in serum free DMEM. Cells were incubated with peptide **7**, **9**, or **10** in serum free DMEM for 4 h at 37°C with 6% CO_2. Following this, cells were again washed with serum free DMEM, then 2 mL of complete DMEM was added and cells incubated for a further 16 h at 37°C with 6% CO_2 in a humidified incubator. For the negative control ('Control') cells were treated as described above but not exposed to bimane-modified peptide. Cells were imaged by confocal microscopy (FV10i) at ex. 405 nm and em. 420-520 nm using identical settings between treatments.

2.5 Acknowledgements

We thank Philip Clements for his assistance in collecting NMR spectra. The research was supported by the ARC Centre of Excellence in Nanoscale BioPhotonics (CNBP) (CE140100003). The facilities of the OptoFab node of the Australian National Fabrication Facility (ANFF) and associated Commonwealth and SA State Government funding are also gratefully acknowledged. Circular Dichroism was carried out at the UniSA Sansom Institute for Health Research Biophysical Characterisation Facility.

REFERENCES

1. A. M. Watkins and P. S. Arora, Structure-based inhibition of protein-protein interactions, *Eur. J. Med. Chem.*, 2015, **94**, 480-488.
2. N. Tsomaia, Peptide therapeutics: Targeting the undruggable space, *Eur. J. Med. Chem.*, 2015, **94**, 459-470.
3. T. A. Hill, N. E. Shepherd, F. Diness and D. P. Fairlie, Constraining cyclic peptides to mimic protein structure motifs, *Angew. Chem. Int. Ed. (English)*, 2014, **53**, 13020-13041.
4. A. Jamieson and N. Robertson, Regulation of protein-protein interactions using stapled peptides, *Rep. Org. Chem*, 2015, **5**, 65-74.
5. V. Azzarito, K. Long, N. S. Murphy and A. J. Wilson, Inhibition of alpha-helix-mediated protein-protein interactions using designed molecules, *Nature Chem.*, 2013, **5**, 161-173.
6. K. L. Keeling, O. Cho, D. B. Scanlon, G. W. Booker, A. D. Abell and K. L. Wegener, The key position: influence of staple location on constrained peptide conformation and binding, *Org. Biomol. Chem.*, 2016, **14**, 9731-9735.
7. K. L. Wegener, A. E. McGrath, N. E. Dixon, A. J. Oakley, D. B. Scanlon, A. D. Abell and J. Bruning, Rational design of a 310-helical PIP-box mimetic targeting PCNA - the human sliding clamp, *Chem. Eur. J.*, 2018, **24**, 11325-11331.
8. K. Estieu-Gionnet and G. Guichard, Stabilized helical peptides: overview of the technologies and therapeutic promises, *Expert Opin. Drug Dicov.*, 2011, **6**, 937-936.
9. J. Yu, J. R. Horsley and A. D. Abell, A controllable mechanistic transition of charge transfer in helical peptides: from hopping to superexchange, *RSC Adv.*, 2017, **7**, 42370-42378.
10. J. R. Horsley, J. Yu, K. E. Moore, J. G. Shapter and A. D. Abell, Unraveling the interplay of backbone rigidity and electron rich side-chains on electron transfer in peptides: the realization of tunable molecular wires, *J. Am. Chem. Soc.*, 2014, **136**, 12479-12488.
11. Ø. Jacobsen, H. Maekawa, N.-H. Ge, C. H. Görbitz, P. I. Rongved, O. P. Ottersen, M. Amiry-Moghaddam and J. Klaveness, Stapling of a 310-Helix with Click Chemistry, *J. Org. Chem.*, 2011, **76**, 1228-1238.
12. A. K. Boal, I. Guryanov, A. Moretto, M. Crisma, E. L. Lanni, C. Toniolo, R. H. Grubbs and D. J. O'Leary, Facile and E-Selective Intramolecular Ring-Closing Metathesis Reactions in 310-Helical Peptides: A 3D Structural Study, *J. Am. Chem. Soc.*, 2007, **129**, 6986-6987.
13. N. E. Shepherd, H. N. Hoang, G. Abbenante and D. P. Fairlie, Single Turn Peptide Alpha Helices with Exceptional Stability in Water, *J. Am. Chem. Soc.*, 2005, **127**, 2974-2983.
14. S. A. Jones, P. M. Neilsen, L. Siew, D. F. Callen, N. E. Goldfarb, B. M. Dunn and A. D. Abell, A template-based approach to inhibitors of calpain 2, 20S proteasome, and HIV-1 protease, *ChemMedChem*, 2013, **8**, 1918-1921.
15. A. D. Pehere, X. Zhang and A. D. Abell, Macrocyclic Peptidomimetics Prepared by Ring-Closing Metathesis and Azide–Alkyne Cycloaddition, *Aust. J. Chem.*, 2017, **70**, 138-151.
16. A. D. Abell, M. A. Jones, J. M. Coxon, J. D. Morton, S. G. Aitken, S. B. McNabb, H. Y. Lee, J. M. Mehrtens, N. A. Alexander, B. G. Stuart, A. T. Neffe and R. Bickerstaffe, Molecular modeling, synthesis, and biological evaluation of macrocyclic calpain inhibitors, *Angew. Chem. Int. Ed. (English)*, 2009, **48**, 1455-1488.
17. K. C. Chua, M. Pietsch, X. Zhang, S. Hautmann, H. Y. Chan, J. B. Bruning, M. Gutschow and A. D. Abell, Macrocyclic protease inhibitors with reduced peptide character, *Angew. Chem. Int. Ed. (English)*, 2014, **53**, 7828-7831.
18. J. W. Taylor, The Synthesis and Study of Side-Chain Lactam-Bridged Peptides, *Pep. Sci.*, 2002, **66**, 49-75.
19. M. J. Klein, S. Schmidt, P. Wadhwani, J. Burck, J. Reichert, S. Afonin, M. Berditsch, T. Schober, R. Brock, M. Kansy and A. S. Ulrich, Lactam-Stapled Cell-Penetrating Peptides: Cell Uptake and Membrane Binding Properties, *J. Med. Chem.*, 2017, **60**, 8071-8082.
20. H. N. Hoang, R. W. Driver, R. L. Beyer, T. A. Hill, A. D. de Araujo, F. Plisson, R. S. Harrison, L. Goedecke, N. E. Shepherd and D. P. Fairlie, Helix Nucleation by the Smallest Known alpha-Helix in Water, *Angew. Chem. Int. Ed. (English)*, 2016, **55**, 8275-8279.
21. Y. W. Kim, T. N. Grossmann and G. L. Verdine, Synthesis of all-hydrocarbon stapled alpha-helical peptides by ring-closing olefin metathesis, *Nat. Protoc.*, 2011, **6**, 761-771.
22. C. E. Schafmeister, J. Po and G. L. Verdine, An All-Hydrocarbon Cross-Linking System for Enhancing the Helicity and Metabolic Stability of Peptides, *J. Am. Chem. Soc.*, 2000, **122**, 5891-5892.
23. H. E. Blackwell, J. D. Sadowsky, R. J. Howard, J. N. Sampson, J. A. Chao, W. E. Steinmetz, D. J. O'Leary and R. H. Grubbs, Ring-Closing Metathesis of Olefinic Peptides: Design, Synthesis, and Structural Characterisation of Macrocyclic Helical Peptides, *J. Org. Chem.*, 2001, **66**, 5291-5302.
24. S. A. Kawamoto, A. Coleska, X. Ran, H. Yi, C. Y. Yang and S. Wang, Design of triazole-stapled BCL9 alpha-helical peptides to target the beta-catenin/B-cell CLL/lymphoma 9 (BCL9) protein-protein interaction, *J. Med. Chem.*, 2012, **55**, 1137-1146.
25. Y. H. Lau, P. de Andrade, N. Skold, G. J. McKenzie, A. R. Venkitaraman, C. Verma, D. P. Lane and D. R. Spring, Investigating peptide sequence variations for 'double-click' stapled p53 peptides, *Org. Biomol. Chem.*, 2014, **12**, 4074-4077.
26. A. D. Pehere, M. Pietsch, M. Gutschow, P. M. Neilsen, D. S. Pedersen, S. Nguyen, O. Zvarec, M. J. Sykes, D. F. Callen and A. D. Abell, Synthesis and extended activity of triazole-containing macrocyclic protease inhibitors, *Chem. Eur. J.*, 2013, **19**, 7975-7981.
27. A. M. Spokoyny, Y. Zou, J. J. Ling, H. Yu, Y. S. Lin and B. L. Pentelute, A perfluoroaryl-cysteine S(N)Ar chemistry approach to unprotected peptide stapling, *J. Am. Chem. Soc.*, 2013, **135**, 5946-5949.

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

29. L. D. Walensky and G. H. Bird, Hydrocarbon-stapled peptides: principles, practice, and progress, *J. Med. Chem.*, 2014, **57**, 6275-6288.

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

31. S. R. Perry, T. A. Hill, A. D. de Araujo, H. N. Hoang and D. P. Fairlie, Contiguous hydrophobic and charged surface patches in short helix-constrained peptides drive cell permeability, *Org. Biomol. Chem.*, 2018, **16**, 367-371.

32. D. Migoń, D. Neubauer and W. Kamysz, Hydrocarbon Stapled Antimicrobial Peptides, *The Protein Journal*, 2018, **37**, 2-12.

33. C. Wu, H. N. Hoang, L. Liu and D. P. Fairlie, Glucuronic acid as a helix-inducing linker in short peptides, *Chem. Commun.*, 2018, **54**, 2162-2165.

34. L. Peraro, T. R. Siegert and J. A. Kritzer, Conformational Restriction of Peptides Using Dithiol Bis-Alkylation, *Methods Enzymol.*, 2016, **580**, 303-332.

35. L. Peraro, Z. Zou, K. M. Makwana, A. E. Cummings, H. L. Ball, H. Yu, Y. S. Lin, B. Levine and J. A. Kritzer, Diversity-Oriented Stapling Yields Intrinsically Cell-Penetrant Inducers of Autophagy, *J. Am. Chem. Soc.*, 2017, **139**, 7792-7802.

36. G. Zhang, F. Barragan, K. Wilson, N. Levy, A. Herskovits, M. Sapozhnikov, Y. Rodriguez, L. Kelmendi, H. Alkasimi, H. Korsmo, M. Chowdhury and G. Gerona-Navarro, A Solid-Phase Approach to Accessing Bisthioether-Stapled Peptides Resulting in a Potent Inhibitor of PRC2 Catalytic Activity, *Angew. Chem. Int. Ed. (English)*, 2018, **57**, 17073-17078.

37. D. P. Fairlie and A. D. de Araujo, Stapling peptides using cysteine crosslinking, *Biopolymers*, 2016, **106**, 843-852.

38. A. Dantas de Araujo, S. R. Perry and D. P. Fairlie, Chemically Diverse Helix-Constrained Peptides Using Selenocysteine Crosslinking, *Org. Lett.*, 2018, **205**, 1453-1456.

39. J.-R. Bertrand, C. Malvy, T. Auguste, G. K. Toth, O. Kiss-Ivankovtis, E. Illyes, M. Hollosi, S. Bottka and I. Lacsko, Synthesis and Studies on Cell-Penetrating Peptides, *Bioconjugate Chem.*, 2009, **20**, 1307-1314.

40. A. Mishra, G. H. Lai, N. W. Schmidt, V. Z. Sun, A. R. Rodriguez, R. Tong, L. Tang, J. Cheng, T. J. Deming, D. T. Kamei and G. C. Wong, Translocation of HIV TAT peptide and analogues induced by multiplexed membrane and cytoskeletal interactions, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 16883-16888.

41. H. H. Szeto, P. W. Schiller, K. Zhao and G. Luo, Fluorescent dyes alter intracellular targeting and function of cell-penetrating tetrapeptides, *The FASEB Journal*, 2005, **19**, 118-120.

42. C. A. Puckett and J. K. Barton, Fluorescein Redirects a Ruthenium-Octaarginine Conjugate to the Nucleus, *J. Am. Chem. Soc.*, 2009, **131**, 8738-8739.

43. A. Szabo, T. Szendi-Szatmari, L. Ujlaky-Nagy, I. Radi, G. Vereb, J. Szollosi and P. Nagy, The Effect of Fluorophore Conjugation on Antibody Affinity and the Photophysical Properties of Dyes, *Biophys. J.*, 2018, **114**, 688-700.

44. S. F. Hedegaard, M. S. Derbas, T. K. Lind, M. R. Kasimova, M. V. Christensen, M. H. Michaelsen, R. A. Campbell, L. Jorgensen, H. Franzyk, M. Cárdenas and H. M. Nielsen, Fluorophore labeling of a cell-penetrating peptide significantly alters the mode and degree of biomembrane interaction, *Sci. Rep.*, 2018, **8**, 6327.

45. E. Jamasbi, G. D. Ciccotosto, J. Tailhades, R. M. Robins-Browne, C. L. Ugalde, R. A. Sharples, N. Patil, J. D. Wade, M. A. Hossain and F. Separovic, Site of fluorescent label modifies interaction of melittin with live cells and model membranes, *Biochim. Biophys. Acta, Biomembr.*, 2015, **1848**, 2031-2039.

46. E. M. Kosower, B. Pazhenchevsky and E. Herschkowitz, 1,5,-Diazabicyclo[3.3.0]octadienediones(9,10-Dioxabimanes). Strong Fluorescent Syn Isomers, *J. Am. Chem. Soc.*, 1978, **100**, 6516-6518.

47. N. S. Kosower, E. M. Kosower, G. L. Newton and H. M. Ranney, Bimane fluorescent labels: Labeling of normal human red cells under physiological conditions, *Proc. Natl. Acad. Sci. U. S. A.*, 1979, **76**, 3382-3386.

48. E. M. Kosower and N. S. Kosower, Bimane Derivative as Fluorescent Probes for Biological Macromolecules, *Methods Enzymol.*, 1985.

49. P. Danielsohn and A. Nolte, Bromobimanes- fluorescent labeling agents for histochemical detection of sulfur containing neuropeptides in semithin sections, *Histochemistry*, 1987, **86**, 281-285.

50. E. M. Kosower and N. S. Kosower, Bromobimane Probes for Thiols, *Methods Enzymol.*, 1995, **251**, 133-148.

51. X. Shen, C. B. Pattillo, S. Pardue, S. C. Bir, R. Wang and C. G. Kevil, Measurement of plasma hydrogen sulfide in vivo and in vitro, *Free Radical Biol. Med.*, 2011, **50**, 1021-1031.

52. L. A. Montoya, X. Shen, J. J. McDermott, C. G. Kevil and M. D. Pluth, Mechanistic investigations reveal that dibromobimane extrudes sulfur from biological sulfhydryl sources other than hydrogen sulfide, *Chem. Sci.*, 2015, **6**, 294-300.

53. B. Krishnan, A. Szymanska and L. M. Gierasch, Site-specific fluorescent labeling of poly-histidine sequences using a metal-chelating cysteine, *Chem. Biol. Drug Des.*, 2007, **69**, 31-40.

54. D. Ho, M. R. Lugo and A. R. Merrill, Harmonic analysis of the fluorescence response of bimane adducts of colicin E1 at helices 6, 7, and 10, *J. Biol. Chem.*, 2013, **288**, 5136-5148.

55. A. M. Jones Brunette and D. L. Farrens, Distance mapping in proteins using fluorescence spectroscopy: tyrosine, like tryptophan, quenches bimane fluorescence in a distance-dependent manner, *Biochemistry*, 2014, **53**, 6290-6301.

56. M. Masureel, Y. Zou, L. P. Picard, E. van der Westhuizen, J. P. Mahoney, J. Rodrigues, T. J. Mildorf, R. O. Dror, D. E. Shaw, M. Bouvier, E. Pardon, J. Steyaert, R. K. Sunahara, W. I. Weis, C. Zhang and B. K. Kobilka, Structural insights into binding specificity, efficacy and bias of a beta2AR partial agonist, *Nat. Chem. Biol.*, 2018, **14**, 1059-1066.

57. L. Konermann, Addressing a Common Misconception: Ammonium Acetate as Neutral pH "Buffer" for Native Electrospray Mass Spectrometry, *J. Am. Soc. Mass Spectrom.*, 2017, **28**, 1827-1835.

58. M. M. Madden, C. I. Rivera Vera, W. Song and Q. Lin, Facile synthesis of stapled, structurally reinforced peptide helices via a photoinduced intramolecular 1,3-dipolar cycloaddition reaction, *Chem. Commun.*, 2009, **1**, 5588-5590.

59. A. Kentsis and T. R. Sosnick, Trifluoroethanol Promotes Helix Formation by Destabilizing Backbone Exposure: Desolvation Rather than Native Hydrogen Bonding Defines the Kinetic Pathway of Dimeric Coiled Coil Folding, *Biochemistry*, 1998, **37**, 14613-14622.

60. D. S. Wishart, Interpreting protein chemical shift data, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2011, **58**, 62-87.

61. D. S. Wishart, B. D. Sykes and F. M. Richards, Simple techniques for the quantification of protein secondary structure by 1H NMR spectroscopy, *FEBS J.*, 1991, **293**, 72-80.

62. G. A. Bermejo and C. D. Schwieters, Protein Structure Elucidation from NMR Data with the Program Xplor-NIH, *Methods Mol. Biol.*, 2018, **1688**, 311-340.

63. C. Schwieters, J. Kuszewski and G. Mariusclore, Using Xplor–NIH for NMR molecular structure determination, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2006, **48**, 47-62.

64. C. D. Schwieters, J. J. Kuszewski, N. Tjandra and G. M. Clore, The Xplor-NIH NMR molecular structure determination package, *J. Magn. Reson.*, 2003, **160**, 65-73.

65. M. S. Cheung, M. L. Maguire, T. J. Stevens and R. W. Broadhurst, DANGLE: A Bayesian inferential method for predicting protein backbone dihedral angles and secondary structure, *J. Magn. Reson.*, 2010, **202**, 223-233.

66. W. F. Vranken, W. Boucher, T. J. Stevens, R. H. Fogh, A. Pajon, M. Llinas, E. L. Ulrich, J. L. Markley, J. Ionides and E. D. Laue, The CCPN data model for NMR spectroscopy: development of a software pipeline, *Proteins: Struct., Funct., Genet.*, 2005, **59**, 687-696.

67. J. Garner and M. M. Harding, Design and synthesis of alpha-helical peptides and mimetics, *Org. Biomol. Chem.*, 2007, **5**, 3577-3585.

68. A. Bhattacharya, R. Tejero and G. T. Montelione, Evaluating protein structures determined by structural genomics consortia, *Proteins: Struct., Funct., Genet.*, 2007, **66**, 778-795.

69. R. A. Laskowski, M. W. MacArthur, D. S. Moss and J. M. Thornton, PROCHECK: A program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr.*, 1992, **26**, 283-291.

70. R. A. Laskowski, J. A. C. Rullman, M. W. MacArthur, R. Kaptein and J. M. Thornton, AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR, *J. Biomol. NMR*, 1996, **8**, 477-486.

71. J. L. Wong, M. S. Brown, K. Matsumoto, R. Oesterlin and H. Rapoport, The Structure of Nisin, *J. Am. Chem. Soc.*, 1971, **93**, 4634-4635.

72. N. L. Daly, B. Callaghan, R. J. Clark, S. T. Nevin, D. J. Adams and D. J. Craik, Structure and activity of alpha-conotoxin PeIA at nicotinic acetylcholine receptor subtypes and GABA(B) receptor-coupled N-type calcium channels, *J. Biol. Chem.*, 2011, **286**, 10233-10237.

73. A. N. Appleyard, S. Choi, D. M. Read, A. Lightfoot, S. Boakes, A. Hoffmann, I. Chopra, G. Bierbaum, B. A. M. Rudd, M. J. Dawson and J. Cortes, Dissecting Structural and Functional Diversity of the Lantibiotic Mersacidin, *Chem. Biol.*, 2009, **16**, 490-498.

74. T. B. Potocky, A. K. Menon and S. H. Gellman, Cytoplasmic and nuclear delivery of a TAT-derived peptide and a beta-peptide after endocytic uptake into HeLa cells, *J. Biol. Chem.*, 2003, **278**, 50188-50194.

75. P. Ramoino, A. Diaspro, M. Fato and C. Usai, in *Molecular Regulation of Endocytosis*, 2012, DOI: 10.5772/46061, ch. Chapter 6.

76. W. S. Hancock and J. E. Battersvy, A new micro-test for the detection of incomplete coupling reactions in solid-phase peptide synthesis using 2,4,6-trinitrobenzene-sulphonic acid, *Anal. Biochem.*, 1976, **71**, 260-264.

77. A. W. Schuttelkopf and D. M. van Aalten, PRODRG: a tool for high-throughput crystallography of protein-ligand complexes, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2004, **60**, 1355-1363.

78. I. Goldberg, K. Bernstein, E. M. Kosower, E. Goldstein and B. Pazhenchevsky, Bimanes. 19. Characteristics of the Molecular Structures of 1,5-Diazabicyclo[3.3.0]octadienediones (9,10-Dioxabimanes) Determined by X-ray Crystallography, *J. Heterocycl. Chem.*, 1983, **20**, 903-912.

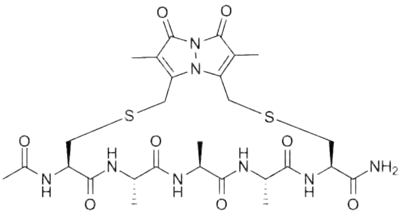
79. M. Heinig and D. Frishman, STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins, *Nucleic Acids Res.*, 2004, **32**, W500-502.

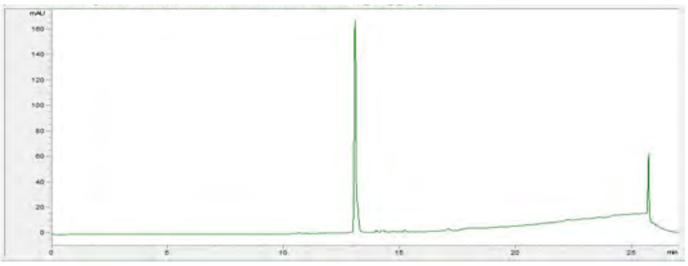
Supplemental Data.

A BIMANE-BASED PEPTIDE STAPLE FOR COMBINED HELICAL INDUCTION AND FLUORESCENT IMAGING

S2.1 Characterisation data

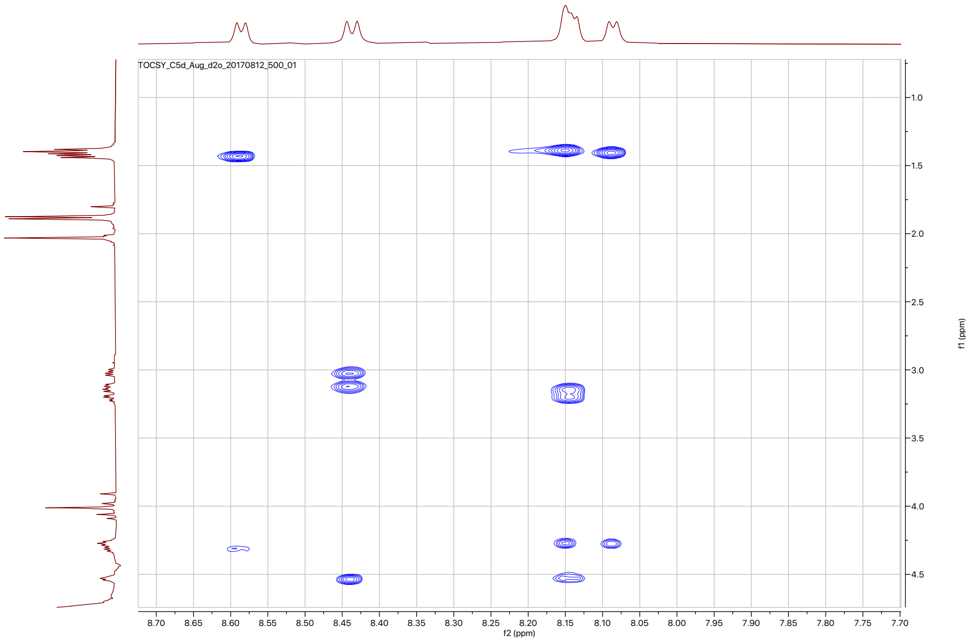
(1,5-cyclo)-Ac-CAAAC-Am: ***1***



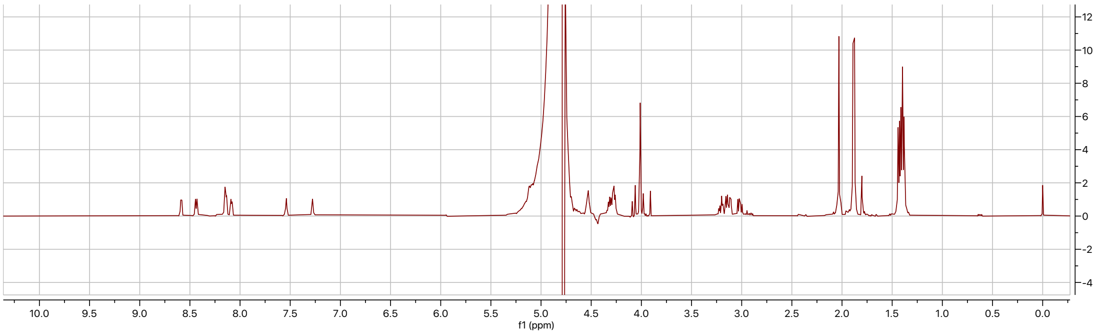


HPLC trace of peptide 1 at 254 nm, run over a linear gradient of 0-100% ACN + 0.1% TFA from 5 to 20 minutes.

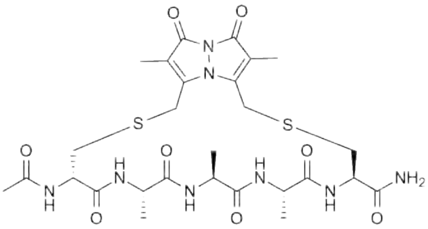
Assignment of H^1 spectrum Table S2. ^1H NMR (10% aq. D_2O, 500 MHz): δ 8.59 (d, J=5.6 Hz, 1 H), 8.44 (d, J=6.8 Hz, 1 H), 8.12 - 8.18 (m, 2 H), 8.09 (d, J=5.6 Hz, 1 H), 7.54 (br. s., 1 H), 7.28 (br. s., 1 H), 4.22 - 4.35 (m, 3 H), 3.87 - 4.11 (m, 4 H), 3.10 - 3.25 (m, 3 H), 2.91 - 3.06 (m, 1 H), 2.03 (s, 3 H), 1.89 (s, 3 H), 1.88 (s, 3 H), 1.31 - 1.49 (m, 9 H) ppm; ^{13}C NMR (10% aq. D_2O, 126 MHz): δ 177.8, 177.7, 177.6, 177.1, 176.8, 174.7, 151.7, 117.6, 117.4, 93.6, 56.3, 56.1, 53.0, 36.5, 36.4, 28.6, 24.4, 19.4, 19.1, 18.4, 9.1 ppm; HRMS (ESI+) Expected $[M+H]^+$ for $C_{27}H_{38}N_8O_8S_2$: 667.2332, observed: $[M+H]^+$ 667.2330.

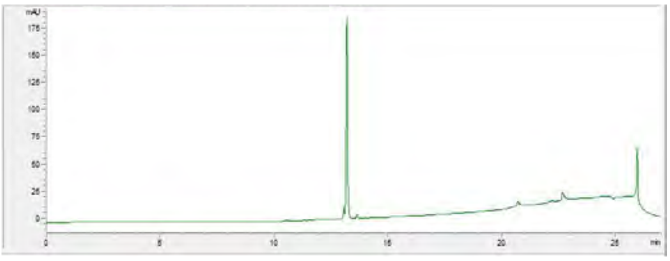


NMR spectra of peptide **1** in 10% aq. D_2O at ~pH 5, zoomed in 2D ^1H homonuclear zTOCSY spectrum and below:full 1D ^1H spectrum



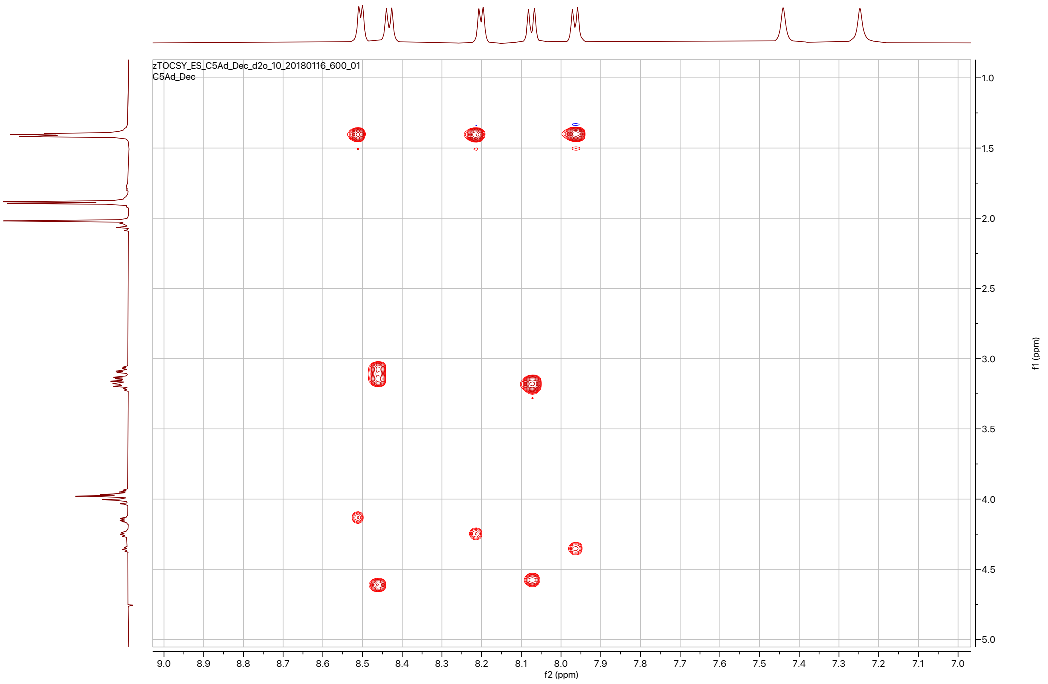
(1,5-cyclo)-Ac-cAAAC-Am: ***2***



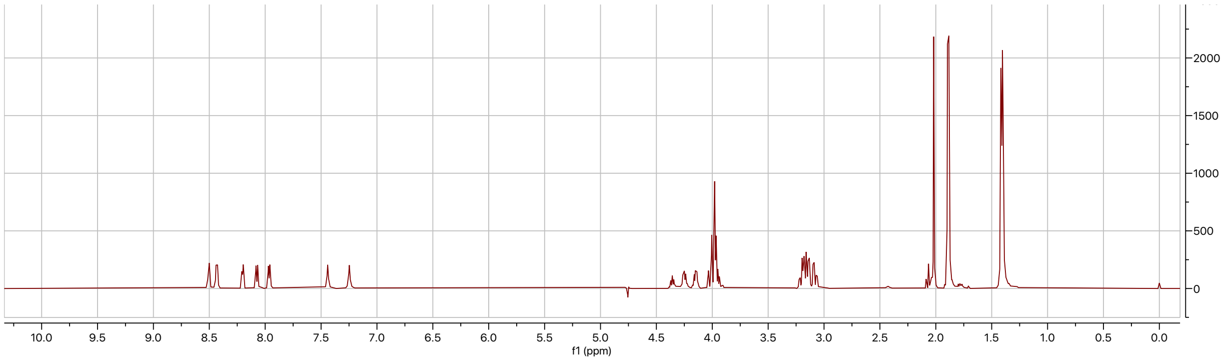


HPLC trace of peptide 2 at 254 nm, run over a linear gradient of 0-100% ACN + 0.1% TFA from 5 to 20 minutes.

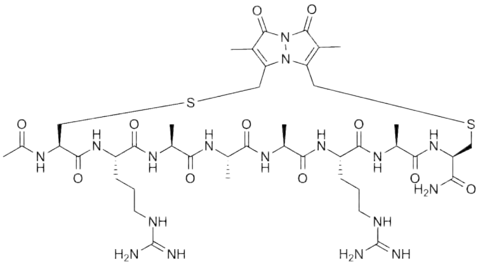
Assignment of H^1 spectrum Table S3. ^1H NMR (599 MHz, 10% aq. D_2O): δ 8.54 (d, *J*=4.7 Hz, 1 H), 8.49 (d, *J*=6.7 Hz, 1 H), 8.25 (d, *J*=5.3 Hz, 1 H), 8.10 (d, *J*=7.6 Hz, 1 H), 7.99 (d, *J*=6.5 Hz, 1 H), 7.36 - 7.54 (m, 1 H), 7.29 (s, 1 H), 4.59 - 4.63 (m, 1 H), 4.55 - 4.59 (m, 1 H), 4.32 - 4.39 (m, 1 H), 4.22 - 4.28 (m, 1 H), 4.10 - 4.16 (m, 1 H), 3.91 - 4.05 (m, 4 H), 3.12 - 3.26 (m, 3 H), 3.04 - 3.10 (m, 1 H), 2.01 (s, 3 H), 1.85 - 1.93 (m, 6 H), 1.40 (s, 9 H) ppm; ^{13}C NMR (151 MHz, 10% aq. D_2O): δ 178.0, 177.8, 177.0, 176.7, 174.4, 165.4, 165.3, 152.2, 151.9, 117.9, 117.3, 55.9, 55.8, 54.1, 53.4, 52.7, 36.4, 35.7, 28.4, 24.5, 18.9, 18.8, 9.1, 9.1 ppm; HRMS (ESI+) Expected $[M+H]^+$ for $C_{27}H_{38}N_8O_8S_2$: 667.2332, observed: $[M+H]^+$ 667.2298.

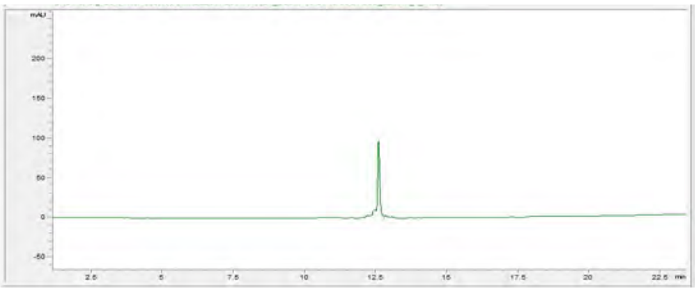


NMR spectra of peptide **2** in 10% aq. D_2O at ~pH 5, zoomed in 2D ^1H homonuclear zTOCSY spectrum and below:full 1D ^1H spectrum



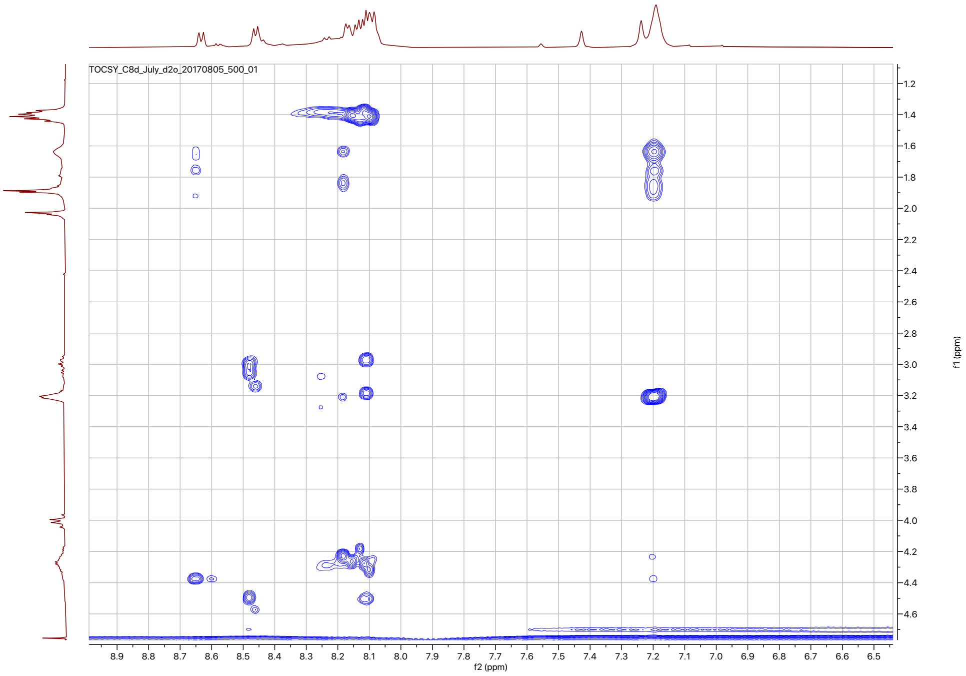
(1,8-cyclo)-Ac-CRAAARAC-Am: ***3***



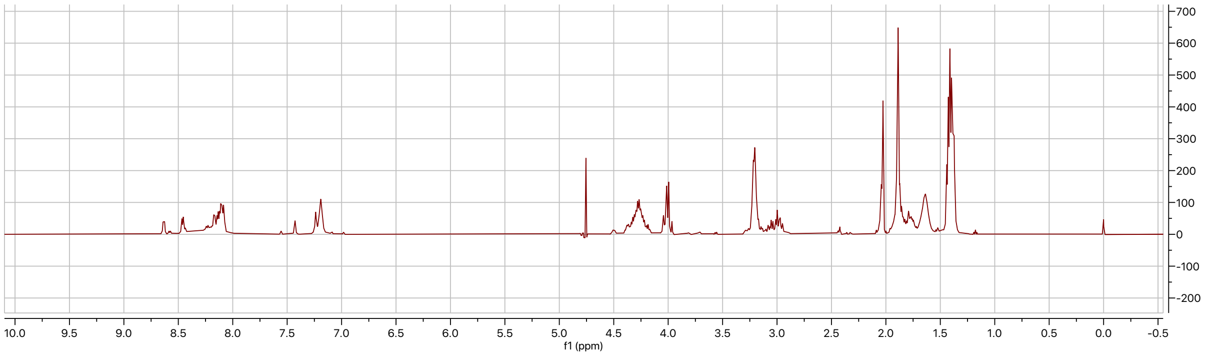


HPLC trace of peptide **3** at 254 nm, run over a linear gradient of 0-100% ACN + 0.1% TFA from 5 to 20 minutes.

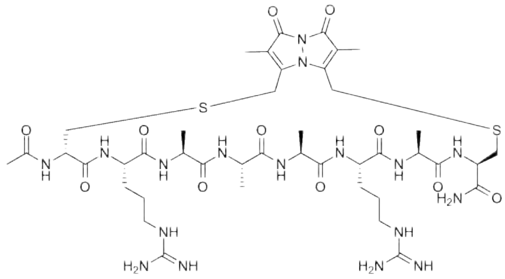
Assignment of H^1 spectrum Table S4. ^{1}H NMR (500 MHz, 10% aq. D_2O) δ 8.65 (d, *J* = 7.3 Hz, 1H), 8.48 (d, *J* = 6.6 Hz, 1H), 8.05 - 8.22 (m, 6H), 7.43 (s, 1H), 7.15 - 7.29 (m, 3H), 4.14 - 4.58 (m, 8H), 3.82 - 4.13 (m, 4H), 3.11 - 3.34 (m, 8H), 2.89 - 3.10 (m, 4H), 1.53 - 2.16 (m, 17H), 1.31 - 1.52 (m, 12H) ppm; ^{13}C NMR (126 MHz, 10% aq. D_2O) δ 178.0, 177.5, 176.9, 176.8, 176.5, 176.1, 175.9, 175.8, 174.7, 165.5, 165.1, 159.7, 151.2, 150.8, 120.2, 117.9, 117.3, 56.3, 56.1, 56.0, 55.7, 53.6, 52.8, 52.7, 52.5, 43.3 (br), 30.3, 27.2, 27.0, 24.4, 19.4, 19.1, 19.0, 18.8, 9.2, 9.1 ppm; HRMS (ESI+) Expected $[M+3H]^{3+}$ for $C_{42}H_{67}N_{17}O_{11}S_2$: 350.8294, observed: $[M+H]^+$ 350.8298.

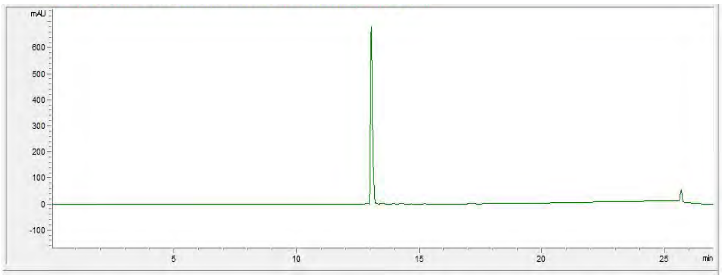


NMR spectra of peptide **3** in 10% aq. D_2O at ~pH 5, zoomed in 2D ^{1}H homonuclear zTOCSY spectrum and below:full 1D ^{1}H spectrum



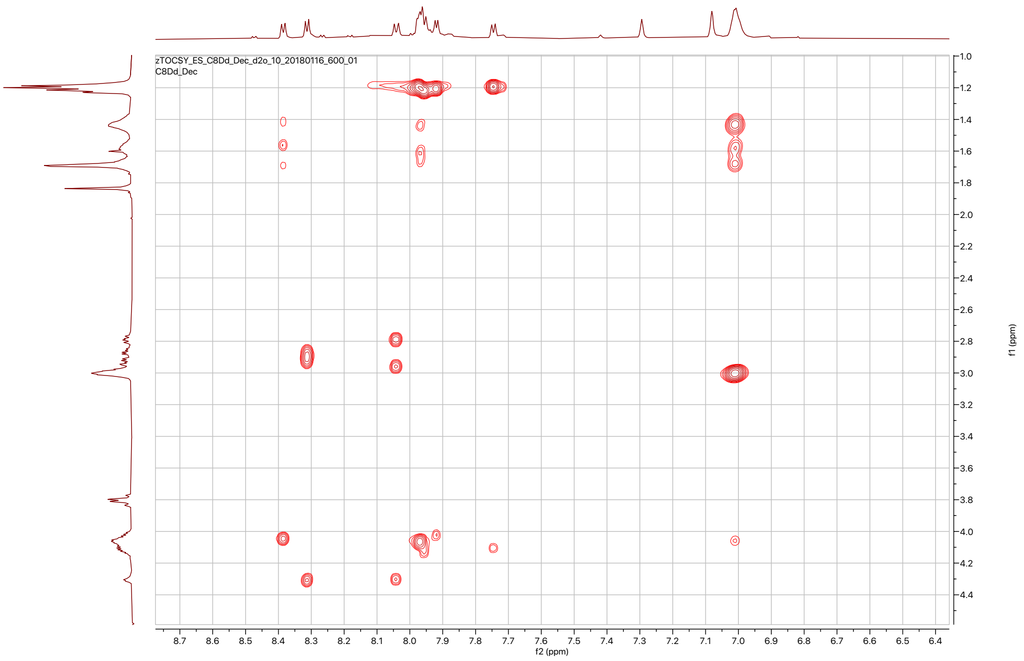
(1,8-cyclo)-Ac-cRAAARAC-Am: ***4***



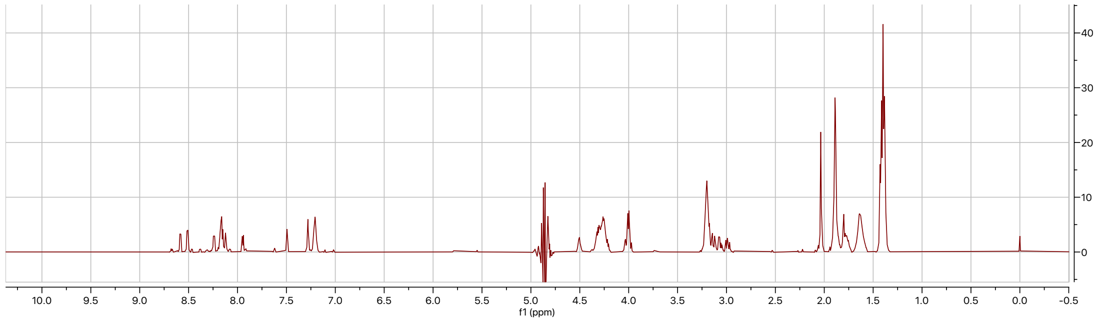


HPLC trace of peptide 4 at 254 nm, run over a linear gradient of 0-100% ACN + 0.1% TFA from 5 to 20 minutes

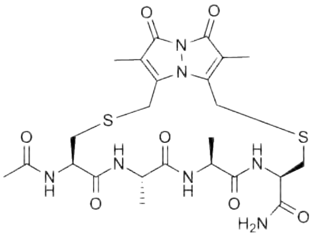
Assignment of H^1 spectrum Table S5. ^{1}H NMR (599 MHz, 10% aq. D$_2$O) δ 8.58 (d, J = 6.2 Hz, 1H), 8.51 (d, J = 6.2 Hz, 1H), 8.24 (d, J = 7.3 Hz, 1H), 8.14 - 8.19 (m, 3H), 8.12 (d, J = 5.3 Hz, 1H), 7.94 (d, J = 6.2 Hz, 1H), 7.49 (s, 1H), 7.28 (s, 1H), 7.17-7.24 (m, 2H), 4.48 - 4.54 (m, 1H), 4.16 - 4.37 (m, 7H), 3.93 - 4.06 (m, 4H), 2.92 - 3.28 (m, 8H), 2.00 - 2.06 (m, 3H), 1.54 - 1.96 (m, 12H), 1.33 - 1.48 (m, 12H) ppm; ^{13}C NMR (151 MHz, 10% aq. D$_2$O) δ 177.9, 177.6, 177.1, 176.8, 176.1, 175.1, 165.6, 165.4, 165.0, 159.4, 151.1, 119.8, 117.9, 117.2, 116.3, 56.9, 56.4, 56.1, 55.5, 53.2, 52.7, 52.6, 52.2, 43.1, 30.3, 30.0, 27.6, 27.2, 26.9, 24.3, 19.0, 18.9, 18.8, 18.7, 9.0, 8.9, 8.4 ppm; HRMS (ESI+) Expected [M+2H]$^{2+}$ for $C_{42}H_{67}N_{17}O_{11}S_2$: 525.7402, observed [M+2H]$^{2+}$ 525.7394.

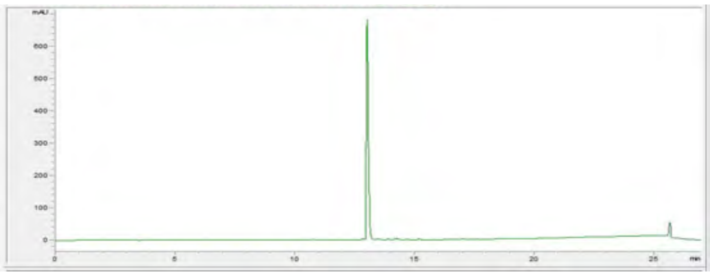


NMR spectra of peptide **4** in 10% aq. D$_2$O at ~pH 5, zoomed in 2D ^{1}H homonuclear zTOCSY spectrum and below: full 1D ^{1}H spectrum



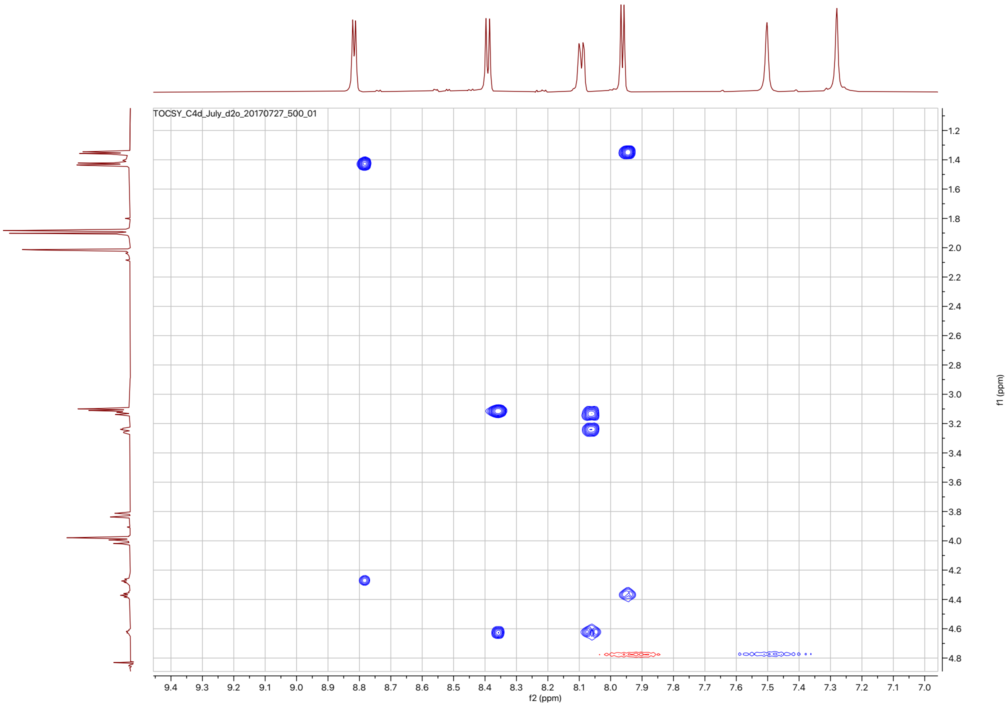
(1,4-cyclo)-Ac-CAAC-Am: ***5***



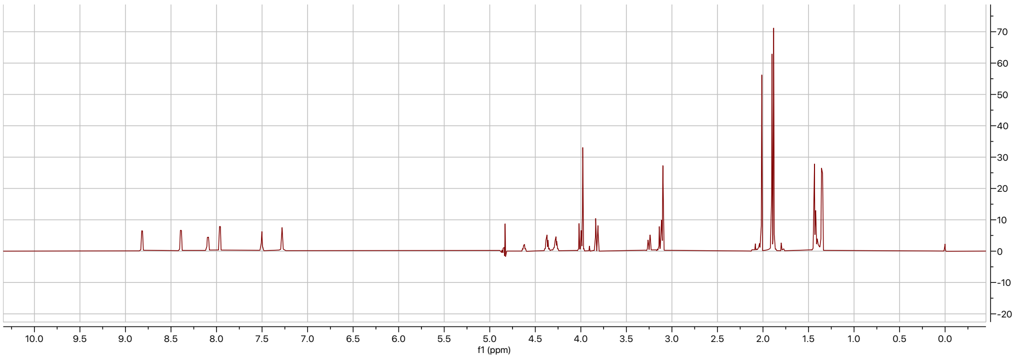


HPLC trace of peptide 5 at 254 nm, run over a linear gradient of 0-100% ACN + 0.1% TFA from 5 to 20 minutes

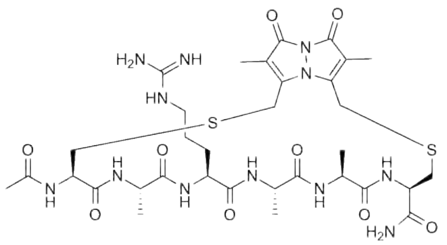
Assignment of H^1 spectrum Table S6. ^{1}H NMR (10% aq. D_2O, 500 MHz): δ 8.78 (d, *J*=5.4 Hz, 1 H), 8.35 (d, *J*=6.8 Hz, 1 H), 8.06 (d, *J*=7.8 Hz, 1 H), 7.94 (d, *J*=6.1 Hz, 1 H), 7.47 (br. s., 1 H), 7.25 (br. s., 1 H), 4.56 - 4.67 (m, 1 H), 4.34 - 4.40 (m, 1 H), 4.22 - 4.30 (m, 1 H), 3.96 - 4.07 (m, 3 H), 3.83 (d, *J*=14.9 Hz, 1 H), 3.21 - 3.27 (m, 1 H), 3.07 - 3.17 (m, 3 H), 2.01 (s, 3 H), 1.90 (s, 3 H), 1.88 (s, 3 H), 1.43 (d, *J*=7.3 Hz, 3 H), 1.35 (d, *J*=7.1 Hz, 3 H) ppm; ^{13}C NMR (10% aq. D_2O, 126 MHz): δ 178.1, 177.1, 176.7, 176.6, 174.8, 165.2, 152.1, 151.6, 117.5, 117.3, 55.6, 54.9, 53.9, 52.7, 36.8, 35.7, 28.7, 27.9, 24.4, 19.2, 18.7, 9.1, 8.8 ppm; HRMS (ESI+) Expected $[M+H]^+$ for $C_{24}H_{33}N_7O_7S_2$: 596.1961, observed: $[M+H]^+$ 596.1956.

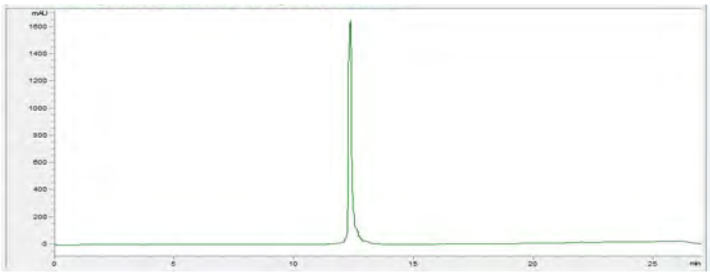


NMR spectra of peptide **5** in 10% aq. D_2O at ~pH 5, zoomed in 2D ^{1}H homonuclear zTOCSY spectrum and below, full 1D ^{1}H spectrum



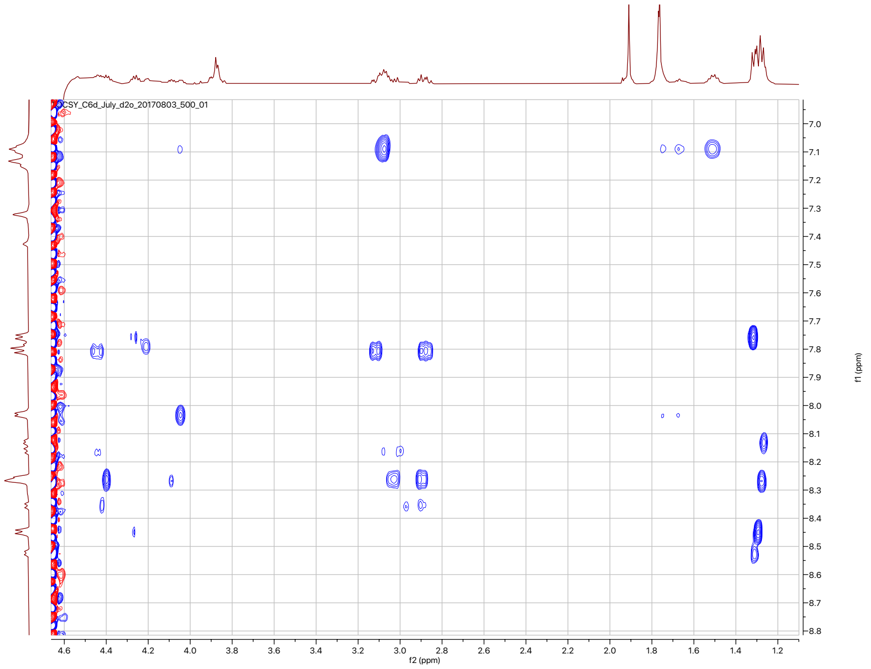
(1,6-cyclo)-Ac-CARAAC-Am: ***6***



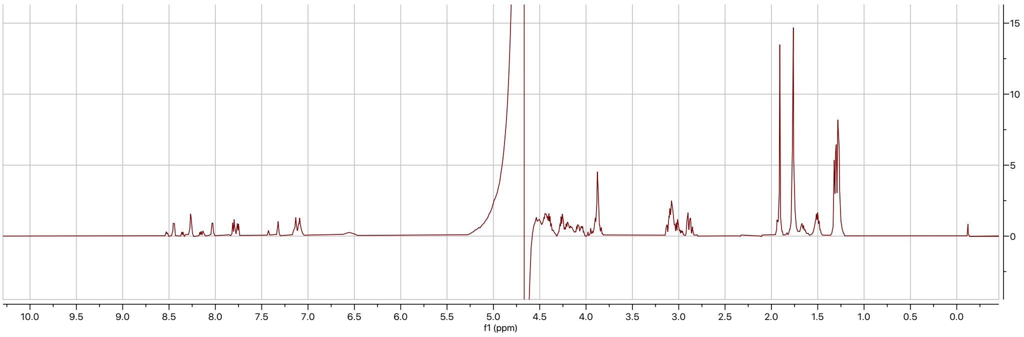


HPLC trace of peptide 6 at 254 nm, run over a linear gradient of 0-100% ACN + 0.1% TFA from 5 to 20 minutes

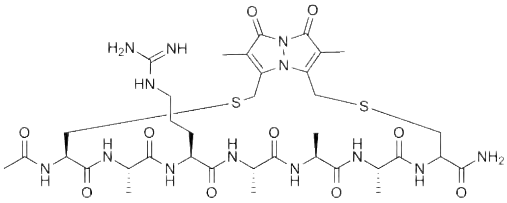
Assignment of H^1 spectrum Table S7. ^{1}H NMR (10% aq. D_2O, 500 MHz): δ 8.57 (d, *J*=6.1 Hz, 1 H), 8.35 - 8.42 (m, 2 H), 8.15 (d, *J*=5.1 Hz, 1 H), 7.93 (d, *J*=8.1 Hz, 1 H), 7.88 (d, *J*=6.8 Hz, 1 H), 7.44 (br. s., 1 H), 7.25 (br. s., 1 H), 7.21 (t, *J*=5.1 Hz, 1 H), 3.83 - 4.66 (m, 12 H), 2.90 - 3.29 (m, 6 H), 2.03 (s, 3 H), 1.84 – 1.93 (m, 10 H), 1.56 – 1.83 (m, 3 H), 1.32 - 1.47 (m, 9 H) ppm; ^{13}C NMR (10% aq. D_2O,126 MHz): δ 177.5, 176.8, 176.7, 174.5, 165.2, 165.1, 151.6, 117.2, 57.3, 55.7, 52.5, 52.3, 43.1, 36.4, 28.5, 27.8, 27.1, 24.3, 18.8, 18.5, 9.0 ppm; HRMS (ESI+) Expected $[M+H]^+$ for $C_{33}H_{50}N_{12}O_9S_2$: 823.3343, observed: $[M+H]^+$ 823.3336.

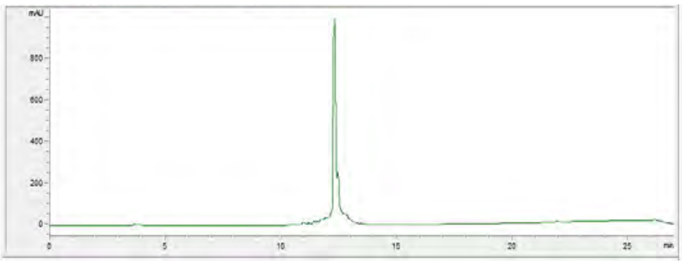


NMR spectra of peptide **6** in 10% aq. D_2O at ~pH 5, zoomed in 2D ^{1}H homonuclear zTOCSY spectrum and below, full 1D ^{1}H spectrum



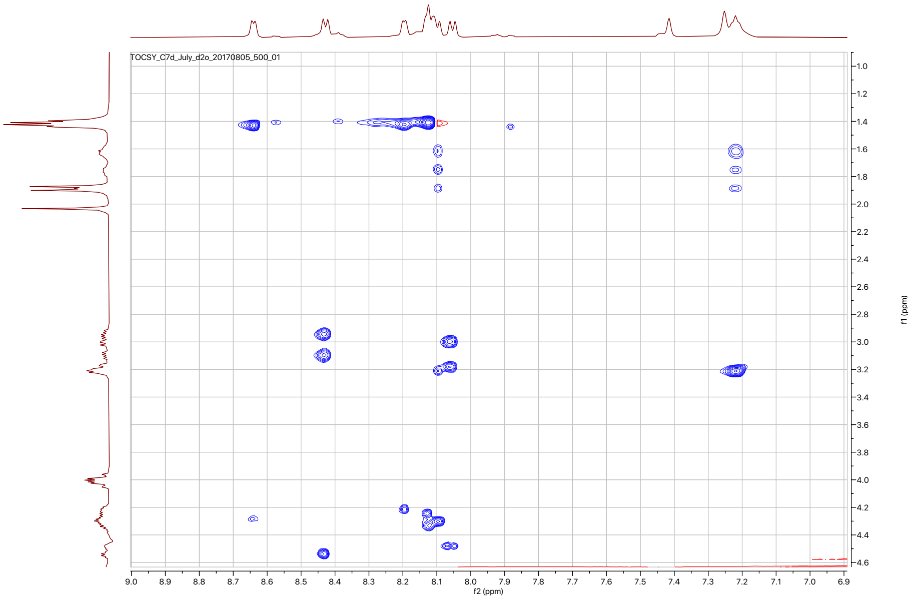
(1,7-cyclo)-Ac-CARAAAC-Am: **7**



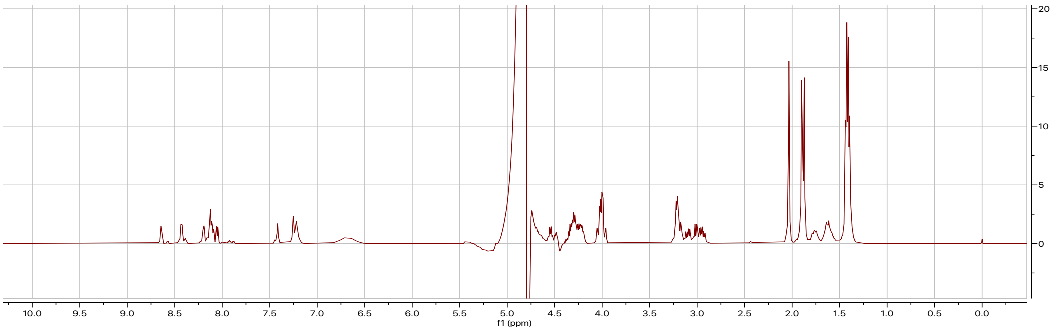


HPLC trace of peptide 7 at 254 nm, run over a linear gradient of 0-100% ACN + 0.1% TFA from 5 to 20 minutes.

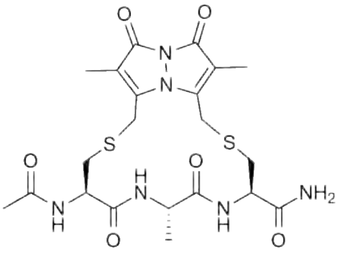
Assignment of H[1] spectrum Table S8. ^{1}H NMR (10% aq. D_2O, 500 MHz): δ 8.64 (d, *J*=4.9 Hz, 1 H), 8.43 (d, *J*=6.6 Hz, 1 H), 8.20 (d, *J*=4.9 Hz, 1 H), 8.08 - 8.15 (m, 3 H), 8.05 (d, *J*=7.6 Hz, 1 H), 7.42 (s, 1 H), 7.25 (br. s., 1 H), 7.22 (d, *J*=5.5 Hz, 1 H), 4.16 - 4.31 (m, 3 H), 3.92 - 4.08 (m, 4 H), 2.87 - 3.25 (m, 6 H), 2.03 (s, 3 H), 1.85 - 1.93 (m, 7 H), 1.52 - 1.81 (m, 3 H), 1.34 - 1.48 (m, 9 H) ppm; ^{13}C NMR (10% aq. D_2O,126 MHz): δ 178.0, 177.6, 176.8, 176.0, 172.5, 165.1, 163.6, 159.5, 152.3, 145.3, 117.3, 117.2, 114.0, 82.4, 53.3, 52.7, 52.5, 43.3, 35.5, 28.9, 28.1, 24.3, 21.3, 18.8, 18.7, 18.6, 9.0 ppm; HRMS (ESI+) Expected $[M+H]^+$ for $C_{36}H_{55}N_{13}O_{10}S_2$: 894.3714, observed: $[M+H]^+$ 894.3840.

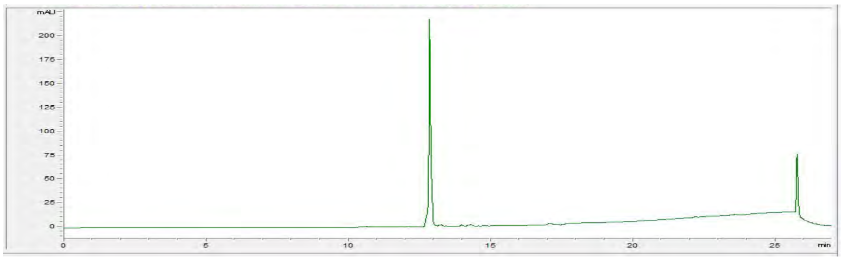


NMR spectra of peptide **7** in 10% aq. D_2O at ~pH 5, zoomed in 2D ^{1}H homonuclear zTOCSY spectrum and below, full 1D ^{1}H spectrum



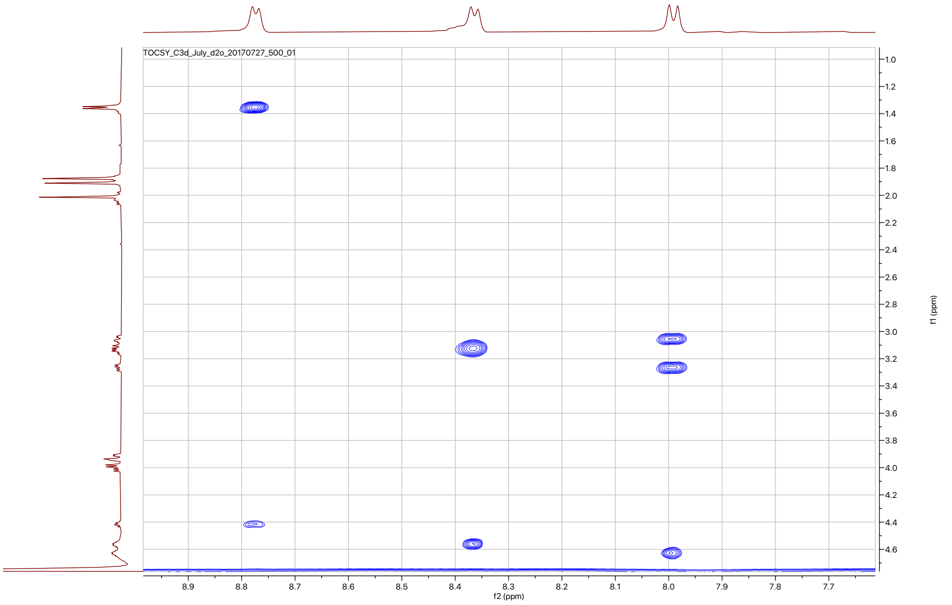
(1,3-cyclo)-Ac-CAC-Am: ***8***



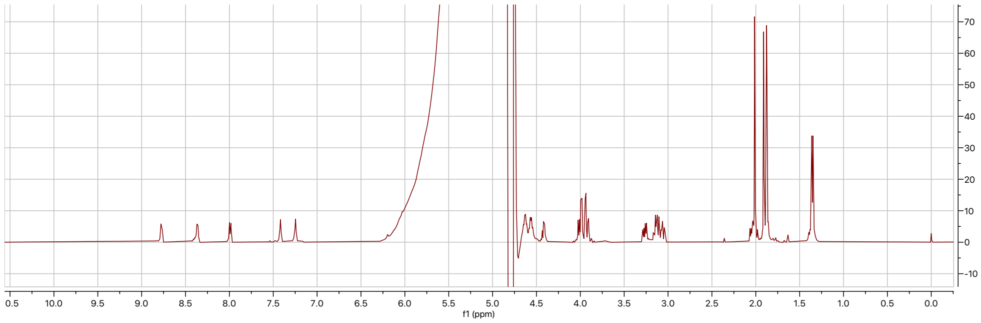


HPLC trace of peptide 8 at 254 nm, run over a linear gradient of 0-100% ACN + 0.1% TFA from 5 to 20 minutes.

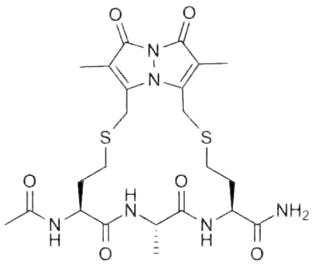
Assignment of H[1] spectrum Table S9. ^{1}H NMR (10% aq. D_2O, 500 MHz): δ 8.77 (d, *J*=6.4 Hz, 1 H), 8.36 (d, *J*=6.8 Hz, 1 H), 7.99 (d, *J*=8.1 Hz, 1 H), 7.42 (br. s., 1 H), 7.25 (br. s., 1 H), 4.51 - 4.59 (m, 1 H), 4.37 - 4.46 (m, 1 H), 3.87 - 4.07 (m, 4 H), 3.27 (dd, *J*=13.8, 6.5 Hz, 1 H), 2.99 - 3.18 (m, 3 H), 2.01 (s, 3 H), 1.91 (s, 3 H), 1.88 (s, 3 H), 1.36 (d, *J*=7.1 Hz, 3 H) ppm; ^{13}C NMR (10% aq. D_2O, 126 MHz): δ 177.2, 177.0, 176.6, 174.7, 165.7, 152.9, 152.5, 118.5, 118.0, 55.6, 55.2, 52.8, 36.4, 35.9, 28.9, 28.3, 24.4, 18.7, 9.4, 9.1 ppm; HRMS (ESI+) Expected $[M+H]^+$ for $C_{21}H_{28}N_6O_6S_2$: 525.1590, observed: $[M+H]^+$ 525.1575.

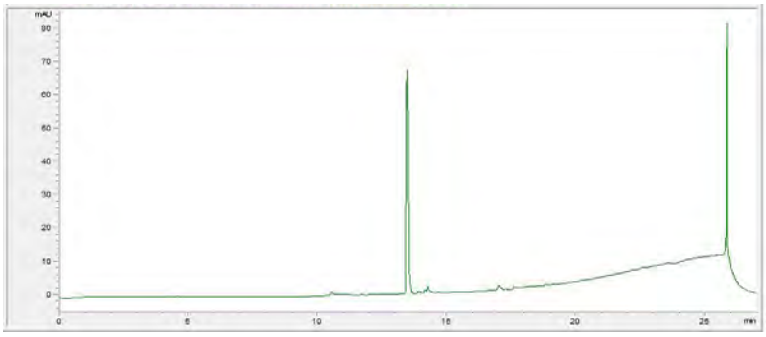


NMR spectra of peptide **8** in 10% aq. D_2O at ~pH 5, zoomed in 2D ^{1}H homonuclear zTOCSY spectrum and below, full 1D ^{1}H spectrum



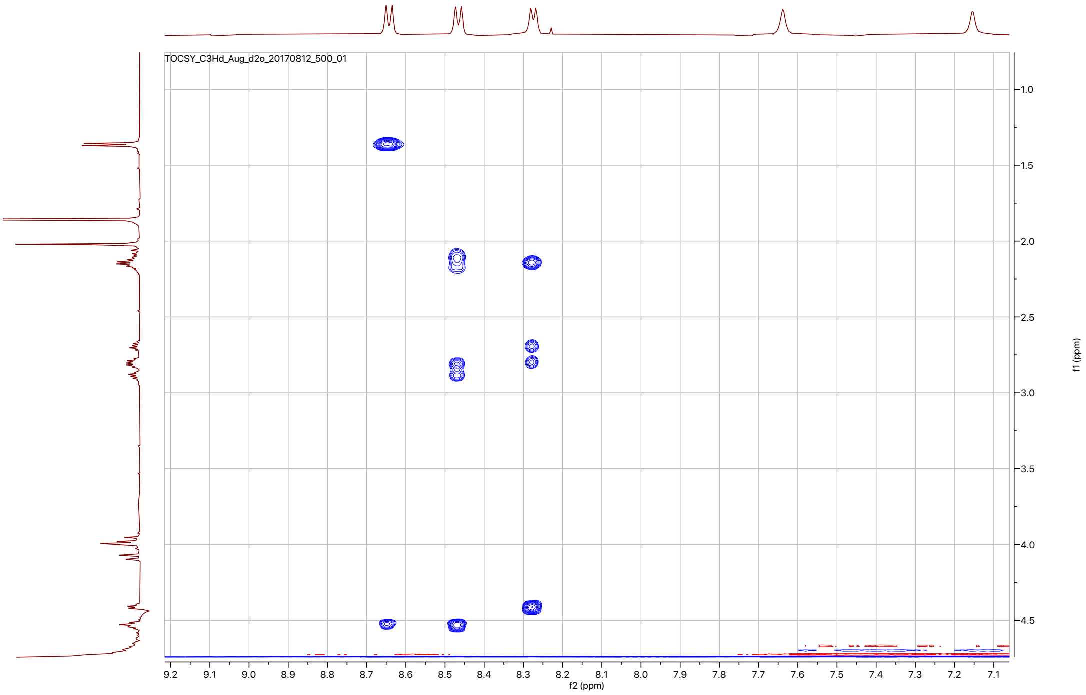
(1,3-cyclo)-Ac-HcyAlaHcy-Am: ***9***



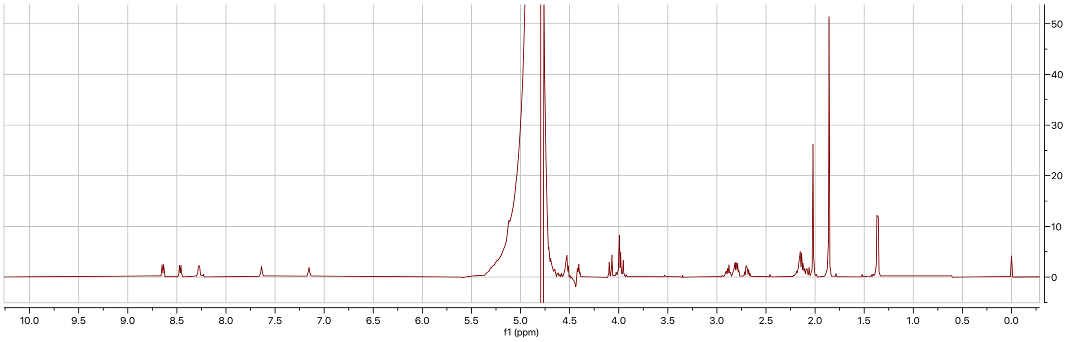


HPLC trace of peptide 9 at 254 nm, run over a linear gradient of 0-100% ACN + 0.1% TFA from 5 to 20 minutes.

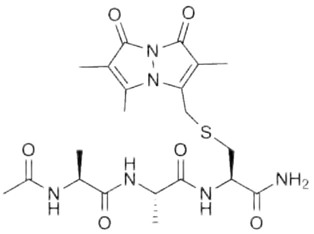
Assignment of H^1 spectrum Table S10. ^{1}H NMR (10% aq. D_2O, 500 MHz): δ 8.64 (d, J=7.8 Hz, 1 H), 8.47 (d, J=7.8 Hz, 1 H), 8.27 (d, J=6.6 Hz, 1 H), 7.64 (br. s., 1 H), 7.15 (br. s., 1 H), 3.84 - 4.16 (m, 4 H), 2.61 - 2.95 (m, 4 H), 2.05 - 2.28 (m, 4 H), 2.02 (s, 3 H), 1.86 (s, 6 H), 1.36 (d, J=7.1 Hz, 3 H) ppm; ^{13}C NMR (126 MHz, 10% aq. D_2O) δ 178.1, 177.0, 176.9, 175.3, 165.4, 150.8, 150.4, 117.0, 55.6, 54.9, 51.9, 34.5, 33.4, 33.2, 32.0, 30.6, 28.2, 27.4, 24.5, 19.1, 8.8, 8.7 ppm; HRMS (ESI+) Expected $[M+H]^+$ for $C_{23}H_{32}N_6O_6S_2$: 553.1903, observed: $[M+H]^+$ 553.1866.

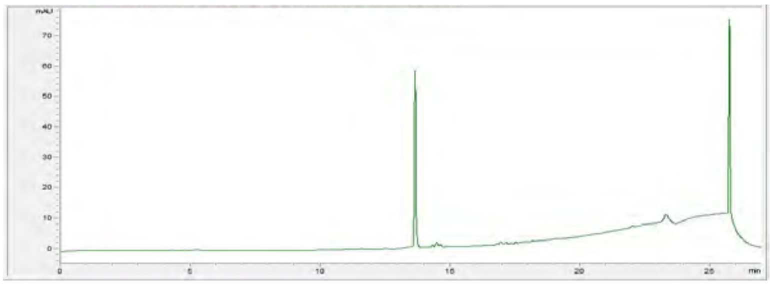


NMR spectra of peptide **9** in 10% aq. D_2O at ~pH 5, zoomed in 2D ^{1}H homonuclear zTOCSY spectrum and below:full 1D ^{1}H spectrum



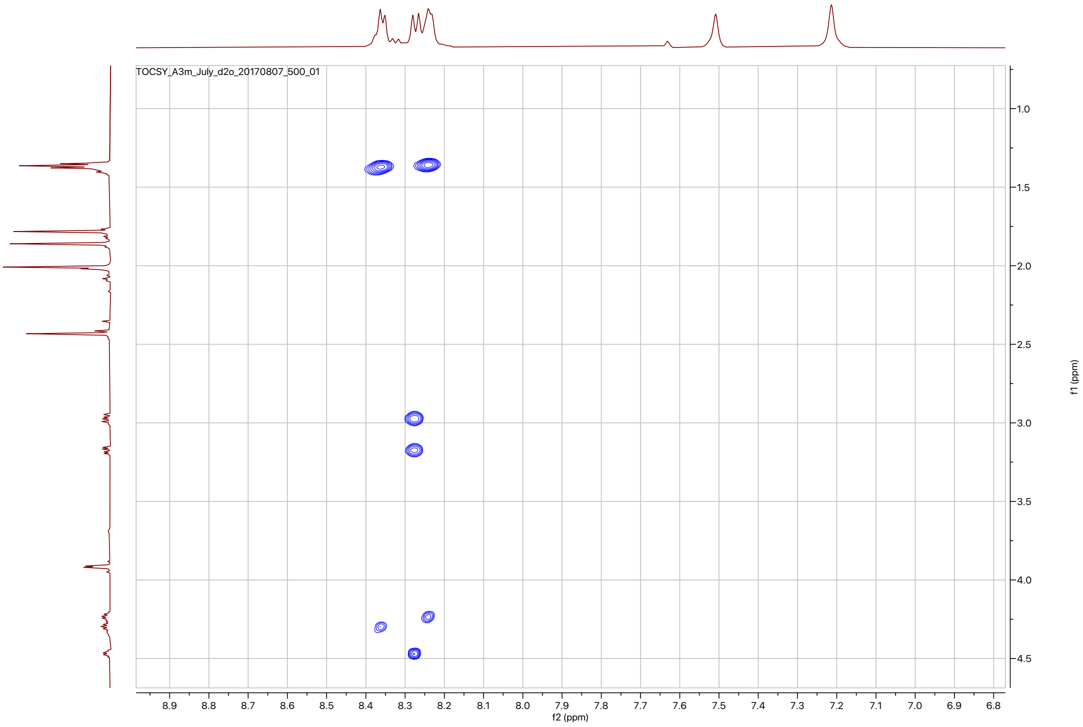
(3-bimane)-Ac-AAC-Am: ***10***



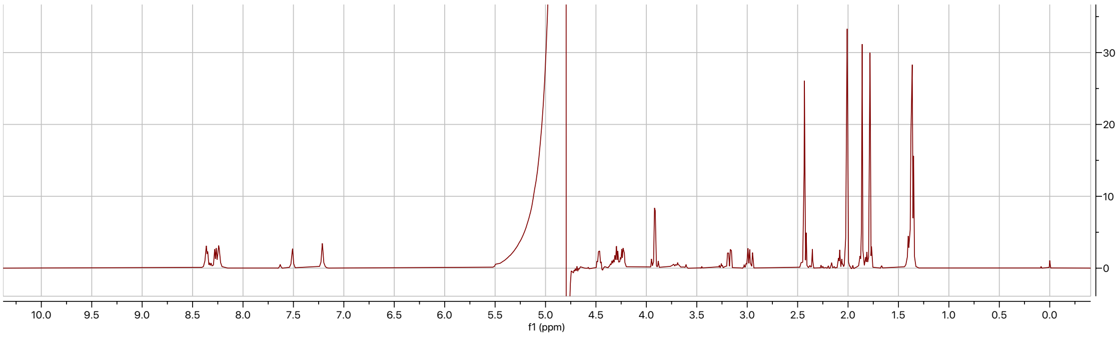


HPLC trace of peptide 10 at 254 nm, run over a linear gradient of 0-100% ACN + 0.1% TFA from 5 to 20 minutes.

Assignment of H^1 spectrum Table S11. ^{1}H NMR (10% aq. D_2O, 500 MHz): δ 8.36 (d, *J*=6.1 Hz, 1 H), 8.19 - 8.31 (m, 2 H), 7.51 (br. s., 1 H), 7.21 (br. s., 1 H), 4.43 - 4.50 (m, 1 H), 4.18 - 4.38 (m, 2 H), 3.85 - 3.98 (m, 2 H), 3.11 - 3.24 (m, 1 H), 2.86 - 3.05 (m, 1 H), 2.43 (s, 3 H), 2.01 (s, 3 H), 1.84 - 1.92 (m, 3 H), 1.78 (s, 3 H), 1.30 - 1.40 (m, 6 H) ppm; ^{13}C NMR (10% aq. D_2O, 126 MHz): δ 178.1, 177.7, 176.9, 176.8, 165.7, 165.0, 152.4, 150.8, 116.2, 114.2, 110.0, 55.3, 52.6, 52.5, 35.6, 27.5, 24.3, 19.1, 18.9, 13.8, 8.8, 8.3 ppm; HRMS (ESI+) Expected [M+H]$^+$ for $C_{21}H_{30}N_6O_6S$: 495.2026, observed: [M+H]$^+$ 495.2019.



NMR spectra of peptide **10** in 10% aq. D_2O at ~pH 5, zoomed in 2D ^{1}H homonuclear zTOCSY spectrum and below:full 1D ^{1}H spectrum



S2.2 %Helicity calculations

Table S1: %Helicity calculations for peptides over 5 amino-acids in length using CD MRE (in mdeg/res/dmol)

	1	2	3	4	6	7
$MRE_{observed}$ @ 207 nm	-6991.13	-10046.4	-5639.6	-8342.8	-5900.88	-6265.42
$MRE_{observed}$ @ 215nm	-3627.432	-6250.4	-2196.405	-3462.95	-3341.64	-3440.566
MRE at 215 nm expected for peptide if k=4[1]	-7550	-7550	-18875	-18875	-12583	-16178
% helicity	53.6	84.6	15.6	22.0	31.4	25.4
Ratio MRE 215/207	0.52	0.62	0.39	0.42	0.57	0.55

From Shepherd/Fairlie 2005[1]

$$MRE@215_{expected} = (-44000 + 250T)(1-k/N_p)$$

Where T = temperature in ºC (here: T=25ºC), k=4 for small peptides and N_p=the number of amino-acid residues

$$\%Helicity=(MRE@215_{observed} - MRE_{Coil})/(MRE@215_{expected} - MRE_{Coil})$$

Where MRE_{coil} = 2220-53(T) and T = temperature in ºC (here: T=25ºC)

S2.3 NMR structure assignment

Atoms on the bimane structure are detailed in Figure S3. Acy represents an acetyl group and Nh3 the *C*-terminal functional group, as defined in CCPNMR Analysis.

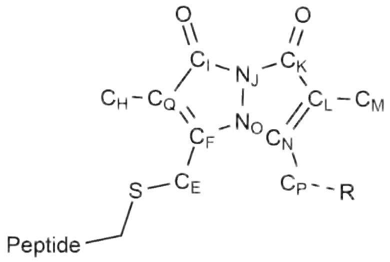


Figure S3: Atom codes for the NMR assignment of the bimane ring. Protons are all labelled with corresponding carbon code and then numbered (non-stereospecifically) sequentially. R= H_{P3} if mono-bimane (BimH) or S-Peptide if di-bimane (BimC)

Table S1: NMR Assignment of peptide **1** from CCPNMR Analysis for NMR structure calculations. BimC atom codes in Figure S3.

1								
Residue	**HN**	**Ac-Me**	**H_α**	**H_β**	**H_E**	**H_H**	**H_M**	**H_P**
0Acy	-	2.03	-	-	-	-	-	-
1BimC	8.44	-	4.53	3.02,3.13	4.08	1.88	1.88	4.01
2Ala	8.59	-	4.31	1.41	-	-	-	-
3Ala	8.08	-	4.27	1.41	-	-	-	-
4Ala	8.15	-	4.27	1.41	-	-	-	-
5Cys	8.14	-	4.53	3.13,3.21	-	-	-	-
6Nh3	7.55	-	-	-	-	-	-	-

Table S2: NMR Assignment of peptide **2** from CCPNMR Analysis for NMR structure calculations. BimC atom codes in Figure S3

2								
Residue	**HN**	**Ac-Me**	**H_α**	**H_β**	**H_E**	**H_H**	**H_M**	**H_P**
0Acy	-	2.01	-	-	-	-	-	-
1BimC	8.50	-	4.61	3.07,3.16	3.98,3.98	1.89	1.89	3.98,3.98
2Ala	8.52	-	4.13	1.4	-	-	-	-
3Ala	8.24	-	4.25	1.4	-	-	-	-
4Ala	7.98	-	4.35	1.4	-	-	-	-
5Dcy	8.10	-	4.58	3.16,3.17	-	-	-	-
6Nh3	7.28,7.46	-	-	-	-	-	-	-

Table S3: NMR Assignment of peptide **3** from CCPNMR Analysis for NMR structure calculations. BimC atom codes in Figure S3

3											
Residue	**HN**	**Ac-Me**	**H_{α}**	**H_{β}**	**H_{δ}**	**H_{ε}**	**H_{γ}**	**H_H**	**H_M**	**H_P**	**H_E**
0Acy	-	2.03	-	-	-	-	-	-	-	-	-
1BimC	8.48	-	4.49	2.99, 3.02	-		-	1.87	1.87	4.01	4.01
2Arg	8.65	-	4.37	1.77, 1.90	3.20, 3.20	7.2	1.63, 1.76	6.66	-	-	-
3Ala	8.12	-	4.27	1.41	-	-	-	-	-	-	-
4Ala	8.15	-	4.27	1.41	-	-	-	-	-	-	-
5Ala	8.14	-	4.18	1.41	-	-	-	-	-	-	-
6Arg	8.17	-	4.23	1.86	3.20	*or* 7.2	1.64	6.66	-	-	-
7Ala	8.12	-	4.32	1.41	-	-	-	-	-	-	-
8Cys	8.12	-	4.50	2.98, 3.20	-	-	-	-	-	-	-

Table S4: NMR Assignment of peptide **4** from CCPNMR Analysis for NMR structure calculations. BimC atom codes in Figure S3

4											
Residue	**HN**	**Ac-Me**	**H_{α}**	**H_{β}**	**H_{δ}**	**H_{ε}**	**H_{γ}**	**H_H**	**H_M**	**H_P**	**H_E**
0Acy	-	2.04	-	-	-	-	-	-	-	-	-
1Dcy	8.51	-	4.51	3.09, 3.10	-	-	-	-	-	-	-
2Arg	8.58	-	4.25	1.78, 1.88	3.20	7.21	1.62	6.69	-	-	-
3Ala	8.17	-	4.31	1.4	-	-	-	-	-	-	-
4Ala	8.12	-	4.23	1.4	-	-	-	-	-	-	-
5Ala	7.94	-	4.30	1.4	-	-	-	-	-	-	-
6Arg	8.16	-	4.27	1.79, 1.87	3.20	*or* 7.21	1.64	6.69	-	-	-
7Ala	8.15	-	4.33	1.41	-	-	-	-	-	-	
8BimC	8.24	-	4.50	3.00, 3.13	-		-	1.88	1.88	4.01	4.01
9Nh3	7.49	-	-	-	-	-	-	-	-	-	-

Table S5: NMR Assignment of peptide **5** from CCPNMR Analysis for NMR structure calculations. BimC atom codes in Figure S3

5								
Residue	**HN**	**Ac-Me**	**H_{α}**	**H_{β}**	**H_E**	**H_H**	**H_M**	**H_P**
0Acy	-	2.01	-	-	-	-	-	-
1BimC	8.36	-	4.63	3.11	3.83,4.01	1.90	1.89	3.98
2Ala	8.78	-	4.27	1.43	-	-	-	-
3Ala	7.94	-	4.36	1.35	-	-	-	-
4Cys	8.06	-	4.63	3.13, 3.24	-	-	-	-
5Nh3	7.25, 7.47	-	-	-	-	-	-	-

Table S6: NMR Assignment of peptide **6** from CCPNMR Analysis for NMR structure calculations. BimC atom codes in Figure S3

6											
Residue	**HN**	**Ac-Me**	**H_{α}**	**H_{β}**	**H_{δ}**	**H_{ε}**	**H_{γ}**	**H_H**	**H_M**	**H_P**	**H_E**
0Acy	-	2.04	-	-	-	-	-	-	-	-	
1BimC	8.39	-	4.52	3.02, 3.14	-		-	1.88	1.88	4.00	4.00
2Ala	8.58	-	4.39	1.42	-	-	-	-	-	-	-
3Arg	8.16	-	4.16	1.80, 1.88	3.21, 3.21	7.21	1.64	6.63, 6.68	-	-	-
4Ala	8.39	-	4.20	1.42	-	-	-	-	-	-	-
5Ala	7.90	-	4.38	1.43	-	-	-	-	-	-	-
6Cys	7.91	-	4.57	3.02, 3.22	-	-	-	-	-	-	-
7Nh3	7.45	-	-	-	-	-	-	-	-	-	-

Table S7: NMR Assignment of peptide **7** from CCPNMR Analysis for NMR structure calculations. BimC atom codes in Figure S3

7											
Residue	**HN**	**Ac-Me**	**H_α**	**H_β**	**H_δ**	**H_ε**	**H_γ**	**H_H**	**H_M**	**H_P**	**H_E**
0Acy	-	2.03	-	-	-	-	-	-	-	-	-
1BimC	8.42	-	4.53	2.95, 3.09	-	-	-	1.89	1.89	4.01	4.01
2Ala	8.64	-	4.28	1.42	-	-	-	-	-	-	-
3Arg	8.08	-	4.30	1.75, 1.89	3.20	7.22	1.62	6.66, 6.68	-	-	-
4Ala	8.1	-	4.24	1.42	-	-	-	-	-	-	-
5Ala	8.2	-	4.21	1.42	-	-	-	-	-	-	-
6Ala	8.1	-	4.33	1.42	-	-	-	-	-	-	-
7Cys	8.07	-	4.48	2.99, 3.20	-	-	-	-	-	-	-
8Nh3	7.42	-	-	-	-	-	-	-	-	-	-

Table S8: NMR Assignment of peptide **8** from CCPNMR Analysis for NMR structure calculations. BimC atom codes in Figure S3

8								
Residue	**HN**	**Ac-Me**	**H_α**	**H_β**	**H_E**	**H_H**	**H_M**	**H_P**
0Acy	-	2.01	-	-	-	-	-	-
1BimC	8.34	-	4.56	3.11	3.92	1.88	1.87	3.97
2Ala	8.75	-	4.43	1.34	-	-	-	-
3Cys	7.97	-	4.64	3.09,3.25	-	-	-	-
4Nh3	7.4	-	-	-	-	-	-	-

Table S9: NMR Assignment of peptide **9** from CCPNMR Analysis for NMR structure calculations. BimC atom codes in Figure S3

9									
Residue	**HN**	**Ac-Me**	**H_α**	**H_β**	**H_δ**	**H_H**	**H_M**	**H_P**	**H_E**
0Acy	-	2.02	-	-	-	-	-	-	-
1Hbim	8.27	-	4.42	2.14	2.70,2.80	1.86	1.86	3.98,4.08	4.00
2Ala	8.64	-	4.53	1.37	-	-	-	-	-
3Hcys	8.46	-	4.54	2.12,2.16	2.82,2.89	-	-	-	-
4Nh3	7.15, 7.64	-	-	-	-	-	-	-	-

Table S10: NMR Assignment of peptide **10** from CCPNMR Analysis for NMR structure calculations. BimC atom codes in Figure S3

10								
Residue	**HN**	**Ac-Me**	**H_α**	**H_β**	**H_E**	**H_H**	**H_M**	**H_P**
0Acy	-	2.01	-	-	-	-	-	-
1Ala	8.24	-	4.23	1.36	-	-	-	-
2Ala	8.37	-	4.30	1.37	-	-	-	-
3BimC	8.27	-	4.47	2.97,3.18	1.86	1.78	2.44	3.92,
4Nh3	7.21, 7.52	-	-	-	-	-	-	-

S2.4 NMR structure calculations

Table S11: Summary of the NMR NOE restraints and chemical shift dihedral angle restraints (DANGLE) used for the NMR structure calculations conducted in Xplor-NIH.[2-4] Sort Energy calculated by Xplor-NIH based on BOND, ANGL, IMPR, noe and CDIH energies.

Peptide	Xplor Sort Energy (± Std Dev)	No. of NOE Restraints Total	No. of unambiguous NOE Restraints	No. of ambiguous NOE Restraints	No. of dihedral angle restraints (DANGLE)[5, 6]
1	3.61 ± 0.01	45	16	29	6
2	10.44 ± 0.77	60	30	30	6
3	10.42 ± 0.47	57	19	38	12
4	18.63 ± 0.88	65	22	43	12
5	9.92 ± 1.11	60	42	18	4
6	36.01 ± 2.24	78	39	39	8
7	23.50 ± 0.59	74	38	36	10
8	3.47 ± 0.78	32	20	12	2
9	3.49 ± 0.78	43	3	40	2
10	3.33 ± 0.75	29	20	9	2

Table S12: Summary of Phi/Psi angles (in degrees) analysed by STRIDE,[7] averaged over the twenty lowest energy NMR structure calculations; and the structure reported for each residue is the most commonly returned for the residue.

Peptide	Residue	Structure	Phi	Psi
1	1-Cys	Coil	-120.60 ± 25.16	3.07 ± 37.89
	2-Ala	3_{10}Helix	-61.34 ± 0.05	-39.99 ± 0.06
	3-Ala	3_{10}Helix	-61.34 ± 0.05	-39.99 ± 0.05
	4-Ala	3_{10}Helix	-61.33 ± 0.06	-40.01 ± 0.05
	5-Cys	Coil	-90.51 ± 11.49	
2	2-Ala	Turn	-61.10 ± 0.61	-40.49 ± 0.70
	3-Ala	Turn	-60.09 ± 0.49	-42.20 ± 0.93
	4-Ala	Turn	-70.22 ± 9.40	-32.92 ± 10.04
	5-Cys	Coil	34.25 ± 177.91	
3	1-Cys	Turn	-100.11 ± 70.06	49.40 ± 20.96
	2-Arg	Turn	-61.47 ± 2.17	-45.26 ± 4.00
	3-Ala	Turn	-68.66 ± 10.91	-39.10 ± 4.03
	4-Ala	Turn	-85.47 ± 4.94	-0.68 ± 5.61
	5-Aka	Turn	-59.86 ± 0.46	-39.94 ± 0.43
	6-Arg	Turn	-60.36 ± 0.24	-46.37 ± 1.10
	7-Ala	Turn	-64.40 ± 6.17	-41.65 ± 6.66
	8-Cys	Coil	-69.28 ± 36.59	
4	1-D-Cys	Coil	98.14 ± 101.58	-119.67 ± 7.54
	2-Arg	α-Helix	-60.05 ± 1.68	-44.56 ± 4.82
	3-Ala	α-Helix	-62.51 ± 0.77	-40.44 ± 0.37
	4-Ala	α-Helix	-81.08 ± 3.14	-9.10 ± 0.82
	5-Aka	α-Helix	-60.07 ± 0.45	-42.11 ± 0.75
	6-Arg	α-Helix	-64.02 ± 1.56	-42.28 ± 2.41
	7-Ala	α-Helix	-63.27 ± 3.46	-36.28 ± 3.98
	8-Cys	Coil	-76.28 ± 18.49	
5	1-Cys	Coil	-98.87 ± 68.65	61.18 ± 15.67
	2-Ala	Coil	-61.69 ± 0.18	-40.33 ± 1.06
	3-Ala	Coil	-61.49 ± 0.97	-40.22 ± 0.17
	4-Cys	Coil	-94.65 ± 18.61	
6	1-Cys	Turn	-73.44 ± 16.27	-14.59 ± 29.73
	2-Ala	Turn	-63.80 ± 11.59	-25.55 ± 25.19
	3-Arg	Turn	-91.78 ± 2.15	-62.88 ± 22.87
	4-Ala	Turn	-64.70 ± 1.10	-52.24 ± 8.13
	5-Ala	Turn	-69.03 ± 9.65	-37.82 ± 1.41
	6-Cys	Coil	-153.40 ± 77.95	
7	1-Cys	Turn	-151.94 ± 3.10	154.23 ± 1.92
	2-Ala	Turn	-60.88 ± 0.17	-40.22 ± 0.39
	3-Arg	Turn	-90.27 ± 1.25	-7.44 ± 0.70
	4-Ala	Turn	-58.15 ± 0.71	-20.42 ± 0.76
	5-Ala	Turn	-59.18 ± 0.33	-42.95 ± 0.88
	6-Ala	Turn	-101.95 ± 0.59	-10.85 ± 0.35
	7-Cys	Coil	-108.23 ± 3.42	

Table S13: Conformer 1 of peptide **6** phi/psi angles in degrees, averaged over the lowest 3 energy NMR structures.

	Top Structures 1-3	
Conformer 6-1	**Phi**	**Psi**
1-Cys	-79.52 ± 1.14	53.13 ± 0.84
2-Ala	-90.62 ± 0.02	-74.12 ± 0.28
3-Arg	-95.96 ± 0.19	-9.83 ± 0.11
4-Ala	-62.93 ± 0.94	-71.01 ± 0.07
5-Ala	-91.33 ± 0.08	-35.19 ± 1.89
6-Cys	-178.66 ± 1.24	

Table S14: Conformer 2 of peptide **6** phi/psi angles in degrees, averaged over the lowest 4th -15th energy NMR structures.

	Top Structures 4-15	
Conformer 6-2	**Phi**	**Psi**
1-Cys	-83.32 ± 0.18	-22.66 ± 0.12
2-Ala	-58.46 ± 0.19	-7.42 ± 0.29
3-Arg	-90.25 ± 0.13	-71.90 ± 0.19
4-Ala	-65.50 ± 0.15	-48.42 ± 0.11
5-Ala	-64.54 ± 0.15	-37.91 ± 0.18
6-Cys	-167.45 ± 0.60	

Table S15: Conformer 3 of peptide **6** phi/psi angles in degrees, averaged over the lowest 16th -20th energy NMR structures.

	Top Structures 16-20	
Conformer 6-3	**Phi**	**Psi**
1-Cys	-46.07 ± 0.02	-35.83 ± 0.08
2-Ala	-60.51 ± 0.03	-39.93 ± 0.07
3-Arg	-92.94 ± 0.04	-73.08 ± 0.24
4-Ala	-63.83 ± 0.08	-50.17 ± 0.05
5-Ala	-66.45 ± 0.05	-39.19 ± 0.02
6-Cys	-175.42 ±0.08	

Table S16: RMSD values for the 20 lowest energy conformers calculated in Chimera, relative to the lowest energy conformer.

Peptide	Backbone RMSD (N,CA,C)	All Heteroatom RMSD
1	0.2914	0.9582
2	0.3692	0.9640
3	0.2625	1.6643
4	0.1385	1.2434
5	0.2021	1.0656
6	0.2827	1.2076
7	0.0497	1.0505

Cyclisation reaction optimisation

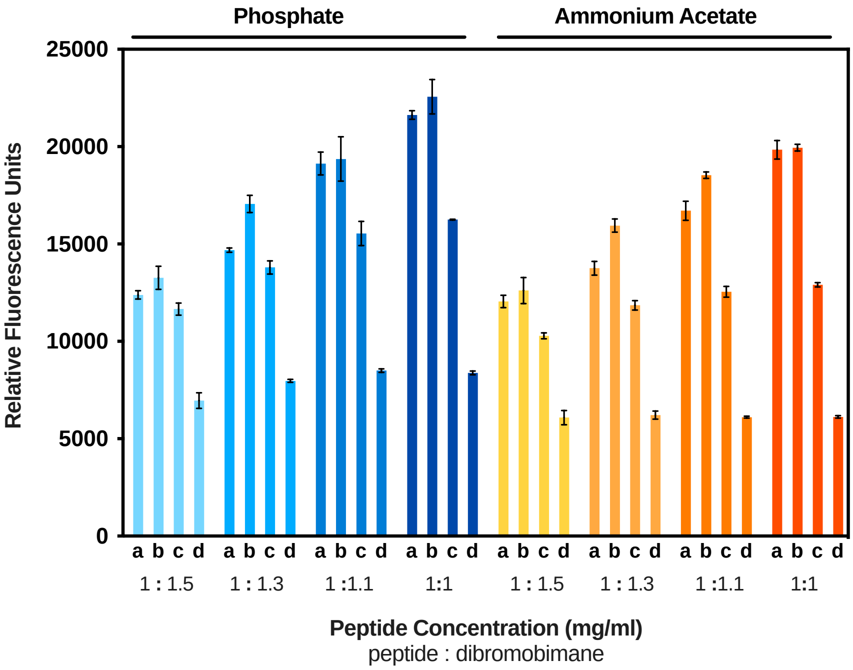


Figure S1: Summary of the peptide concentration and dibromobimane equivalent optimisation, and comparison between two buffer systems after 15 min. Optimisation conducted on linear precursor peptide of **1**. Higher fluorescence correlates to more cyclised product. 'a' 1 mg/ml, 'b' 0.5 mg/ml, 'c' 0.2 mg/ml, 'd' 0.1 mg/ml peptide concentration. Peptide:dibromobimane ratio decreases (left to right) from 1.5 to 1.0 as labelled for each buffer. Buffers used were 10 mM phosphate pH 7.8 and 10 mM ammonium acetate buffer at pH 7.8.

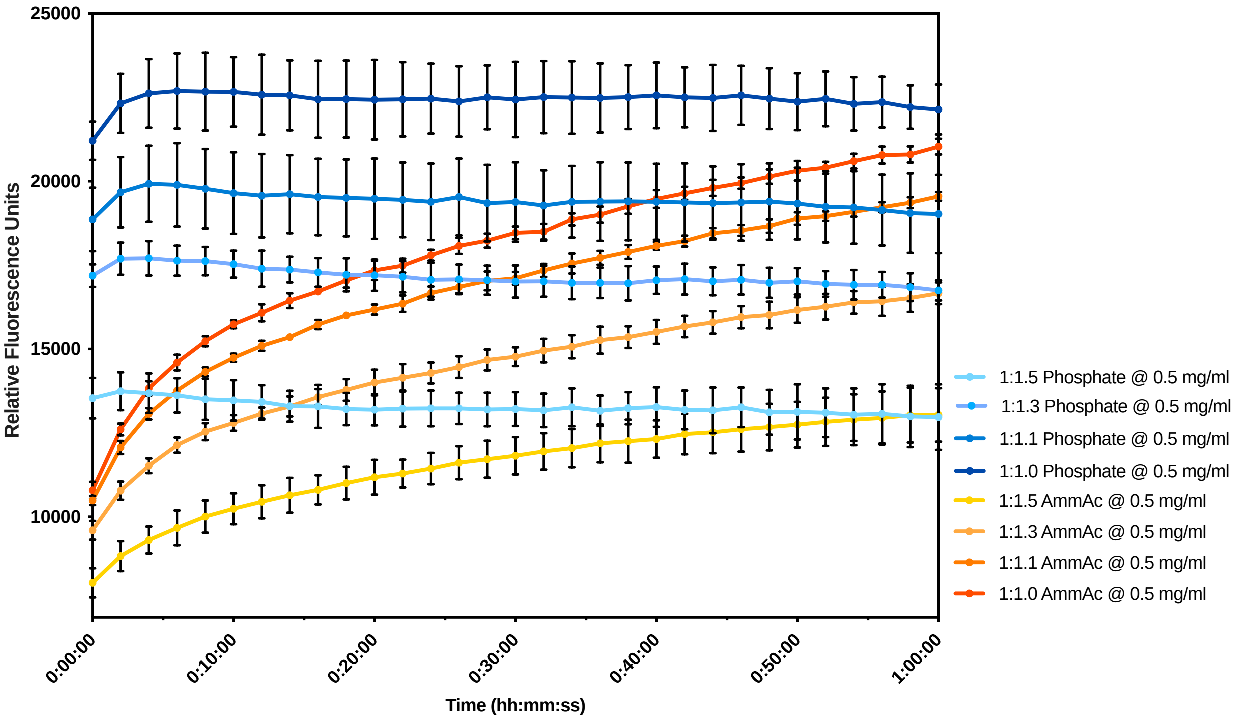


Figure S2: Time monitored fluorescence of optimisation experiments at a peptide concentration of 0.5 mg/ml. The reaction in phosphate buffer (blue) is almost complete by the first fluorescence scan (scanned as soon after addition of dibromobimane as possible, <2 min). No further increase in fluorescence indicates no further dibromobimane reaction/product formation. Ammonium acetate (orange) reaction proceeds notably slower.

Structural data

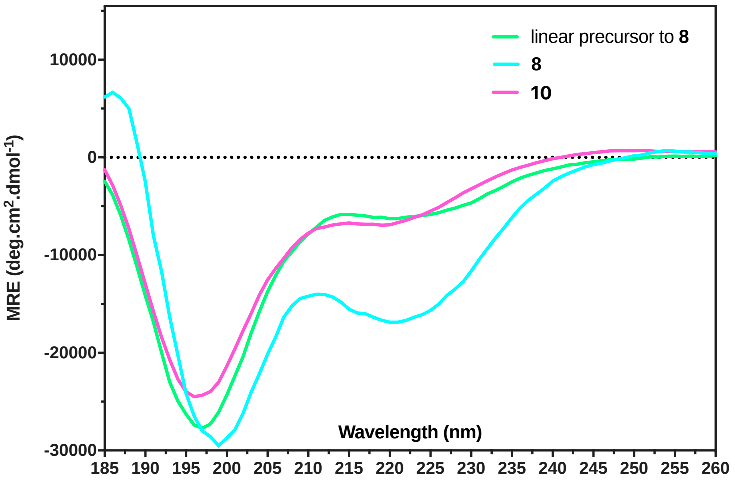


Figure S3: CD spectrum of peptide **8** (light blue), the linear precursor to **8** (green, i.e. no bimane moiety) and acyclic peptide **10** (pink) collected in 10 mM phosphate at pH 7.2. No additional features or no significant change in features are observed between the linear precursor to **8** and **10**, indicating that the bimane moiety does not absorb within this window to adversely influence the CD spectrum.

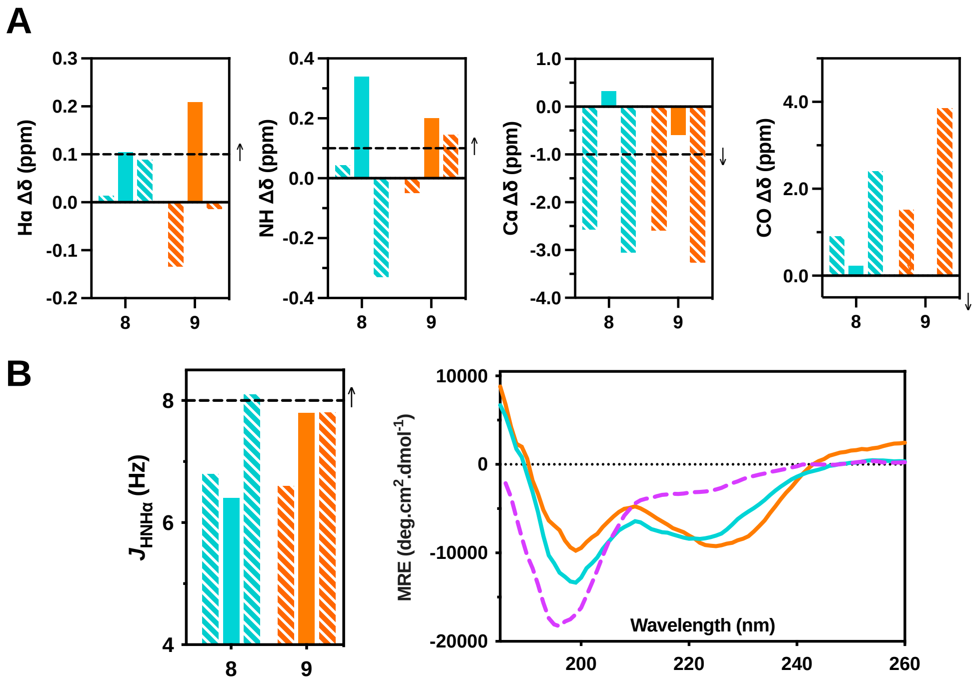


Figure S4: Structural data for **8** (light blue), **9** (orange) and **10** (pink dashed). **A)** NMR structural data (10% aq. D_2O, pH ~5): Hα, Cα and carbonyl carbon secondary shifts are calculated relative to the random coil literature values reported by Wishart (2011) and nearest neighbour corrections applied[8], each column represents a consecutive residue. Dashed columns represent cysteine residues, whose secondary shifts were calculated relative to random coil cysteine data, since more appropriate reference values were not available. Three consecutive residues should surpass the horizontal dashed lines to indicate the presence of β-strand structure. **B)** CD spectrum collected in 10 mM phosphate at pH 7.2, 298 K (Pink dashed line 10 indicates bimane itself does not introduce structure into the peptides).

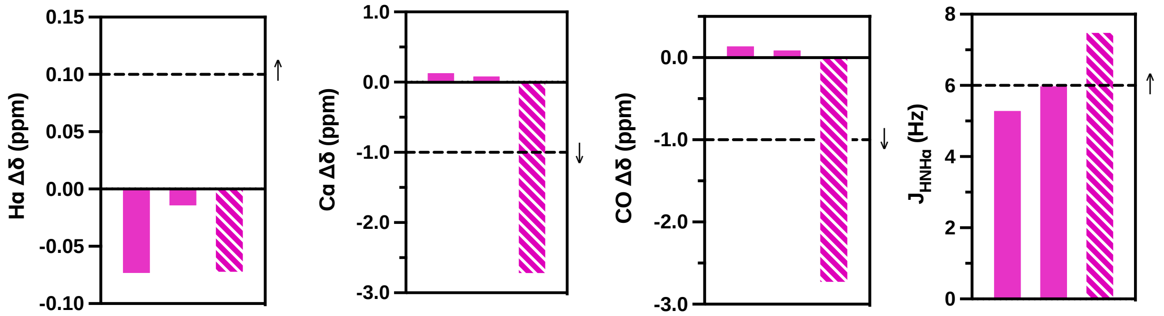


Figure S5: NMR data for peptide **10**, secondary shifts are calculated relative to random coil literature values reported by Wishart (2011),[8] support a lack of secondary structure for **10**, with small secondary shifts for both alanine α-protons (Hα), and carbonyl (CO) carbons. Dashed columns represent cysteine residues, whose secondary shifts were calculated relative to random coil cysteine data, since more appropriate reference values were not available.

Fluorescence characterisation

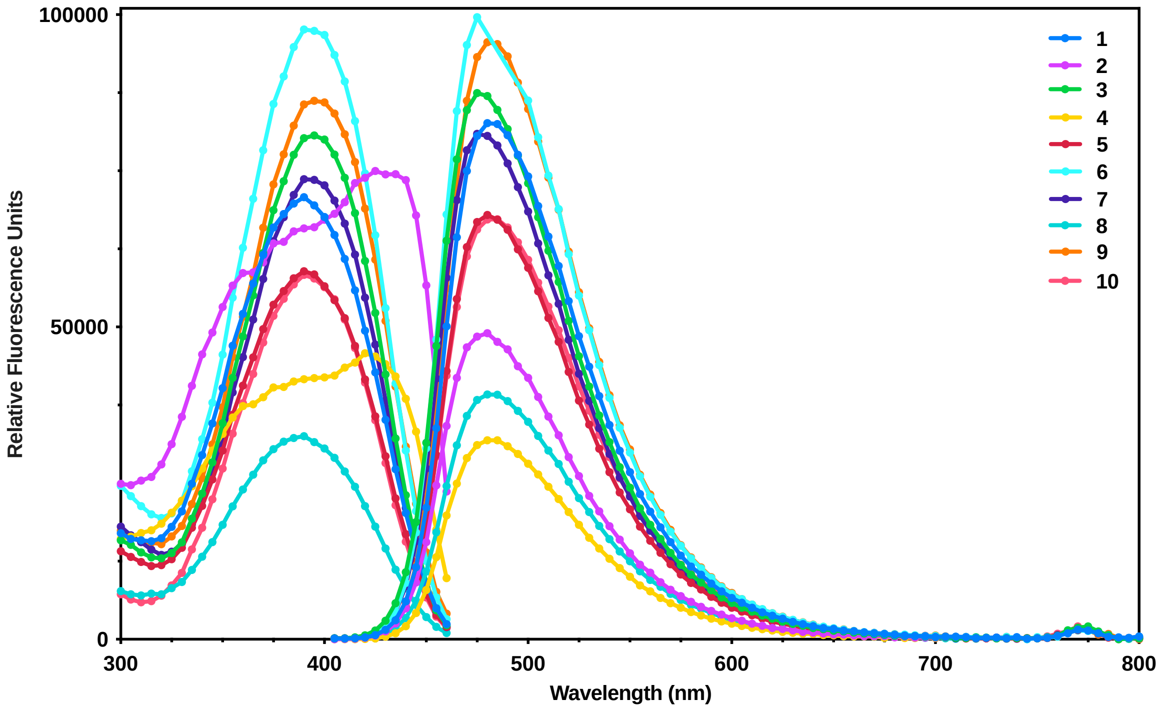


Figure S6: Fluorescence spectra of peptide **1**-**10** at 50 μM concentration in water. Fluorescence excitation spectra when emission is collected statically at 480 nm and excitation is varied from 300 nm to 460 nm over 5 nm increments (left); and the fluorescence emission spectra where excitation remains static at 385 nm and emission is collected at 5 nm increments from 400 nm to 800 nm (right). The two graphs are scaled relative to one another.

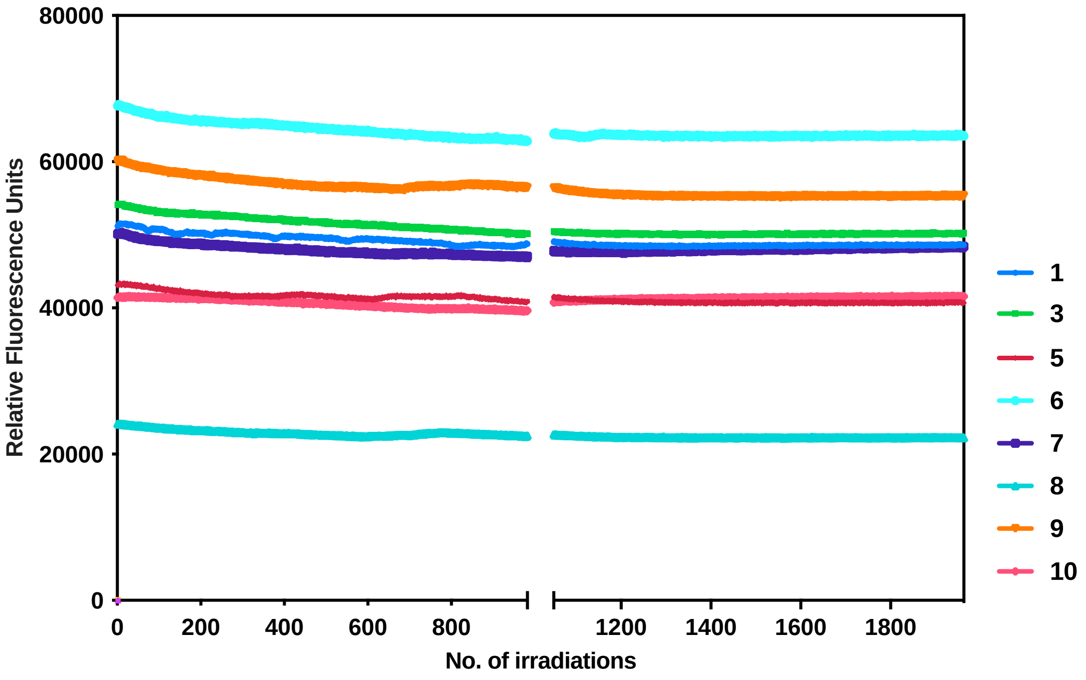


Figure S7: Fluorescence (ex. 385 nm, em. 484 nm) of peptide **1**, **3**, **5**-**10** at 50 μM concentration in water. Each well was scanned every 11 seconds for 3 hours, and then repeated for a further three hours – totalling 1963 scans for each well; using high Xenon Flash irradiation in a H4 Synergy Plate reader, as detailed in the fluorescence characterisation section. The largest decrease in fluorescence was observed for 8 with a decrease of 7.9% over 6 hours.

Table S18: Photobleaching assessment: % change in fluorescence of peptides **1**, **3**, **5**-**10** at 50 μM concentration in water. Where the initial fluorescence is in relative fluorescence units, the absolute change in fluorescence is the difference between the initial and final fluorescence, and the %change is calculated relative to the initial fluorescence for that peptide ([Absolute Change/Initial Fluorescence]*100)

Peptide	**1**	**3**	**5**	**6**	**7**	**8**	**9**	**10**
Initial Fluorescence	51188.00	54074.66	43085.33	67628.00	50114.00	24137.66	59957.33	41375
Absolute change in Fluorescence	2562.00	3957.00	2423.66	3983.33	1778.00	1912.66	4724.00	224.66
% change relative to Initial	5.01	7.32	5.63	5.89	3.55	7.92	7.88	0.54

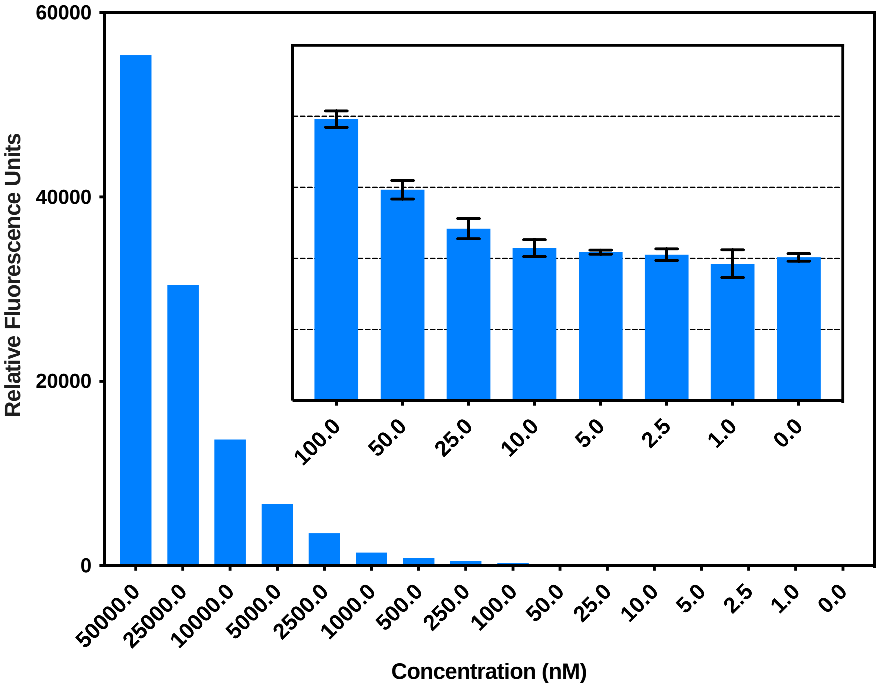


Figure S8: Fluorescence (ex. 385 nm, em. 477 nm) of peptide **1** over a concentration gradient from 50 μM to 1 nM (stock concentration determined by weight) in 10 mM phosphate at pH 7. Inset is a zoom in of the low concentration range, which was scanned separately to the higher concentrations. Peptide **1** could be clearly distinguished from the blank (0 nM) at 25 nM.

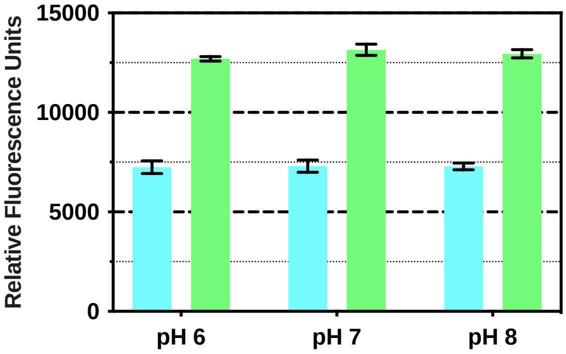


Figure S9: Fluorescence (ex. 385 nm, em. 477 nm) of peptide **1** (green) and 1**0** (blue) at a concentration of 10 μM (determined by weight) in 10 mM phosphate at pH 6, 7 and 8. No significant difference in fluorescence is observed over this biologically relevant pH range.

REFERENCES

1. N. E. Shepherd, H. N. Hoang, G. Abbenante and D. P. Fairlie, Single Turn Peptide Alpha Helices with Exceptional Stability in Water, *J. Am. Chem. Soc.*, 2005, **127**, 2974-2983.
2. G. A. Bermejo and C. D. Schwieters, Protein Structure Elucidation from NMR Data with the Program Xplor-NIH, *Methods Mol. Biol.*, 2018, **1688**, 311-340.
3. C. Schwieters, J. Kuszewski and G. Mariusclore, Using Xplor–NIH for NMR molecular structure determination, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2006, **48**, 47-62.
4. C. D. Schwieters, J. J. Kuszewski, N. Tjandra and G. M. Clore, The Xplor-NIH NMR molecular structure determination package, *J. Magn. Reson.*, 2003, **160**, 65-73.
5. M. S. Cheung, M. L. Maguire, T. J. Stevens and R. W. Broadhurst, DANGLE: A Bayesian inferential method for predicting protein backbone dihedral angles and secondary structure, *J. Magn. Reson.*, 2010, **202**, 223-233.
6. W. F. Vranken, W. Boucher, T. J. Stevens, R. H. Fogh, A. Pajon, M. Llinas, E. L. Ulrich, J. L. Markley, J. Ionides and E. D. Laue, The CCPN data model for NMR spectroscopy: development of a software pipeline, *Proteins: Struct., Funct., Genet.*, 2005, **59**, 687-696.
7. M. Heinig and D. Frishman, STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins, *Nucleic Acids Res.*, 2004, **32**, W500-502.
8. D. S. Wishart, Interpreting protein chemical shift data, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2011, **58**, 62-87.

Chapter 3.

APPROACHES TO INTRODUCE HELICAL STRUCTURE IN CYSTEINE-CONTAINING PEPTIDES WITH A BIMANE GROUP

Research Article: ChemBioChem, **2021**, 22(17), 2711-2720

Approaches to introduce helical structure in cysteine-containing peptides with a bimane group

Aimee J. Horsfall,[1-3] Daniel P. McDougal,[3,4] Denis B. Scanlon,[2,3] John B. Bruning[1-4] and Andrew D. Abell[1-3*]

[1] ARC Centre of Excellence for Nanoscale BioPhotonics (CNBP)

[2] School of Physical Sciences, The University of Adelaide, Adelaide, South Australia, 5005, Australia

[3] Institute of Photonics and Advanced Sensing (IPAS), University of Adelaide, Adelaide SA 5005

[4] School of Biological Sciences, University of Adelaide, Adelaide SA 5005

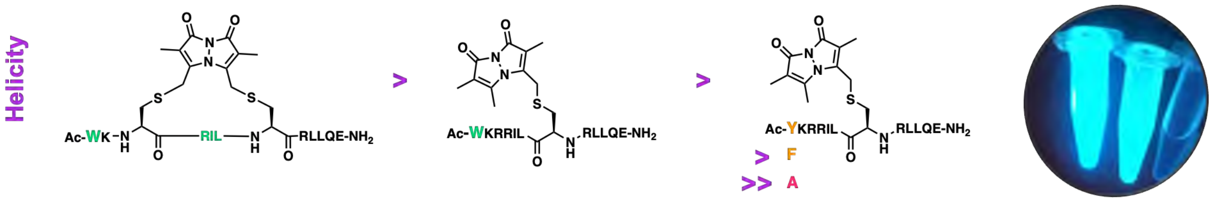
*corresponding author

First published online: 9th June 2021

© 2021 ChemBioChem, Wiley

ABSTRACT

An *i*-*i*+4 or *i*-*i*+3 bimane-containing linker was introduced into a peptide known to target Estrogen Receptor alpha (ERα) in order to stabilise an α-helical geometry. These macrocycles were studied by CD and NMR to reveal the *i*-*i*+4 constrained peptide adopts a 3_{10}-helical structure in solution, and an α-helical conformation on interaction with the ERα coactivator recruitment surface *in silico*. An acyclic bimane-modified peptide is also helical, when it includes a tryptophan or tyrosine residue; but is significantly less helical with a phenylalanine or alanine residue, which indicates such a bimane modification influences peptide structure in a sequence dependent manner. The fluorescence intensity of the bimane appears influenced by peptide conformation, where helical peptides displayed a fluorescence increase when TFE was added to phosphate buffer, compared to a decrease for less helical peptides. This study presents the bimane as a useful modification to influence peptide structure as an acyclic peptide modification, or as a side-chain constraint to give a macrocycle.



STATEMENT OF AUTHORSHIP

Title of Paper	Understanding the influence of a bimane modification on peptide secondary structure
Publication Status	☐ Published ☐ Accepted for Publication ☒ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	**Research Article:** A. J. Horsfall, D. P McDougal, D. B. Scanlon, J. B. Bruning, and A. D. Abell, *Chemistry A European Journal,* 2021

Principal Author

Name of Principal Author (Candidate)	Aimee J Horsfall		
Contribution to the Paper	Designed and synthesised peptides, NMR assignment and analysis, CD collection and analysis, fluorescence assays, analysed & discussed results, wrote and edited manuscript		
Overall percentage (%)	70%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	15/02/2021

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Daniel P McDougal		
Contribution to the Paper	Computational modelling studies, discussed results, edited manuscript		
Signature		Date	01-03-2021

Name of Co-Author	Denis B Scanlon		
Contribution to the Paper	Peptide synthesis, discussed results, and edited manuscript.		
Signature		Date	01/03/2021

Name of Co-Author	John B Bruning		
Contribution to the Paper	Supervised AJH & DPM, discussed results, and edited manuscript.		
Signature		Date	18/2/21

Name of Co-Author	Andrew D Abell		
Contribution to the Paper	Supervised AJH, discussed results, and edited manuscript.		
Signature		Date	15/2/2021

3.1 Introduction

The syn-9,10-(CH_3,CH_3)bimane (**Bim** Figure 1, referred to herein simply as a bimane group) is a versatile fluorophore that is commonly introduced to thiol-containing species by reaction with a non-fluorescent reagent such as dibromobimane (**dBB**, Figure 1) or monobromobimane (**mBB**, Figure 1).[1-13] The bimane fluorophore has found many applications in cell biology and protein structure studies, where bimane derivatives (such as **dBB** or **mBB**) are used to detect hydrogen sulphide and glutathione,[4-8] or react with cysteines found in peptides and proteins as a basis for protein structure determination assisted by metal ion FRET[9] or proteomics.[10, 11] Here we present a study on how the bimane group influences secondary structure in short peptides, as a basis for developing inherently fluorescent peptidomimetics to study protein interactions. There are two possible mechanisms for a bimane modification, in a peptide, to influence the peptide secondary structure – as a covalent constraint where the bimane linker defines a macrocycle, or via a non-covalent interaction with an aromatic residue where the bimane is appended to a single cysteine side-chain.

Recently, we reported bis-alkylation of two cysteine residues in short peptides with **dBB** to introduce a covalent linker that defines secondary structure, while also providing a fluorescent tag for imaging by fluorescence microscopy.[14] The bimane linker was installed on-resin following assembly of the peptide sequence by solid-phase peptide synthesis, and also introduced off-resin in a phosphate buffered solution. The bimane linker was introduced as an *i*-*i*+2, through to an *i*-*i*+7 constraint, to give rise to macrocycles that ranged from 16 to 31 atoms in size. The *i*-*i*+4 constrained pentapeptides, and a *i*-*i*+3 constrained tetrapeptide displayed α-helical structure as defined by CD and NMR studies. These short bimane-linked macrocyclic peptides were cell permeable, and were imaged using the bimane fluorescence.[14] The bimane linker provides a means to study macrocyclic peptides in cells, without the need for attaching an auxiliary fluorophore (e.g. fluorescein) which can interfere with secondary structure, cell permeability and also interaction with the target protein.[15-18]

mBB has been used extensively to label cysteine residues in proteins to study protein structure dynamics via distance mapping experiments.[19-22] The bimane fluorescence is quenched by a short-range interaction (< 11 Å) with tryptophan, in a distance dependent manner. This provides a basis to assess conformational changes in proteins using intramolecular fluorescence quenching, where the bimane is introduced at a site such that it will encounter a tryptophan following a conformational change in the protein.[21] This fluorescence quenching effect is also observed, to a lesser extent, for tyrosine and phenylalanine residues. The nature of the interaction between the bimane and aromatic amino-acid that gives rise to the fluorescence quenching is unclear.

Here we investigate, in extension to our previous study, the secondary structure induced in an *i*-*i*+3 or *i*-*i*+4 linked bimane macrocycle. Additionally, the secondary structure influence of a tryptophan, tyrosine, phenylalanine or alanine residue in the presence of a bimane-modified cysteine is investigated, as interactions between the aromatic residues and the bimane have been noted in literature. Here we explore

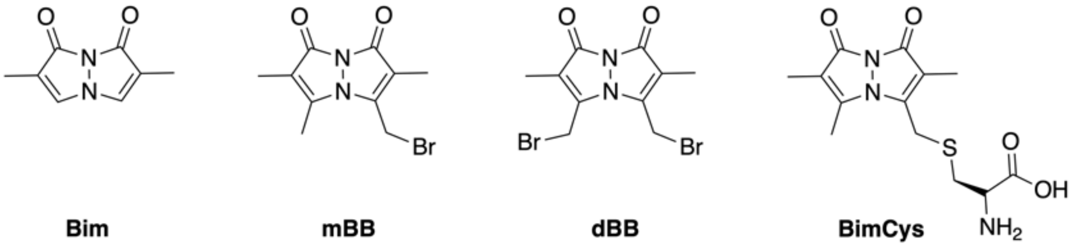


Figure 1: Bimane fluorophore derivatives. Monobromobimane (mBB) and dibromobimane (dBB) are non-fluorescent and are decorated with bromine group/s to allow reaction with thiols.

the influence of the bimane on peptide structure in a sequence known to target Estrogen Receptor Alpha (ERα).[23-27] Activity of the nuclear receptor, ERα, is inhibited by helical stapled peptides derived from ER coactivator proteins. Such peptides bind through three key conserved leucine residues, arranged as a LxxLL motif, that embeds into a hydrophobic cleft on the activation function-2 (AF-2) surface of the ERα ligand binding domain (LBD), and thereby blocks coactivator protein recruitment.[23-29] Peptide inhibitors of the ERα—coactivator interaction may have therapeutic utility in diseases with overexpression of ERα, such as ER-positive cancers.[27, 30] This study defines the bimane as a useful peptide modification to introduce helical secondary structure in biologically important peptides, and simultaneously introduce a fluorophore which provides a handle to enable fluorescence imaging studies of the bimane-modified peptides. Additionally, the resulting peptides may find use as inherently fluorescent protein interaction probes or inhibitors of ERα.

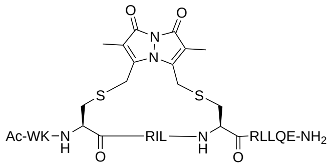
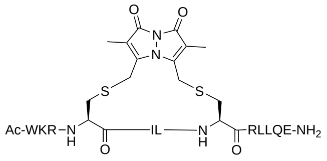
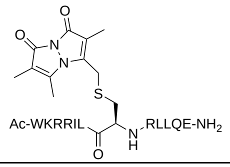
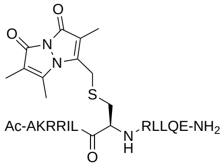
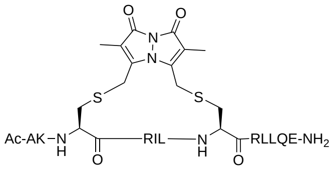
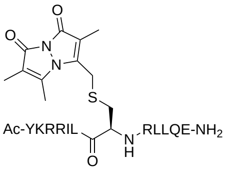
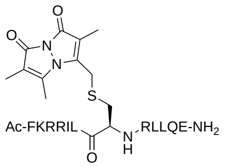
3.2 Results and discussion:

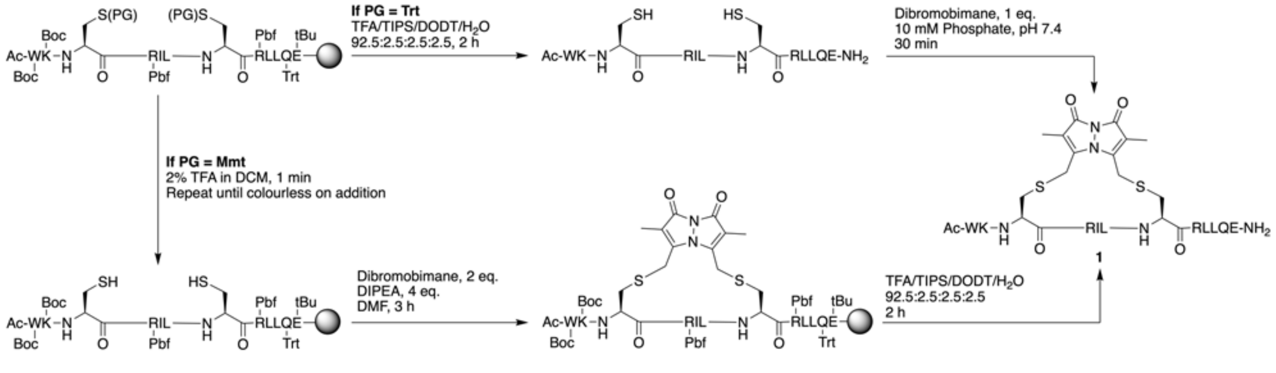
Design and synthesis

The peptide sequence, Ac-WKARILARLLQE-NH$_2$, is known to bind ERα on the AF-2 surface, as an α-helix where three conserved leucines (LxxLL) are embedded onto the protein surface.[23] This sequence was chosen to study the ability of the bimane modification to preorganise a peptide backbone into an α-helical conformation. Cysteines at positions 3 and 7, or 4 and 7 were introduced (**1a** and **2a**, Table 1) for reaction with **dBB**, to give macrocycles **1** (*i*-*i*+4) and **2** (*i*-*i*+3). The acyclic bimane peptide (**3**) was also prepared, where a cysteine was included at position 7 (peptide **3a**) and reacted with **mBB**. An arginine was included at position 3, where alanine is known to favour helical structure. Peptide **3** was initially prepared to serve as a control for Circular Dichroism (CD) experiments, to ensure the bimane did not absorb in the window relevant to secondary structure determination (190-260 nm).

Resin-bound precursors to bimane-modified peptides **1**, **2** and **3** were synthesised by Fmoc solid-phase peptide synthesis as described in the experimental section. The resin-bound peptides were treated with 2% TFA to remove the 2-methoxytrityl (Mmt) cysteine protecting groups and then reacted with **dBB** (2 equiv, in the case of macrocyclic **1** and **2**) or **mBB** (2 equiv, in the case of acyclic **3**), and DIPEA (4 equiv) on resin (Scheme 1, PG = Mmt). The peptides were cleaved from the resin and globally deprotected by acidolysis with TFA, to give the cyclised peptides **1** and **2**, and acyclic peptide **3**. Analysis by RP-HPLC revealed a single major peak, with m/z signals corresponding to **1** and **2**, and acyclic **3**, without evidence

Table 1: Bimane modified peptides and the precursor peptides where C* indicates cysteine residues that are reacted with the bimane group

Code	Sequence	Modification	Structure
1a	Ac-WKCRILCRLLQE-NH2	unmodified	
1	Ac-WKC*RILC*RLLQE-NH2	*i*-*i*+4 cyclised at C* with bimane	
2a	Ac-WKRCILCRLLQE-NH2	unmodified	
2	Ac-WKRC*ILC*RLLQE-NH2	*i*-*i*+3 cyclised at C* with bimane	
3a	Ac-WKRRILCRLLQE-NH2	unmodified	
3	Ac-WKRRILC*RLLQE-NH2	bimane reacted at C*	
4	Ac-AKRRILC*RLLQE-NH2	bimane reacted at C*, W1A mutant	
5	Ac-AKC*RILC*RLLQE-NH2	*i*-*i*+4 cyclised at C* with bimane, W1A mutant	
6	Ac-YKRRILC*RLLQE-NH2	bimane reacted at C*, W1Y mutant	
7	Ac-FKRRILC*RLLQE-NH2	bimane reacted at C*, W1F mutant	


If PG = Trt
TFA/TIPS/DODT/H2O
92.5:2.5:2.5:2.5, 2 h
Dibromobimane, 1 eq.
10 mM Phosphate, pH 7.4
30 min
If PG = Mmt
2% TFA in DCM, 1 min
Repeat until colourless on addition
Dibromobimane, 2 eq.
DIPEA, 4 eq.
DMF, 3 h
TFA/TIPS/DODT/H2O
92.5:2.5:2.5:2.5
2 h
1

Scheme 1: Representative synthesis of bimane cyclised peptide **1**. Cyclisation was performed on a fully protected peptide on-resin (PG = Mmt) or on an unprotected peptide in solution (PG = Trt).

of unreacted peptide (**1a**, **2a** or **3a**). Each peptide was purified by RP-HPLC and characterised by MS and 2D ^{1}H homonuclear NMR (zTOCSY and ROESYAD).

Peptides **1**, **2** and **3** were also prepared in solution by reaction of the fully unprotected precursor peptides (**1a**, **2a** and **3a**, Table 1) with the required bromobimane (**dBB** to give **1** and **2**, or **mBB** to give **3**), to demonstrate that the cysteine thiol/s react selectively with bromobimane in the presence of the amine (Lys), amide (Gln), carboxylic acid (Glu), guanadinium (Arg) and indole (Trp) groups of the amino-acid side-chains. The precursor peptides **1a**, **2a** and **3a** were synthesised by solid-phase peptide synthesis, then cleaved from the resin by acidolysis with TFA, which simultaneously removed all side-chain protecting groups. The crude peptides were dissolved in phosphate buffer at pH 7.4, and then **dBB** (1 equiv, in the case of **1a** and **2a**) or **mBB** (1 equiv, in the case of **3a**) was added (Scheme 1, PG = Trt). Each reaction solution was stirred for 30 min, lyophilised, and subsequently analysed by MS to reveal masses corresponding to the desired peptides **1**, **2** and **3**, respectively. RP-HPLC revealed one major peak in each reaction mixture. Additionally, the major peak of the crude reaction mixture displayed the same retention time as the bimane-modified peptides that were prepared on-resin (Figure S3). The crude lyophilised bimane-reacted peptides were then subjected to an Ellman's reagent test to confirm the reaction had reached completion and determine if the **dBB**/**mBB** had reacted with only the cysteine thiols. Only one equivalent of **dBB** or **mBB** (relative to the peptide) was added to the reaction mixture, hence cysteine thiols would remain if **dBB**/**mBB** had instead reacted with an alternate functional group (Lys/Gln/Glu/Arg/Trp). A sample of each **1-3**, and also **1a** (positive thiol-containing control), were dissolved in Tris.HCl pH 7.5 and 5,5'-dithio-bis-(2-nitrobenzoic acid) (DTNB, Ellman's reagent) was added (see experimental). A second sample of each peptide was prepared without added DTNB. All samples were incubated at rt for two min and then analysed, in triplicate, in a 96-well plate. The absorbance spectrum, measured between 300-700 nm, was collected. The difference between the absorbance spectra with and without DTNB (averaged across the three replicates) was calculated for each peptide to remove the contribution of the bimane absorbance, centred around 380 nm, in each spectrum. The positive-control peptide **1a** displayed higher absorbance at 412 nm (0.085, Figure S1, yellow) owing to the liberated 2-nitro-5-thiobenzoate (TNB), compared to the blank sample containing no peptide (0.009, Figure S1, grey). The absorbance at 412 nm of crude peptides **1**, **2** and **3** was comparable to the blank absorbance (0.013, 0.000 and 0.004 respectively; Figure S1, purple, green and blue), thus no free thiols were evident. The linear precursor peptides (**1a**, **2a** or **3a**) were not observed by MS in the crude reaction mixtures, also in support of complete reaction of **dBB** or **mBB** with the respective peptide. Each sample was then purified by RP-HPLC and the major peak was verified to correspond to the desired bimane-modified peptide (**1**, **2** or **3**) by NMR. A NOE crosspeak observed between the cysteine C_β protons and bimane methylene protons indicated the bimane was appended to the cysteine side-chain/s (see characterisation data in the SI). **dBB** or **mBB** thus reacts selectively with cysteine thiols in an unprotected peptide sequence.

Structural analysis:

The secondary structures of linear peptide **1a** and macrocycles **1**, **2** and **3** were investigated by CD spectroscopy (Figure 2) and NMR secondary shift analysis (Figure 3), to determine if the peptide backbone of the bimane-constrained macrocycles was preorganised into an α-helical conformation. Peptide **3** was expected to be unstructured and prepared to ensure the fluorophore did not absorb in the CD window relevant to secondary structure determination (190-260 nm), nor significantly alter the NMR chemical shift of the modified cysteine residue. The peptide main-chain ^{1}H resonances were assigned using ROESYAD and zTOCSY spectra (Table S2). The α-proton (αH) and amide proton (NH) secondary shifts were calculated as the difference (in ppm) of the experimentally observed resonance from a literature random coil value for the corresponding residue type (Table S2).[31, 32] The secondary shifts of the first three *N*-terminal residues in the linear peptide **1a** suggest helical structure, where these values are negative (Figure 3, grey), but the remaining secondary shifts are variable in magnitude which suggests a lack of structure. In support of an undefined secondary structure, the CD spectrum of **1a** showed a single deep minimum at 200 nm (Figure 2A, grey), as expected.[33] Any defined structure in the bimane-modified peptides (**1**, **2** and **3**) is thus attributed to introduction of the bimane, as the unmodified peptide **1a** is unstructured. The CD data for *i*-*i*+3 cyclised peptide **2** also revealed a lack of secondary structure with a single minimum at 203 nm (Figure 2A, orange). This is supported by the secondary shifts of **2** which were highly variable (Figure 3, orange). It is thus clear that an *i*-*i*+3 bimane staple does not stabilise helical structure in this sequence.

The CD spectrum of *i*-*i*+4 cyclised peptide **1** revealed a major minimum at 206 nm and a second minimum at 218 nm, accompanied by a strong maximum at 188 nm (Figure 2, purple), consistent with helical content.[33] The NMR secondary shifts of **1** are almost uniformly negative (Figure 3, purple), further supporting helical structure. The helicity of peptide **1** is enhanced relative to unconstrained peptide **1a**, with calculated helicities of 66%-73% (based on the MRE intensity at 222 nm),[34] and 42% helicity respectively (Figure 2, Table S1). The ratio between the minima at 208 and 222 nm in the CD spectrum of **1** was calculated to be 0.66, suggesting 3_{10}-helical structure (expected ratio of 0.6),[34-37] as opposed to the 0.9 ratio expected for α-helical structure (Figure 2A, Table S1). The *i*-*i*+4 bimane staple thus stabilises 3_{10}-helical structure, rather than α-helical structure, in peptide **1**. The *i*-*i*+4 bimane staple of **1** clearly enhances α-helicity, relative to the linear peptide **1a**. Additionally, a qualitative comparison of the peptide **1** CD spectrum to the analogous peptide sequence constrained with an *m*-xylene, perfluoroaryl or vinyl sulphide linkers,[23] suggests that **1** is more helical than these macrocycles, indicated by deeper 222 and 208 minima. The helicity of **1** is somewhat reduced relative to the analogous all-hydrocarbon, lactam or triazole linked peptides.[23]

The CD spectrum of acyclic bimane peptide **3** (Figure 2A, light blue) unexpectedly gave minima at 204 and 218 nm and a maximum below 190 nm again indicating helical structure. Large negative proton secondary shifts were observed for **3**, also consistent with helical structure (Figure 3, light blue). In particular, the negative secondary proton shifts are large for the *N*-terminal residues (1-5 αH and 2-4 NH, Figure 3, light blue), which suggests a stable helical region. It is surprising that the acyclic peptide **3** reveals a defined secondary structure given that the peptide is not constrained. This secondary structure stabilisation may be the result of an interaction between the bimane and tryptophan side-chain, which has been observed elsewhere by quenching of the bimane fluorescence at up to 10 Å separation. However, the influence of such an interaction on local structure has not been well characterised.[9, 20, 22] This type of interaction is analogous to amino-acid side chain interactions where sequence specific interactions are known to stabilise helicity.[38-43] A fourth peptide (**4**, Table 1) was prepared where the tryptophan of **3** was substituted for alanine to investigate this (Table 1). The absence of the tryptophan indole would reveal whether the secondary structure in peptide **3** was sequence specific and indicate if tryptophan was important to this stabilisation. The CD spectrum of peptide **4** revealed a single deep minimum at 200 nm (Figure 2A, yellow), consistent with a lack of defined secondary structure. Additionally, the NMR secondary shifts (Figure 3, yellow) were highly variable in magnitude in support of a lack of defined structure. Hence this suggests a key role of the tryptophan at position 1 in stabilising the helical structure observed for bimane-modified peptide **3**.

An analogue of the cyclic peptide **1**, peptide **5**, was also prepared to ensure that the helicity of **1** was also not dependent on the tryptophan residue, as seen for acyclic peptide **3** compared to peptide **4**. The CD spectrum of peptide **5** (Figure 2B, green) displayed a minimum at 220 nm and a deeper minimum at 203 nm, which is consistent with some helical structure. The CD spectrum suggests a slightly lower helicity of **5** at 48%, c.f. 73% for **1** (Figure 2B, green c.f. purple). The NMR secondary shifts displayed a near identical pattern to the secondary shifts of peptide **1**, consistent with helical structure for peptide **5**. This indicates that the W1A modification in **5** did not significantly impact the peptide helicity, compared to **1**.

Analogues of peptide **3**, with a tyrosine (W1Y, peptide **6**) or phenylalanine (W1F, peptide **7**) modification at position 1, were next prepared to investigate if the helicity of the acyclic peptide **3** could be influenced by these aromatic residues. Tyrosine and phenylalanine are known to interact with the bimane moiety in unrelated protein systems.[20] The CD spectrum of peptide **6** revealed a minimum at 215 nm consistent with some helical structure (61%, Figure 2B, blue). The NMR secondary shifts of peptide **6** were all negative, consistent with helical structure (Figure 3, blue). The CD spectrum of peptide **7** displayed a smaller minimum at 200 nm (Figure 2B, red), compared to peptide **6**, and a shallow minimum at 215 nm, which corresponds to a helicity of 41%. The majority of the NMR secondary shifts for peptide **7** were negative, however as consecutive residues do not have a large negative shift this is not consistent with stabilised helical structure (Figure 3, red). In summary, the helicity of the acyclic peptides that contained a bimane appended to a cysteine at position 6, was **3** (W) > **6** (Y) > **7** (F) >> **4** (A) where the letter in brackets denotes the residue at position 1. This suggests an aromatic residue at position 1 enhances helical structure when a bimane modified cysteine is positioned six residues away.

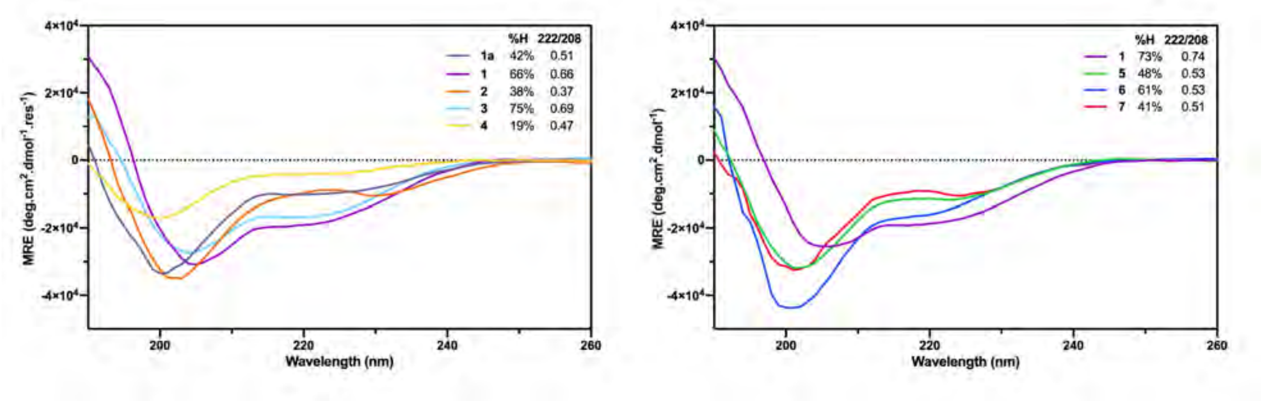


Figure 2: Circular dichroism (CD) spectra. MRE – molar residue ellipticity. All spectra were fitted with an in-built Savitzky-Golay smoothing function. Percentage helicity (%H) was calculated based on the 222 nm minimum.[34] 222/208 represents the ratio of the intensity of the MRE at 222 and 208 nm. All values (Table S1) and calculations are provided in the SI. **A** CD spectra of peptides **1a**, **1**, **2**, **3**, **4** (50 μM) collected in 10 mM phosphate buffer at pH 7.2. **B** A separate CD experiment, spectra of peptides (50 μM) collected in 10 mM phosphate buffer at pH 7.2 with the addition of 20% 2,2,2-trifluoroethanol (TFE).

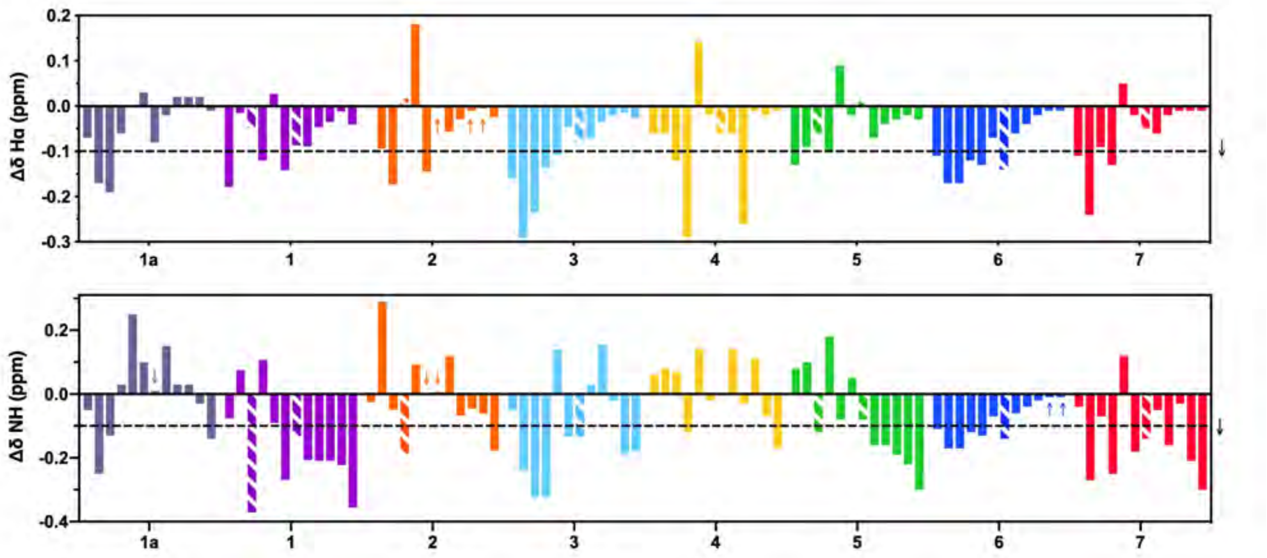


Figure 3: NMR secondary shift analysis calculated relative to random coil shifts reported by Wishart (2011) and nearest neighbour corrections applied,[31] for the α-proton (αH) and amide proton (NH). Values are tabulated in Table S2. Each bar represents a subsequent amino-acid in the peptide sequence. Arrows indicate small secondary shifts, on the side of zero which they lie. Bimane reacted cysteine residues were calculated relative to Cys (reduced) and are indicated by striped bars. The dashed line (at -0.1 ppm) indicates the threshold which should be surpassed to indicate helicity; values close to zero are indicative of random coil structure. NMR spectra were collected in 10% aq. D_2O at pH ~5, 298 K.

Alpha-helical ERα binding conformation

Peptides **1** and **3** were modelled onto the ERα surface *in silico* to see if this reveals an α-helical peptide conformation required to facilitate interaction with ERα. These peptides were chosen for study given they adopt a helical structure in solution, as determined by CD and NMR. Peptides **1** and **3** were manually docked onto the coactivator binding site of the ERα LBD (PDB ID: 6CBZ)[44] *in silico* and consequently energy minimised and annealed (n=20) in *ICM-Pro* using the inbuilt function.[45] The conserved hydrophobic binding motif, LxxLL, of both peptides **1** and **3** clearly inserts into the hydrophobic cleft of the AF-2 surface (Figure 4 and S4-6), as per other published crystal structures of short coactivator peptides bound to ERα.[27, 28, 46] Peptides **1** and **3** both adopt a classical α-helical conformation when bound to ERα (Figure 4).

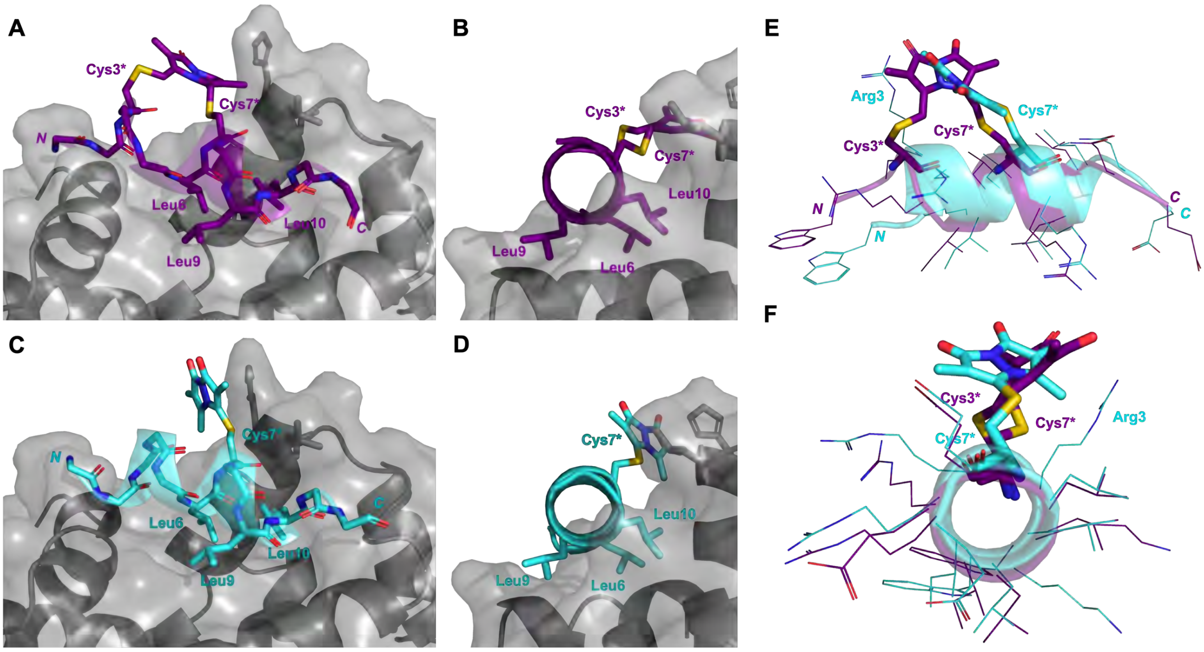


Figure 4: Recognition of the AF2 surface by the conserved LxxLL motif is not perturbed by incorporation of the bimane group. Peptides **1** (**A** and **B**, purple), **3** (**C** and **D**, light blue) presented as cartoon loop bound to ERα in grey with the chain shown as cartoon loop and and transparent surface representation. Shown in sticks are the conserved Leu residues for AF-2 surface recognition and bimane modified cysteine residues. **A** and **C** the mainchain of the peptide is shown in sticks. **E** and **F** show peptides **1** (purple) and **3** (light blue) in cartoon representation overlaid with the side-chains shown as lines and the bimane-modified cysteine residues as sticks. Atoms: carbons (colour coded to compound), oxygen (red), nitrogen (blue), sulfur (yellow). Peptide **1a** was also computationally modelled as a control and is shown in in Figure S4.

Peptide **1** was determined to adopt a 3_{10}-helix in solution, which is known to precede α-helical structure in proteins,[47] and may help preorganise the peptide backbone *en route* to adopting an α-helical structure. Importantly, the bimane staple of **1** allows sufficient flexibility of the peptide backbone to adopt the α-helix conformation necessary to interact with ERα and peptide **1** is not fixed in a 3_{10}-helical conformation. This is supported by a classical α-helical CD spectrum of **1** in 10 mM phosphate with 20% TFE (Figure S4). An interaction between the tryptophan side-chain and bimane was hypothesised to stabilise the helical structure in peptide **3**, however this interaction was not observed *in silico* during molecular modelling, but does not preclude the hypothesised interaction from occurring in solution, or when bound to ERα.

Fluorescence characterisation

The bimane fluorescence in the modified peptides were characterised in order to determine whether the peptide structure altered the bimane fluorescence. All peptides were dissolved in phosphate buffer, and analysed in triplicate. The absorbance and fluorescence emission spectra (ex. 380 nm) were collected (Figure S8), where the fluorescence spectra were normalised to the 380 nm absorbance and are shown in Figure 5. All peptides displayed a fluorescence maximum around 480 nm and the fluorescence intensity of peptides **3**, **4**, and **6** was identical. The fluorescence intensity of peptide **7** was modestly higher than the other acyclic peptides (**3**, **4** and **6**). The interaction between the bimane and tryptophan of **3**, hypothesised to stabilise the helical structure, might be expected to quench the bimane fluorescence.[20] However, that was not observed here where the fluorescence of peptide **3** and the unstructured W1A modified peptide **4**, were of identical intensity. This suggests that any interaction between the bimane and tryptophan (or phenylalanine or tyrosine), is not sufficiently stabilised to result in fluorescence quenching of the bimane. The cyclic peptides **1** and **5** displayed lower fluorescence intensity than the acyclic peptides (**3**, **4**, **6** and **7**) which suggests the bimane fluorescence may be influenced by the sequence or peptide structure, however in this case it is not clear what effect is responsible.

Fluorescence spectra were also collected for peptide **1**, and **3**-**7** in 20% TFE in phosphate (ex. 380 nm) where TFE is known to promote helical folding of peptides, in order to determine if the difference in bimane fluorescence intensity was related to peptide structure. The emission maximum of these peptides is blue-shifted slightly to 470-475 nm (Figure 5A, Table S3). Further, the fluorescence intensity of all the peptides in 20% TFE in phosphate buffer was different than the fluorescence intensity of the corresponding peptide in phosphate buffer without TFE (Figure 5B). The fluorescence intensity for peptides **1**, **3**, **5** and **6** was greater in 20% TFE in phosphate buffer, than phosphate buffer without TFE. Interestingly, peptides **1**, **3**, **5** and **6** are the peptides with the greatest helical character in phosphate buffer. Conversely, the fluorescence intensity was lower for peptides **3** and **7** in 20% TFE in phosphate buffer, than in phosphate buffer without TFE. Peptides **3** and **7** show the least helical character and are largely random coil in phosphate buffer. This may indicate that changes in the bimane fluorescence intensity, between phosphate buffer with or without 20% TFE, are likely related to the peptide structure and folding, where it is known the peptides adopt a more α-helical conformation in 20% TFE. The fluorescence of a bimane modified cysteine residue (**BimC**, Figure 1) is greater in phosphate buffer with 20% TFE compared to without (Figure S10D). There was also small blue-shift of the **BimC** fluorescence in 20% TFE in phosphate, relative to phosphate buffer, as was seen for peptides **1**, and **3**-**7**. The fluorescence properties of the bimane, when incorporated into a protein, are reported to be influenced by conformational changes of the protein that in turn alter the local microenvironment of the fluorophore.[48, 49]

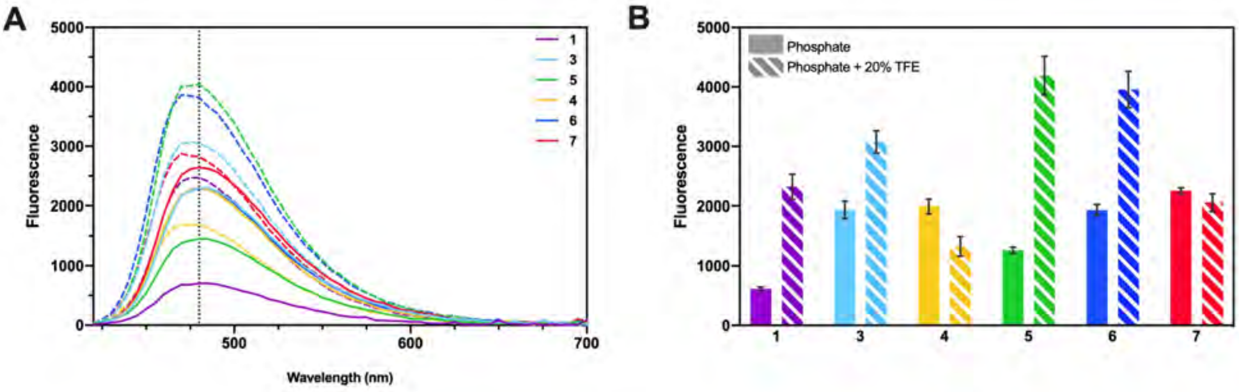


Figure 5: Fluorescence characterisation of the bimane peptides excited at 380 nm in 10 mM phosphate pH 7.2 (solid line/bar) and in 10 mM phosphate pH 7.2 with 20% TFE (striped line/bar). Spectra have been corrected for the absorbance at 380 nm, see full spectra in SI(Figures S6 and 7). **A** Fluorescence emission spectra. Dotted black line *i* marks 480 nm emission **B** Comparison of the emission intensity at 480 nm.

3.3 Conclusions

Seven bimane-containing ERα co-activator peptides (**1**-**7**) were prepared and the resulting secondary structures characterised by CD and NMR secondary shift analysis. The *i*-*i*+4 constrained peptide **1** shows 3_{10}-helical structure in solution at pH 7.2 with a classical α-helical geometry on the addition of 20% TFE, conditions known to promote helicity. Interestingly, the acyclic bimane-modified peptide **3** that contains a tryptophan at position 1, and a bimane-modified cysteine at position 6, also adopts a helical structure. Helicity was also observed for the W1Y modified peptide **6** and weakly in the W1F modified peptide **7**, but not the W1A modified peptide **4**. The helicity of the acyclic bimane-modified peptides is summarised as **3** (W) > **6** (Y) > **7** (F) >> **4** (A), where the amino-acid denoted in parentheses is 6 residues away from the bimane modified cysteine at position 6. NMR secondary shift analysis of peptide **3** and **6** indicates that helicity is largely localised to the *N*-terminal residues of the peptide.

The *i*-*i*+4 constrained peptide **1** and acyclic bimane-modified peptide **3**, which adopt a 3_{10}-helix in solution, adopt α-helical structure on binding ERα *in silico*, consistent with the bioactive helical conformation requisite for the LxxLL motif to be positioned into the hydrophobic cleft on the ERα surface. Peptide sequences that contain a tryptophan (or tyrosine) and a judiciously placed bimane-modification, may provide useful scaffolds to develop helical protein-protein interaction inhibitors inspired by native protein sequences, where aromatic residues are commonly found at these interfaces. An investigation into the ability of the peptides presented here to bind and inhibit ERα *in vitro* will be subject of a future study.

We thus present the bimane group (as part of a linker, or as a stand alone entity) as a versatile peptide modification unit that introduces helicity as an *i*-*i*+4 peptide staple; or stabilises helicity in a sequence dependent manner when appended to an appropriately positioned cysteine side-chain. In addition, the bimane is fluorescent which allows biological imaging (e.g. fluorescence microscopy or fluorescence-assisted cell sorting (FACS)) without further modification as is normally required for other peptide constraint methodologies.[14] Additionally, the bimane modification provides a fluorophore to exploit in fluorescence polarisation assays to determine the binding affinity between a peptidomimetic and protein of interest. The bimane fluorescence intensity appears to be influenced by the peptide conformation, which may provide a basis to detect conformational changes in short peptides, or study protein interactions. Furthermore, these ERα coactivator-derived peptides may themselves serve as a therapeutic lead for ERα-dependent disorders, and will be the subject of further investigation.

3.4 Experimental section

3.4.1 Synthesis and characterisation

Unless otherwise indicated, all starting materials were purchased from commercial sources and used without further purification. All peptides were synthesised by the Fmoc solid-phase peptide synthesis protocol detailed below, with all L-amino-acids (unless otherwise specified), and then *N*-terminally acetylated before reaction with bromobimane on-resin, where appropriate. Peptides were subsequently cleaved from the resin (and simultaneously globally deprotected). Purification was carried out by semi-preparative HPLC on a Gilson GX-Prep RP-HPLC system on a Phenomenex Aeris Peptide C18 (10 x 250 mm), over a gradient as specified for each compound in the Supplementary

Information. RP-HPLC solvents were buffer A: H_2O with 0.1% TFA, and buffer B: ACN with 0.1% TFA; 0.2 µm filtered. Peptide identity was confirmed by High-resolution mass spectra were collected using an Agilent 6230 ESI-TOF via direct injection in ACN with 0.1% formic acid as the running buffer; and 2D NMR (zTOCSY and ROESYAD) provided in the Supplementary Information. Purity of all compounds was confirmed by analytical RP-HPLC on an Agilent 1260 HPLC equipped with a Phenomenex Luna C18(2) column (250 x 4.6 mm) over a gradient of 5-50% B (15 min). All graphs were generated using GraphPad Prism 9 software and docking structures visualised with PyMOL 2.0.

General solid-phase peptide synthesis:

The following Fmoc-protected amino-acids were used for all peptides unless otherwise specified. All amino-acids were L- and purchased from Chem-Impex Int'l.: Fmoc-Trp(Boc)-OH Fmoc-Lys(Boc)-OH, Fmoc-Cys(Mmt)-OH, Fmoc-Arg(Pbf)-OH, Fmoc-Ile-OH, Fmoc-Leu-OH, Fmoc-Gln(Trt)-OH, Fmoc-Glu(tBu)-OH. Rink Amide PL resin (0.2 mmol, 644 mg, 0.31 mmol/g, Agilent) was swollen in 1:1 DMF/DCM (15 mL) for 15 min. The Fmoc-protecting group was removed by treatment of the resin with a solution of 20% piperidine and 0.1 M HOBt in DMF (8 mL) for 15 min. The solution was drained and the resin washed with DMF (3 x 8 mL). Amino-acid couplings were achieved by addition of a solution of Fmoc-protected amino-acid (5 equiv), HATU (5 equiv) and DIPEA (10 equiv) in DMF (8 mL), to the resin and stirred intermittently for 1 h. The solution was drained and the resin washed with DMF (5 x 8 mL). The *N*-terminal Fmoc-protecting group was removed by treatment of the resin with a solution of 20% piperidine and 0.1 M HOBt in DMF (8 mL) for 10 min, the solution was drained and the resin washed with DMF (5 x 8 mL). Successive couplings and Fmoc-deprotections were repeated to achieve the desired sequence. After the final Fmoc-deprotection, the *N*-terminus was protected with an acetyl functionality by reaction with acetic anhydride (470 µL) and DIPEA (870 µL) in DMF (10 mL) for 15 min. The resin was washed with DMF (3 x 8 mL), and DCM (5 x 8 mL).

Bimane on-resin attachment:

Following linear peptide, the cysteine side-chain/s were selectively deprotected: Mmt groups were removed by repetitive treatment of the resin with 2% TFA in DCM (8 mL) for 1 min, followed by washing with DCM (3 x 8 mL). Treatments were repeated until the solution no longer turned yellow on addition to the resin (~ 150-200 mL total). The resin was then further washed with DCM (5 x 8 mL) and DMF (5 x 8 mL). With cysteines freshly deprotected, resin (0.1 mmol) was treated with a solution of dibromobimane (2 equiv, 70 mg) or monobromobimane (2 equiv, 54 mg), and DIPEA (4 equiv, 174 µL) in DMF (8 mL), and reacted for 3 h with intermittent stirring. The solution was then removed and the resin washed with DMF (5 x 8 mL) and DCM (5 x 8 mL), then dried with diethyl ether (3 x 8 mL).

General Procedure for Cleavage and Isolation:

The peptide was cleaved from the resin by addition of 92.5:2.5:2.5:2.5 TFA/TIPS/DODT/H_2O (10 mL) and rocked for 1 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 -20°C. The precipitate was pelleted by centrifugation (7600 rpm, 10 min), the supernatant decanted. The pellet was dried under a nitrogen stream, and then dissolved in 1:1 ACN/H_2O, before syringe filtering (0.2 µm) and lyophilised.

Linear peptide 1a: The desired peptide sequence was assembled as described in Solid-*phase Peptide Synthesis of linear precursor*, then peptide was cleaved from the resin by *General Procedure for Cleavage and Isolation*. The peptide was purified by semi-preparative RP-HPLC and Phenomenex Aeris Peptide C18 Column (10 x 250 mm) over a gradient of 25-50% buffer B (25 min). Pure fractions were combined and lyophilised to give the final purified peptide **1a** as a white fluffy powder. HRMS (ESI+) Expected $[M+3H]^{3+}$ for $C_{71}H_{120}N_{22}O_{16}S_2$ (1600.8694): 534.6309, observed: $[M+3H]^{3+}$ 534.6279.

Cyclic peptide 1: The desired peptide sequence was assembled as described in *Solid-phase Peptide Synthesis of linear precursor*, the cysteine Mmt-protected side-chains were then selectively deprotected and cyclised by reaction with dibromobimane as described in *Bimane on-resin attachment.* The peptide was then cleaved from the resin by *General Procedure for Cleavage and Isolation*. The cyclic peptide was purified by semi-preparative RP-HPLC and

Phenomenex Aeris Peptide C18 Column (10 x 250 mm) over a gradient of 30-60% buffer B (15 min). Pure fractions were combined and lyophilised to give the final purified peptide **1** as a pale yellow fluffy powder. HRMS (ESI+) Expected $[M+3H]^{3+}$ for $C_{81}H_{128}N_{24}O_{18}S_2$ (1788.9280): 597.3171, observed: $[M+3H]^{3+}$ 597.3161.

Linear peptide 2a: The desired peptide sequence was assembled as described in *Solid-phase Peptide Synthesis of linear precursor*, then peptide was cleaved from the resin by *General Procedure for Cleavage and Isolation*. The peptide was purified by semi-preparative RP-HPLC and Phenomenex Aeris Peptide C18 Column (10 x 250 mm) over a gradient of 25-50% buffer B (25 min). Pure fractions were combined and lyophilised to give the final purified peptide **2a** as a white fluffy powder. HRMS (ESI+) Expected $[M+3H]^{3+}$ for $C_{71}H_{120}N_{22}O_{16}S_2$ (1600.8694): 534.6309, observed: $[M+3H]^{3+}$ 534.6338.

Cyclic peptide 2: The desired peptide sequence was assembled as described in *Solid-phase Peptide Synthesis of linear precursor*, the cysteine Mmt-protected side-chains were then selectively deprotected and cyclised by reaction with dibromobimane as described in *Bimane on-resin attachment.* The peptide was then cleaved from the resin by *General Procedure for Cleavage and Isolation*. The cyclic peptide was purified by semi-preparative RP-HPLC and Phenomenex Aeris Peptide C18 Column (10 x 250 mm) over a gradient of 30-60% buffer B (15 min). Pure fractions were combined and lyophilised to give the final purified peptide **2** as a pale yellow fluffy powder. HRMS (ESI+) Expected $[M+3H]^{3+}$ for $C_{81}H_{128}N_{24}O_{18}S_2$ (1788.9280): 597.3171, observed: $[M+3H]^{3+}$ 597.3163.

Linear peptide 3a: The desired peptide sequence was assembled as described in *Solid-phase Peptide Synthesis of linear precursor*, then peptide was cleaved from the resin by *General Procedure for Cleavage and Isolation*. The peptide was purified by semi-preparative RP-HPLC and Phenomenex Aeris Peptide C18 Column (10 x 250 mm) over a gradient of 25-50% buffer B (25 min). Pure fractions were combined and lyophilised to give the final purified peptide **3a** as a white fluffy powder. HRMS (ESI+) Expected $[M+4H]^{4+}$ for $C_{74}H_{127}N_{25}O_{16}S$ (1653.9613): 414.4981, observed: $[M+4H]^{4+}$ 414.4973.

Acyclic peptide 3: The desired peptide sequence was assembled as described in *Solid-phase Peptide Synthesis of linear precursor*, the cysteine Mmt-protected side-chain was then selectively deprotected and bimane attached by reaction with monobromobimane as described in *Bimane on-resin attachment.* The peptide was then cleaved from the resin by *General Procedure for Cleavage and Isolation.* The cyclic peptide was purified by semi-preparative RP-HPLC and Phenomenex Aeris Peptide C18 Column (10 x 250 mm) over a gradient of 25-50% buffer B (25 min). Pure fractions were combined and lyophilized to give the final purified peptide **3** as a pale yellow fluffy powder. HRMS (ESI+) Expected $[M+3H]^{3+}$ for $C_{84}H_{137}N_{27}O_{18}S$ (1844.0356): 615.6863, observed: $[M+3H]^{3+}$ 615.6852.

Acyclic peptide 4: The desired peptide sequence was assembled as described in *Solid-phase Peptide Synthesis of linear precursor*, the cysteine Mmt-protected side-chain was then selectively deprotected and bimane attached by reaction with monobromobimane as described in *Bimane on-resin attachment.* The peptide was then cleaved from the resin by *General Procedure for Cleavage and Isolation.* The cyclic peptide was purified by semi-preparative RP-HPLC and Phenomenex Aeris Peptide C18 Column (10 x 250 mm) over a gradient of 25-50% buffer B (25 min). Pure fractions were combined and lyophilised to give the final purified peptide **4** as a pale yellow fluffy powder. HRMS (ESI+) Expected $[M+3H]^{3+}$ for $C_{76}H_{132}N_{26}O_{18}S$ (1728.9934): 577.3389, observed: $[M+3H]^{3+}$ 577.3392.

Cyclic peptide 5: The desired peptide sequence was assembled as described in *Solid-phase Peptide Synthesis of linear precursor*, the cysteine Mmt-protected side-chains were then selectively deprotected and cyclised by reaction with dibromobimane as described in *Bimane on-resin attachment.* The peptide was then cleaved from the resin by *General Procedure for Cleavage and Isolation*. The cyclic peptide was purified by semi-preparative RP-HPLC and Phenomenex Aeris Peptide C18 Column (10 x 250 mm) over a gradient of 25-50% buffer B (15 min). Pure fractions were combined and lyophilised to give the final purified peptide **5** as a pale yellow fluffy powder. HRMS (ESI+) Expected $[M+3H]^{3+}$ for $C_{73}H_{123}N_{23}O_{18}S_2$ (1673.8858) 558.9697, observed: $[M+3H]^{3+}$ 558.9687.

Acyclic peptide 6: The desired peptide sequence was assembled as described in *Solid-phase Peptide Synthesis of linear precursor*, the cysteine Mmt-protected side-chain was then selectively deprotected and bimane attached by reaction with monobromobimane as described in *Bimane on-resin attachment.* The peptide was then cleaved from the resin by *General Procedure for Cleavage and Isolation*. The cyclic peptide was purified by semi-preparative RP-HPLC and Phenomenex Aeris Peptide C18 Column (10 x 250 mm) over a gradient of 25-50% buffer B (25 min). Pure fractions were combined and lyophilised to give the final purified peptide **6** as a pale yellow fluffy powder. HRMS (ESI+) Expected $[M+3H]^{3+}$ for $C_{82}H_{136}N_{26}O_{19}S$ (1821.0196) 608.0143, observed: $[M+3H]^{3+}$ 608.0139.

Acyclic peptide 7: The desired peptide sequence was assembled as described in *Solid-phase Peptide Synthesis of linear precursor*, the cysteine Mmt-protected side-chain was then selectively deprotected and bimane attached by reaction with monobromobimane as described in *Bimane on-resin attachment.* The peptide was then cleaved from the resin by *General Procedure for Cleavage and Isolation*. The cyclic peptide was purified by semi-preparative RP-HPLC and Phenomenex Aeris Peptide C18 Column (10 x 250 mm) over a gradient of 25-50% buffer B (25 min). Pure fractions were combined and lyophilised to give the final purified peptide **7** as a pale yellow fluffy powder. HRMS (ESI+) Expected $[M+3H]^{3+}$ for $C_{82}H_{136}N_{26}O_{18}S$ (1805.0247) 602.6827: observed: $[M+3H]^{3+}$ 602.6828.

3.4.2 NMR analysis

^{1}H and ^{13}C NMR 1D and 2D spectra were recorded on an a Varian Inova 600 MHz instrument in 10% aq. D_2O at 298K, pH ~5, and referenced to DSS at 0 ppm. ES suppression sequences were used for all H^1 1D and 2D homonuclear spectra. ROESYAD and zTOCSY were obtained for all compounds. Spectra were analysed in CCPNMR Analysis.[50] The BimC residue file (cysteine modified with the bimane moiety) for assignment in Analysis was built and supplied Wayne Boucher. Chemical shifts are reported in ppm (δ). Full H^1 assignments of the backbone resonances are included in Supplementary Table S2. Secondary shifts were calculated relative to the random coil shifts reported by Wishart 2011[31] and nearest neighbour corrections applied, these values are summarised in Table S2 and represented graphically in Figure 1.

3.4.3 Bimane attachment of unprotected peptide in solution and Ellman reagent test

Solution-phase bimane attachment

Linear peptide (30 mg [crude], 0.5 mg/mL final) was dissolved in 1:1 ACN/H_2O (6 mL, 10% total reaction volume), and added to 100 mM phosphate pH 7.8 (4 mL), [TFE (6 mL, 10% total reaction volume) for peptides **1** and **2**] and made up to 54 mL with water before dibromobimane (peptides **1** and **2**, 1 equiv) – or monobromobimane in the case of **3** – dissolved in methanol (6 mL) was added and the solution rocked for 30 min. The solution was then further diluted with water and lyophilized to give the crude bimane-peptide as a pale yellow fluffy powder.

Sample preparation for Ellman reagent test

A sample of lyophilised reaction mixture was re-dissolved in water, and peptide concentration determined by absorbance at 205 nm (A_{205}), baseline corrected to 700 nm absorbance. Three consecutive readings were used to calculate an average A_{205} for each peptide. The ε_{205} for each peptide was calculated using an online calculator (http://nickanthis.com/tools/a205.html) as outlined in Anthis 2013.[51] The peptide stock solution concentration was then calculated per $c = A_{205}/\varepsilon_{205}.\ell$ where concentration is in molar, A_{205} is absorbance at 205 nm calculated as an average of three readings, ℓ is the pathlength in cm (1 mm for Nanodrop2000), and ε_{205} is the molar absorptivity at 205 nm.

The peptide samples, **1a** (positive control), **1**, **2** and **3**, were then further diluted to 1 mM (100 µL) with water. The peptides (1 mM, 10 µL) were each diluted into 100 mM Tris.HCl (pH 7.5, 90 µL, 0.1 mM peptide). A negative control with no peptide (vehicle) was prepared in the same manner.

Ellman reagent test

Ellman reagent test was adapted from G-Biosciences.[52] Dithiobisnitrobenzoic acid (DTNB) in DMSO (10 mM, 500 µL) was diluted into 100 mM Tris.HCl (pH 7.5, 49.5 mL) to give a solution of DTNB in 100 mM Tris.HCl (0.1 mM DTNB). Peptide (or vehicle) in Tris.HCl (0.1 mM peptide, 25 µL) was added to DTNB in Tris (475 µL) and the mixtures incubated at rt, for 2 min and gently mixed on the rocker. Samples without DTNB were also prepared in the same manner.

Following the incubation period, the samples were plated in triplicate (100 µL) in a 96-well clear bottomed plate. The absorbance was collected at 412 nm (DTNB) (and 380 nm (bimane)). The absorbance at 412 nm of the peptide without DTNB was subtracted from the peptide with DTNB to correct for the influence of the bimane spectrum overlapping, and the results plotted (Figure S1). The full spectra before and after correction are provided in Figure S2.

3.4.4 Computational modelling of bimane-peptide conformation by energy minimisation

Peptides **1a**, **1**-**4** were constructed in PyMOL and bimane modifications were added to cysteine residues in ICM-Pro Molsoft. Chemical bonds between bimane modification and cysteine residue sidechain were written in ICM-Pro. Following construction, peptides were subsequently energy minimised (n=20) to settle secondary structure in a simulated native conformation. The bimane-peptides were manually docked to the AF-2 surface of the ERα LBD (PDB: 6CBZ) in WinCoot by aligning the conserved Leu residues (LxxLL) to those of the 6CBZ coactivator peptide, which was subsequently deleted and the bimane-peptide:ERα structure PDB coordinates were merged. The bimane-peptide:ERα complex was prepared for energy minimisation by removal of all additional chains, chemicals (excluding ligand) and waters. Protein complexes were transferred to ICM-PRO Molsoft and energy minimised (n=20) to identify the most favourable binding conformation and side-chain interactions.[53] As a control, the linear unmodified peptide was aligned to the 6CBZ crystal structure to validate the modelling method.

3.4.5 Fluorescence characterisation

All fluorescence values were measured (from the bottom) in a black polystyrene, clear bottomed, tissue culture treated polystyrene 96-well plate (Corning Inc. costar® 3605) at an excitation wavelength of 385 nm and emission wavelength of 480 nm unless otherwise specified, on an H4 Synergy Plate Reader, with Xenon light source and a slit width of 9.0.

3.4.6 Circular dichroism

Precise peptide concentrations were determined by NMR spectroscopy on an Agilent 500 MHz spectrometer. A water_ES 1D ^{1}H spectrum, in 10% aq. D_2O, with 256 scans and pulse angle of 90° was acquired for each sample, and integration of the aliphatic signals below 4 ppm used to determine the sample concentration by comparison to a caffeine standard using the 'determine concentration' function in vNMRj. The samples were then diluted into 10 mM phosphate buffer at pH 7.2 (Figure 2) to a final concentration of 50 µM, or 20% TFE in 10 mM phosphate buffer pH 7.2 (Figure S3). Data were collected on a Jasco J-810 spectropolarimeter in a 1 mm cuvette (UniSA Biophysical Characterisation Lab). Each spectrum is the average of 8 scans with a scan rate of 50 nm/min and pitch of 1 nm; the final spectra were smoothed with the in-built Savitzky-Golay function at a convolution width of 7.

3.5 Acknowledgments

The research was supported by the ARC Centre of Excellence in Nanoscale BioPhotonics (CNBP) (CE140100003). The facilities of the OptoFab node of the Australian National Fabrication Facility (ANFF) and associated Commonwealth and SA State Government funding are also gratefully acknowledged. A.J.H. and D.P.M are supported by an Australian Government Research Training Program Stipends (RTPS). Circular Dichroism was carried out at the UniSA Sansom Institute for Health Research Biophysical Characterisation Facility. We thank Philip Clements for his assistance in collecting NMR spectra.

3.6 Conflicts of interest

The authors declare no conflicts of interest

REFERENCES

1. N. S. Kosower, E. M. Kosower, G. L. Newton and H. M. Ranney, Bimane fluorescent labels: Labeling of normal human red cells under physiological conditions, *Proc. Natl. Acad. Sci. U. S. A.*, 1979, **76**, 3382-3386.

2. N. S. Kosower, G. L. Newton, E. M. Kosower and H. M. Ranney, Bimane Fluorescent Labels - Characterization of the Bimane Labeling of Human Hemoglobin, *Biochim. Biophys. Acta*, 1980, **622**, 201-209.

3. E. M. Kosower and N. S. Kosower, Bimane Derivative as Fluorescent Probes for Biological Macromolecules, *Methods Enzymol.*, 1985.

4. A. E. Radkowsky and E. M. Kosower, Bimanes. 17. (Haloalky1)- 1,5-diazabicyclo[3.3.0loctadienediones (Halo-9,lO-dioxabimanes): Reactivity toward the Tripeptide Thiol, Glutathione, *J. Am. Chem. Soc.*, 1986, **108**, 4527-4531.

5. P. Danielsohn and A. Nolte, Bromobimanes- fluorescent labeling agents for histochemical detection of sulfur containing neuropeptides in semithin sections, *Histochemistry*, 1987, **86**, 281-285.

6. E. M. Kosower and N. S. Kosower, Bromobimane Probes for Thiols, *Methods Enzymol.*, 1995, **251**, 133-148.

7. X. Shen, C. B. Pattillo, S. Pardue, S. C. Bir, R. Wang and C. G. Kevil, Measurement of plasma hydrogen sulfide in vivo and in vitro, *Free Radical Biol. Med.*, 2011, **50**, 1021-1031.

8. L. A. Montoya, X. Shen, J. J. McDermott, C. G. Kevil and M. D. Pluth, Mechanistic investigations reveal that dibromobimane extrudes sulfur from biological sulfhydryl sources other than hydrogen sulfide, *Chem. Sci.*, 2015, **6**, 294-300.

9. J. W. Taraska, M. C. Puljung and W. N. Zagotta, Short-distance probes for protein backbone structure based on energy transfer between bimane and transition metal ions, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 16227-16232.

10. E. V. Petrotchenko, K. Xiao, J. Cable, Y. Chen, N. V. Dokholyan and C. H. Borchers, BiPS, a photocleavable, isotopically coded, fluorescent cross-linker for structural proteomics, *Mol. Cell. Proteomics*, 2009, **8**, 273-286.

11. A. Sinz and K. Wang, Mapping Protein Interfaces with a Fluorogenic Cross-Linker and Mass Spectrometry: Application to Nebulin-Calmodulin Complexes, *Biochemistry*, 2001, **40**, 7903-7913.

12. B. Krishnan, A. Szymanska and L. M. Gierasch, Site-specific fluorescent labeling of poly-histidine sequences using a metal-chelating cysteine, *Chem. Biol. Drug Des.*, 2007, **69**, 31-40.

13. M. C. Puljung and W. N. Zagotta, Fluorescent labeling of specific cysteine residues using CyMPL, *Curr. Protoc. Protein Sci.*, 2012, **Chapter 14**, 14.11-14.10.

14. A. J. Horsfall, K. R. Dunning, K. L. Keeling, D. B. Scanlon, K. L. Wegener and A. D. Abell, A bimane ‐ based peptide staple for combined helical induction and fluorescent imaging, *ChemBioChem*, 2020, **21**, 3423-3432.

15. H. H. Szeto, P. W. Schiller, K. Zhao and G. Luo, Fluorescent dyes alter intracellular targeting and function of cell-penetrating tetrapeptides, *The FASEB Journal*, 2005, **19**, 118-120.

16. D. Birch, M. V. Christensen, D. Staerk, H. Franzyk and H. M. Nielsen, Fluorophore labeling of a cell-penetrating peptide induces differential effects on its cellular distribution and affects cell viability, *Biochim. Biophys. Acta. Biomembr.*, 2017, **1859**, 2483-2494.

17. M. P. Luitz, A. Barth, A. H. Crevenna, R. Bomblies, D. C. Lamb and M. Zacharias, Covalent dye attachment influences the dynamics and conformational properties of flexible peptides, *PLoS One*, 2017, **12**, e0177139.

18. S. F. Hedegaard, M. S. Derbas, T. K. Lind, M. R. Kasimova, M. V. Christensen, M. H. Michaelsen, R. A. Campbell, L. Jorgensen, H. Franzyk, M. Cárdenas and H. M. Nielsen, Fluorophore labeling of a cell-penetrating peptide significantly alters the mode and degree of biomembrane interaction, *Sci. Rep.*, 2018, **8**, 6327.

19. E. Sato, M. Sakashita, Y. Kanaoka and A. M. Kosower, XIV. Novel Fluorogenic Substrates for Microdetermination of Chymotrypsin and Aminopeptidase: Bimane Fluorescence Appears after Hydrolysis, *Bioorg. Chem.*, 1988, **16**, 298-306.

20. A. M. Jones Brunette and D. L. Farrens, Distance mapping in proteins using fluorescence spectroscopy: tyrosine, like tryptophan, quenches bimane fluorescence in a distance-dependent manner, *Biochemistry*, 2014, **53**, 6290-6301.

21. S. E. Mansoor, H. S. Mchaourab and D. L. Farrens, Mapping Proximity within Proteins Using Fluorescence Spectroscopy. A Study of T4 Lysozyme Showing That Tryptophan Residues Quench Bimane Fluorescence, *Biochemistry*, 2002, **41**.

22. D. Ho, M. R. Lugo and A. R. Merrill, Harmonic analysis of the fluorescence response of bimane adducts of colicin E1 at helices 6, 7, and 10, *J. Biol. Chem.*, 2013, **288**, 5136-5148.

23. Y. Tian, Y. Jiang, J. Li, D. Wang, H. Zhao and Z. Li, Effect of Stapling Architecture on Physiochemical Properties and Cell Permeability of Stapled alpha-Helical Peptides: a Comparative Study, *ChemBioChem*, 2017, **18**, 2087-2093.

24. Y. Jiang, K. Hu, X. Shi, Q. Tang, Z. Wang, X. Ye and Z. Li, Switching substitution groups on the in-tether chiral centre influences backbone peptides' permeability and target binding affinity, *Org. Biomol. Chem.*, 2017, **15**, 541-544.

25. X. Qin, H. Zhao, Y. Jiang, F. Yin, Y. Tian, M. Xie, X. Ye, N. Xu and Z. Li, Development of a potent peptide inhibitor of estrogen receptor alpha, *Chin. Chem. Lett.*, 2018, **29**, 1160-1162.

26. A.-M. Leduc, J. O. Trent, J. L. Wittliff, K. S. Bramlett, S. L. Briggs, N. Y. Chirgadze, Y. Wang, T. P. Burris and A. F. Spatola, Helix-stabilized cyclic peptides as selective inhibitors of steroid receptor– coactivator interactions, *Biochemistry*, 2003, **100**, 11273-11278.

27. T. E. Speltz, S. W. Fanning, C. G. Mayne, C. Fowler, E. Tajkhorshid, G. L. Greene and T. W. Moore, Stapled Peptides with gamma-Methylated Hydrocarbon Chains for the Estrogen Receptor/Coactivator Interaction, *Angew. Chem. Int. Ed. (English)*, 2016, **55**, 4252-4255.

28. K. W. Nettles, J. B. Bruning, G. Gil, J. Nowak, S. K. Sharma, J. B. Hahm, K. Kulp, R. B. Hochberg, H. Zhou, J. A. Katzenellenbogen, B. S. Katzenellenbogen, Y. Kim, A. Joachimiak and G. L. Greene, NFκB selectivity of estrogen receptor ligands revealed by comparative crystallographic analyses, *Nature Chemical Biology*, 2008, **4**, 241-247.

29. C.-Y. Chang, J. D. Norris, H. Grøn, L. A. Paige, P. T. Hamilton, D. J. Kenan, D. Fowlkes and D. P. Mcdonnell, Dissection of the LXXLL Nuclear Receptor-Coactivator Interaction Motif Using Combinatorial Peptide Libraries: Discovery of Peptide Antagonists of Estrogen Receptors Alpha and Beta, *Mol. Cell. Biol.*, 1999, **19**, 8226-8239.

30. P. Ravindranathan, T. K. Lee, L. Yang, M. M. Centenera, L. Butler, W. D. Tilley, J. T. Hsieh, J. M. Ahn and G. V. Raj, Peptidomimetic targeting of critical androgen receptor-coregulator interactions in prostate cancer, *Nat. Comm.*, 2013, **4**, 1923.

31. D. S. Wishart, Interpreting protein chemical shift data, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2011, **58**, 62-87.

32. D. S. Wishart, B. D. Sykes and F. M. Richards, The Chemical Shift Index: A Fast and Simple Method for the Assignment of Protein Secondary Structure through NMR Spectroscopy, *Biochemistry*, 1992, **31**, 1647-1651.

33. R. W. Woody, The development and current state of protein circular dichroism, *Biomed. Spectrosc. Imaging*, 2015, **4**, 5-34.

34. N. E. Shepherd, H. N. Hoang, G. Abbenante and D. P. Fairlie, Single Turn Peptide Alpha Helices with Exceptional Stability in Water, *J. Am. Chem. Soc.*, 2005, **127**, 2974-2983.

35. J. Yu, J. R. Horsley and A. D. Abell, Turning electron transfer 'on-off' in peptides through side-bridge gating, *Electrochim. Acta*, 2016, **209**, 65-74.

36. C. Toniolo, A. Polese, F. Formaggio, M. Crisma and J. Kamphuis, Circular Dichroism Spectrum of a Peptide 310-Helix, *J. Am. Chem. Soc.*, 1996, **118**, 2744-2745.

37. F. Formaggio, M. Crisma, P. Rossi, P. Scrimin, B. Kaptein, Q. B. Broxterman, J. Kamphuis and C. Toniolo, The First Water-Soluble 310-Helical Peptides, *Chem. Eur. J.*, 2000, **6**, 4498-4504.

38. P. C. Lyu, M. I. Liff, L. A. Marky and N. R. Kallenbach, Side Chain Contribution to the Stability of Alpha-Helical Structure in Peptides, *Science*, 1990, **250**, 669-673.

39. P. C. Lyu, J. C. Sherman, A. Chen and N. R. Kallenbach, a-Helix stabilzation by natural and unnatural amine acids with alkyl side chains, *Proc. Natl. Acad. Sci. U. S. A.*, 1991, **88**, 5317-5320.

40. Z. Shit, C. A. Olson, A. J. B. Jr. and N. R. Kallenbach, Stabilization of alpha-Helix Structure by Polar Side- Chain Interactions: Complex Salt Bridges, Cation-pi Interactions, and C–H . . . O H-Bonds, *Biopolymers* 2001, **60**, 366-380.

41. A. E. Garcia and K. Y. Sanbonmatsu, Alpha-helical stabilization by side chain shielding of backbone hydrogen bonds, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 2782-2787.

42. S. E. Miller, A. M. Watkins, N. R. Kallenbach and P. S. Arora, Effects of side chains in helix nucleation differ from helix propagation, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 6636-6641.

43. C. D. Andrew, S. Bhattacharjee, N. Kokkoni, J. D. Hirst, G. R. Jones and A. J. Doig, Stabilizing Interactions between Aromatic and Basic Side Chains in alpha-helical Peptides and Proteins. Tyrosine Effects on Helix Circular Dichroism, *J. Am. Chem. Soc.*, 2002, **124**, 12706-12714.

44. A. M. Brzozowski, A. C. W. Pike, Z. Dauter, R. E. Hubbard, T. Bonn, O. Engström, L. Öhman, G. L. Greene, J.-Å. Gustafsson and M. Carlquist, Molecular basis of agonism and antagonism in the oestrogen receptor, *Nature*, 1997, **389**, 753-758.

45. R. Abagyan, M. Totrov and D. Kuznetsov, ICM - New Method for Protein Modeling and Design: Applications to Docking and Structure Prediction from the Distorted Native Conformation, *J. Comput. Chem.*, 1994, **15**, 488-506.

46. P. Y. Maximov, B. Abderrahman, S. W. Fanning, S. Sengupta, P. Fan, R. F. Curpan, D. M. Q. Rincon, J. A. Greenland, S. S. Rajan, G. L. Greene and V. C. Jordan, Endoxifen, 4-Hydroxytamoxifen and an Estrogenic Derivative Modulate Estrogen Receptor Complex Mediated Apoptosis in Breast Cancer, *Mol. Pharmacol.*, 2018, **94**, 812-822.

47. D. J. Barlow and J. M. Thronton, ￼￼Helix Geometry in Proteins, *J. Mol. Biol.*, 1988, **201**, 601-619.

48. E. M. Kosower, R. Giniger, A. Radkowsky, D. Hebel and A. Shusterman, Bimanes 22. Flexible fluorescent molecules. Solvent effects on the photophysical properties of syn-bimanes (1,5-diazabicyclo[3.3.0]octa-3,6-diene-2,8-diones), *J. Phys. Chem.*, 1986, **90**, 5552-5557.

49. S. E. Mansoor, H. S. Mchaourab and D. L. Farrens, Determination of Protein Secondary Structure and Solvent Accessibility Using Site-Directed Fluorescence Labeling. Studies of T4 Lysozyme Using the Fluorescent Probe Monobromobimane, *Biochemistry*, 1999, **38**, 16383-16393.

50. W. F. Vranken, W. Boucher, T. J. Stevens, R. H. Fogh, A. Pajon, M. Llinas, E. L. Ulrich, J. L. Markley, J. Ionides and E. D. Laue, The CCPN data model for NMR spectroscopy: development of a software pipeline, *Proteins: Struct., Funct., Genet.*, 2005, **59**, 687-696.

51. N. J. Anthis and G. M. Clore, Sequence-specific determination of protein and peptide concentrations by absorbance at 205 nm, *Protein Sci.*, 2013, **22**, 851-858.

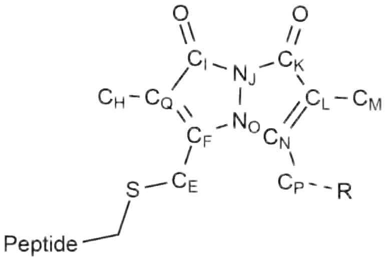
52. GBiosciences, Ellman's Reagent, https://www.gbiosciences.com/Protein-Research/Ellman-s-Reagent, (accessed 3/8/2020).

53. R. Abagyan and M. Totrov, Biased Probability Monte Carlo Conformational Searches and Electrostatic Calculations for Peptides and Proteins, *Journal of Molecular Biology*, 1994, **235**, 983-1002.

Supplemental Data.

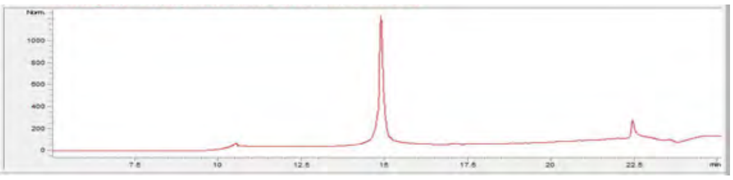
APPROACHES TO INTRODUCES HELICAL STRUCTURE IN CYSTEINE-CONTAINING PEPTIDES WITH A BIMANE GROUP

S3.1 Additional characterisation data

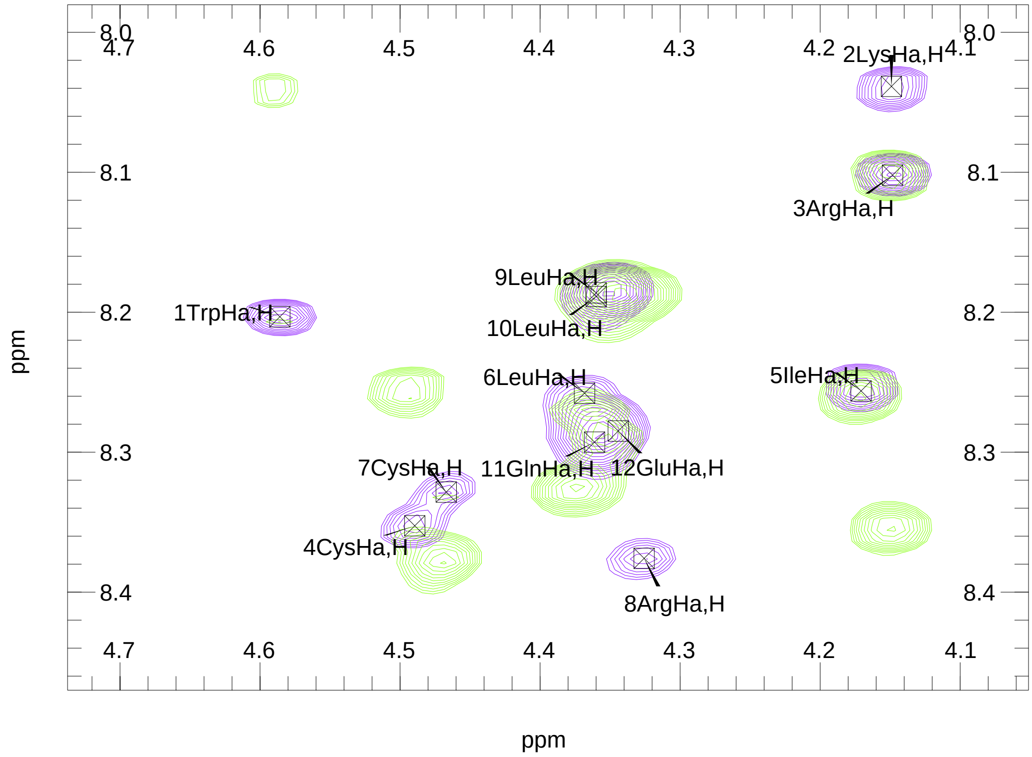


Atom codes for the bimane appended cysteine residue, termed BimC, used in CCPNMR Analysis for assignment. Cp is a methylene and R=cysteine thioether when di-bimane introduced, or Cp is methyl when mono-bimane introduced. Residue built by Wayne Boucher.

*Peptide **1a***

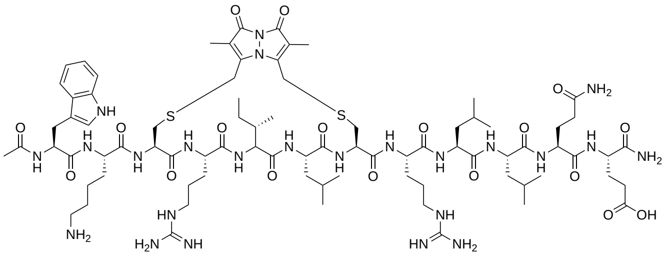


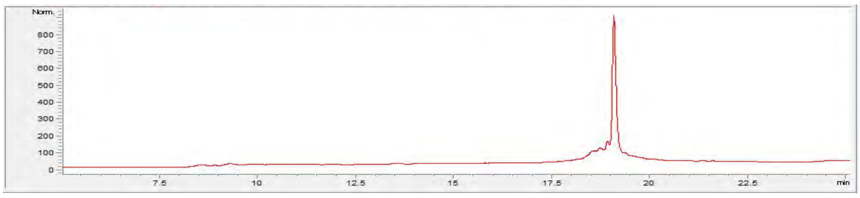
Analytical HPLC spectrum of **1a**, over a gradient of 0-100% buffer B on a Phenomenex Luna C18(2) 250 x 4.6 mm 5 µm, visualised at 220 nm



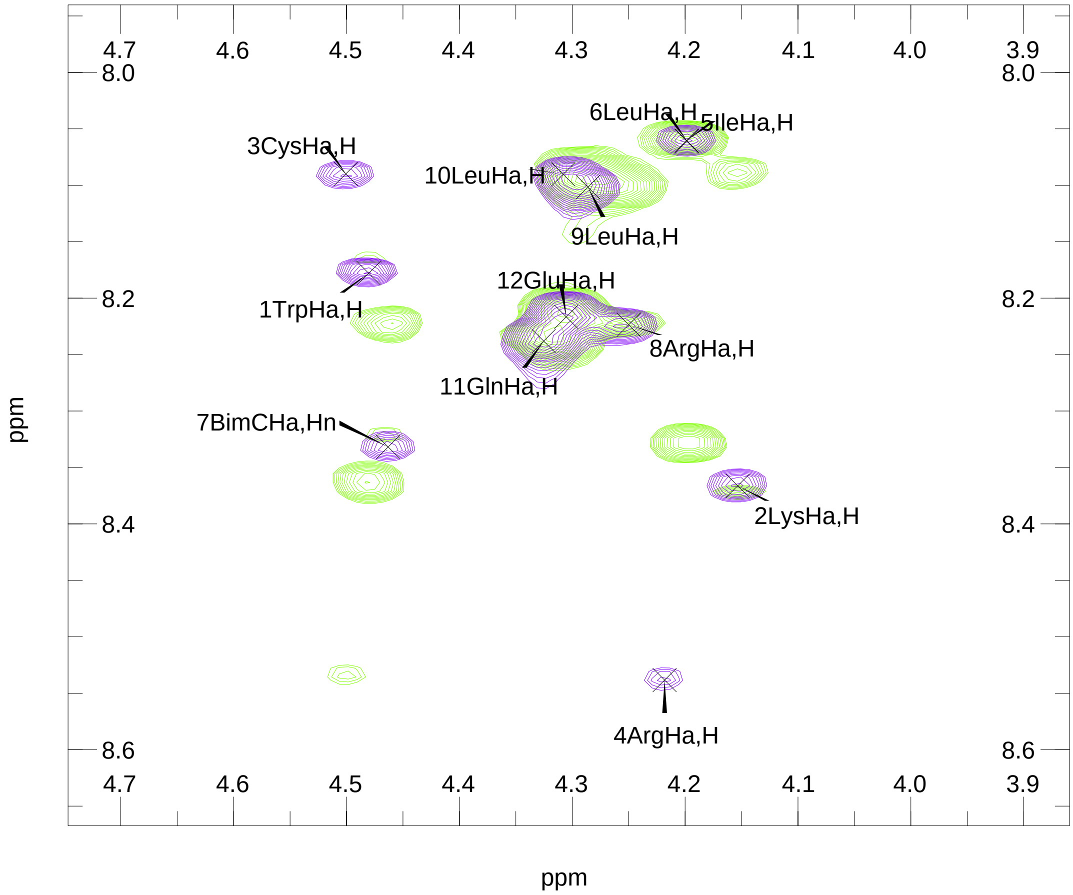
2D H^1 homonuclear TOCSY (purple) and ROESY (green) spectrum of **1a** in the main-chain region

*Peptide **1***

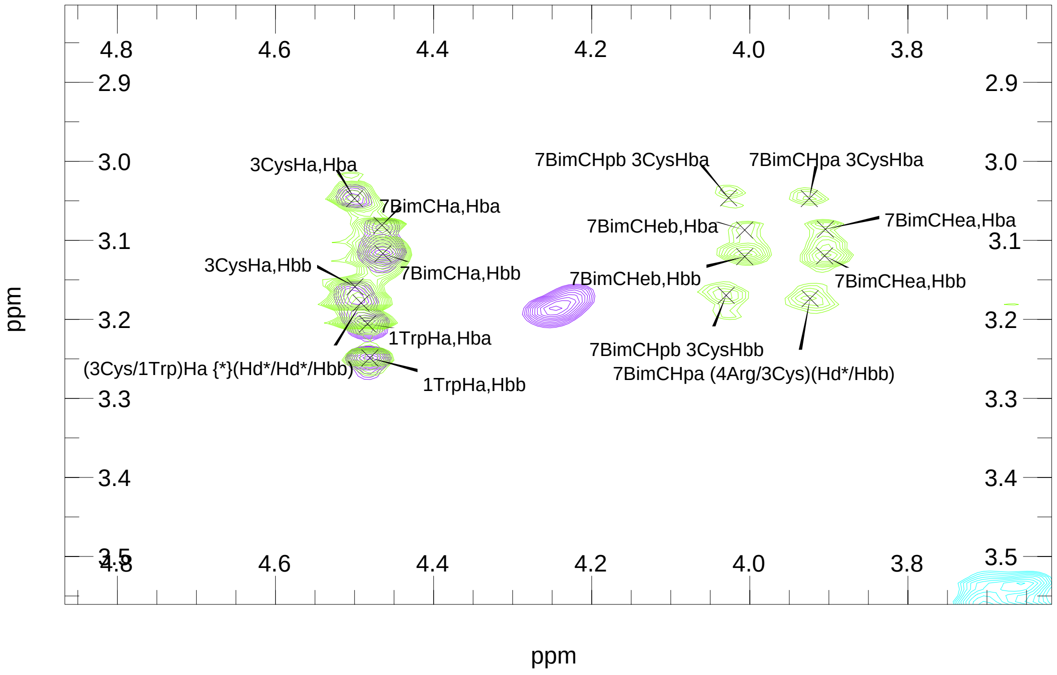




Analytical HPLC spectrum of **1**, over a gradient of 0-100% buffer B on a Phenomenex Luna C18(2) 250 x 4.6 mm 5 µm, visualised at 220 nm

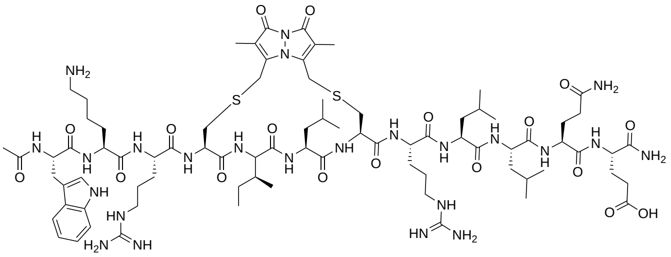


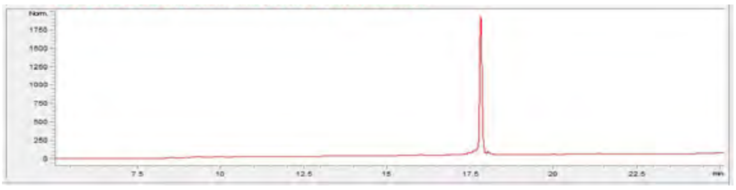
2D H^1 homonuclear TOCSY (purple) and ROESY (green) spectrum of **1** in the main-chain region



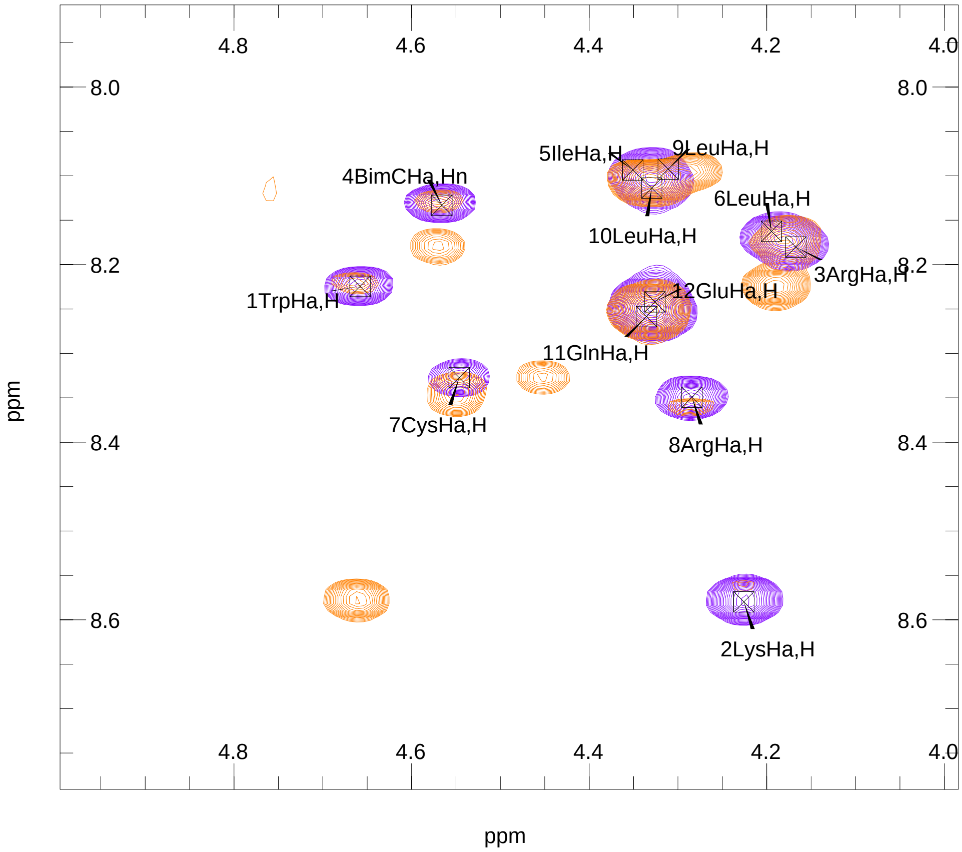
2D H^1 homonuclear TOCSY (purple) and ROESY (green) spectrum of **1** showing the side-chain crosspeaks indicative of cysteine alkylation with the bimane. All bimane atoms are part of the BimC block, where Hea and Heb are the bimane methylene protons adjacent to the 7BimC thiol and Hpa and Hpb are the bimane methylene protons adjacent to the 3Cys thiol.

Peptide **2**



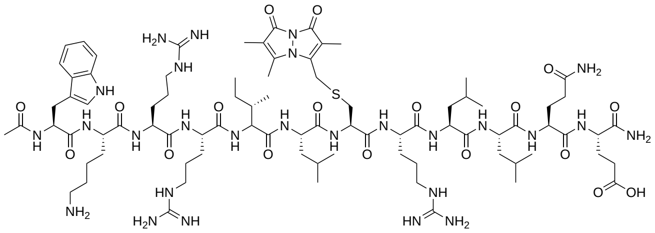


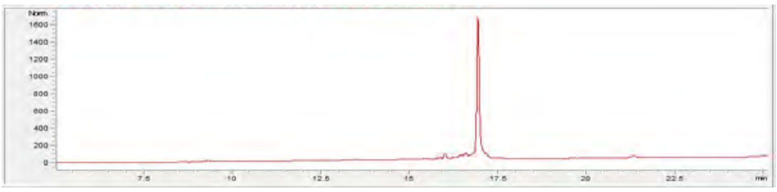
Analytical HPLC spectrum of **2**, over a gradient of 0-100% buffer B on a Phenomenex Luna C18(2) 250 x 4.6 mm 5 µm, visualised at 220 nm



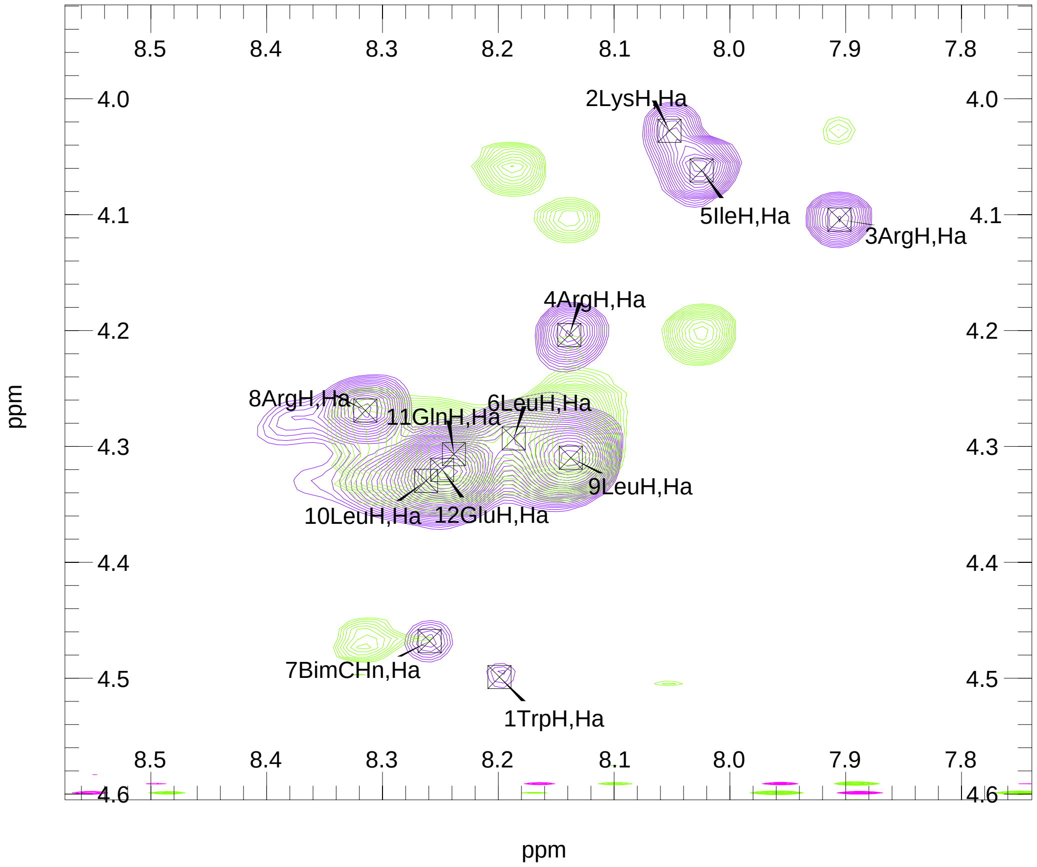
2D H^1 homonuclear TOCSY (purple) and ROESY (orange) spectrum of **2** in the main-chain region

*Peptide **3***



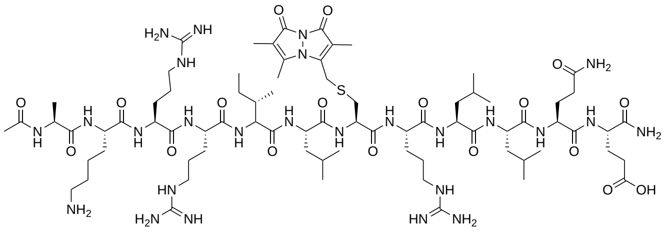


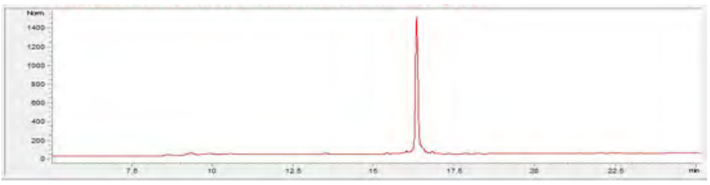
Analytical HPLC spectrum of **3**, over a gradient of 0-100% buffer B on a Phenomenex Luna C18(2) 250 x 4.6 mm 5 µm, visualised at 220 nm



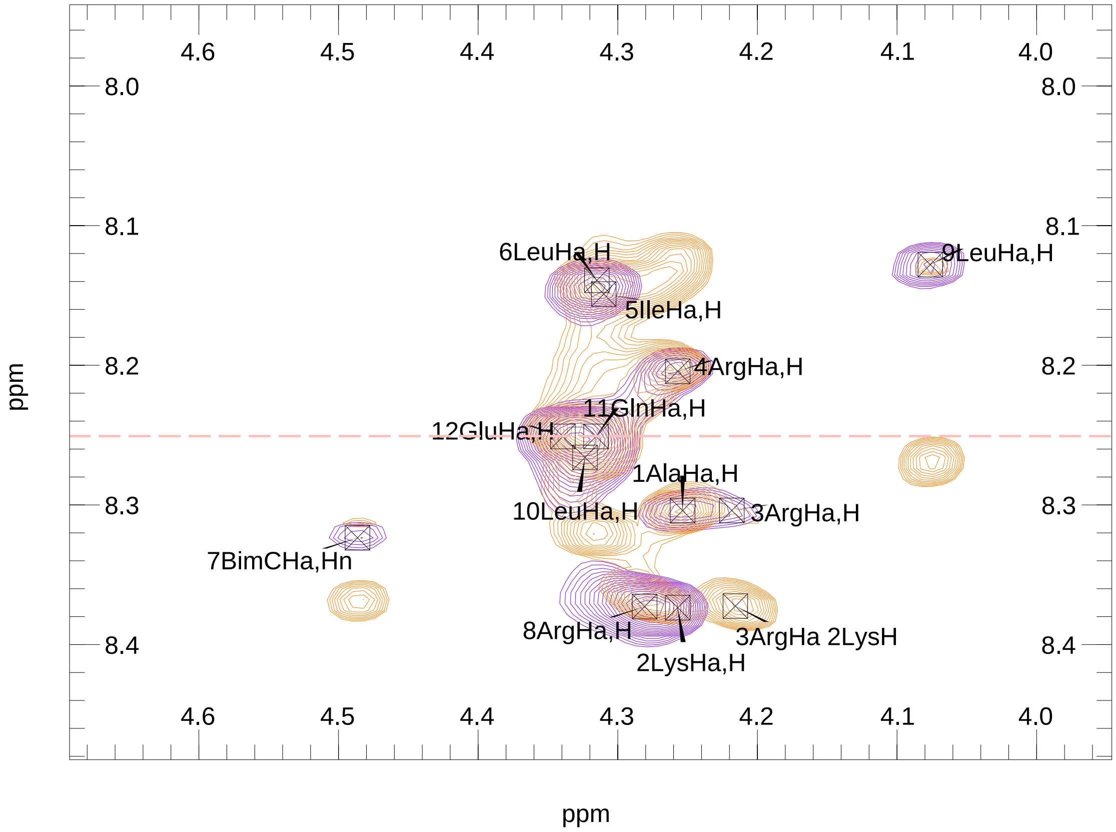
2D H^1 homonuclear TOCSY (purple) and ROESY (green) spectrum of **3** in the main-chain region

*Peptide **4***



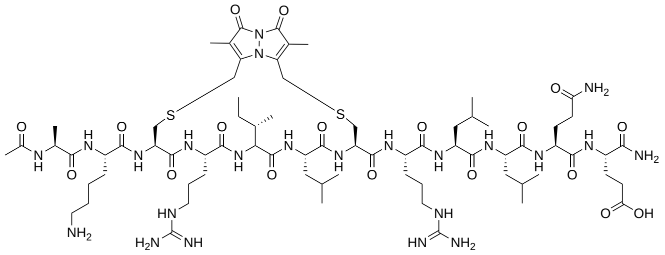


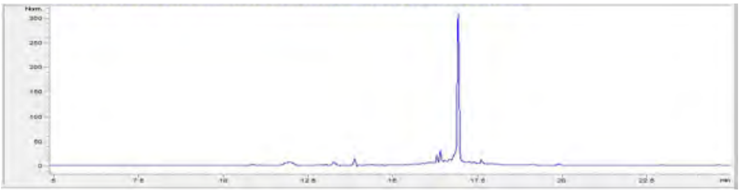
Analytical HPLC spectrum of **4**, over a gradient of 0-50% buffer B on a Phenomenex Luna C18(2) 250 x 4.6 mm 5 μm, visualised at 220 nm



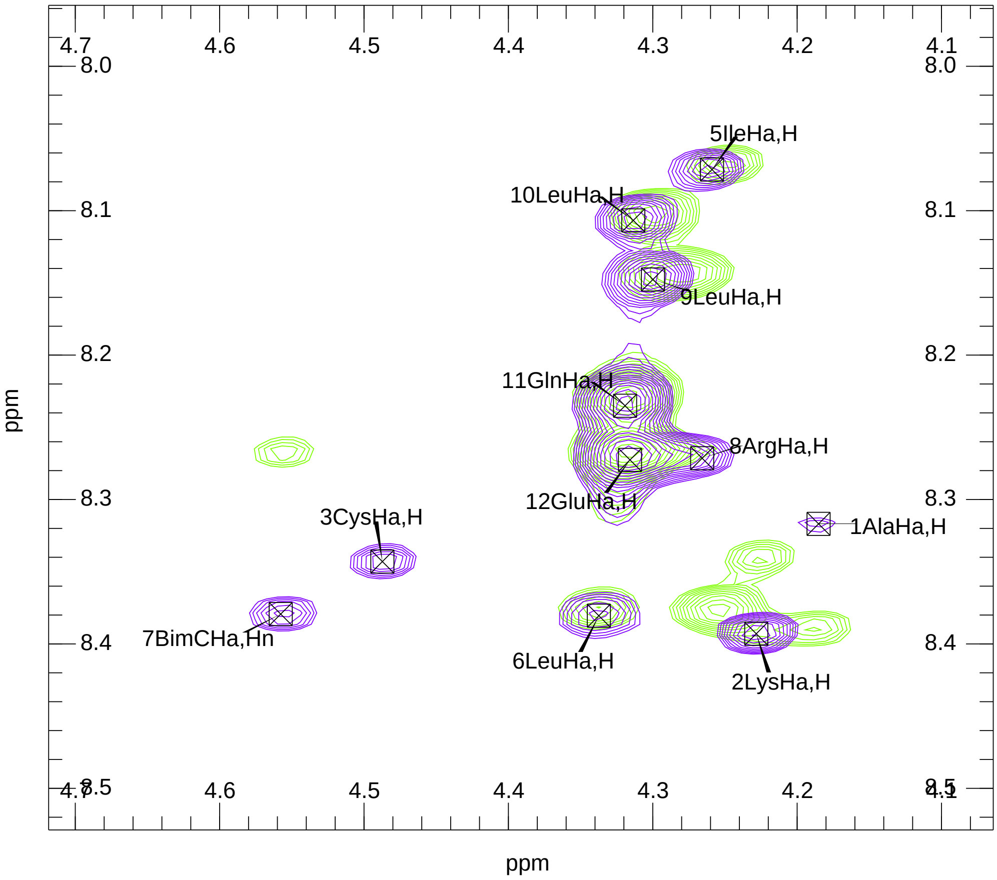
2D H^1 homonuclear TOCSY (purple) and ROESY (orange) spectrum of **4** in the main-chain region

*Peptide **5***



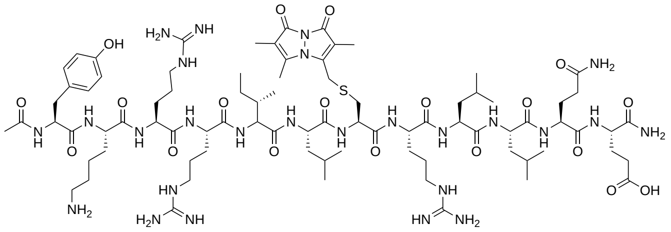


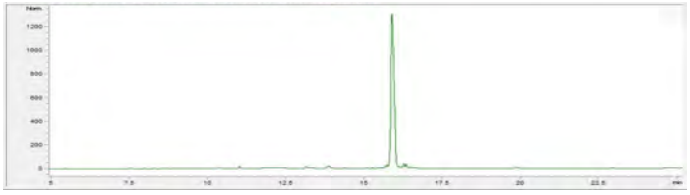
Analytical HPLC spectrum of **5**, over a gradient of 0-50% buffer B on a Phenomenex Luna C18(2) 250 x 4.6 mm 5 µm, visualised at 220 nm



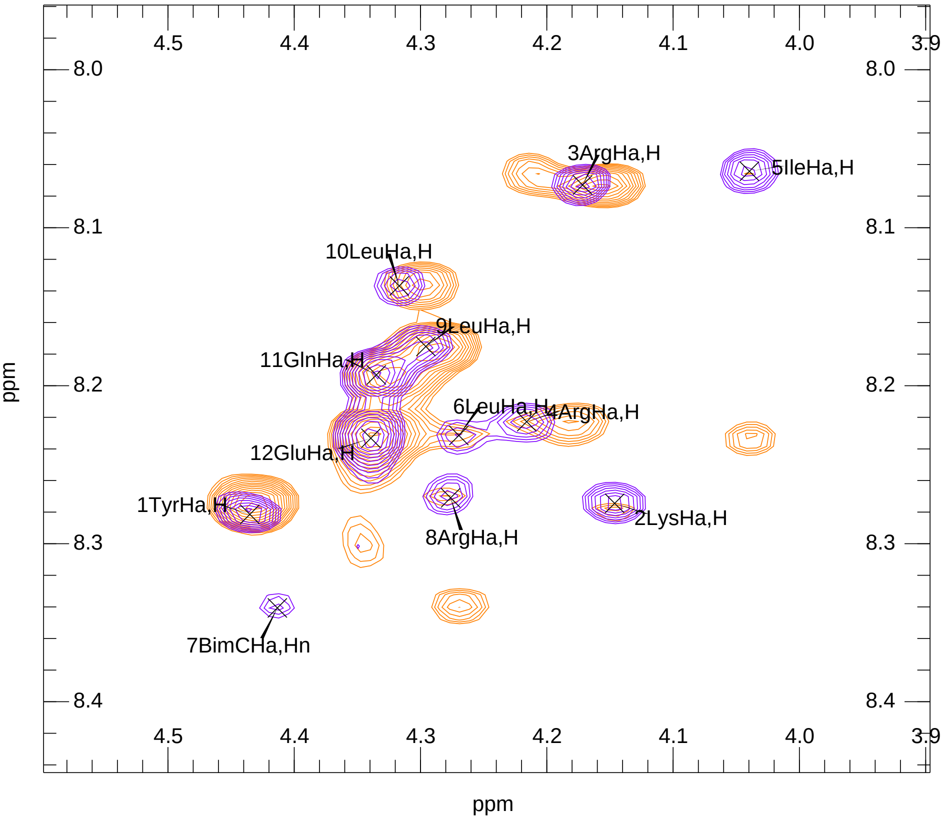
2D H¹ homonuclear TOCSY (purple) and ROESY (green) spectrum of **5** in the main-chain region

*Peptide **6***


OH
H_2N
NH
NH
O
N
S
NH_2
NH_2
HN
H_2N
NH
HN
NH_2
OH

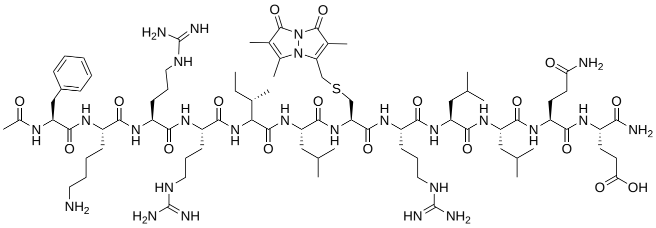


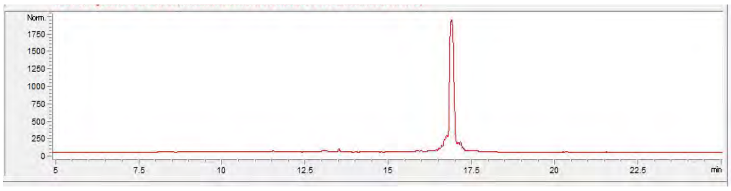
Analytical HPLC spectrum of **6**, over a gradient of 0-50% buffer B on a Phenomenex Luna C18(2) 250 x 4.6 mm 5 µm, visualised at 220 nm



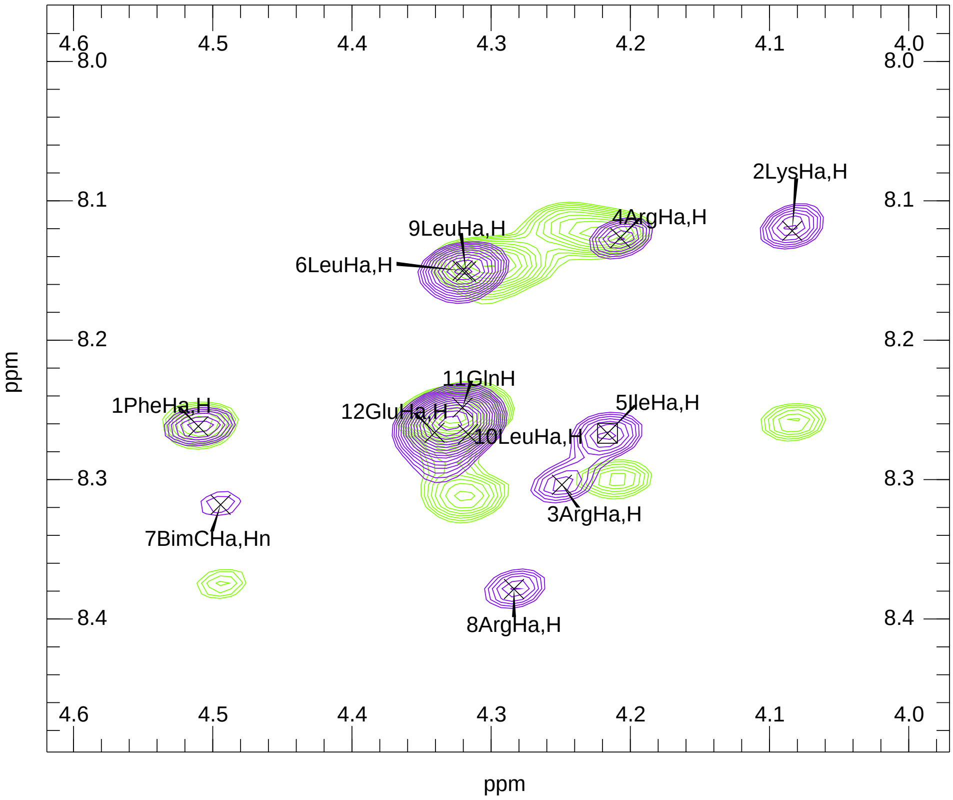
2D H^1 homonuclear TOCSY (purple) and ROESY (orange) spectrum of **6** in the main-chain region

Peptide **7**





Analytical HPLC spectrum of **7**, over a gradient of 0-50% buffer B on a Phenomenex Luna C18(2) 250 x 4.6 mm 5 µm, visualised at 220 nm



2D H^1 homonuclear TOCSY (purple) and ROESY (green) spectrum of **7** in the main-chain region

S3.2 Ellman reagent test absorbance spectra

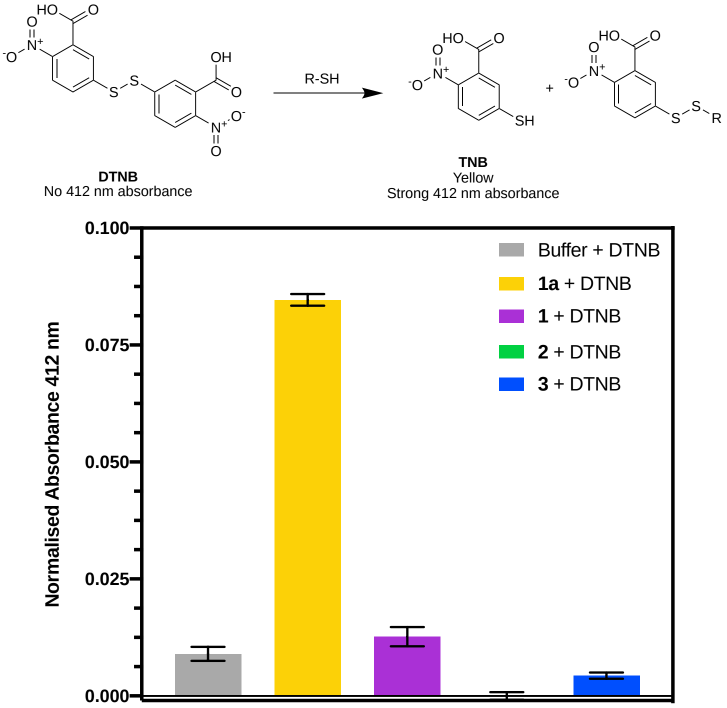


Figure S1: Ellman test. Peptides (5 µM) in 100 mM Tris.HCl buffer pH 7.5 are combined with 0.1 mM *bis*-dinitrothiobenzene (DTNB), if free thiols are present dinitrothiobenzene is liberated giving rise to a characteristic 412 nm absorbance. Bimane reacted peptides **1**, **2** and **3** show no indication of free thiol (within error of negative control – buffer + DTNB), in comparison to the free thiol containing positive control peptide **1a**. DTNB - 5,5'-dithio-*bis*-(2-nitrobenzoic acid), TNB - 5-thio(2-nitrobenzoic acid)

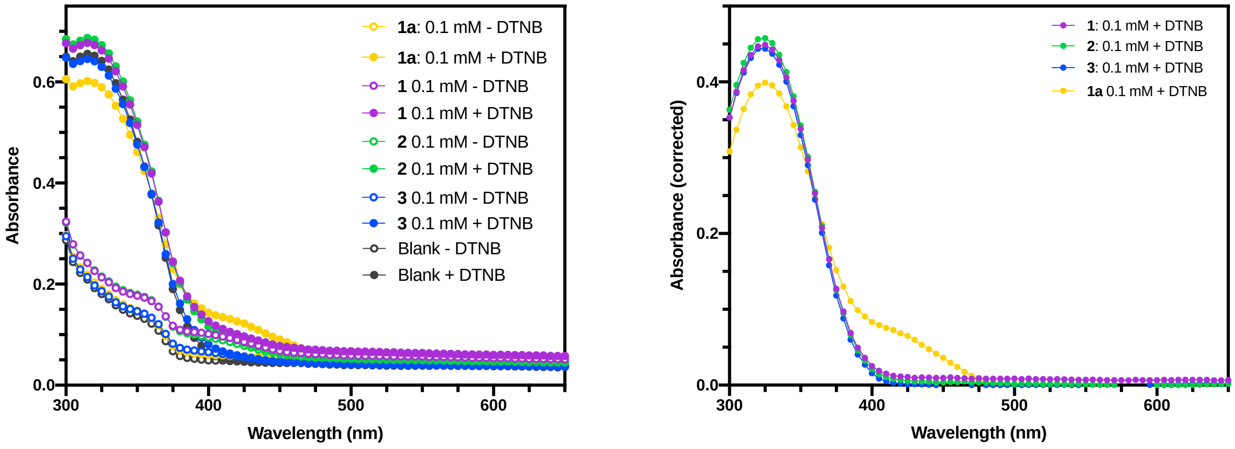


Figure S2: Absorbance spectra for Ellman Reagent Test. Peptides (**1a**, **1-4**) were analysed after a 2 min incubation with (+) or without (-) *bis*-dinitrothiobenzene – DTNB. All data are shown on the left. Due to the absorbance of the bimane group at 380 nm, the -DTNB spectrum was subtracted from the respective +DTNB spectrum resulting in the spectra displayed on the left.

S3.3 CD spectra

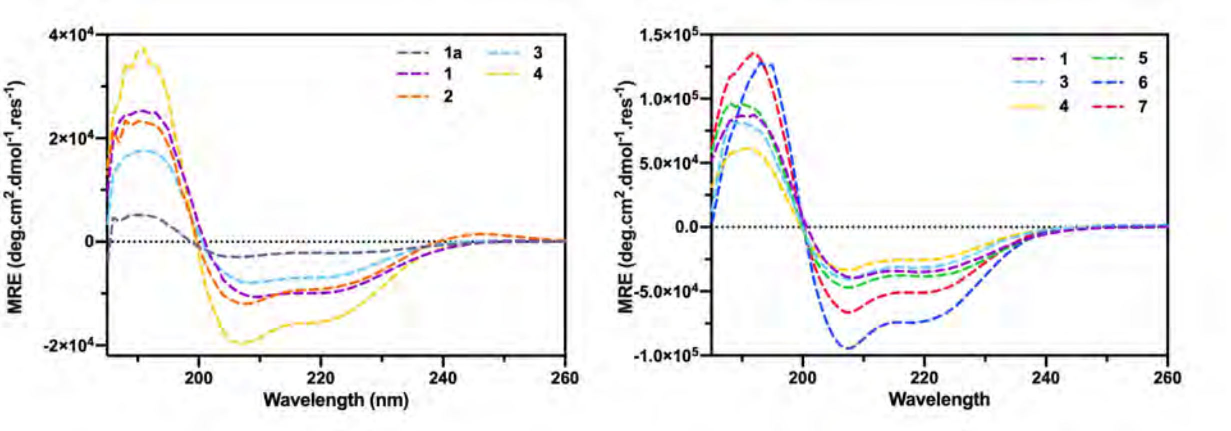


Figure S3: CD spectra collected in 10 mM phosphate buffer pH 7.2 with 20% TFE and smoothed with a Savitsky-Golay function **1a**, yellow; **1**, purple; **2**, green; **3**, blue. Left and right panels represent data from two separate experiments.

%Helicity calculations

Table S1: Summary of CD data. % Helicity calculated based on the 222 nm minimum.[1] MRE – molar residue ellipticity. MRE at 222 nm expected for an ideal helix is -25166 mdeg dmol^{-1} res^{-1}

	MRE$_{observed}$			
Peptide	**208 nm**	**222 nm**	**%helicity**	**208/222 ratio**
1a	-19630	-10050	42	0.51
1	-28163	-18617	66	0.66
2	-24472	-8990	38	0.37
3	-23827	-16421	75	0.69
4	-8638	-4086	19	0.47
Round 2				
1	-24650	-18159	73	0.74
5	-22129	-11427	48	0.53
6	28467	-16069	61	0.53
7	-19766	-9176	41	0.51

From Shepherd/Fairlie 2005[1]

$$\mathbf{MRE@222_{expected} = (-44000 + 250T)(1-k/N_p)}$$

$$= (-44000 + 250(25))(1-4/12)$$

$$= -25166.67 \text{ mdeg dmol}^{-1} \text{ res}^{-1}$$

Where: T = temperature in °C (here, T=25°C),

k = 4 for small peptides

N_p = the number of amino-acid residues in the peptide (here, N_p = 12)

$$\mathbf{\%Helicity = (MRE@222_{observed} - MRE_{Coil})/(MRE@222_{expected} - MRE_{Coil})}$$

Where MRE_{coil} = 2220-53(T), and T = temperature in °C (here: T=25°C)

S3.4 NMR chemical shift analysis tables

Table S2: Summary of the assigned NMR chemical shifts of the alpha-proton (Hα) and amide proton (NH). The random coil secondary shifts (RC) for each residue are reported with nearest neighbour corrections applied (NN).[2] The resulting secondary shift ($\Delta\delta$) are reported as the difference between the experimental chemical shift and random coil literature shift (Hα-RC(NN) = $\Delta\delta$, or NH-RC(NN) = $\Delta\delta$). No nearest neighbour corrections were necessary for the Hα in this particular sequence.

Peptide	Residue	Hα	RC*	RC(NN)	$\Delta\delta_{H\alpha}$		Residue	NH	RC*	RC(NN)	$\Delta\delta_{NH}$
1a	Trp	4.59	4.66	4.66	-0.07		Trp	8.2	8.25	8.25	-0.05
	Lys	4.15	4.32	4.17	-0.17		Lys	8.04	8.29	8.29	-0.25
	Arg	4.15	4.34	4.34	-0.19		Arg	8.1	8.23	8.37	-0.13
	Cys	4.49	4.55	4.55	-0.06		Cys	8.35	8.32	8.47	0.03
	Ile	4.17	4.17	4.17	0.00		Ile	8.25	8.00	8.20	0.25
	Leu	4.37	4.34	4.34	0.03		Leu	8.26	8.16	8.33	0.10
	Cys	4.47	4.55	4.55	-0.08		Cys	8.33	8.32	8.46	0.01
	Arg	4.32	4.34	4.34	-0.02		Arg	8.38	8.23	8.43	0.15
	Leu	4.36	4.34	4.34	0.02		Leu	8.19	8.16	8.31	0.03
	Leu	4.36	4.34	4.34	0.02		Leu	8.19	8.16	8.30	0.03
	Gln	4.36	4.34	4.34	0.02		Gln	8.29	8.32	8.46	-0.03
	Glu	4.34	4.35	4.35	-0.01		Glu	8.28	8.42	8.57	-0.14
1	Trp	4.48	4.66	4.66	-0.18		Trp	8.17	8.25	8.25	-0.08
	Lys	4.16	4.32	4.17	-0.01		Lys	8.36	8.29	8.29	0.07
	Cys*	4.50	4.55	4.55	-0.05		Cys*	8.09	8.32	8.46	-0.37
	Arg	4.22	4.34	4.34	-0.12		Arg	8.54	8.23	8.43	0.11
	Ile	4.20	4.17	4.17	0.03		Ile	8.06	8.00	8.15	-0.09
	Leu	4.20	4.34	4.34	-0.14		Leu	8.06	8.16	8.33	-0.27
	Cys*	4.46	4.55	4.55	-0.09		Cys*	8.33	8.32	8.46	-0.13
	Arg	4.25	4.34	4.34	-0.09		Arg	8.22	8.23	8.43	-0.21
	Leu	4.29	4.34	4.34	-0.05		Leu	8.10	8.16	8.31	-0.21
	Leu	4.31	4.34	4.34	-0.04		Leu	8.09	8.16	8.30	-0.21
	Gln	4.33	4.34	4.34	-0.01		Gln	8.24	8.32	8.46	-0.22
	Glu	4.31	4.35	4.35	-0.04		Glu	8.21	8.42	8.57	-0.36
2	Trp	4.66	4.66	4.66	0.00		Trp	8.23	8.25	8.25	-0.03
	Lys	4.23	4.32	4.17	-0.09		Lys	8.58	8.29	8.29	0.29
	Arg	4.17	4.34	4.34	-0.17		Arg	8.18	8.23	8.37	-0.05
	Cys*	4.57	4.55	4.55	0.02		Cys*	8.13	8.32	8.47	-0.19
	Ile	4.35	4.17	4.17	0.18		Ile	8.09	8.00	8.20	0.09
	Leu	4.20	4.34	4.34	-0.15		Leu	8.16	8.16	8.33	0.00
	Cys*	4.55	4.55	4.55	0.00		Cys*	8.33	8.32	8.46	0.01
	Arg	4.28	4.34	4.34	-0.06		Arg	8.35	8.23	8.43	0.12
	Leu	4.31	4.34	4.34	-0.03		Leu	8.09	8.16	8.31	-0.07
	Leu	4.33	4.34	4.34	-0.01		Leu	8.11	8.16	8.30	-0.05
	Gln	4.34	4.34	4.34	0.00		Gln	8.26	8.32	8.46	-0.06
	Glu	4.33	4.35	4.35	-0.02		Glu	8.24	8.42	8.57	-0.18
3	Trp	4.50	4.66	4.66	-0.16		Trp	8.20	8.25	8.25	-0.05
	Lys	4.03	4.32	4.17	-0.29		Lys	8.05	8.29	8.29	-0.24
	Arg	4.11	4.34	4.34	-0.23		Arg	7.91	8.23	8.37	-0.32
	Arg	4.21	4.34	4.34	-0.14		Arg	8.14	8.23	8.38	-0.32
	Ile	4.06	4.17	4.17	-0.11		Ile	8.03	8.00	8.15	0.14
	Leu	4.29	4.34	4.34	-0.05		Leu	8.19	8.16	8.33	-0.13
	Cys*	4.47	4.55	4.55	-0.08		Cys*	8.26	8.32	8.46	-0.13
	Arg	4.27	4.34	4.34	-0.07		Arg	8.32	8.23	8.43	0.03
	Leu	4.31	4.34	4.34	-0.04		Leu	8.14	8.16	8.31	0.15
	Leu	4.32	4.34	4.34	-0.02		Leu	8.13	8.16	8.3	-0.02
	Gln	4.33	4.34	4.34	-0.01		Gln	8.24	8.32	8.46	-0.19
	Glu	4.32	4.35	4.35	-0.03		Glu	8.25	8.42	8.57	-0.18
4	Ala	4.26	4.32	4.32	-0.06		Ala	8.3	8.24	8.24	0.06
	Lys	4.26	4.32	4.32	-0.06		Lys	8.37	8.29	8.29	0.08
	Arg	4.22	4.34	4.34	-0.12		Arg	8.3	8.23	8.37	0.07
	Arg	4.26	4.55	4.55	-0.29		Arg	8.2	8.32	8.47	-0.12
	Ile	4.31	4.17	4.17	0.14		Ile	8.14	8.00	8.20	0.14
	Leu	4.32	4.34	4.34	-0.02		Leu	8.14	8.16	8.33	-0.02
	Cys*	4.49	4.55	4.55	-0.06		Cys*	8.32	8.32	8.46	0.00
	Arg	4.28	4.34	4.34	-0.06		Arg	8.37	8.23	8.43	0.14
	Leu	4.08	4.34	4.34	-0.26		Leu	8.13	8.16	8.31	-0.03
	Leu	4.33	4.34	4.34	-0.01		Leu	8.27	8.16	8.30	0.11
	Gln	4.32	4.34	4.34	-0.02		Gln	8.25	8.32	8.46	-0.07
	Glu	4.34	4.35	4.35	-0.01		Glu	8.25	8.42	8.57	-0.17

Table S2: cont….

5	Ala	4.19	4.32	4.32	-0.13	Ala	8.32	8.24	8.24	0.08
	Lys	4.23	4.32	4.32	-0.09	Lys	8.39	8.29	8.29	0.10
	Cys*	4.49	4.55	4.55	-0.06	Cys*	8.34	8.32	8.46	-0.12
	Arg	4.24	4.34	4.34	-0.10	Arg	8.61	8.23	8.43	0.18
	Ile	4.26	4.17	4.17	0.09	Ile	8.07	8.00	8.15	-0.08
	Leu	4.32	4.34	4.34	-0.02	Leu	8.38	8.16	8.33	0.05
	Cys*	4.56	4.55	4.55	0.01	Cys*	8.38	8.32	8.46	-0.08
	Arg	4.27	4.34	4.34	-0.07	Arg	8.27	8.23	8.43	-0.16
	Leu	4.3	4.34	4.34	-0.04	Leu	8.15	8.16	8.31	-0.16
	Leu	4.31	4.34	4.34	-0.03	Leu	8.11	8.16	8.30	-0.19
	Gln	4.32	4.34	4.34	-0.02	Gln	8.24	8.32	8.46	-0.22
	Glu	4.32	4.35	4.35	-0.03	Glu	8.27	8.42	8.57	-0.30
6	Tyr	4.44	4.55	4.55	-0.11	Tyr	8.26	8.12	8.12	0.14
	Lys	4.15	4.32	4.32	-0.17	Lys	8.27	8.29	8.29	-0.02
	Arg	4.17	4.34	4.34	-0.17	Arg	8.07	8.23	8.37	-0.30
	Arg	4.22	4.34	4.34	-0.12	Arg	8.22	8.23	8.38	-0.16
	Ile	4.04	4.17	4.17	-0.13	Ile	8.06	8.00	8.15	-0.09
	Leu	4.27	4.34	4.34	-0.07	Leu	8.23	8.16	8.33	-0.10
	Cys*	4.41	4.55	4.55	-0.14	Cys*	8.34	8.32	8.46	-0.12
	Arg	4.28	4.34	4.34	-0.06	Arg	8.27	8.23	8.43	-0.16
	Leu	4.3	4.34	4.34	-0.04	Leu	8.18	8.16	8.31	-0.13
	Leu	4.32	4.34	4.34	-0.02	Leu	8.14	8.16	8.30	-0.16
	Gln	4.33	4.34	4.34	-0.01	Gln	8.19	8.32	8.46	-0.27
	Glu	4.35	4.35	-0.01	4.34	Glu	8.23	8.42	8.57	-0.34
7	Phe	4.51	4.62	4.62	-0.11	Phe	8.26	8.30	8.30	-0.04
	Lys	4.08	4.32	4.32	-0.24	Lys	8.12	8.29	8.39	-0.27
	Arg	4.25	4.34	4.34	-0.09	Arg	8.3	8.23	8.37	-0.07
	Arg	4.21	4.34	4.34	-0.13	Arg	8.13	8.23	8.38	-0.25
	Ile	4.22	4.17	4.17	0.05	Ile	8.27	8.00	8.15	0.12
	Leu	4.32	4.34	4.34	-0.02	Leu	8.15	8.16	8.33	-0.18
	Cys*	4.50	4.55	4.55	-0.05	Cys*	8.32	8.32	8.46	-0.14
	Arg	4.28	4.34	4.34	-0.06	Arg	8.38	8.23	8.43	-0.05
	Leu	4.32	4.34	4.34	-0.02	Leu	8.15	8.16	8.31	-0.16
	Leu	4.33	4.34	4.34	-0.01	Leu	8.27	8.16	8.30	-0.03
	Gln	4.33	4.34	4.34	-0.01	Gln	8.25	8.32	8.46	-0.21
	Glu	4.34	4.35	4.35	-0.01	Glu	8.27	8.42	8.57	-0.30

S3.5 Supplementary computational modelling images

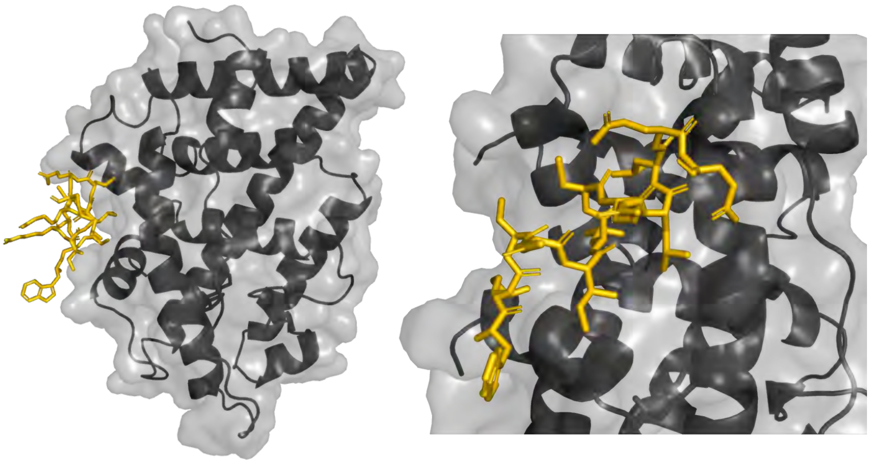


Figure S4: Control peptide **1a** docked to the AF2 surface of the ERα LBD shown as sticks. The binding mode of conserved Leu residues is conserved with other published crystal structures (data not shown).[3-5]

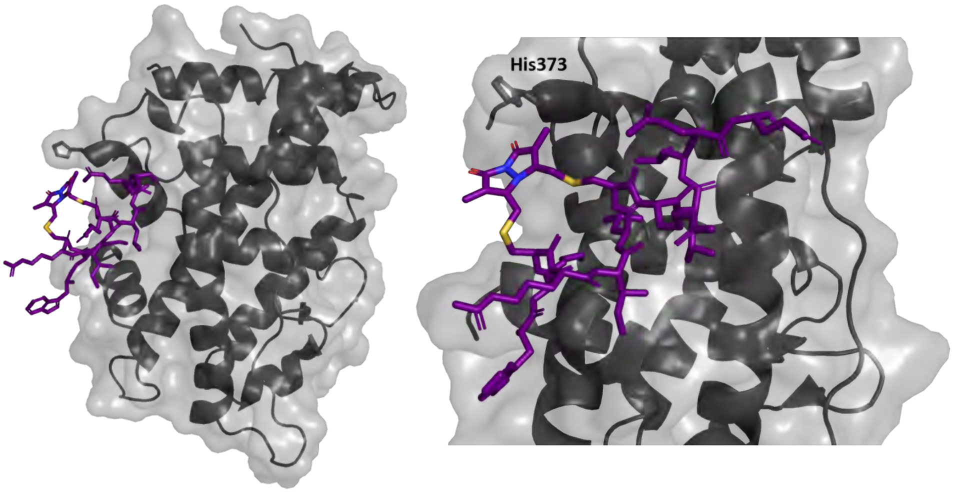


Figure S5: Peptide **1** docked to the AF2 surface of the ERα LBD shown as sticks.

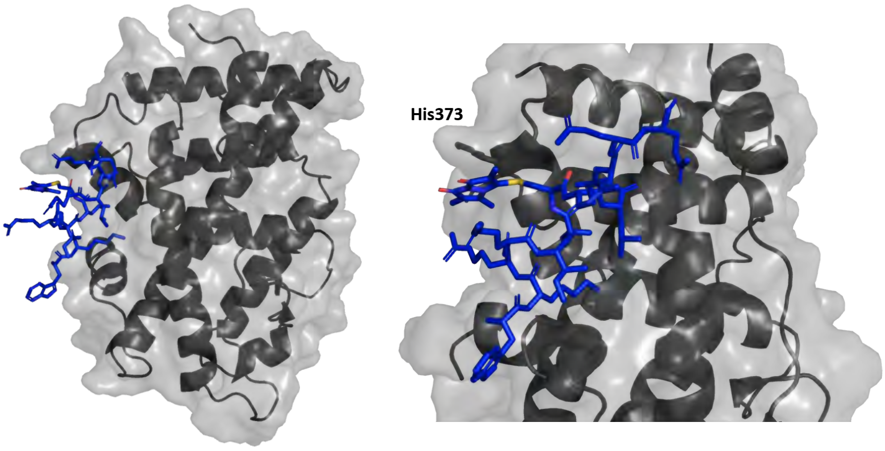


Figure S6: Peptide **3** docked to the AF2 surface of the ERα LBD shown as sticks. Potential interaction of the bimane modification with the H373 of the ERα LBD is shown (right).

S3.6 Fluorescence/absorbance characterisation

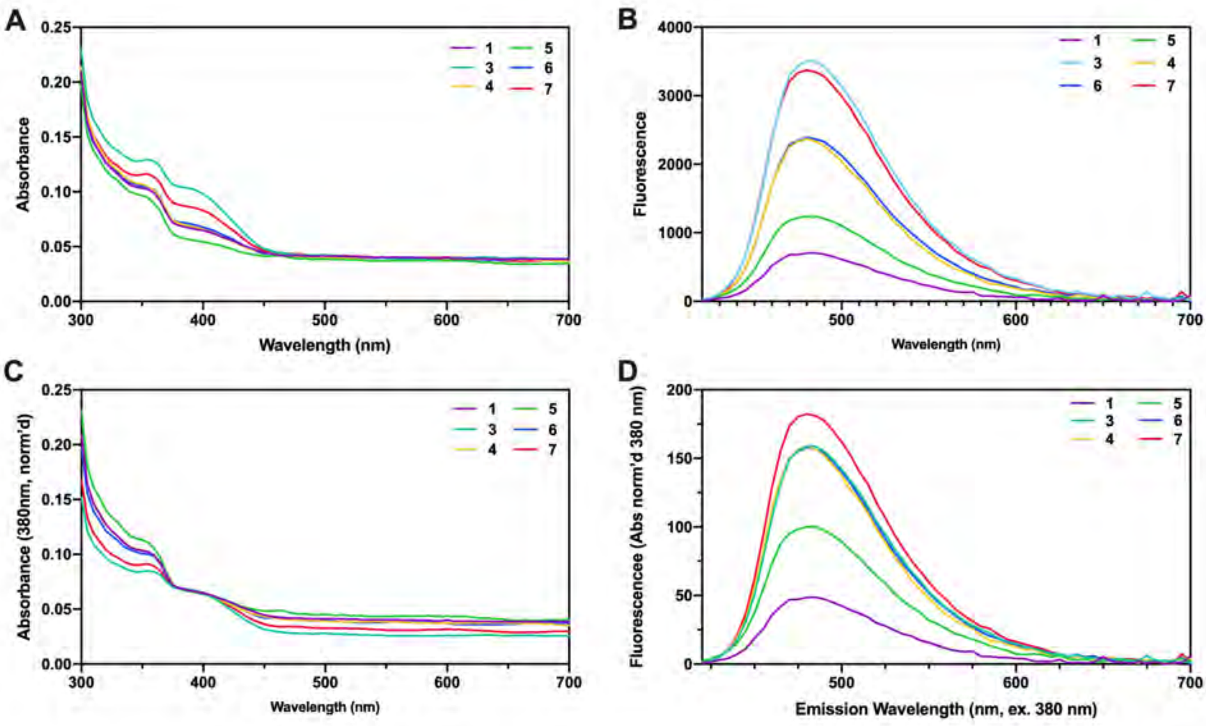


Figure S7: Raw Absorbance (**A**) and fluorescence spectra (**B**, ex. 380 nm) and absorbance (**C**) and fluorescence (**D**, ex. 380 nm) spectra normalised to absorbance at 380 nm, collected in 10 mM phosphate buffer at pH 7.2.

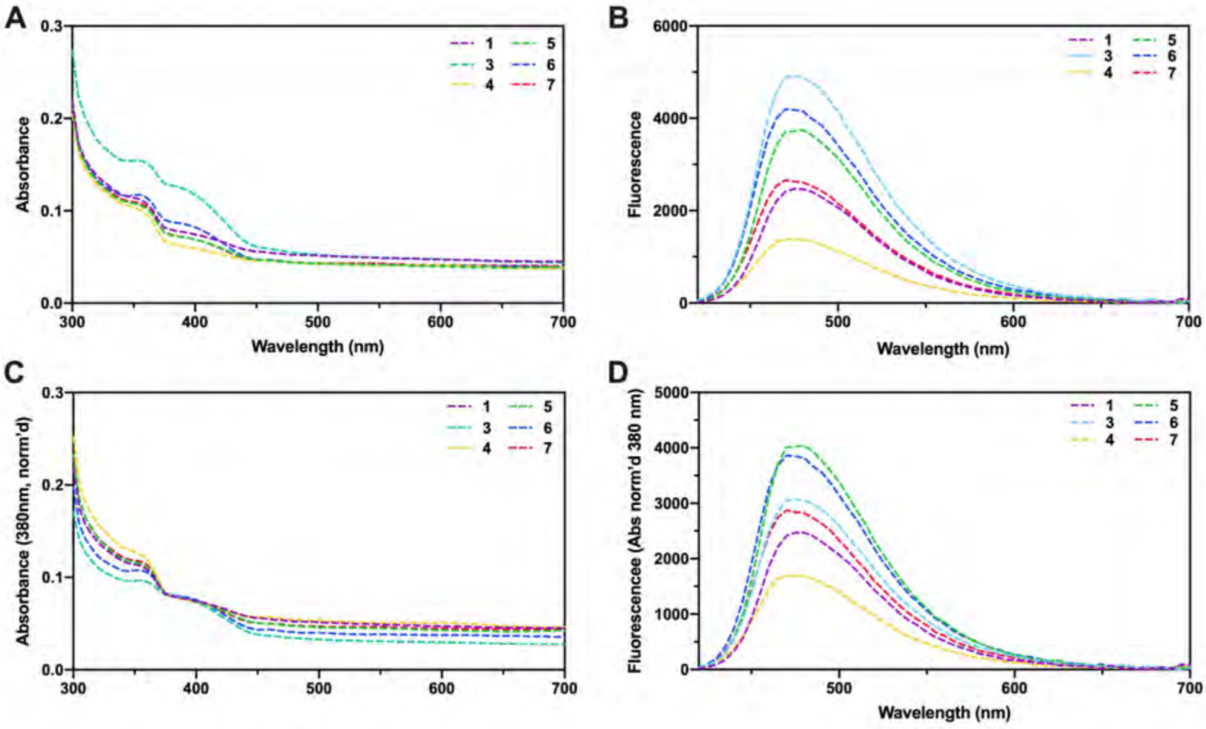


Figure S8: Raw Absorbance (**A**) and fluorescence spectra (**B**, ex. 380 nm) and absorbance (**C**) and fluorescence (**D**, ex. 380 nm) spectra normalised to absorbance at 380 nm, collected in 10 mM phosphate buffer at pH 7.2 with 20% TFE.

Table S3: Fluorescence analysis: Spectra collected in 10 mM phosphate pH 7.2 (Phos) or 10 mM phosphate pH 7.2 with 20% TFE (20% TFE).

	Emission Max (nm)			**Intensity @ Emission Max**			
Peptide	**Phos**	**20% TFE**	**Shift**	**Phos**	**20%**	**Change (Ratio)**	
1	485	475	-10	705	2472	+1767	(+3.51)
3	485	475	-10	3503	4910	+1407	(+1.40)
4	480	470	-10	2369	1380	-989	(-0.58)
5	480/485	480	0 – -5	1241	3742	+2501	(+3.02)
6	480	470	-10	2391	4200	+1809	(+1.76)
7	480	470	-10	3371	2661	-710	(-0.71)

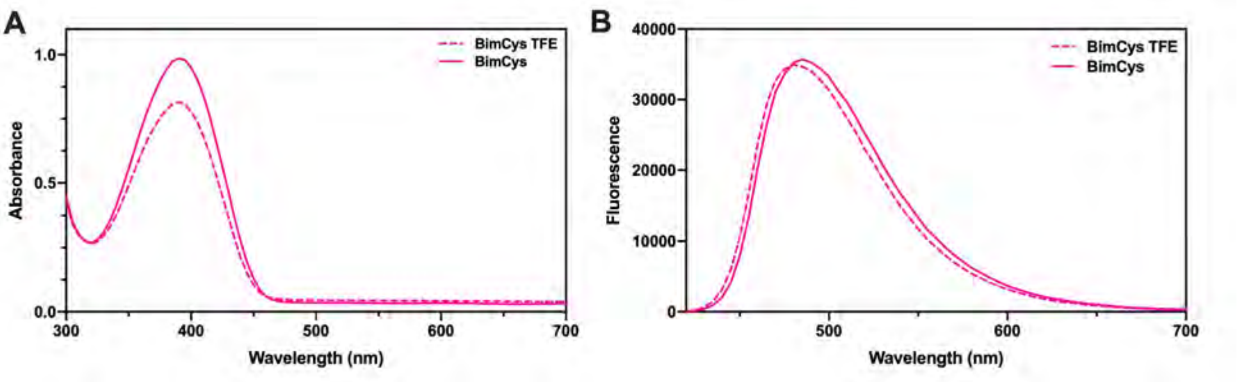


Figure S9: Raw absorbance (**A**) and fluorescence spectra (**B**, ex. 380 nm) collected in 10 mM phosphate buffer (solid line) and in 10 mM phosphate buffer at pH 7.2 with 20% TFE (dashed line). **BimCys synthesis:** Cysteine hydrochloride (2.0 mg) was dissolved in water (200 µL), and combined with 100 mM Phosphate pH 7.2 (60 uL) and a solution of monobromobimane (1 equiv) in MeOH (400 µL). The solution was placed on an orbital rocker for 2 h, and then lyophilised. The resulting powder was resuspended in water (300 uL) and centrifuged to remove any particulate (2 min, 10000 rpm). An SPE cartridge (150 mg, Supelco C18-Lt) was soaked with MeOH for 10 min, then washed with water (3 x 3 mL). The bimane solution was added to the cartridge, and loaded onto the top of the silica, then washed with water (2 x 3 mL). The BimCys was then eluted with 9% aq. ACN (300 uL) and the solution lyophilised to give BimCys as a pale yellow powder.

REFERENCES

1. N. E. Shepherd, H. N. Hoang, G. Abbenante and D. P. Fairlie, Single Turn Peptide Alpha Helices with Exceptional Stability in Water, *J. Am. Chem. Soc.*, 2005, **127**, 2974-2983.
2. D. S. Wishart, Interpreting protein chemical shift data, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2011, **58**, 62-87.
3. K. W. Nettles, J. B. Bruning, G. Gil, J. Nowak, S. K. Sharma, J. B. Hahm, K. Kulp, R. B. Hochberg, H. Zhou, J. A. Katzenellenbogen, B. S. Katzenellenbogen, Y. Kim, A. Joachmiak and G. L. Greene, NFkappaB selectivity of estrogen receptor ligands revealed by comparative crystallographic analyses, *Nat. Chem. Biol.*, 2008, **4**, 241-247.
4. T. E. Speltz, S. W. Fanning, C. G. Mayne, C. Fowler, E. Tajkhorshid, G. L. Greene and T. W. Moore, Stapled Peptides with gamma-Methylated Hydrocarbon Chains for the Estrogen Receptor/Coactivator Interaction, *Angew. Chem. Int. Ed. (English)*, 2016, **55**, 4252-4255.
5. P. Y. Maximov, B. Abderrahman, S. W. Fanning, S. Sengupta, P. Fan, R. F. Curpan, D. M. Q. Rincon, J. A. Greenland, S. S. Rajan, G. L. Greene and V. C. Jordan, Endoxifen, 4-Hydroxytamoxifen and an Estrogenic Derivative Modulate Estrogen Receptor Complex Mediated Apoptosis in Breast Cancer, *Mol. Pharmacol.*, 2018, **94**, 812-822.

Chapter 4.

TARGETING PCNA WITH PEPTIDE MIMETICS FOR THERAPEUTIC PURPOSES

Minireview: ChemBioChem **2019**, 21(4), 442-450

Targeting PCNA with peptide mimics for therapeutic purposes

Aimee J. Horsfall,[1] and Andrew D. Abell,[1] John B. Bruning[2*]

[1] ARC Centre of Excellence for Nanoscale BioPhotonics, Institute for Photonics and Advanced Sensing (IPAS) Department of Chemistry, University of Adelaide, Nth Tce, Adelaide, 5005

[2] Institute of Photonics and Advanced Sensing (IPAS) School of Biological Sciences, University of Adelaide Nth Tce, Adelaide, 5005

*corresponding author

© 2019 ChemBioChem, Wiley

First published online: 27th June 2019

ABSTRACT

Proliferating cell nuclear antigen (PCNA) is an excellent inhibition target to shut down highly proliferative cells and thereby develop a broad spectrum cancer therapeutic. It interacts with a wide variety of proteins through a conserved motif referred to as the PCNA-interacting protein (PIP) box. There is large sequence diversity between high affinity PCNA binding partners, with conservation of the binding structure - a well-defined 3_{10}-helix. Here, all current PIP-box peptides crystallised with human PCNA are collated to reveal common trends between binding structure and affinity. Key intra- and inter-molecular hydrogen bonding networks which stabilise the 3_{10}-helix of PIP-box partners are highlighted, and related back to the canonical PIP-box motif. High correlation with the canonical PIP-box sequence does not directly afford high affinity. Instead, we summarise key interactions which stabilise the binding structure that lead to enhanced PCNA binding affinity. These interactions also implicate the ‘non-conserved’ residues within the PIP-box that have previously been overlooked. Such insights will allow a more directed approach to develop therapeutic PCNA inhibitors.

An intricate understanding of peptide partner interactions is required to design new inhibitors of PCNA, the human sliding clamp. Insights made here will enable efficient design of new, broad spectrum cancer therapeutics.

STATEMENT OF AUTHORSHIP

Title of Paper	Targeting PCNA with peptide mimetics for therapeutic purposes
Publication Status	☒ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	**MiniReview:** A. J. Horsfall, A. D. Abell and J. B. Bruning, *ChemBioChem*, 2019, **21**, 442-450

Principal Author

Name of Principal Author (Candidate)	Aimee J Horsfall		
Contribution to the Paper	Researched, devised structure and wrote manuscript.		
Overall percentage (%)	70%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	15/02/2021

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);
ii. permission is granted for the candidate in include the publication in the thesis; and
iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Andrew D Abell		
Contribution to the Paper	Supervised AJH and edited manuscript.		
Signature		Date	15/2/2021

Name of Co-Author	John B Bruning		
Contribution to the Paper	Supervised AJH and edited manuscript.		
Signature		Date 18/2/21	

4.1 Introduction

Healthy cell proliferation is dependent on the ability to faithfully replicate the genome. This requires a host of proteins which range from chromatin remodelling proteins, helicases, and polymerases, to DNA-repair proteins and ligases. The regulation of this highly complex process and its precise execution remains of significant interest. Proliferating cell nuclear antigen (PCNA) was discovered in 1978 as an essential factor for proliferation.[1] However it wasn't until 1996 that the first co-crystal structure of human PCNA was elucidated.[2] PCNA was initially identified as a processivity factor for polymerase δ,[3] however there are now over 200 proteins identified as PCNA-binding partners which interact dynamically with PCNA throughout cell cycle regulation, DNA-replication and -repair processes.[4-8] PCNA acts as a mobile scaffold to encircle DNA and slide unidirectionally with the progressing replication fork. This provides an anchoring point for requisite proteins to bind and gain proximity to DNA.[5-7]

PCNA is upregulated in a large number of cancer cell lines to present as an important target for development of cancer therapeutics.[9, 10] A large majority of PCNA partners bind to the surface of PCNA through a conserved motif, referred to as the PCNA-Interacting Protein (PIP) box. The cell-cycle control protein p21 binds PCNA through a PIP-box with the highest known affinity, and upon binding shuts down DNA-replication. p21 is thus a key template for therapeutic PCNA inhibitor design. New trends can be observed within the growing library of PCNA-partner co-crystal structures. These developments reveal how dramatic differences in PIP-box sequence and subtle differences in binding conformation impact binding affinity. A number of related binding motifs have been identified including the AlkB homolog 2 PCNA Interacting Motif (APIM), the binding region of which overlaps with the PIP-box binding site.[8, 11, 12] This Minireview focuses only on bona fide human PIP-box proteins and their interaction with the human sliding clamp, PCNA. The peptides which have been co-crystallised with PCNA are summarised in Table 1.

4.2 PCNA structure and function

PCNA is a toroidal-shaped homotrimer. Each monomer consists of two similar domains connected by a strand with no defined secondary structure, known as the inter-domain connecting loop (IDCL, residues 119-134) (Figure 1). The IDCL is 16 residues long and spans approximately forty angstroms over the surface of the ring, nearest the progressing replication fork.

Each monomer associates in a head-to-tail fashion (A and B domains alternate) resulting in six-fold pseudo-symmetry. The first crystal structure of human PCNA was solved in 1996 in complex with a 22 amino-acid peptide from p21 (Figure 1A).[2] However, it wasn't until 2005 that the apo crystal structure of PCNA was elucidated.[13] The ring exterior is composed primarily of anti-parallel β-strands which form negatively charged, curved β-sheets. The inner diameter of the ring spans approximately thirty angstroms, allowing sufficient room for dsDNA to be encircled.[14, 15] The internal surface of the ring comprises four arginine-rich α-helices per monomer. The positive charge facilitates non-specific interactions between the sliding clamp and DNA phosphodiester backbone.[14] PCNA cannot spontaneously assemble around DNA and must first be loaded over primer-template junctions by the RFC complex in an ATP-dependent

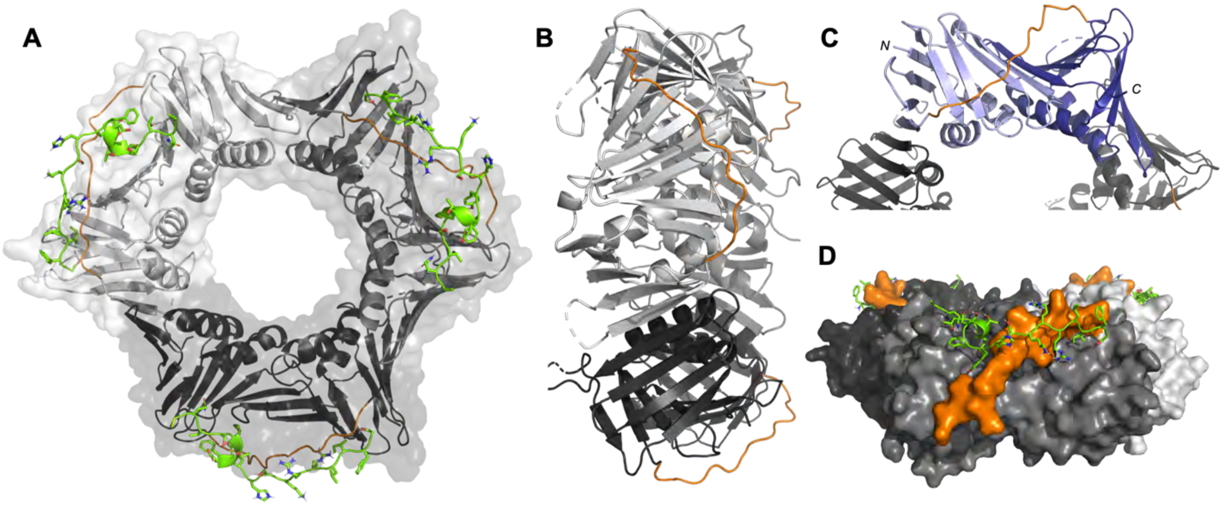


Figure 1. PCNA structure: PCNA (grey, 1AXC), PCNA-Interdomain Connecting Loop (IDCL, orange), $p21_{139\text{-}160}$ (green, 1AXC) **A)** Front view of PCNA, which has toroidal shape comprised of three identical monomers indicated in shades of grey. The PIP-box peptide p21 binds next to the IDCL. **B)** Side-view of PCNA (cartoon). The orange IDCLs face the direction of replication as PCNA slides along the DNA. **C)** Each PCNA monomer is comprised of two similar domains distinguished in light and dark purple, connected by the IDCL (orange). **D)** The 3_{10}-helix of p21 sits in a hydrophobic cleft while the *C*-terminal tail interacts with the IDCL.

fashion.[16] The PCNA-DNA interaction is mediated (at two of three surfaces) by the peptidic arms of $p15^{PAF}$, which reach into the clamp barrel from protein bound at the PIP-box, thereby fastening DNA within the PCNA ring.[15, 17]

A large number of PCNA-interaction partners bind near the IDCL through a PIP-box motif.[6] Although PCNA itself has no enzymatic function, association with PCNA is crucial for enzymatic activity or processivity of many binding partners.[18, 19] Other partners utilise PCNA as a common meeting ground to provide a defined steric orientation for post-translational modifications (such as ubiquitination) to be introduced.[20] Some of these interactions require PCNA to be DNA-bound, whereas others rely on post translationally modified PCNA to enhance recognition or binding affinity.[21-23] For example the TLS polymerases preferentially bind mono-ubiquitinated PCNA.[17, 20, 24, 25] There is not believed to be a cancer-related mutant, despite the clear link between PCNA and cancerous cells.[26]

4.3 The PCNA-interacting protein (PIP) box and p21

The PIP-box consensus sequence is defined as Qxxφxxψψ where Q is glutamine, x is any amino-acid, φ is a hydrophobic residue and ψ an aromatic residue (commonly Phe or Tyr).[11] This motif binds a hydrophobic patch on the PCNA surface, flanked by the protein *C*-terminus and the IDCL. Henceforth the first residue of the PIP-box is referred to as position 1, the first ‘x’ place as position 2, and so-on, until the second ψ residue at position 8. This may also be indicated after an amino-acid identifier by the superscript P1, P2 etc. Position 1 Gln residue protrudes into the aptly named ‘Q-pocket’ and commonly hydrogen bonds with the main-chain of PCNA residues Ala252, Ala208 and Val45. The φ, ψ, ψ residue combination forms a three-pronged hydrophobic plug which inserts into a groove on the PCNA surface, and consequently twists the peptide backbone into a 3_{10}-helix from position 4-8. This 3_{10}-helix is observed in all known structures of PIP-peptides/proteins bound to PCNA (Figure 2C). A key example of the PIP-box interaction with PCNA is a 22-mer peptide derived from the $p21^{WAF/CIP1}$ protein.[2] The p21 peptide

(^{139}GRKRRQTSMTDFYHSKRRLIFS^{160}), with PIP-box QTSMTDFY, fulfils all consensus sequence requirements and was the first crystal structure to be elucidated in complex with human PCNA. The crystal structure (PDB: 1AXC, Figure 1A) shows three p21 peptides, one bound to each PCNA monomer, and resolves all bar the 5 *N*-terminal residues of each peptide. These positively charged residues, though unresolved, are positioned near the negatively charged PCNA *C*-terminus and are thought to act like a lid to hold Gln144^{P1} in the Q-pocket through ion pairing. The glutamine side-chain of p21 makes two hydrogen bonds within the Q-pocket; to the main-chain carbonyls of Ala252 and Pro253 in PCNA (Figure 3B). The Met147^{P4} side-chain binds into a deep hydrophobic pocket under the IDCL, in combination with Tyr151^{P8} which hydrogen bonds to Gln131 of PCNA (Figure 3A and B). The Tyr151^{P8} interaction plays a key role in enhancing binding affinity between p21 and PCNA, as a p21 Tyr151Phe^{P8} mutant binds with 3-fold lower affinity.[27] The p21 peptide adopts a 3$_{10}$-helical conformation from Ser146^{P3} to Tyr151^{P8} which is stabilised by a number of intramolecular hydrogen bonds. Both Thr148^{P5} and Asp149^{P6} hydrogen bond to the hydroxyl of Ser146^{P3}, and the Thr145^{P2} amide nitrogen hydrogen bonds to the Asp149^{P6} side-chain (Figure 3B). This network stabilises the 3$_{10}$-helix and positions Met147^{P4} and Tyr151^{P8} within the hydrophobic cleft. Establishing the 3$_{10}$-helix and insertion into the hydrophobic cleft is a cooperative process. The positively charged residues *C*-terminal to the PIP-box form an anti-parallel β-sheet with the negatively charged IDCL (Figure 2C, green). p21 has the highest binding affinity of any reported PIP-box:PCNA interaction; however, it is not fully understood why this is so. The hydrogen bonding networks made by p21, and its PCNA binding conformation will be contrasted to alternate PIP-box partners to uncover factors important in enhancing binding affinity.

4.4 PCNA-binding partners display structural variation

Sequence variation exists between different PIP-boxes, as well as differences in binding conformation. Only a few partners exhibit the IDCL β-sheet interaction observed in p21, which appears to be correlated to PIP-box location within the protein sequence. Both p21 and FEN-1 make extended β-sheet contacts with the IDCL, with both PIP-boxes located in the *C*-terminus of the full-length protein.[2, 28] PIP-boxes positioned at the extreme protein *C*-terminus do not form a defined IDCL interaction, such as in TRAIP where only two residues proceed the PIP-box.[29] In comparison, neither the *N*-terminally located pol δ p12 PIP-box, nor the internally located pol ι PIP-box, make extended contacts with the IDCL, even in peptides with sufficient length to do so.[30] The residues *N*-terminal to the PIP-box also display a variety of interactions across PCNA partners. For example, residues *N*-terminal to the p15 PIP-box form a β-turn allowing the preceding sequence to extend down into the centre of the clamp and interact with the encircled DNA (Fig 2, blue).[15] Residues *N*-terminal to the RNaseH2B 3$_{10}$-helix form a β-strand with the PCNA *C*-terminus.[31] The 3$_{10}$-helix of PCNA-bound partners is conserved in all known structures (Figure 2). This is particularly interesting as the helix is not a pre-organised motif and the large majority of PIP-boxes are found in intrinsically disordered protein domains.[32, 33] Though regions flanking the PIP-box are likely to enhance binding affinity, it is probable that the key dictator of binding affinity lies with stabilisation of the 3$_{10}$-helix.

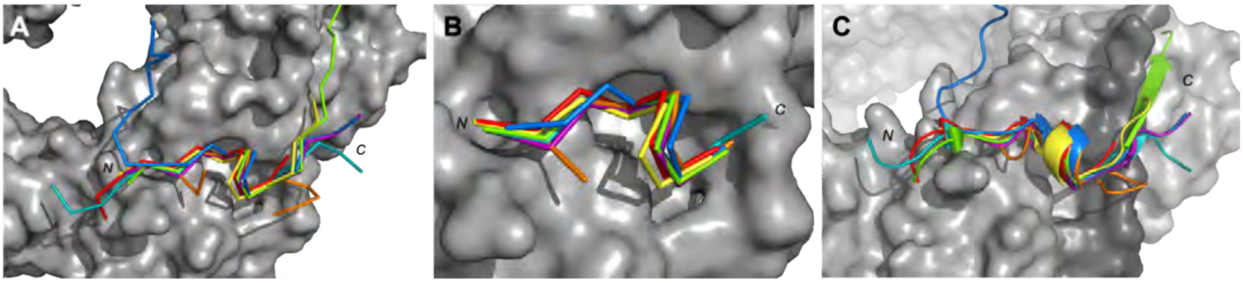


Figure 2. Similarities and differences in PIP-box partners, all peptides are orientated *N*-terminal (left) to *C*-terminal (right). PCNA (grey, 1AXC, surface representation), p21 (green, 4RJF), pol δ $_{p66}$ (red, 1U76), pol η (blue, 2ZVK), FEN-1 (yellow, 1U7B), pol ι (orange, 2ZVM), PARG (purple, 5MAV), Pogo Ligase (light blue, 1UYJ). **A)** Peptides in ribbon representation. The PIP-box sequences bind in a similar manner, however there is structural variation in regions *N*- and *C*-terminal to the PIP-box. **B)** Ribbon representation of PIP-box only. The PIP-box binding structure is highly conserved. **C)** Cartoon representation of peptides. All PIP-box peptides display a 3_{10}-helix upon binding PCNA. IDCL coloured dark grey.

4.5 The non-canonical PIP-box and sequence variation

A conserved PCNA-binding motif was identified in 1998,[34] with a large number of proteins subsequently shown to bind PCNA at the same site through increasingly divergent sequences.[8, 12] Necessity of the position 1 Gln residue in the PIP-box consensus sequence has been questioned, in part due to the prevalence of non-canonical PIP-boxes and similar motifs which bind near the PIP-box site.[12, 35] A non-canonical PIP-box does not prescribe Gln at position 1 and most often contains Lys or Met in its place, however Asp and Ser have also been observed. The most notable non-canonical PIPs belong to the Y-family TLS polymerases. Peptides derived from pol η and pol ι (PIP-box: ^{701}MQTLESFF and ^{421}KGLIDYYL, respectively) were shown to bind with comparable (if not higher) affinity than the canonical PIP-box peptides from FEN-1 (^{337}QGRLDDFF) and pol δ p66 (^{456}QVSITGFF).[30, 36] The pol η Met701^{P1} binds in the Q-pocket but is only capable of making hydrophobic contacts, not the characteristic hydrogen bonds observed for the canonical Gln. Interestingly, but perhaps predictably, a mutation from Met to Gln improves the binding affinity of the pol η peptide. By comparison, the pol ι Lys241^{P1} is somewhat disordered and doesn't enter the Q-pocket, and instead folds back into a "β-bend-like" structure to interact with residues at position 6 and 7. Interestingly the requirement of a position 8 aromatic residue is not met in pol ι with incorporation of a Leu residue instead. The only remaining similarities between the pol ι PIP-box and classical PIP-box are a hydrophobic residue at position 4, an aromatic residue at position 7, and formation of a 3_{10}-helix between these positions, yet binding affinity for PCNA remains in the low micromolar range (0.39 μM). Therefore, the primary sequence prescribed by the PIP-box consensus does not directly correlate to enhanced PCNA binding affinity. The growing pool of structural data allows for a more detailed discussion on how sequence variation and subtle structural differences impact binding affinity. This information will assist in understanding the role and regulation of each protein during DNA-replication and -repair processes. Furthermore, discerning the important interactions and sequence combinations which stabilise the PIP-box binding conformation will allow accelerated, rational therapeutic design.

4.6 PIP-box divergence

A large diversity between PIP-box sequences exists, with no clear correlation between those which conform closely to the canonical binding motif and those with high binding affinity. For example, peptides

Table 1. PIP-box peptide-PCNA co-crystal structure data including reported PCNA binding affinities.

PDB ID	Year	Authors[a]	Peptide	Sequence[b]	PIP-box	Affinity	Binding Assay
1AXC	1996	Gulbis, Kuriyan	**p21** (139-160)	GRKRR**QTSMTDFY**SKRRLIFS	**QTSMTDFY**	2.5 nM[c]	Biosensor Analysis
6QC0	2018	Hayashi, Nishitani	**Cdt2** (704-717)	SS**MRKICTYF**HRKS	**MRKICTYF**	57 nM	FP
1VYJ	2005	Kontopidis, Walkinshaw	**Pogo-Ligase** (mutant, 1-16)	SAVL**QKKITDYF**HPKK	**QKKITDYF**	100 nM	ITC
4RJF	2015	Kroker, Bruning	**p21-151F** (139-160)	GRKRR**QTSMTDFF**HSKRRLIFS	**QTSMTDFF**	250 nM	ITC
2ZVM	2009	Hishiki, Sato	**pol ι** (415-437)	ALNTAK**KGLIDYYL**MPSLSTTSR	**KGLIDYYL**	0.39 μM	SPR
2ZVK	2009	Hishiki, Sato	**pol η** (693-713)	CKRPRPEG**MQTLESFF**KPLTH	**MQTLESFF**	0.4 μM	SPR
6CBI	2018	Wegener, Bruning	**p21 mimetic** (mutant, 139-152)	GRKRRQ**X**SMT**X**FYH[c]		1.61 μM	FP
5MAV	2017	Kaufmann, Slade	**PARG** (402-420)	QHGKK**DSKITDHF**MRLPKA	**DSKITDHF**	3.3 μM	ITC
2ZVL	2009	Hishiki, Sato	**pol κ** (861-874 +PLTH)	P**KHTLDIFF**KPLTH	**KHTLDIFF**	4.9 μM	SPR
5MLO	2017	Sebesta, Ahel	**ZRANB3** (515 – 529)	EKEK**QHDIRSFF**VPQ	**QHDIRSFF**	4.80 μM	ITC
4D2G	2015	De Biasio, Blanco	**p15^{Paf}** (50-70)	APVCVRPTPKW**QKGIGEFF**AA	**QKGIGEFF**	5.56 μM	ITC[d]
6QCG	2018	Hayashi, Nishitani	**Cdt1** (2-15)	ME**QRRVTDFF**ARRR	**QRRVTDFF**	7.20 μM	FP
1U76	2004	Bruning, Shamoo	**Pol δ p66** (452-466)	KANR**QVSITGFF**QRK	**QVSITGFF**	15.6 μM	ITC
5IY4	2016	Wang, Jiang	**DVC1** (321-336)	SNSH**QNVLSNYF**PRVS	**QNVLSNYF**	15.6 μM	ITC[e]
5YCO	2017	Chen, Zang	**UHRF2** (784-800)	NEIL**QTLLDLFF**PGYSK	**QTLLDLFF**	25.7 μM	ITC
4ZTD	2016	Hoffmann, Mailand	**TRAIP** (458-469)	AF**QAKLDTFL**WS	**QAKLDTFL**	30.7 μM	ITC[f]
6GWS	2018	De March, De Biasio	**p15** (41-72)	RKAENKYAGGNPVCVRPTP KW**QKGIGEFF**RLS	**QKGIGEFF**	35 μM	NMR
6HVO	2019	Gonzalez-Magana, De Biasio	**pol δ p12** (1-19)	MGR**KRLITDSY**PVVKRREG	**KRLITDSY**	38 μM	ITC
1U7B	2004	Bruning, Shamoo	**FEN1** (331-350)	SRQGST**QGRLDDFF**KVTGSL	**QGRLDDFF**	59.9 μM	ITC
3P87	2011	Bubeck, Jackson	**RNaseH2B** (290-312)	DKSG**MKSIDTFF**GVKNKKKIGKV	**MKSIDTFF**	-	

Abbreviations: FP – fluorescence polarisation (Competition assay); ITC – isothermal titration calorimetry; SPR – surface plasmon resonance; NMR – Nuclear Magnetic Resonance (Titration experiment)

[a] Only the surname of first and last authors are displayed for brevity, (First, Last).

[b] Sequence of peptide used in crystallisation experiment listed. PIP-box is indicated in bold, and residues resolved in crystal structure are underlined.

[c] X residues indicate the lactam cyclised positions.

[d] p15^{PAF} (50-77) peptide used in binding assay

[e] KRR sequence added *N*-terminally to DVC1(321-336) peptide used in binding assay to aid in solubility

[f] TRAIP (447-469) peptide used in binding assay

from p21 and Cdt2 both bind with low nanomolar affinity, however the p21 peptide possesses a canonical PIP-box whereas the Cdt2 PIP-box is highly divergent and non-canonical.[2, 20] There have also been a number of alternate definitions proposed to predict PCNA binding partners, however like the canonical definition, these classifications do not indicate which partner will bind with higher affinity than another. Here the nuances of PCNA interactions are discussed in order to develop a set of predictors for high binding affinity.

pol ι (^{421}KGLIDYYL) presents as a particularly interesting example of the divergent PIP-box with limited resemblance to the consensus sequence. However, it binds with low micromolar affinity (0.39 μM). Features that contribute to the high affinity of pol ι are not inherently unique and are observed in a number of other partners. Pol ι does not meet the requirement for two aromatic residues at position 7/8 of the PIP-box. Leu428^{P8} contacts the hydrophobic pocket of PCNA, particularly with the highly conserved Leu126, however it does not protrude deep into the pocket like the analogous Tyr151^{P8} in p21. Furthermore, the position 1 Gln requirement is not met, as the substitute Lys421^{P1} does not occupy the Q-pocket (Fig 3C, D). Pol ι forms an intramolecular network to overcome these sequence modifications. This stabilises the bound complex to maintain low micromolar affinity. Most notable is a hydrogen bond between the main-chain carbonyl of Lys421^{P1} and the phenol of Tyr427^{P7}. It was demonstrated that a Tyr427Phe^{P7} mutant no longer binds PCNA, indicating the Tyr427-Lys421 interaction is essential to pol ι binding affinity. In addition, there is an intramolecular main-chain hydrogen bond between Lys420 and Leu423^{P3}, and an intermolecular hydrogen bond between Asp425^{P5} and His44 of PCNA (Figure 3D). The hydrogen bond between PCNA His44 and an acidic residue at PIP-box position 5 is a common trait of these highly divergent PIP-boxes, and undoubtedly contributes to the higher than anticipated binding affinity. Pol κ contains a position 5 Asp866, and pol η includes the acidic Glu705^{P5,} both of these residues interact with the PCNA His44 sidechain.[30] A number of other PCNA-peptide structures display a key interaction with PCNA His44, including a UHRF2 peptide where no binding was observed when assayed against a His44Ala PCNA mutant.[37] Formation of the 3_{10}-helix is necessary to orientate the acidic side-chain of a residue at position 5 toward PCNA His44. Interestingly in p21, the Thr148 at position 5 does not make any notable hydrogen bonds to PCNA. It is presumably too removed from His44 to make contact.

It has been proposed that the ThrAsp combination at position 5/6 in the PIP-box is a critical contributor to p21's high binding affinity for PCNA. This is in part due to three intramolecular hydrogen bonds that stabilise the 3_{10}-helix (Figure 3B).[5] Similar interactions are also observed for the PogoLigase peptide which possesses the ThrAsp motif. The Asp10^{P6} in PogoLigase does not make any contacts with PCNA in the co-crystal structure. However, mutation at position 6 to Ala or even Ser decreases ability of a PogoLigase peptide to inhibit CDK4/cyclin D1 kinase activity. This indicates the carboxylic acid group at position 6 is important for stabilising the binding motif.[13] Similarly, the position 6 Asp412 in a PARG peptide does not contact PCNA however, molecular dynamic simulations suggest that Asp412^{P6} aids 3_{10}-helix stabilisation.[38] pol δ p12 and Cdt1 also contain a ThrAsp motif at PIP-box position 5/6. However, these peptides bind with micromolar affinity which is notably weaker than the nanomolar affinity of p21 or PogoLigase. Although the

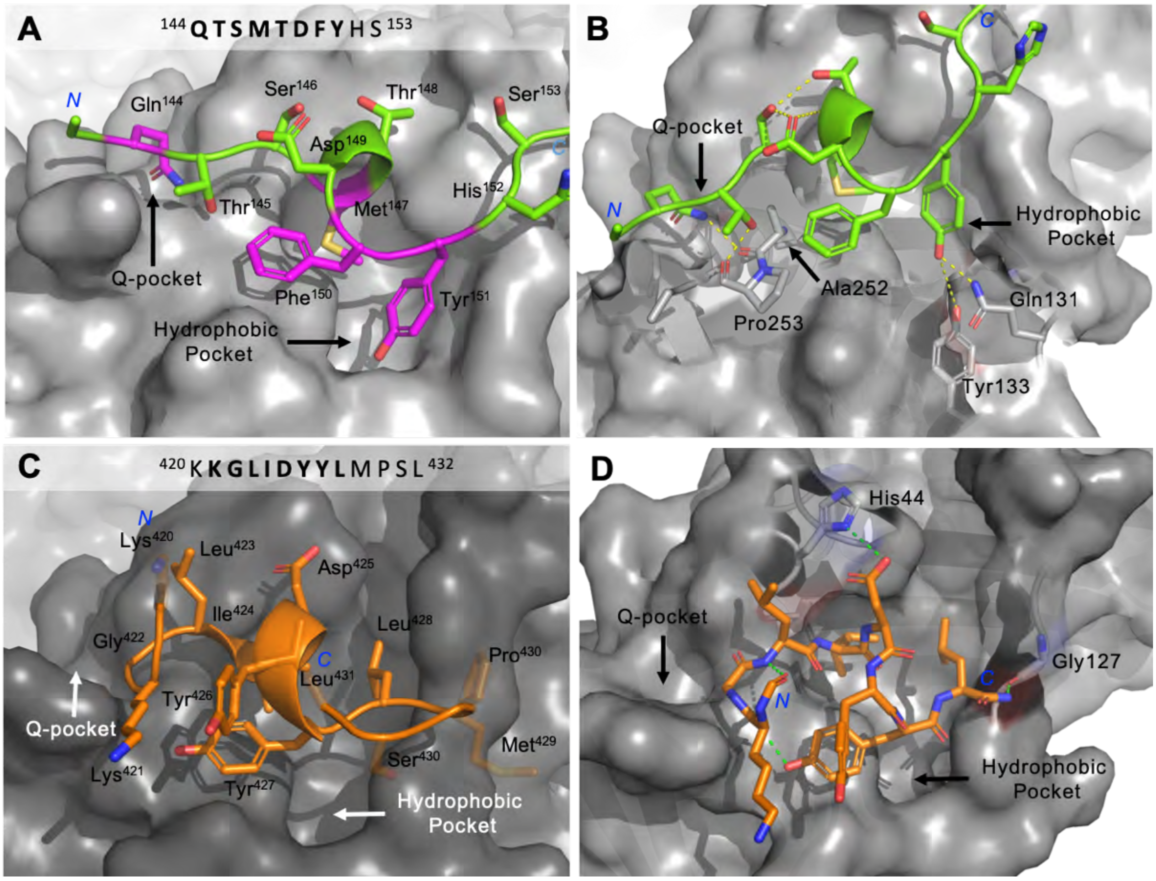

A
^{144}QTSMTDFYHS^{153}
N
Gln^{144}
Ser^{146}
Thr^{148}
Ser^{153}
Asp^{149}
C
His^{152}
Thr^{145}
Met^{147}
Q-pocket
Phe^{150}
Tyr^{151}
Hydrophobic Pocket
B
C
Q-pocket
N
Hydrophobic Pocket
Ala252
Gln131
Pro253
Tyr133
C
^{420}KKGLIDYYLMPSL^{432}
N
Lys^{420}
Leu^{423}
Asp^{425}
Ile^{424}
Leu^{428}
Gly^{422}
C
Leu^{431}
Pro^{430}
Tyr^{426}
Q-pocket
Met^{429}
Ser^{430}
Lys^{421}
Tyr^{427}
Hydrophobic Pocket
D
His44
Q-pocket
Gly127
C
N
Hydrophobic Pocket

Figure 3. Differences in canonical and non-canonical binding partners. PCNA (grey, 1AXC) in surface representation. **A)** p21 (green, 1AXC, cartoon/sticks) and conserved residues highlighted in purple. The position 1 Gln binds the 'Q-pocket' and three hydrophobic residues form a plug, fitting in a cavity on the surface. **B)** Three stabilising intramolecular hydrogen bonds, as well as hydrogen bonds to the Q-pocket and by Tyr151^{P8} of p21 are shown (yellow, dashed**)**. **C)** Lys421^{P1} of pol ı (orange, 2ZVM, stick/cartoon) does not bind in the Q-pocket, nor is the hydrophobic PCNA pocket filled. Residues flanking the PIP-box are transparent for clarity**. D)** Two key intermolecular PCNA-pol ı interactions, and two intramolecular hydrogen bonds (green, dashed) stabilise the pol ı binding structure. Only PIP-box residues shown for clarity.

ThrAsp combination has a stabilising effect on the binding conformation, it is evident this motif alone cannot enhance binding affinity to reach the nanomolar range.[20, 33] The position 5/6 ThrAsp motif, in combination with a basic residue four positions after the PIP-box forms a PIP-degron motif recognised by CRL4^{Cdt2}, an E3 ligase, which ubiquitinates the substrate thereby tagging it for degradation.[39]

The importance of an intramolecular hydrogen bond to a position 7 tyrosine is recurrent. As mentioned above, PCNA binding of a pol ı Tyr427Phe^{P7} mutant is abolished. This is likely due to inability to form a hydrogen bond between the Tyr427^{P7} phenol and Lys421^{P1} and stabilise the integral 3_{10}-helix, as seen in the native peptide. Similarly, mutation of Cdt2 Tyr712^{P7} to Ala,[20] or DVC1 Tyr331^{P7} mutation to Phe, eliminates peptide-PCNA binding affinity.[40] Tyr331^{P7} of the DVC1 peptide forms an intramolecular hydrogen bond with Asn326^{P2} to stabilise the 3_{10}-helix. Interestingly, the binding affinity for a DVC1 Tyr331Phe^{P7}/Phe332Tyr^{P8} mutant peptide was also undeterminable. This mutant would be expected to be a tight binder as the position 7/8 PheTyr motif is observed only in the high affinity binding partners p21 and PogoLigase. The importance of position 8 Tyr in p21 to binding has been confirmed by a 3-fold decrease in affinity observed for a Tyr151Phe^{P8} mutant peptide which was co-crystallised with PCNA.[27] This illustrates the difficulty in predicting binding affinity based solely on amino-acid sequence and emphasises there may not be only one solution to the “highest-affinity PIP-box” puzzle.

A hydrophobic residue at position 4 is critical to binding, in contrast to the positions 7/8 aromatic residues where deletion of at least one can be tolerated (when other stabilising factors are in place). It has been shown that mutation at position 4 to a smaller residue eliminates binding to PCNA.[20, 41] This residue plays an integral role in anchoring the 3_{10}-helix to the PCNA binding surface.

A wide variety of residues are incorporated at PIP-box positions 2 and 3, however a bias appears against negatively charged residues. Lys409^{P3} in the PARG peptide plays a critical role in binding, with an Ala mutant showing no interaction with PCNA.[38] The importance of this residue is attributed to both its positive charge, as a Lys409Arg^{P3} mutant recovered one-third of the interaction with PCNA; as well as hydrophobicity of the alkyl chain, as a Lys409Leu^{P3} mutant maintained PCNA binding affinity. Many peptides exhibit key intramolecular interactions between hydrogen bond donor residues in the first half of the PIP-box and acceptor residues located later in the PIP-box. In the previous discussion on 3_{10}-helix stabilisation such intramolecular interactions have involved conserved residues or the highlighted TD motif. Here this effect arises from cooperative interaction of a variety of ‘x’ residues which has been given little emphasis in literature to date. For example, an intramolecular hydrogen bond between Lys6 at position 2 and Asp10 at position 6 of the PogoLigase peptide aids stabilisation of the 3_{10}-helix.[13] The only current crystal structure of a peptide which includes a negatively charged amino-acid at position 2 or 3 of the PIP-box is the ZRANB3 peptide with inclusion of Asp521^{P3}.[42] This peptide unusually contains an Arg523 at position 5, where positively charged residues are usually disfavoured. The charge complementarity between residues 3 and 5, placed either side of the 3_{10}-helix, is likely a stabilising factor for the low micromolar binding affinity (4.80 µM) of the ZRANB3 peptide.

It is important to note that binding affinities reported for PCNA-peptide are likely influenced by regions flanking the PIP-box, and are not solely dictated by PIP-box interactions. In support of this, the binding affinity of pol κ peptide (856-870), or (851-870), to PCNA could not be determined; however, addition of the sequence PLTH (from pol ι) *C*-terminal to the pol κ (856-870) peptide gave rise to 4.9 µM affinity.[30] Similarly, removal of residues 417-420 (LPKA) from PARG peptide (402-420) resulted in 14-fold lower affinity,[38] and a pol δ p66 peptide showed a 10-fold increase in affinity when the *C*-terminal tail of p21 (RRLIFS) was appended.[36] It would be of interest to investigate the impact of the PIP-box independently of these influences.

Together these data suggest that although there is a preference for particular residues at set positions in the PIP-box, considering one position independently of the others does not allow prediction of a high affinity partner. This is further evident as non-canonical PIP-boxes, those without the requisite position 1 Gln, are not inherently low affinity partners. The PIP-box sequence must be considered as a whole, however the resultant hydrogen bond network upon folding into the binding conformation emerges as a primary predictor of affinity. To summarise, the absence of a Gln residue and its essential hydrogen bonding in the Q-pocket can be overcome with an acidic residue at position 5 which interacts with PCNA His44. The position 4 residue must be hydrophobic for the peptide to initiate 3_{10}-helix formation. However, completing the 'hydrophobic plug' to affix the peptide to the PCNA surface only requires one of position 7 or 8 to be aromatic. Tyr at position 8 of the PIP-box often forms important contacts with PCNA. Conversely, Tyr at position 7 frequently interacts intramolecularly with acceptor residues at positions 1, 2 or 3 and stabilises the 3_{10}-helix. Similarly, acidic or hydrogen bond donor residues at position 6 can stabilise the helix by bonding to positively charged residues at positions 2 and 3. However these roles can be reversed to no significant detriment. Additional factors such as the difference in free energy between the free and bound states, as well as the sequences flanking the PIP-box are likely to have a notable contribution on the binding affinity. However, we have limited this discussion to PIP-box contributions. Consequently, we cannot yet define a set of rules by which to predict affinity but have highlighted common factors which appear to contribute to 3_{10}-helix stabilization and enhanced PCNA affinity. As more data emerges these predictors can be exercised and validated.

4.7 Chemical modification of the PIP-box and therapeutic design

A large number of cancer cell lines show upregulation of PCNA which likely contributes to the high proliferation rates of these cells. It has been shown that cancer cells can contain up to 5-6 fold more PCNA than healthy cells.[26] p21 is a key cell-cycle regulator and consequently was identified as a point of control for proliferative cells. p21-derived peptides are known to selectively inhibit malignant cell-growth by blocking access to the PIP-box binding site on PCNA.[43, 44] These results highlight PCNA as an excellent inhibition target for development of a broad spectrum cancer therapeutic, evident by patents in this area.[45-49] Some studies on small molecule inhibitors of PCNA have been reported. The small molecule 3,3',5-triiodothyronine (T3) was crystallised in complex with PCNA and shown to bind at the PIP-box site on PCNA in the hydrophobic cleft normally occupied by an aromatic Phe or Tyr.[50-52] The T3 molecule was shown to bind with an IC_{50} of 3 μM by a fluorescence polarisation competition assay with the PogoLigase peptide. However, T3 is a potent thyroid hormone and consequently its utility as a therapeutic was not pursued. A derivative of T3 has emerged as its successor with an improved IC_{50} of 1 μM. This analogue, T2AA, was shown to limit p21 binding in a competition assay and inhibit the PCNA-Pol δ p66 interaction in cells, thereby demonstrating its utility as a viable replication inhibitor.[50] Further compound development in this area has been pursued.

Small molecules are often unable to make sufficient surface area contact to displace a native protein partner from a protein-protein interaction. Short peptides are considered an excellent comprise between the specificity of proteins and the size of small molecules.[53, 54] High levels of specificity can also be maintained by mimicking known substrates. Peptides often lack defined structure, however this can be reintroduced by incorporation of structural constraints as has been demonstrated in a short p21 peptide.[32]

Peptides are not commonly viewed as suitable drug candidates due to the susceptibility to protease degradation resulting in short *in vivo* half-life, ill-defined structure in solution and in some cases poor membrane permeability. The introduction of a specific chemical constraint into a peptide can address these shortcomings.[55-58] As noted, a PIP-box peptide must organise itself into a 3_{10}-helix for binding to PCNA.

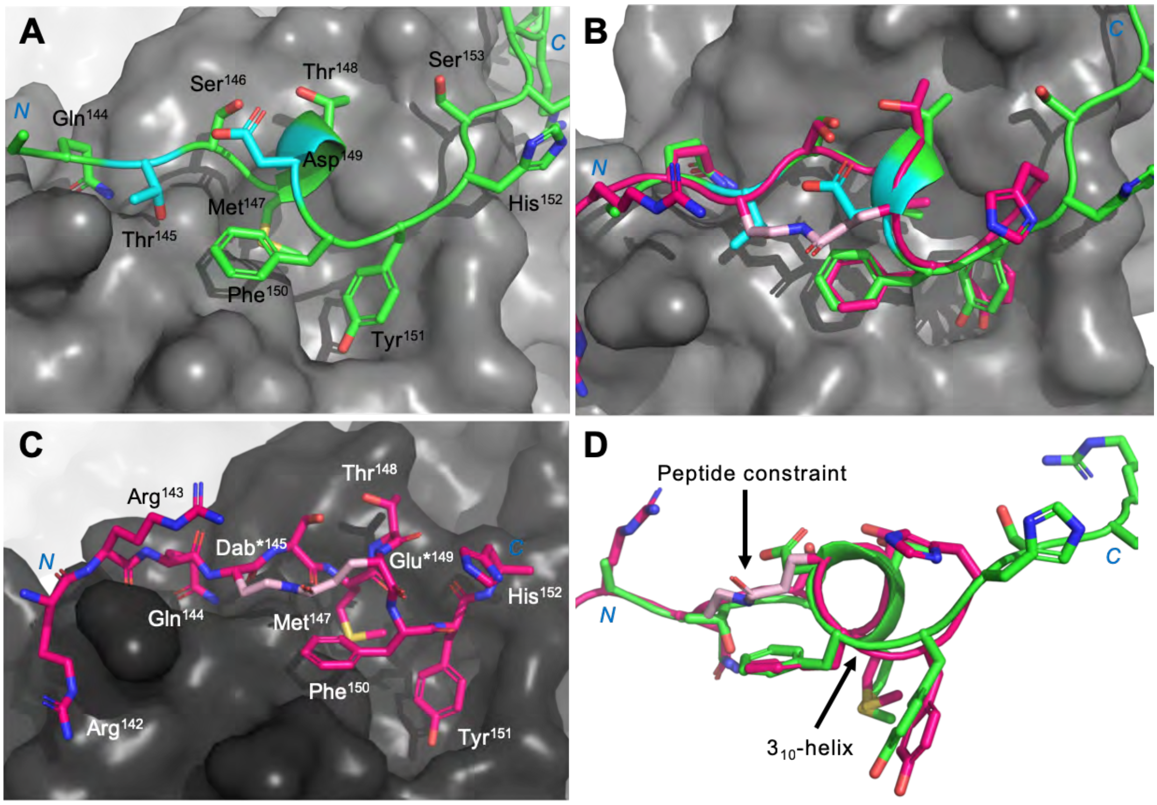


Figure 4: Chemical stabilisation of the p21 structure bound to PCNA **A)** p21 (green, 1AXC) bound to PCNA (grey, 1AXC) and key non-conserved residues (light blue) identified as a point for cyclisation. **B)** Overlay of p21 (green, 1AXC) and p21-lactam mimetic (pink, 6CBI) bound to PCNA (grey, 1AXC) shows the structure of the binding mode is preserved in the mimetic. The covalent constraint is shown in light pink. **C**) p21-lactam mimetic (pink, 6CBI) bound to PCNA (grey, 6CBI). **D)** Overlay of p21-lactam mimetic illustrates the key 3$_{10}$-helical structure is maintained as in p21 peptide (green, 1AXC).

The introduction of a covalent lactam constraint between residues 146^{P2} and 149^{P6} of a p21-derived peptide has been shown to preorganise it into this binding conformation (Figure 4D).[32] Two such peptidomimetics were synthesised and NMR structure calculations were conducted to reveal increased 3$_{10}$-helical structure in both cases compared to the unmodified sequence. These peptides bound to PCNA with the same conformation as the native p21 peptide (Figure 4.) Interestingly, the less helical peptide bound PCNA displayed higher affinity (601 nM) than the more helical analogue (1.61 μM). This may reflect sub-optimal interactions between the constraint and PCNA. Despite a slight decrease in affinity relative to the unconstrained peptide, these peptidomimetics still bind tighter than the large majority of known PIP-box partners (Table 1). Furthermore, high binding affinity of these peptidomimetics demonstrates tolerance of the PIP-box to amino-acid substitution, assuming these substitutions lead to structural stabilisation of the binding motif.

A covalent peptide constraint can be chemically altered to optimise interactions with the protein surface and also aid in extending the peptide half-life by physically inhibiting protease degradation.[59, 60] However, these properties are yet to be explored in a p21-derived peptidomimetic. The large size of peptide-based drugs can limit cellular uptake, though with careful design this can be mitigated. A number of p21 peptides conjugated to cell-penetrating peptides have been developed.[44, 61, 62] p21-based peptides have also been successfully targeted to the nucleus where they are required to effect their function.[61, 62] Currently no

PIP-box inhibitors have reached clinical trial status but many other therapeutic peptides are in trials or on the market. This confirms inherent disadvantages of peptides can be harnessed or overcome.[54, 56, 63, 64] Furthermore, with the continual emergence of new techniques and strategies, designing viable peptide-based therapeutics has not only become possible but an efficient approach.[53, 54, 65, 66]

4.8 Summary and outlook

Significant progress has been made on understanding the underlying mechanisms of DNA-replication and -repair, and associated implications for the potential treatment of malignant growths.[10] PCNA is clearly a central player in these processes and an intricate understanding of the structure based mechanisms for PIP-box interaction is progressing rapidly, with the regular emergence of new co-crystal structures and binding studies. Twelve of the twenty co-crystal structures discussed here were published in the past five years, which is more than the previous fifteen years combined. This demonstrates a concerted effort to elucidate mechanisms of DNA regulation due to its importance in wide-spread diseases such as cancer.

A detailed understanding of the structural basis for interaction with PCNA is fundamental to the development of viable therapeutics to selectively inhibit PCNA. Here some of the subtler, but common, interactions made by PIP-box partners both intramolecularly, and intermolecularly with PCNA, have been collated and discussed. This has highlighted the importance of structure-based design, in particular the importance of 3_{10}-helix stabilisation for PIP-box partners. The primary amino-acid sequence of the PIP-box alone does not provide sufficient information to determine binding affinity, (nor do alternate PCNA-binding definitions). Furthermore, a canonical PIP-box sequence does not guarantee higher binding affinity, as particular sequence combinations work cooperatively to overcome PIP-box divergence and enhance PCNA affinity. These subtleties also implicate the ‘x’ amino-acids which separate the defined consensus residues. Although interactions of such ‘x’ residues intramolecularly within the PIP-box or with PCNA have been acknowledged, their true impact and utility may not yet have been fully realised. As more co-crystal structures emerge it would be of interest to present a full review in which the amino-acid character at each PIP-box position is tabulated. Such analysis may reveal a strict set of determinants with which PCNA affinity can be predicted. From this brief analysis, it is evident that the 3_{10}-helix is essential to the binding mode of PIP-box peptides and enhanced stabilisation of this motif strongly correlates with binding affinity.

Disparate works to improve PCNA affinity, cell permeability, structural and biological stability of p21 and other PIP-box peptides have been carried out, but the next essential steps require the unity of these studies if a viable clinical therapeutic is to be achieved.

4.9 Acknowledgements

The work was supported by the ARC Centre of Excellence in Nanoscale BioPhotonics (CNBP) (CE140100003).

REFERENCES

1. K. Miyachi, M. J. Fritzler and E. M. Tan, Autoantibody to a Nuclear Antigen in Proliferating Cells, *J. Immunol.*, 1978, **121**, 2228-2234.

2. J. M. Gulbis, Z. Kelman, J. Hurwitz, M. O'Donnell and J. Kuriyan, Structure of the C-Terminal Region of p21 WAF1/CIP1 Complexed with Human PCNA, *Cell*, 1996, **87**, 297-306.

3. X. Lu, C. K. Tan, J. Q. Zhou, M. You, L. M. Carastro, K. M. Downey and A. G. So, Direct interaction of proliferating cell nuclear antigen with the small subunit of DNA polymerase delta, *J. Biol. Chem.*, 2002, **277**, 24340-24345.

4. T. S. R. Krishna, X.-P. Kong, S. Gary, P. M. Burgers and J. Kuriyan, Crystal Structure of the Eukaryotic DNA Polymerase Processtivity Factor PCNA, *Cell*, 1994, **79**, 1233-1243.

5. K. N. Choe and G.-L. Moldovan, Forging Ahead through Darkness: PCNA, Still the Principal Conductor at the Replication Fork, *Mol. Cell*, 2016, **65**, 380-392.

6. A. De Biasio and F. J. Blanco, Proliferating cell nuclear antigen structure and interactions: too many partners for one dancer?, *Adv. Prot. Chem. Struct. Bio.*, 2013, **91**, 1-36.

7. E. M. Boehm, M. S. Gildenberg and M. T. Washington, The Many Roles of PCNA in Eukaryotic DNA Replication, *Enzymes*, 2016, **39**, 231-254.

8. K. M. Gilljam, E. Feyzi, P. A. Aas, M. M. Sousa, R. Muller, C. B. Vågbø, T. C. Catterall, N. B. Liabakk, G. Slupphaug, F. Drabløs, H. E. Krokan and M. Otterlei, Identification of a novel, widespread, and functionally important PCNA-binding motif, *J. Cell Biol.*, 2009, **186**, 645-654.

9. T. Abbas and A. Dutta, p21 in cancer: Intricate networks and multiple activities, *Nat. Rev. Cancer*, 2009, **9**, 400-414.

10. I. Stoimenov and T. Helleday, PCNA on the crossroad of cancer, *Biochem. Soc. Trans.*, 2009, **37**, 605-613.

11. H. Xu, P. Zhang, L. Liu and M. Y. W. T. Lee, A Novel PCNA-Binding Motif Identified by the Panning of a Random Peptide Display Library, *Biochemistry*, 2001, **40**, 4512-4520.

12. E. M. Boehm and M. T. Washington, R.I.P. to the PIP: PCNA-binding motif no longer considered specific, *BioEssays*, 2016, **38**, 1117-1122.

13. G. Kontopidis, S.-Y. Wu, D. I. Zheleva, P. Taylor, C. McInnes, D. P. Lane, P. M. Fischer and M. D. Walkinshaw, Structural and biochemical studies of human proliferating cell nuclear antigen complexes provide a rationale for cyclin association and inhibitor design, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 1871-1876.

14. M. De March, N. Merino, S. Barrera-Vilarmau, R. Crehuet, S. Onesti, F. J. Blanco and A. De Biasio, Structural basis of human PCNA sliding on DNA, *Nat. Comm.*, 2017, **8**, 13935.

15. A. De Biasio, A. I. de Opakua, G. B. Mortuza, R. Molina, T. N. Cordeiro, F. Castillo, M. Villate, N. Merino, S. Delgado, D. Gil-Cartón, I. Luque, T. Diercks, P. Bernadó, G. Montoya and F. J. Blanco, Structure of p15(PAF)-PCNA complex and implications for clamp sliding during DNA replication and repair, *Nat. Comm.*, 2015, **6**, 6439.

16. L. M. Podust, V. N. Podust, J. M. Sogo and U. Hübscher, Mammalian DNA Polymerase Auxillary Proteins: Analysis of Replication Factor C-Catalyzed Proliferating Cell Nuclear Antigen Loading onto Circular Double-Stranded DNA, *Mol. Cell. Biol.*, 1995, **15**, 3072-3081.

17. M. De March, S. Barrera-Vilarmau, E. Crespan, E. Mentegari, N. Merino, A. Gonzalez-Magana, M. Romano-Moreno, G. Maga, R. Crehuet, S. Onesti, F. J. Blanco and A. De Biasio, p15PAF binding to PCNA modulates the DNA sliding surface, *Nucleic Acids Res.*, 2018, **46**, 9816–9828.

18. E. Johansson, P. Garg and P. M. Burgers, The Pol32 subunit of DNA polymerase delta contains separable domains for processive replication and proliferating cell nuclear antigen (PCNA) binding, *J. Biol. Chem.*, 2004, **279**, 1907-1915.

19. O. Chilkova, P. Stenlund, I. Isoz, C. M. Stith, P. Grabowski, E. B. Lundström, P. M. Burgers and E. Johansson, The eukaryotic leading and lagging strand DNA polymerases are loaded onto primer-ends via separate mechanisms but have comparable processivity in the presence of PCNA, *Nucleic Acids Res.*, 2007, **35**, 6588-6597.

20. A. Hayashi, N. N. Giakoumakis, T. Heidebrecht, T. Ishii, A. Panagopoulos, C. Caillat, M. Takahara, R. G. Hibbert, N. Suenaga, M. Stadnik-Spiewak, T. Takahashi, Y. Shiomi, S. Taraviras, E. von Castelmur, Z. Lygerou, A. Perrakis and H. Nishitani, Direct binding of Cdt2 to PCNA is important for targeting the CRL4Cdt2 E3 ligase activity to Cdt1, *Life Sci. Alliance*, 2018, **1**, e201800238.

21. G. L. Moldovan, B. Pfander and S. Jentsch, PCNA, the maestro of the replication fork, *Cell*, 2007, **129**, 665-679.

22. Z. Zhang, S. Zhang, S. H. Lin, X. Wang, L. Wu, E. Y. Lee and M. Y. Lee, Structure of monoubiquitinated PCNA: implications for DNA polymerase switching and Okazaki fragment maturation, *Cell Cycle*, 2012, **11**, 2128-2136.

23. S. Y. Park, M. S. Jeong, C. W. Han, H. S. Yu and S. B. Jang, Structural and Functional Insight into Proliferating Cell Nuclear Antigen, *J. Microbiol. Biotechnol.*, 2016, **26**, 637-647.

24. L. M. Dieckman, B. D. Freudenthal and M. T. Washington, PCNA structure and function: insights from structures of PCNA complexes and post-translationally modified PCNA, *Subcell. Biochem.*, 2012, **62**, 281-299.

25. Y. Masuda, R. Kanao, K. Kaji, H. Ohmori, F. Hanaoka and C. Masutani, Different types of interaction between PCNA and PIP boxes contribute to distinct cellular functions of Y-family DNA polymerases, *Nucleic Acids Res.*, 2015, **43**, 7898-7910.

26. S. N. Naryzhny and H. Lee, Characterization of proliferating cell nuclear antigen (PCNA) isoforms in normal and cancer cells: There is no cancer-associated form of PCNA, *FEBS Lett.*, 2007, **581**, 4917-4920.

27. A. J. Kroker and J. B. Bruning, p21 Exploits Residue Tyr151 as a Tether for High-Affinity PCNA Binding, *Biochemistry*, 2015, **54**, 3483-3493.

28. S. Sakurai, K. Kitano, H. Yamaguchi, K. Hamada, K. Okada, K. Fukuda, M. Uchida, E. Ohtsuka, H. Morioka and T. Hakoshima, Structural basis for recruitment of human flap endonuclease 1 to PCNA, *EMBO J.*, 2005, **24**, 683-693.

29. S. Hoffmann, S. Smedegaard, K. Nakamura, G. B. Mortuza, M. Räschle, A. Ibañez de Opakua, Y. Oka, Y. Feng, F. J. Blanco, M. Mann, G. Montoya, A. Groth, S. Bekker-Jensen and N. Mailand, TRAIP is a PCNA-binding ubiquitin ligase that protects genome stability after replication stress, *J. Cell Biol.*, 2016, **212**, 63-75.

30. A. Hishiki, H. Hashimoto, T. Hanafusa, K. Kamei, E. Ohashi, T. Shimizu, H. Ohmori and M. Sato, Structural basis for novel interactions between human translesion synthesis polymerases and proliferating cell nuclear antigen, *J. Biol. Chem.*, 2009, **284**, 10552-10560.

31. D. Bubeck, M. A. Reijns, S. C. Graham, K. R. Astell, E. Y. Jones and A. P. Jackson, PCNA directs type 2 RNase H activity on DNA replication and repair substrates, *Nucleic Acids Res.*, 2011, **39**, 3652-3666.

32. K. L. Wegener, A. E. McGrath, N. E. Dixon, A. J. Oakley, D. B. Scanlon, A. D. Abell and J. Bruning, Rational design of a 310-helical PIP-box mimetic targeting PCNA - the human sliding clamp, *Chem. Eur. J.*, 2018, **24**, 11325-11331.

33. A. Gonzalez-Magaña, A. Ibáñez de Opakua, M. Romano-Moreno, J. Murciano-Calles, N. Merino, I. Luque, A. L. Rojas, S. Onesti, F. J. Blanco and A. De Biasio, The p12 subunit of human polymerase δ uses an atypical PIP-box for molecular recognition of proliferating cell nuclear antigen (PCNA), *J. Biol. Chem.*, 2019, DOI: 10.1074/jbc.RA118.006391.

34. E. Warbrick, PCNA binding through a conserved motif, *BioEssays*, 1998, **20**, 195-199.

35. K. Hara, M. Uchida, R. Tagata, H. Yokoyama, Y. Ishikawa, A. Hishiki and H. Hashimoto, Structure of proliferating cell nuclear antigen (PCNA) bound to an APIM peptide reveals the universality of PCNA interaction, *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.*, 2018, **74**, 214-221.

36. J. B. Bruning and Y. Shamoo, Structural and thermodynamic analysis of human PCNA with peptides derived from DNA polymerase-delta p66 subunit and flap endonuclease-1, *Structure*, 2004, **12**, 2209-2219.

37. W. Chen, M. Wu, T. Hang, C. Wang, X. Zhang and J. Zang, Structure insights into the molecular mechanism of the interaction between UHRF2 and PCNA, *Biochem. Biophys. Res. Commun.*, 2017, **494**, 575-580.

38. T. Kaufmann, I. Grishkovskaya, A. A. Polyansky, S. Kostrhon, E. Kukolj, K. M. Olek, S. Herbert, E. Beltzung, K. Mechtler, T. Peterbauer, J. Gotzmann, L. Zhang, M. Hartl, B. Zagrovic, K. Elsayad, K. Djinovic-Carugo and D. Slade, A novel non-canonical PIP-box mediates PARG interaction with PCNA, *Nucleic Acids Res.*, 2017, **45**, 9741-9759.

39. C. G. Havens, N. Shobnam, E. Guarino, R. C. Centore, L. Zou, S. E. Kearsey and J. C. Walter, Direct role for proliferating cell nuclear antigen in substrate recognition by the E3 ubiquitin ligase CRL4Cdt2, *J. Biol. Chem.*, 2012, **287**, 11410-11421.

40. Y. Wang, M. Xu and T. Jiang, Crystal structure of human PCNA in complex with the PIP box of DVC1, *Biochem. Biophys. Res. Commun.*, 2016, **474**, 264-270.

41. E. Warbrick, D. P. Lane, D. M. Glover and L. S. Cox, A small peptide inhibitor of DNA replication defines the site of interaction between the cyclin-dependent kinase inhibitor p21 WAF1 and proliferating cell nuclear antigen, *Curr. Biol.*, 1995, **5**, 275-282.

42. M. Sebesta, C. D. O. Cooper, A. Ariza, C. J. Carnie and D. Ahel, Structural insights into the function of ZRANB3 in replication stress response, *Nat. Comm.*, 2017, **8**, 15847.

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

44. I. Massodi, G. L. Bidwell and D. Raucher, Evaluation of cell penetrating peptides fused to elastin-like polypeptide for drug delivery, *J. Controlled Release*, 2005, **108**, 396-408.

45. D. P. Lane, L. S. Cox, E. Warbrick and D. M. Glover, WO 96/14334, *Identifcation of the p21/WAF1-PCNA Interaction Site and Therapeutic Applications Thereof,* 1996.

46. D. Zheleva, P. Fischer, C. McInnes, M. Andrews, W. Chan and G. Atkinson, WO 2005/040802 A2, *Peptides,* 2005.

47. G. Kontopidis, D. Zheleva, C. McInnes, P. Fischer and M. Walkinshaw, WO 2005/108421 A1, *Crystal Structure of Human Proliferating Cell Nuclear Antigen (PCNA) and Uses Thereof,* 2005.

48. D. Zhongyun, M. Wortman, Z. Tan and K. Dillehay, WO 2012/033938 A2, *Identification of PCNA Targeting Compounds for Cancer Therapy and PCNA Function Regulation,* 2012.

49. V. J. Davisson and M. D. Bartolowitis, WO 2017/100154 A1, *Inhibitors for Proliferating Cell Nuclear Antigen and Uses,* 2017.

50. C. Punchihewa, A. Inoue, A. Hishiki, Y. Fujikawa, M. Connelly, B. Evison, Y. Shao, R. Heath, I. Kuraoka, P. Rodrigues, H. Hashimoto, M. Kawanishi, M. Sato, T. Yagi and N. Fujii, Identification of small molecule proliferating cell nuclear antigen (PCNA) inhibitor that disrupts interactions with PIP-box proteins and inhibits DNA replication, *J. Biol. Chem.*, 2012, **287**, 14289-14300.

51. M. Actis, A. Inoue, B. Evison, S. Perry, C. Punchihewa and N. Fujii, Small molecule inhibitors of PCNA/PIP-box interaction suppress translesion DNA synthesis, *Bioorg. Med. Chem.*, 2013, **21**, 1972-1977.

52. B. J. Evison, M. L. Actis, S. Z. Wu, Y. Shao, R. J. Heath, L. Yang and N. Fujii, A site-selective, irreversible inhibitor of the DNA replication auxiliary factor proliferating cell nuclear antigen (PCNA), *Bioorg. Med. Chem.*, 2014, **22**, 6333-6343.

53. S. Marqus, E. Pirogova and T. J. Piva, Evaluation of the use of therapeutic peptides for cancer treatment, *J. Biomed. Sci.*, 2017, **24**, 21.

54. C. Morrison, Constrained peptides' time to shine?, *Nat. Rev. Drug Discovery*, 2018, **17**, 531-533.

55. K. Estieu-Gionnet and G. Guichard, Stabilized helical peptides: overview of the technologies and therapeutic promises, *Expert Opin. Drug Dicov.*, 2011, **6**, 937-936.

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

57. T. A. Hill, N. E. Shepherd, F. Diness and D. P. Fairlie, Constraining cyclic peptides to mimic protein structure motifs, *Angew. Chem. Int. Ed. (English)*, 2014, **53**, 13020-13041.

58. S. R. Perry, T. A. Hill, A. D. de Araujo, H. N. Hoang and D. P. Fairlie, Contiguous hydrophobic and charged surface patches in short helix-constrained peptides drive cell permeability, *Org. Biomol. Chem.*, 2018, **16**, 367-371.

59. L. D. Walensky and G. H. Bird, Hydrocarbon-stapled peptides: principles, practice, and progress, *J. Med. Chem.*, 2014, **57**, 6275-6288.

60. M. Klein, Stabilized helical peptides: overview of the technologies and its impact on drug discovery, *Expert Opin. Drug Dicov.*, 2017, **12**, 1117-1125.

61. J. S. Ryu and D. Raucher, Anti-tumor efficacy of a therapeutic peptide based on thermo-responsive elastin-like polypeptide in combination with gamcitabine, *Cancer Letters*, 2014, **384**, 177-184.

62. Y. Luo, J. Hurwitz and J. Massague, Cell-cycle inhibition by independent CDK and PCNA binding domains in p21Cip1, *Nature*, 1995, **375**, 159-161.

63. N. Tsomaia, Peptide therapeutics: Targeting the undruggable space, *Eur. J. Med. Chem.*, 2015, **94**, 459-470.

64. K. Fosgerau and T. Hoffmann, Peptide therapeutics: current status and future directions, *Drug Discov. Today*, 2015, **20**, 122-128.

65. A. M. Ali, J. Atmaj, N. Van Oosterwijk, M. R. Groves and A. Dömling, Stapled Peptides Inhibitors: A New Window for Target Drug Discovery, *Comput. Struct. Biotechnol. J.*, 2019, **17**, 263-281.

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

Chapter 5.

UNLOCKING THE PIP-BOX: A PEPTIDE LIBRARY REVEALS INTERACTIONS THAT DRIVE HIGH-AFFINITY BINDING TO HUMAN PCNA

Research Article: Journal of Biological Chemistry, **2021**, 296 100773

Unlocking the PIP-box: A peptide library reveals interactions that drive high-affinity binding to human PCNA

Aimee J. Horsfall[1], Beth A. Vandborg[2], Wioleta Kowalczyk[3], Theresa Chav[1], Denis B. Scanlon[1], Andrew D. Abell[1*] and John B. Bruning[2*]

[1] ARC Centre of Excellence for Nanoscale BioPhotonics, Institute of Photonics and Advanced Sensing, School of Physical Sciences, The University of Adelaide, Adelaide, South Australia, 5005, Australia

[2] Institute of Photonics and Advanced Sensing, School of Biological Sciences, The University of Adelaide, Adelaide, South Australia, 5005, Australia

[3] CSIRO Manufacturing, Clayton South, Victoria, 3169, Australia

*corresponding author

© 2021 Journal of Biological Chemistry, ASBMB Publications

First published online: May 2021

ABSTRACT

The human sliding clamp (Proliferating Cell Nuclear Antigen, PCNA) interacts with over 200 proteins through a conserved binding motif, the PIP-box, to orchestrate DNA-replication and repair. Here we present a systematic study of each position within the PIP-box to reveal how PCNA-interacting peptides bind with drastically varied affinities, and begin to understand how affinity for PCNA is fine-tuned. A series of 27 peptides with sequence modifications, inspired by native PIP-box sequences; and another series of 19 peptides containing native PCNA-binding motifs, were synthesised based on a p21-derived peptide scaffold to enable direct comparison. The PCNA-binding affinity of all peptides was characterised by surface plasmon resonance (SPR) where K_D values spanned a 4000-fold range, from 1.83 nM to 7.59 μM. The PCNA-bound peptide structures were studied by X-ray crystallography and computational modelling to reveal inter- and intramolecular interaction networks that correlate with high PCNA affinity. These data informed rational-design of three new PIP-box sequences, that revealed the highest affinity PCNA binding partner to date, at 1.12 nM, with PIP-box QTRITEYF. This work showcases the sequence nuances within the PIP-box that are responsible for high-affinity PCNA binding, which in turn underpins our understanding of how nature tunes PCNA-affinity to regulate DNA-replication and -repair processes. Additionally, these insights will be useful to future design of PCNA-inhibitors.

STATEMENT OF AUTHORSHIP

Title of Paper	Unlocking the PIP-box: Understanding factors for high affinity binding to human PCNA
Publication Status	☐ Published ☐ Accepted for Publication ☒ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	**Research Article:** A. J. Horsfall, B. A. Vandborg, W. Kowalczyk, T. Chav, D. B. Scanlon, A. D. Abell and J. B. Bruning, *Journal of Biological Chemistry,* 2021

Principal Author

Name of Principal Author (Candidate)	Aimee J Horsfall		
Contribution to the Paper	Designed and synthesised all peptides, SPR assays, analysed structure results, discussed results, wrote and edited manuscript		
Overall percentage (%)	60%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	15/02/2021

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Beth A Vandborg		
Contribution to the Paper	Protein synthesis & purification, protein crystallography, structure analysis, discussed results and edited manuscript		
Signature		Date	01/03/2020

Name of Co-Author	Wioleta Kowalczyk		
Contribution to the Paper	Peptide synthesis, discussed results and edited manuscript		
Signature		Date	01/03/2021

Name of Co-Author	Theresa Chav		
Contribution to the Paper	SPR assays, discussed results and edited manuscript		
Signature		Date	

Name of Co-Author	Denis B Scanlon		
Contribution to the Paper	Peptide synthesis, discussed results, and edited manuscript		
Signature	D Scanlon	Date	01/03/2021

Name of Co-Author	Andrew D Abell		
Contribution to the Paper	Supervised AJH, discussed results, and edited manuscript.		
Signature		Date	15/2/2021

Name of Co-Author	John B Bruning		
Contribution to the Paper	Supervised AJH, discussed results, and edited manuscript.		
Signature		Date	18-2-21

5.1 Introduction

Human Proliferating Cell Nuclear Antigen (PCNA) is a member of the sliding clamp family of proteins and acts as an essential processivity factor and mediator of DNA-replication and repair.[1-5] It is upregulated in the majority of cancers in order to cope with the increased demand for DNA-replication. A PCNA knockout is lethal,[6, 7] which reflects its importance in cell cycle progression. PCNA is a toroidal-shaped homotrimer with six-fold pseudosymmetry that forms by association of three 27 kDa subunits, with each subunit containing two, nearly symmetrical, domains connected by an unstructured loop termed the interdomain connecting loop (IDCL, Figure 1A and B). The ring-shaped sliding clamp is loaded onto primer-template junctions of DNA and encircles the double-strand, sliding with the progressing replication fork, to act as a moving docking-platform for enzymes to bind and maintain proximity to the DNA.[4, 5, 8, 9]

More than 200 proteins are known to interact with PCNA during DNA-replication, DNA-repair and cell cycle regulation. However, factors that control regulation and recruitment of an appropriate sliding clamp binding partner, at the correct time and location, are not well understood. We and others have suggested that a large portion of this control arises from differences in the binding affinity of proteins for PCNA, which span four orders of magnitude.[3, 10-12] Such proteins compete to bind PCNA to gain access to the replication fork, where successful binding is ultimately dictated by PCNA affinity. The cell-cycle regulator protein, p21$^{CIP/WAF1}$ (referred to herein as p21), is the highest affinity PCNA-interacting protein known (K_D 2.5-90 nM)[13-15] and upon binding PCNA, shuts down replication. This consequently stalls cell cycle progression to provide a necessary checkpoint for healthy proliferation. p21 requires high affinity for PCNA to fulfil this role, and outcompete other proteins from interacting with PCNA and hence the replication fork. In contrast, the Y-family translesion polymerases pol λ, pol κ, and pol ι bind PCNA with micromolar affinity,[14, 16] and a peptide derived from the major processive polymerase (pol $\delta_{p66\ 452\text{-}466}$) has an affinity for PCNA of 15.6 μM.[14] A fundamental understanding of the of the molecular level nuances that dictate PCNA binding affinity is required to understand the regulation of DNA-replication and repair processes. Here we begin to unpack subtle sequence changes that influence PCNA-binding affinity with an investigation of the structure activity relationship of peptides that bind PCNA.

The majority of PCNA partners bind on the surface of a PCNA subunit located between the two subunit domains, nestled under the IDCL (Figure 1C). PCNA-interacting proteins (or peptides) interact with PCNA through a consensus sequence aptly named the PCNA-Interacting Protein (PIP) box motif. The PIP-box motif is defined as Qxxφxxψψ, where Q is glutamine, 'x' is any amino-acid, 'φ' is a hydrophobic residue and 'ψ' is an aromatic residue (commonly Phe or Tyr). Each of the eight positions of the PIP-box are referred to here as position 1 through 8, or indicated by a superscript P1 through to P8. It has been argued that the PIP-box definition is part of a larger subset of PCNA-binding motifs as many non-canonical PIP-box peptides, which are defined by the absence of a glutamine at position 1, bind PCNA with low micromolar or nanomolar affinity.[11, 17] The non-canonical definition will be used here for simplicity of discussion. The highly canonical PIP-box motif of p21, ^{144}QTSMTDFY, was initially defined following an alanine mutation scan from Lys141 to Ser160, and revealed Gln144^{P1}, Met147^{P4} and Phe150^{P7} and Tyr151^{P8} as the

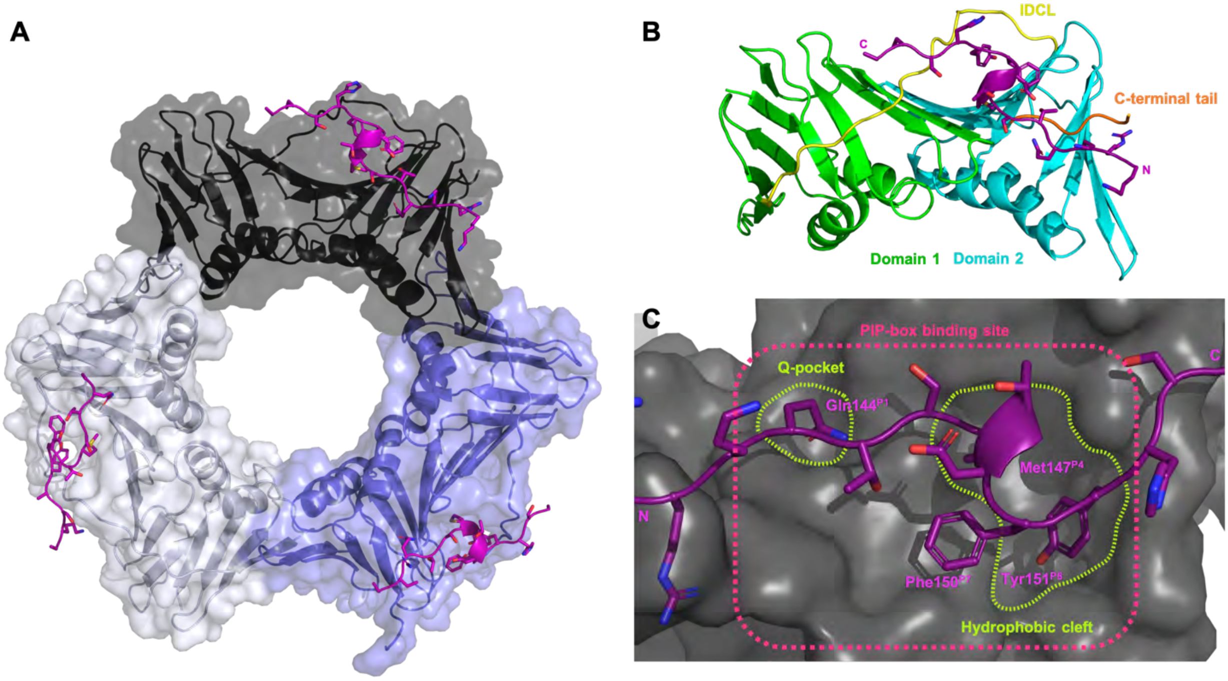


Figure 1. Co-crystal structure of **p21**$_{\mu}$ (PDB: 7KQ1) bound to PCNA solved by X-ray crystallography. **p21**$_{\mu}$ peptide (purple) shown in cartoon form with side-chains as sticks, bound to each PCNA subunit. **A** PCNA shown in cartoon form with transparent surface representation with each subunit indicated in white, grey and pale purple. **B** A PCNA subunit shown in cartoon ribbons, where the two domains are indicated in green and blue, the IDCL in yellow and the *C*-terminal tail of PCNA in orange. **C** A zoom in of a **p21**$_{\mu}$ peptide (purple) bound to PCNA (grey, surface) where the PIP-box binding site is indicated in pink, the key binding pockets are shown in green, and the conserved PIP-box residues bound are labelled in purple.

conserved binding residues (positions 1, 4, 7 and 8).[18, 19] A 22-mer peptide, derived from p21 (residues 139-160) which contains this PIP-box sequence, binds to PCNA as a single well-defined 3_{10}-helical turn that is anchored to the PCNA surface by a hydrophobic triad (Met, Phe, Tyr). Despite these insights, the scope and tolerance at each position within the PIP-box is not well understood. A number of point mutations in p21, and modifications in p21-derived peptides have been investigated, however such modifications have been confined to alanine mutations or simple functional modifications.[18-22] For example, a Tyr151Phe^{P8} modified peptide demonstrated the importance of the tyrosine phenol group that makes a hydrogen bond to Gln131 on PCNA.[22] The non-conserved PIP-box residues (positions 2, 3, 5 and 6) of **p21**$_{\mathbf{139\text{-}160}}$ only participate in three 3_{10}-helical stabilising hydrogen bonds (146-149, 147-150 and 147-151),[10, 15] which has led to the belief that these residues play a limited role in the binding affinity for PCNA. A thorough investigation is required to develop a nuanced understanding of the key interactions to which the p21 PIP-box sequence owes its high affinity, in order to begin to understand how affinity for PCNA plays a role in the regulation of DNA-replication and -repair processes.

Here, a series of short p21-derived peptides is reported, with either one- or two-point modifications made in the PIP-box sequence to study the structure-activity relationship of PCNA-binding peptides, and determine the secondary interactions that fine-tune affinity for PCNA. The amino-acids incorporated at each PIP-box position were chosen to reflect those observed at the respective position in native protein PIP-box sequences. A second series of peptides was prepared with native PIP-box sequences from different PCNA-binding proteins, inserted within the sequence that flanks the p21 PIP-box sequence. This series of peptides allowed a wider array of sequence combinations to be investigated compared to point modifications, and allows direct comparison of the effect of the PIP-box sequence on PCNA-affinity as the peptides contain the same sequence flanking the PIP-box. The PCNA binding affinity of each peptide was determined by a surface plasmon resonance (SPR). The changes in structure of the peptides bound to PCNA, were studied by X-ray crystallography and computational modelling studies, and correlated to the PCNA-binding affinity to uncover interactions responsible for the differences in affinity. These studies inform on the design of three new PIP-box sequences to investigate whether cooperative interactions could be predicted from the interactions observed in the high PCNA affinity modified p21 peptides. This comprehensive and systematic structure-activity investigation advances the ability to predict PCNA binding affinity from PIP-box sequence. Additionally, it provides insight into how nature has fine-tuned affinity of PIP-box sequences for PCNA, which can be leveraged to further investigate how PCNA regulates the DNA-replication and -repair process, or develop inhibitors of these interactions.

5.2 Results

Fifty-one peptides were synthesised by solid-phase peptide synthesis (SPPS, see Methods) to allow systematic study of the structure-activity relationship of peptides that bind PCNA (Table S1). These peptides comprise two control p21 peptides known to bind PCNA; a series of 27 point-modified p21 peptides; a second series of 19 peptides, where the PIP-box sequence from a number of PCNA-binding proteins was included between the sequence that flanks the p21 PIP-box; and a final set of three peptides containing entirely new, rationally designed PIP-box sequences. A short p21 sequence 141-155 (referred to herein as **p21$_{\mu}$**) was chosen as it has been reported the p21$_{139\text{-}160}$ sequence can be shortened without drastically impacting affinity.[13] Interestingly, the preparation of this shorter sequence gave less aspartimide formation compared to the longer 22mer, with a M-18 peak in the mass spectrum (MS) not apparent for p21$_{\mu}$, in contrast to **p21$_{139\text{-}160}$** (unpublished work), thereby improving synthetic yields. All subsequent peptides were based on this shorter **p21$_{\mu}$** peptide (residues 141-155).

Five p21$_{\mu}$ peptides with a 144^{P1} point mutation were prepared: **p21$_{\mu}$-Q144K**, **p21$_{\mu}$-Q144M**, **p21$_{\mu}$-Q144D**, **p21$_{\mu}$-Q144S**, and **p21$_{\mu}$-Q144N**. Four of these were inspired by the following non-canonical P1 residues: lysine (pol ι, pol κ), methionine (Cdt2, pol η and RNaseH2B), aspartic acid (poly(ADP-ribose) glycohydrolase, PARG) or serine (pol λ) at PIP-box position 1. In addition, asparagine with the amidated, but shorter side-chain than glutamine was included (**p21$_{\mu}$-Q144N**). Five p21$_{\mu}$ peptides, with a single amino-acid modification at the conserved position 4 of the PIP-box were prepared, and include valine (**p21$_{\mu}$-M147V**), isoleucine (**p21$_{\mu}$-M147I**), or leucine (**p21$_{\mu}$-M147L**), which are commonly observed in native PIP-box sequences; as well as the smaller but still hydrophobic alanine (**p21$_{\mu}$-M147A**), and the non-polar aromatic tryptophan (**p21$_{\mu}$-M147W**). Aromatic residues, such as phenylalanine and tyrosine, are commonly observed at conserved positions 7 and 8 of the PIP-box. **p21$_{\mu}$** contains the Phe^{P7}Tyr^{P8} combination and consequently three peptides, each with a different permutation of phenylalanine and tyrosine, were prepared (**p21$_{\mu}$-Y151F**, **p21$_{\mu}$-F150Y** and **p21$_{\mu}$-FY150/151YF**). Additionally, a **p21$_{\mu}$-F150H** peptide was prepared, inspired by the PARG PIP-box. Positively-charged residues are commonly observed in native

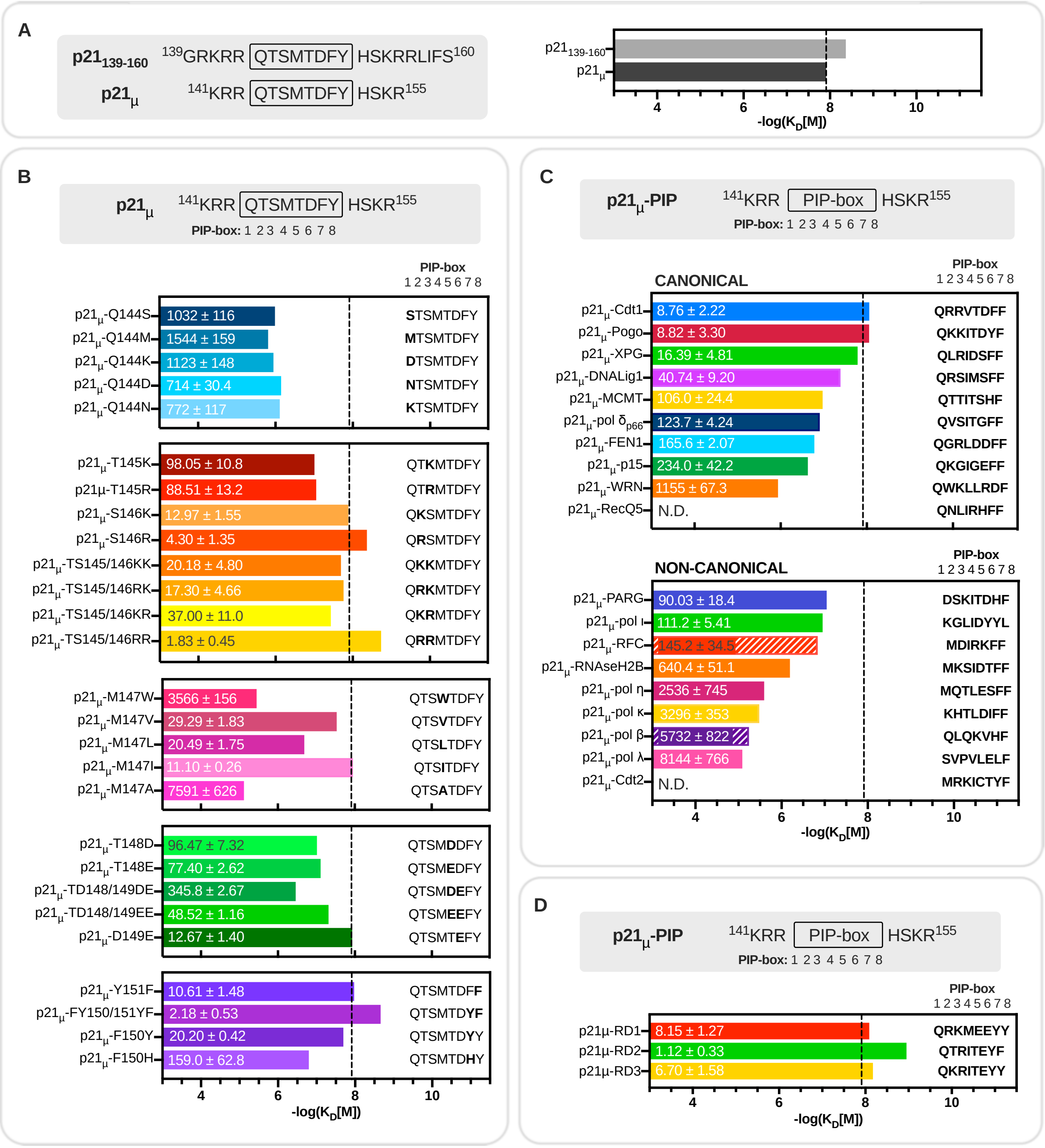


Figure 2. Binding affinity (K_D) of p21 peptides to PCNA determined by SPR. K_D values were calculated from a single experiment using the in-built Biacore Evaluation S200 software and are shown in nM over the respective bar with the corresponding standard error. The bars graphically represent the K_D values (in molar, M) as -log(K_D[M]). All experiments were repeated to ensure reproducibility. The dashed line on all panels represents the affinity of $p21_\mu$ (12.3 nM, -log(K_D[M]) = 7.91) to which all modified $p21_\mu$ peptides are compared throughout the discussion. **A** Binding affinity of the native p21 peptides. **B** Rationally mutated p21 peptides with single- or double-point modifications are introduced into the p21 PIP-box. **C** Binding affinity of peptides containing a native canonical PIP-box from an alternate PCNA binding partner (top) or native non-canonical PIP-box from an alternate PCNA binding partner (bottom) including two PIP-box sequences which only contain seven amino-acids (indicated by the striped bars). The PIP-box sequences are flanked by the same sequence that flanks the p21 PIP-box **D** Binding affinity of the rationally designed PIP-box sequences.

PIP-box sequences at non-conserved positions 2 and 3 (e.g. XPG, FEN1, WRN, PARG, p15, Cdt1). Consequently, eight $p21_\mu$ peptides with arginine or lysine residues at PIP-box position 2 and/or 3 were prepared (**$p21_\mu$-T145K**, **$p21_\mu$-T145R**, **$p21_\mu$-S146K**, **$p21_\mu$-S146R**, **$p21_\mu$-TS145/146KK**, **$p21_\mu$-TS145/146KR**, **$p21_\mu$-TS145/146RK**, and **$p21_\mu$-TS145/146RR**). Conversely, negatively-charged residues are commonly observed at non-conserved positions 5 and 6 of native PIP-box sequences (e.g. pol ι, pol λ, pol κ, p15, Cdt1, XPG, RNaseH2B) and five peptides containing aspartic and/or glutamic acid were prepared (**p21$_{\mu}$-T148D**, **p21$_{\mu}$-T148E**, **p21$_{\mu}$-TD148/149DE**, **p21$_{\mu}$-TD148/149EE**, **p21$_{\mu}$-D149E**). Lastly, in order to investigate a variety of PIP-box sequence combinations, a variety of native PIP-box sequences replaced the p21 PIP-box in a p21$_{\mu}$ peptide. Ten canonical PIP-box sequences from PCNA-binding proteins Cdt1, Pogo, XPG, DNALig1, MCMT, pol δ$_{p66}$, FEN1, p15, WRN and RecQ5 were chosen, along with eight peptides with non-canonical human PIP-box sequences from pol ι, pol κ, Cdt2, pol η, RNaseH2B, PARG, pol λ and RFC$_{p14}$. The PIP-box motif in many of these proteins is located at the protein terminus and does not have a sequence extending *C*- or *N*-terminal from the PIP-box (e.g. pol δ$_{p66}$). Consequently, our design allows direct comparison of PIP-box sequence influence on PCNA affinity and is not influenced by the native flanking sequence.

The binding affinity (K_D) of each peptide for PCNA was determined by SPR and the results are summarised in Figure 2 and Table S2. A representative sample of SPR sensorgrams are included in Figure S1. **p21$_{139-160}$**, prepared as a positive control, bound to PCNA with a K_D value of 4.32 nM (Figure 2A, Table S2), which agrees with previous literature.[13-15] **p21$_{\mu}$** (residues 141-155) bound PCNA with 12.3 nM affinity, which indicates the truncated sequence was well tolerated, even though the peptide is 7 amino-acids shorter than **p21$_{139-160}$** (Figure 2A, Table S2). All modified p21 peptides are compared to **p21$_{\mu}$** throughout the discussion, unless otherwise indicated. The K_D value of **p21$_{\mu}$** is marked as a dashed line on all panels of Figure 2 to provide a benchmark for comparison. The K_D values for the modified p21$_{\mu}$ peptides spanned 4-orders of magnitude, with the best binding peptide of these two series having a K_D value of 1.83 nM (**p21$_{\mu}$-TS145/146RR**) and the lowest affinity was 8.14 μM (**p21$_{\mu}$-pol λ**) (Figure 2B and C, Table S2). The binding affinity of **p21$_{\mu}$-Cdt2** and **p21$_{\mu}$-RecQ5** could not be determined due to non-specific binding to the sensor chip. Interestingly, the analysis revealed seven peptides with higher affinity than **p21$_{\mu}$**: **p21$_{\mu}$-TS145/146RR** (1.83 nM), **p21$_{\mu}$-S146R** (4.30 nM), **p21$_{\mu}$-Cdt1** (8.76 nM), **p21$_{\mu}$-Pogo** (8.82 nM), **p21$_{\mu}$-Y151F** (10.6 nM), and **p21$_{\mu}$-M147I** (11.1 nM) (Figure 2B and C). The affinity of canonical PIP-box p21$_{\mu}$ peptides (with Gln144^{P1}) here range from 1.83 nM to 3.57 μM (Figure 2C 'canonical'). In general, the non-canonical peptides as a group, unsurprisingly, bound PCNA with lower affinity than the canonical p21$_{\mu}$ peptides (see Figure 2C).

The conformation of the peptides bound to PCNA was investigated to uncover potential reasons for differences in affinity. Co-crystal structures of **p21$_{\mu}$** (PDB: 7KQ1, 3.30 Å) and **p21$_{\mu}$-F150Y** bound to PCNA (PDB: 7KQ0, 2.40 Å) were solved. Both served as important controls for our modelling studies: **p21$_{\mu}$** showed that the shorter peptide binds PCNA in a similar manner (position on PCNA surface and secondary structure) as the 22mer **p21$_{139-160}$** peptide (PDB: 1AXC), and **p21$_{\mu}$-F150Y** demonstrated a proof of concept that our docking studies represented the crystal structures accurately. The resulting structures are shown in Figures 3, S1 and S2, and the data collection and refinement statistics are summarised in Table S3. The co-crystal structure of **p21$_{\mu}$** bound to PCNA (Figure 1) revealed 6 intermolecular and 3 intramolecular hydrogen bonds (Figure 3C, Figure S2 and Table S4). The PIP-box residues 144-151 all retain similar conformations to **p21$_{139\text{-}160}$** (PDB: 1AXC, Figure 3A and B), which is represented by an RMSD value of 0.51 Å (Table S18). This is also reflected in a high similarity of the buried surface area (%BSA) of the PIP-box residues of **p21$_{\mu}$** compared to **p21$_{139\text{-}160}$** in 1AXC (Table S19). Gln144^{P1} of **p21$_{\mu}$** makes a 3.3 Å hydrogen bond with Ala252 and Pro253 within the Q-pocket, but not Ala208, in contrast to Gln144^{P1} of p21$_{139\text{-}160}$ that formed 3 hydrogen bonds. The side-chain of Met147^{P4} was packed slightly different to the analogous Met147^{P4} in **p21$_{139\text{-}160}$** (PDB: 1AXC). The *C*-terminal residues of PCNA (Lys254-Asp257) and *N*-terminal residues of **p21$_{\mu}$** (141-143) gave well-defined electron density, in contrast to the p21$_{139\text{-}160}$ structure (PDB: 1AXC), that revealed an electrostatic interaction between Lys141 and Arg142 of **p21$_{\mu}$**, and Glu256 and Asp257 of PCNA (Figure 3C).

Following this, a larger collection of the modified p21$_{\mu}$ peptides were selected for further investigation by computational modelling of the PCNA-bound peptides, to allow high throughput analysis of more structures. The crystal structure of **p21$_{\mu}$** (PDB: 7KQ1) bound to PCNA was used as a starting point for modelling as this peptide is the most similar to the modelled peptides and the structure was overall similar to the **p21$_{139\text{-}160}$** structure bound to PCNA (PDB: 1AXC). The peptides chosen for computational analysis were **p21$_{\mu}$-S146R**, **p21$_{\mu}$-M147I**, **p21$_{\mu}$-D149E**, and **p21$_{\mu}$-FY150/151YF** as these modifications resulted in the largest increase in affinity for the respective PIP-box position. Additionally, a representative group containing native PIP-box sequences was also modelled bound to PCNA: two canonical PIP-box peptides, **p21$_{\mu}$-Pogo** and **p21$_{\mu}$-pol δ$_{p66}$**; two non-canonical PIP-box peptides **p21$_{\mu}$-pol ι** and **p21$_{\mu}$-PARG**, as well as 7 amino-acid PIP-box peptide **p21$_{\mu}$-RFC** (Figures S2-15). The resulting computationally modelled peptide structures

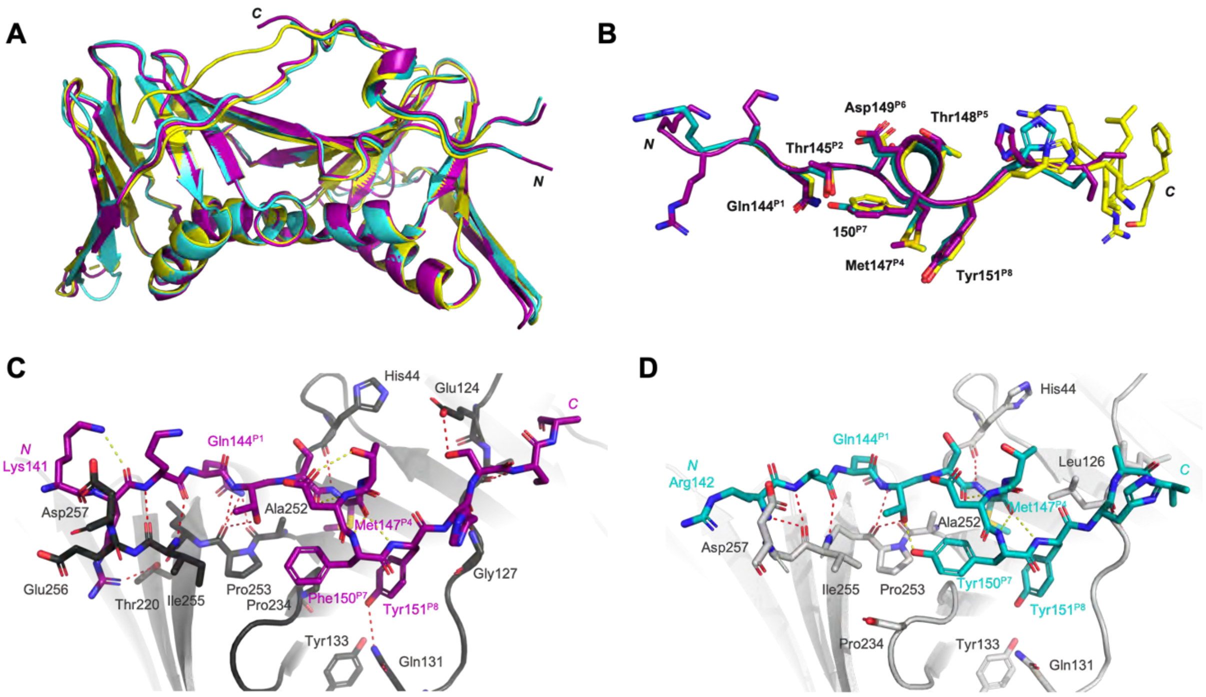

A
C
N
B
Asp149^{P6}
Thr148^{P5}
N
Thr145^{P2}
C
Gln144^{P1}
150^{P7}
Met147^{P4}
Tyr151^{P8}
C
His44
Glu124
C
N
Lys141
Gln144^{P1}
Asp257
Ala252
Met147^{P4}
Glu256
Thr220
Ile255
Pro253
Pro234
Phe150^{P7}
Tyr151^{P8}
Gly127
Tyr133
Gln131
D
His44
Gln144^{P1}
N
Arg142
Leu126
C
Ala252
Met147^{P4}
Asp257
Ile255
Pro253
Tyr150^{P7}
Tyr151^{P8}
Pro234
Tyr133
Gln131

Figure 3. Co-crystal structures of **p21$_{\mu}$** (7KQ1, purple) and **p21$_{\mu}$-F150Y** (7KQ0, blue) bound to PCNA solved by X-ray crystallography, compared to **p21$_{139\text{-}160}$** (1AXC, yellow). **A** Overall binding mode and structure of p21 peptides bound to PCNA is the same. Peptide and PCNA shown are cartoons. **B** Overlay of peptide structures shows the PIP-box residues (labelled) adopt similar orientations. **C** & **D** Intramolecular (yellow dashed lines) and intermolecular (red dashed lines) polar interactions of **p21$_{\mu}$** (purple, **C**) and **p21$_{\mu}$-F150Y** (blue, **D**) shown as sticks bound to PCNA in grey. Conserved residues of PIP-box residues, and *N*-terminal residue, are labelled in the corresponding peptide colour.

bound to PCNA were compared to the co-crystal structure of **p21$_\mu$** bound to PCNA, which revealed the complexed structures were similar overall (Figure 4A), represented by an average RMSD value of 0.223 Å (Table S18). This in turn suggests that the difference in affinity for PCNA is due to subtle structural changes. The conserved residues all adopted similar conformations between all peptide structures (Figure 4B); however, the non-conserved residues (positions 2, 3, 5 and 6) and residues flanking the PIP-box were much more varied (Figure 4C and D). Intramolecular (peptide-protein) and intermolecular (peptide-peptide) interactions for each PCNA-bound peptide structure are summarised in Tables S4, S5-17.

Additionally, the PCNA-bound peptides **p21$_\mu$**, **p21$_\mu$-pol δ$_{p66}$**, **p21$_\mu$-pol ι**, **p21$_\mu$-PARG**, and **p21$_\mu$-Pogo**, were compared to the co-crystal structure reported for the analogous native peptide bound to PCNA, to investigate how the sequence flanking the PIP-box motif influences binding conformation. These structures are **p21$_{139-160}$** (PDB: 1AXC),[15] **pol δ$_{p66\ 452-466}$** (PDB: 1U76),[14] **pol ι$_{415-437}$** (PDB: 2ZVM),[16] **PARG$_{402-420}$** (PDB: 5MAV),[23] and the Pogo PIP-box included in the mutant PogoLigase (**PL**) peptide (PDB: 1VYJ).[24] The PCNA-bound peptide co-crystal structures were overlaid (Figure S16) and the overall peptide conformation (RMSD values in Table S20) and secondary interactions compared (Table S21). These data reveal the PIP-box of peptides that contain the p21, pol δ$_{p66}$, PARG or Pogo PIP-box sequences all adopt similar conformations to the native counterpart on binding PCNA with RMSD values of 0.51, 0.69, 0.49 and 0.69 Å, respectively (Table S20, Figure S16). However, the pol ι PIP-box containing peptides bound to PCNA adopt very different conformations, represented by an RMSD of 1.13 Å (Figure S16 and Table S20).

Lastly, the affinity and structural information for these two series of peptides was harnessed to design three new PIP-box sequences in an attempt test our understanding of the requirements for high affinity PCNA binding. These new PIP-box sequences were designed to mimic favourable secondary interactions observed for native PIP-box sequences or point modified peptides (Figure 5). These sequences were designed with particular primary sequence combinations in mind that correlated with high PCNA affinity, or particular secondary interactions that appeared important to stabilise the binding conformation which are highlighted throughout the discussion. These sequences were then synthesised within the p21$_\mu$ scaffold, as before to give **p21$_\mu$-RD1** – **3**. The binding affinity to PCNA was determined by SPR (Figure 2D), to reveal that all three peptides bind with higher affinity than p21$_\mu$ and helped inform and refine the interaction map shown in Figure 6. The peptides were then computationally modelled bound to the PCNA surface (Figure 7) which indicates that these peptides all bind PCNA in a similar conformation to **p21$_\mu$** (RMSD 0.181-0.270, Table S18), where a 3.0-3.1 Å main-chain hydrogen bond between residues 149^{P6} and 146^{P3} to define the 3$_{10}$-helical binding conformation (Figure 7A). The best performing peptide **p21$_\mu$-RD2** binds PCNA with 1.12 nM affinity, which is remarkably the highest affinity PCNA-binding peptide or protein reported (Figure 2, Table S2).

5.3 Discussion

The binding affinities of the p21$_\mu$ peptides for PCNA, and where appropriate, the structure-based information were compared and contrasted in order to highlight interactions that correlated to changes in PCNA binding affinity. Individual modifications at each PIP-box position were first examined for simplicity, then the cooperative interactions that arise were summarised, and lastly how these observations pertain to the design of the peptides **p21$_{\mu}$-RD1** – **3** are discussed below.

5.3.1 Conserved glutamine position 1: Glutamine enhances PCNA affinity.

Gln^{P1} is known to contribute significantly to PCNA binding affinity, where a reported p21 Gln144Ala^{P1} mutant was not able to effectively inhibit SV40 DNA replication *in vitro.*[18] However, its importance has been disputed due to the prevalence of non-canonical PIP-box sequences.[11, 17] Here, the contribution of Gln^{P1} to PCNA-binding affinity is quantified in comparison to other common position one residues.

Gln144^{P1} is binds in the Q-pocket and the orientation was unchanged between the canonical computationally-modelled or co-crystallised PCNA-bound peptide structures (Figure 4B), with hydrogen bonds to PCNA residues Ala252 and Pro253 within the Q-pocket ranging from 3.0-3.6 Å. The Q-pocket of PCNA went unfilled in the PCNA-bound non-canonical peptide structures (**p21$_{\mu}$-PARG**, **p21$_{\mu}$-pol ι**, **p21$_{\mu}$-RFC**), where the position one residue extended over the Q-pocket to make diverse contacts with the PCNA surface (see Figures S10, S11 and S12). The inclusion of longer and bulkier side-chains at PIP-box position 1, such as lysine (**p21$_{\mu}$-Q144K**) and methionine (**p21$_{\mu}$-Q144M**), resulted in peptides with the lowest affinity for PCNA (K_D values of 1.12 and 1.54 μM, respectively). **p21$_{\mu}$-pol η**, also with Met144^{P1}, bound with similar affinity to **p21$_{\mu}$-Q144M** at 2.54 μM. **p21$_{\mu}$-RNaseH2B**, containing a Met144^{P1}, shows 2.4-fold improved affinity compared to **p21$_{\mu}$-Q144M** (640 nM c.f. 1.54 μM). The Lys144^{P1} containing peptide, **p21$_{\mu}$-pol κ**, bound PCNA with 3.30 μM affinity whereas **p21$_{\mu}$-pol ι** that also contains a Lys144^{P1}, in contrast, gave a K_D value for PCNA of 111 nM.

Peptides with the point modifications of aspartic acid (**p21$_{\mu}$-Q144D**), asparagine (**p21$_{\mu}$-Q144N**) and serine (**p21$_{\mu}$-Q144S**) at position one of the PIP-box displayed significantly lower affinity (50 to 80-fold) than the **p21$_{\mu}$** peptide, with K_D values of 714, 772 nM and 1.03 μM, respectively (see Figure 2B, Table S2). **p21$_{\mu}$-pol λ**, containing Ser144^{P1}, gave a K_D value of 8.14 μM, compared to **p21$_{\mu}$-PARG** with Asp144^{P1} that binds PCNA with significantly higher affinity at 90.0 nM. Asp144^{P1} of **p21$_{\mu}$-PARG** does not, however, make any clear secondary interactions to explain the higher PCNA affinity.

p21$_{\mu}$-RFC, with a seven amino-acid PIP-box, was one of the highest affinity non-canonical PIP-box peptides, with a similar K_D value to **p21$_{\mu}$-PARG** and **p21$_{\mu}$-pol ι**, at 145 nM (Figure 2C, 'non-canonical'). It was assumed here that Ile, Phe, Phe of **p21$_{\mu}$-RFC** formed a hydrophobic triad (and hence PIP-box positions 4, 7 and 8) to insert into the hydrophobic cleft of PCNA, which results in an arginine positioned near the Q-pocket (Figure S12). **p21$_{\mu}$-RFC** was computationally modelled on the PCNA surface and indicated that Arg^{P1} extends over the Q-pocket to make hydrophobic surface contacts with Val45, which is analogous to those made by Lys144^{P1} in **p21$_{\mu}$-pol ι**. The other seven amino-acid PIP-box in **p21$_{\mu}$-pol β**, in contrast, was one of the worst performing peptides at 5.73 μM (Figure 2C, Table S2), where all potential binding modes likely force a polar/charged residue into one of the hydrophobic pockets to create unfavourable interactions and lower binding affinity of **p21$_{\mu}$-pol β** for PCNA.

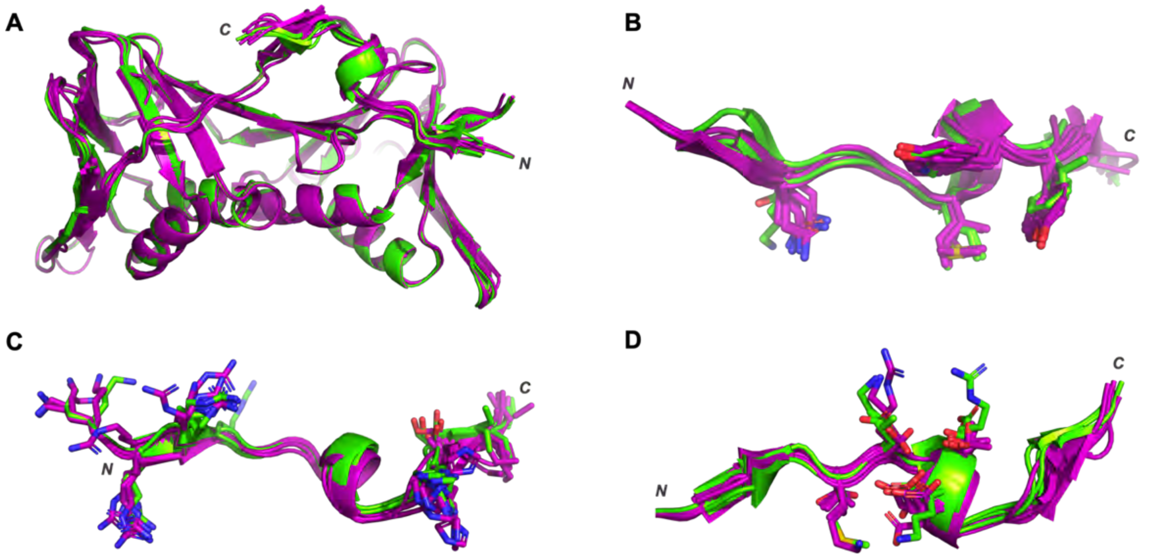


Figure 4. Superimposition of structures of co-crystallised or computationally modelled peptide:PCNA structures. Canonical PIP-box peptides are shown in purple, and non-canonical PIP-box peptide structures are shown in green. Canonical: **p21$_{\mu}$**, **p21$_{\mu}$-F150Y**, **p21$_{\mu}$-S146R**, **p21$_{\mu}$-M147I**, **p21$_{\mu}$-D149E**, **p21$_{\mu}$-FY150/151YF**, **p21$_{\mu}$-Pogo** and **p21$_{\mu}$-pol δ$_{p66}$**. Non-canonical: **p21$_{\mu}$-pol ι**, **p21$_{\mu}$-PARG**, and **p21$_{\mu}$-RFC**. **A** All peptides adopt a single 3$_{10}$-helical turn upon binding PCNA. **B** Conserved PIP-box residues shown with sticks show a high degree of similarity. **C** Flanking residues shown with sticks are orientated in similar directions but are still flexible. **D** Non-conserved PIP-box residues shown with sticks and adopt a large variety of orientations.

5.3.2 Conserved hydrophobic position 4: Not too big and not too small.

The central hydrophobic PIP-box residue plays a key role to stabilise the PCNA-bound binding structure, and anchors the peptide to the PCNA surface. This is evident by the high number (> 5) of interactions the PIP-box position 4 residue makes in all peptides, which are largely hydrophobic interactions (Table S2, S4-13). The position 4 PIP-box residue, in all peptides here, either computationally modelled or co-crystallised with PCNA, is entirely buried within the hydrophobic pocket indicated by a %BSA of 100% (Table S19).

The PCNA-bound peptide structures analysed here all contain a methionine or isoleucine at position four of the PIP-box. There was no significant difference between the overall structures of **p21$_{\mu}$** and **p21$_{\mu}$-M147I**, confirmed by the RMSD value of 0.191 Å (Table S18). **p21$_{\mu}$-M147I** has a marginally higher affinity for PCNA than **p21$_{\mu}$**, at 11.1 nM (Figure 2B). The orientation of Met147^{P4} differed only by slight rotations of the side-chain, owing to its flexibility and large pocket size and made side-chain interactions with hydrophobic residues Val45, Leu47, and Val48 of PCNA. A main-chain hydrogen bond interaction between peptide residue Met147^{P4} and PCNA residue His44 was evident in all eight of the Met147^{P4}-containing modelled peptides bound to PCNA. The Ile147^{P4} peptides **p21$_{\mu}$-pol ι** and **p21$_{\mu}$-PARG** similarly make this contact.

Hydrophobic amino-acid substitutions in **p21$_{\mu}$-M147V** and **p21$_{\mu}$-M147L** were well tolerated and bound PCNA with 29.3 nM and 20.5 nM affinity, respectively. **p21$_{\mu}$-M147W**, with the bulkier Trp147^{P4}, revealed lower affinity for PCNA at 3.57 μM. However, the peptide containing alanine (**p21$_{\mu}$-M147A**), with the smallest side-chain of this subset, resulted in even weaker affinity for PCNA, at 7.59 μM. These results reiterate the necessity for a hydrophobic residue at position 4, with optimal size to fill the hydrophobic pocket and anchor the peptide to the PCNA surface.

5.3.3 *Conserved aromatic positions 7/8:* Hydrogen bonds create tighter interactions between peptide and PCNA.

These aromatic residues complete the PIP-box hydrophobic triad that anchors a PIP-box sequence onto the PCNA surface and stabilises the single 3_{10}-helical turn characteristic of the PCNA-partner binding conformation (Figure 4A). The **p21$_{\mu}$** PIP-box position 8 residue inserts into a hydrophobic pocket on the PCNA surface formed by Gln131, Ile128, Pro234, Tyr133 and Tyr250 (Figure 3C). This is exemplified by the Phe150^{P7}Tyr151^{P8} combination observed in **p21$_{\mu}$**, where the tyrosine phenol forms a 3.4 Å hydrogen bond to PCNA Gln131. This interaction is also seen for Tyr151^{P8} in **p21$_{\mu}$-S146R** and **p21$_{\mu}$-M147I** that hydrogen bond to Gln131 at 2.8 Å and bind with PCNA affinities of 4.20 and 11.1 nM, respectively.

The **p21$_{\mu}$-Y151F** peptide has comparable affinity to **p21$_{\mu}$** (12.3 nM) at 10.6 nM. However, previous literature suggested a Tyr151Phe^{P8} modified p21 peptide resulted in a 3-fold decrease in binding affinity compared to the analogous native peptide (residues 139-160).[22] This may be due to a difference in binding affinity assay (SPR v. ITC), difference in the assay conditions (e.g. buffer), or a series of small conformational changes that allow **p21$_{\mu}$** to bind with relatively high affinity, despite being notably shorter than **p21$_{139\text{-}160}$**. A large number of native PIP-box sequences contain this Phe^{P7}Phe^{P8} combination and give K_D values that range from 8.76 nM for **p21$_{\mu}$-Cdt1**, to 3.30 µM for **p21$_{\mu}$-pol κ** (Figure 2C).

p21$_{\mu}$-F150Y displayed a slight decrease in affinity for PCNA, compared to p21$_{\mu}$, with a K_D value of 20.2 nM (Figure 2B, Table S2). The co-crystal structure of **p21$_{\mu}$-F150Y** bound to PCNA indicates the Tyr150^{P7} side-chain of **p21$_{\mu}$-F150Y** is shifted closer to Pro253 (3.7 Å), that is located at the edge of the hydrophobic cleft (Figure 3D, Figure 5 pink), relative to the Phe^{P7} in p21$_{\mu}$ at 3.9 Å (Figure 3C). The phenol group introduced from the Phe150Tyr^{P7} modification participates in a 3.6 Å hydrogen bond with Thr145^{P2} of the peptide (Figure 3D), that stabilises the 3_{10}-helical binding conformation of the peptide and may enhance PCNA binding affinity. Tyr^{P8} often makes a hydrogen bond with Gln131, which is not observed for **p21$_{\mu}$-F150Y**, suggesting hydrogen bonds between Tyr151^{P8}—Gln131 and Tyr150^{P7}—Thr145^{P2} cannot both occur at once. Interestingly, there are no human PIP-box proteins reported to contain the Tyr^{P7}Tyr^{P8} motif, though one has been identified in the E2F transcription factor from *Drosophila melanogaster.*[11]

The remaining aromatic permutation of Tyr^{P7}Phe^{P8} resulted in the greatest improvement in affinity for this subset, with a K_D value of 2.18 nM for **p21$_{\mu}$-FY150/151YF** (Figure 2B). The computationally modelled structure of **p21$_{\mu}$-FY150/151YF** bound to PCNA shows Phe151^{P8} is lifted slightly out of the pocket, as evidenced by an increased distance from Phe151^{P8} to Gln131 at 4.8 Å (c.f. 3.4 Å in **p21$_{\mu}$**) and there is also no hydrogen bond evident between Tyr150^{P7}—Thr145^{P2}. However, a strong intermolecular hydrogen bond from Tyr150^{P7} to the Ala232 carbonyl may improve PCNA affinity. The high affinity **p21$_{\mu}$-Pogo** (8.82 nM) also displays the Tyr^{P7}Phe^{P8} combination.

A sequence modification to include the aromatic histidine, **p21$_{\mu}$-F150H**, displayed lower affinity for PCNA at 159 nM, compared to **p21$_{\mu}$** with a K_D value of 4.32 nM (Figure 2B). The computationally modelled structure of **p21$_{\mu}$-PARG** bound to PCNA shows His150^{P7} positioned similarly to the analogous Phe150^{P7} in **p21$_{\mu}$**. His150^{P7} is angled toward to *N*-terminus of the peptide, and makes a secondary interaction with Ser145^{P2} (Figure 5, red) which may improve the affinity of the non-canonical PIP-box sequence. Other non-aromatic residues are occasionally observed at PIP-box positions 7 and 8 such as the WRN PIP-box with an aspartic acid at position 7 in place of the conserved aromatic residue. **p21$_{\mu}$-WRN** was the worst performing native canonical PIP-box peptide with an affinity for PCNA of 1.15 μM (Figure 2C), whereas **p21$_{\mu}$-pol ι** which contains a Leu151^{P8}, bound PCNA with 111 nM affinity. The computationally modelled structure of **p21$_{\mu}$-pol ι** bound to PCNA suggests Leu151^{P8} is lifted out of the hydrophobic pocket, compared to Tyr151^{P8} in **p21$_{\mu}$**, indicated by an increased distance to Gln131 (3.2 Å in **p21$_{\mu}$**, to 6.4 Å in **p21$_{\mu}$-pol ι**). This may explain the 9-fold lower affinity of **p21$_{\mu}$-pol ι** (111 nM) compared to **p21$_{\mu}$** (12.3 nM).

5.3.4 Non-conserved position 2/3: A positively charged residue at P3 increases affinity.

No studies to date have explicitly looked at the role of the position 2 and 3 PIP-box residues as the initial alanine scan that identified the key residues in the PIP-box revealed that mutation of residues Thr145^{P2} and Ser146^{P3} in p21 did not significantly impact the PCNA affinity.[18] Ser146^{P3} in the co-crystal structure of **p21$_{\mu}$** bound to PCNA is orientated toward the PCNA surface and makes a 3.0 Å hydrogen bond with the main-chain carbonyl of His44 in 1 of the 3 monomer repeats. Thr145^{P2} makes a 3.3 Å main-chain hydrogen bond to Pro253 (Figure 3C) in all three repeats (PDB: 7KQ1). Inclusion of lysine or arginine at 145^{P2} (**p21$_{\mu}$-T145K** and **p21$_{\mu}$-T145R**) resulted in 8- and 7-fold decreased affinity for PCNA respectively, relative to **p21$_{\mu}$** (Figure 2B) whereas **p21$_{\mu}$-S146K**, with Lys146^{P3}, revealed affinity for PCNA comparable to **p21$_{\mu}$**. However, an Arg146^{P3} instead resulted in a 2.8-fold improvement in the K$_D$ value, to 4.30 nM for **p21$_{\mu}$-S146R** (Figure 2B). The computationally modelled structure of **p21$_{\mu}$-S146R** bound to PCNA indicates Arg146^{P3} main-chain amide makes a 3.2 Å intramolecular hydrogen bond to the carbonyl of the Asp149^{P6} side-chain (Figure S6B), and a 2.8 Å intermolecular hydrogen bond to Ser43 on the PCNA surface (Figure 5, yellow). These interactions may together enhance the PCNA-binding affinity of **p21$_{\mu}$-S146R**. The high affinity peptides **p21$_{\mu}$-XPG**, and **p21$_{\mu}$-FEN1** also contain an Arg146^{P3} and bind PCNA with 16.4 and 166 nM affinity, respectively (Figure 2C). **p21$_{\mu}$-Cdt1** binds with even higher affinity at 8.76 nM, and contains an Arg^{P2}Arg^{P3} motif that results in a significant improvement to PCNA binding affinity when included in the **p21$_{\mu}$** peptide, **p21$_{\mu}$-TS145/146RR**, which has a K$_D$ value of 1.83 nM.

p21$_{\mu}$-Pogo binds PCNA with slightly higher affinity (8.82 nM) than **p21$_{\mu}$-TS145/146KK**, which contains the same Lys^{P2}Lys^{P3} motif (20.2 nM, Figure 2C). The computationally modelled structure of **p21$_{\mu}$-Pogo** bound to PCNA shows Lys146^{P3} make an analogous 3.2 Å intramolecular hydrogen bond, to Arg146^{P3} in p21$_{\mu}$-S146R, with the 149^{P6} residue, but no intermolecular interactions (Figure 5, dark blue). The two other permutations of arginine and lysine as dual modifications at positions 145 and 146 (**p21$_{\mu}$-TS145/146RK**, **p21$_{\mu}$-TS145/146KR**) both resulted in lower K$_D$ values for PCNA compared to **p21$_{\mu}$** at 17.3 and 37.0 nM (Figure 2B).

Negatively charged residues aren't commonly observed at positions 2 and 3, however Asp146^{P3} is present in the RFC PIP-box sequence. The computationally modelled **p21$_{\mu}$-RFC** peptide bound to PCNA indicates

Asp146^{P3} makes a 3.2 Å intramolecular interaction with the positively charged Lys149^{P6} (Figure S12). This complementary interaction is equivalent to the interaction seen for peptides such as **p21$_{\mu}$-S146R** where Arg146^{P3} and Asp149^{P6} interact. Such intramolecular interactions may contribute to the unexpectedly high affinity of **p21$_{\mu}$-RFC** (145 nM) by stabilising the 3$_{10}$-helical binding conformation.

The PIP-box sequences of pol ι and pol δ$_{p66}$ contain neutral residues at PIP-box positions 2 and 3, ^{421}GlyLeu and ^{457}ValSer, respectively. Both **p21$_{\mu}$-pol ι** and **p21$_{\mu}$-pol δ$_{p66}$** bind PCNA with lower affinity than **p21$_{\mu}$** at 111 and 124 nM, respectively, and lack the ability to make a PIP-box position 2 side-chain hydrogen bond with PCNA, as in **p21$_{\mu}$**. This observation reinforces the conclusion that positively charged or polar residues at these non-conserved positions lead to enhanced PCNA binding affinity.

5.3.5 Non-conserved position 5/6: Side-chains with hydrogen bond donor and acceptor character increase affinity.

The Thr^{P5}Asp^{P6} motif in p21 and Pogo is thought to be responsible, in part, for the high PCNA affinity observed, owing to the resultant hydrogen bonds.[25] The initial report of an alanine scan of p21 that identified the PIP-box motif noted Asp149^{P6} as an important residue for inhibition of SV40 DNA replication.[18] The co-crystal structure of **p21$_{139\text{-}160}$** (PDB: 1AXC) highlights two intramolecular interactions of Asp149^{P6} and one of Thr145^{P5}.[10, 15] The co-crystal structure of **p21$_{\mu}$** bound to PCNA (PDB: 7KQ1) here shows the Asp149^{P6} side-chain (hydrogen bond acceptor) makes a 2.9 Å hydrogen bond with the Thr145^{P2} side-chain in 1 of the 3 monomer repeats. Residues Asp149^{P6} and Ser146^{P3} in **p21$_{\mu}$** participate in a 3.2 Å intramolecular main-chain hydrogen bond to define a 3$_{10}$-helix (Figure 3C). These two hydrogen bonds are the only interactions of the PIP-box position 5 or 6 residues in **p21$_{\mu}$**. The analogous interaction is made in the computationally-modelled structure of **p21$_{\mu}$-PARG** and **p21$_{\mu}$-Pogo** bound to PCNA. The Asp149^{P6} side-chain in **p21$_{\mu}$-S146R** acts as a hydrogen bond donor to interact with the main-chain of Arg145^{P2}.

Incorporation of a negatively charged residue at position 5 of the PIP-box (**p21$_{\mu}$-T148E** and **p21$_{\mu}$-T148D**) resulted in ~6 – 8-fold lower affinity for PCNA, than **p21$_{\mu}$**. The K$_D$ value of **p21$_{\mu}$-TD148/149DE** was notably worse than **p21$_{\mu}$**, at 28-fold lower affinity. However, this was recovered to only 4-fold lower than **p21$_{\mu}$** for a Glu^{P5}Glu^{P6} modification (for **p21$_{\mu}$-TD148/149EE** (Figure 2B) and highlights that modifications must be considered in the context of the whole sequence. A number of peptides display an intermolecular hydrogen bond with His44 side-chain of PCNA where an Asp^{P5} or Glu^{P5} acts as a hydrogen bond donor (e.g. Figure 5, light green) and is more common in non-canonical than canonical PIP-box sequences. A negatively charged residue at position 6, in contrast, is more likely to form intramolecular interactions, rather than intermolecular. This can be reasoned by the resulting side-chain positions when the 3$_{10}$-helical binding conformation is adopted. When negatively charged residues are not included at positions 5 and 6, high affinity partners generally contain polar, rather than hydrophobic residues at these PIP-box positions (see DNALig1 or MCMT). In **p21$_{\mu}$-pol ι**, the unusual Tyr149^{P6} points away from the surface in the computationally modelled structure bound to PCNA, and the 3$_{10}$-helix stabilising 3.3 Å intramolecular main-chain hydrogen bond is instead made between position 6 and Ile147^{P4} (Figure S11).

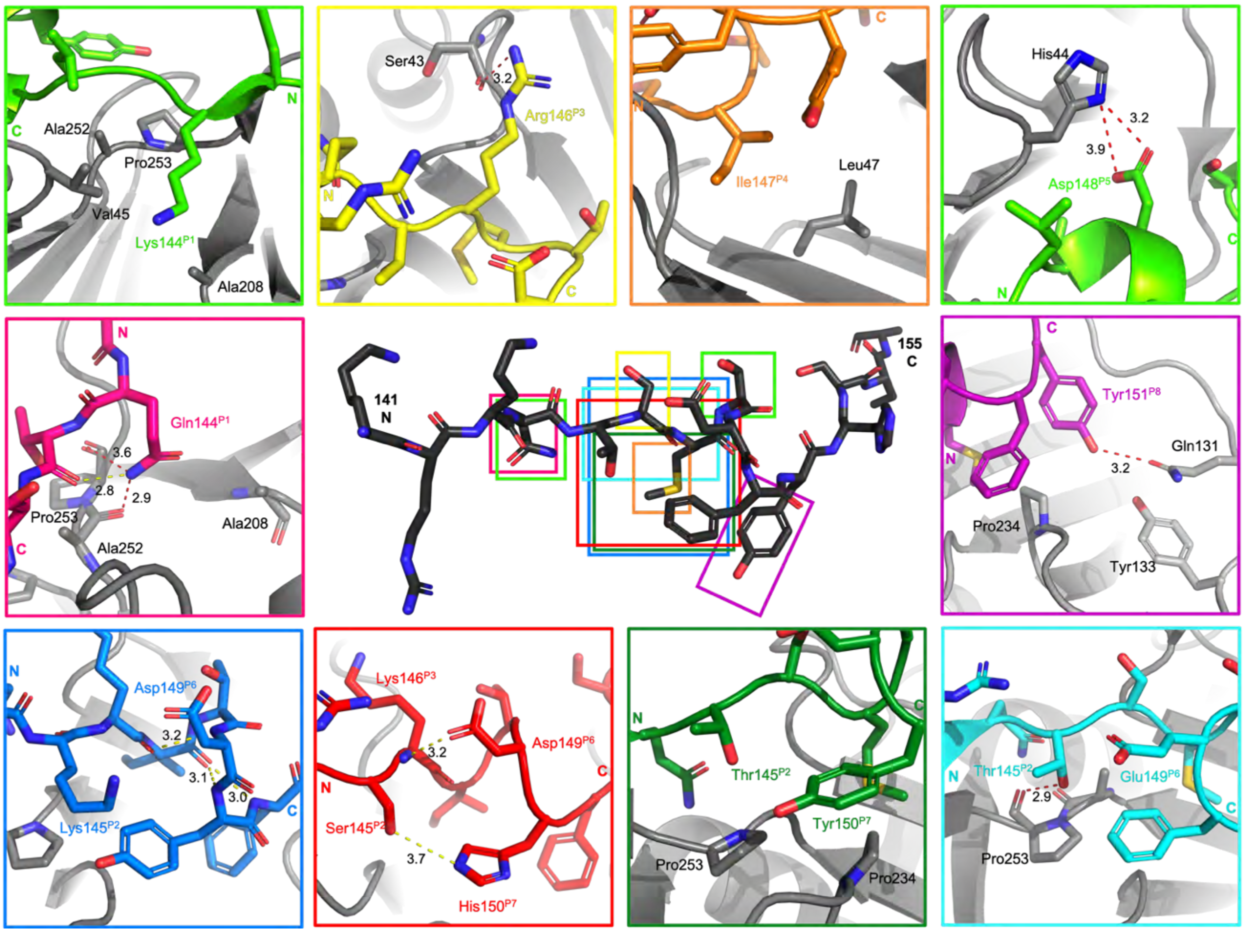

C
Ala252
Pro253
N
Val45
Lys144^{P1}
Ala208
Ser43
3.2
Arg146^{P3}
N
C
C
N
Ile147^{P4}
Leu47
His44
3.2
3.9
Asp148^{P6}
N
C
N
Gln144^{P1}
3.6
2.8
2.9
Pro253
Ala208
C
Ala252
141
N
155
C
C
Tyr151^{P8}
N
Gln131
3.2
Pro234
Tyr133
N
Asp149^{P6}
3.2
3.1
3.0
C
Lys145^{P2}
Lys146^{P3}
3.2
Asp149^{P6}
N
C
Ser145^{P2}
3.7
His150^{P7}
N
C
Thr145^{P2}
Tyr150^{P7}
Pro253
Pro234
N
Thr145^{P2}
Glu149^{P6}
2.9
C
Pro253

Figure 5. Representative examples of key inter- and intramolecular interactions. Structures shown in cartoon format with side-chains as sticks. **p21$_{\mu}$** structure shown in centre in black and area of key interactions shown in coloured rectangles. PCNA is shown in grey. Computationally modelled peptides bound to PCNA: **p21$_{\mu}$-pol ι**, light green; **p21$_{\mu}$-S146R**, yellow; **p21$_{\mu}$-M147I**, orange; **p21$_{\mu}$-pol δ$_{p66}$**, purple; **p21$_{\mu}$-PARG**, red; **p21$_{\mu}$-D149E**, light blue; **p21$_{\mu}$-Pogo**, dark blue; **p21$_{\mu}$-FY150/151YF**, dark green. Co-crystal structures of peptides bound to PCNA: **p21$_{\mu}$-F150Y**, pink; **p21$_{\mu}$**, purple. Intramolecular hydrogen bonds indicated as yellow dashed line, and intermolecular hydrogen bonds as a red dashed line. Distances are indicated in Angstroms. Elemental colouring: nitrogen, blue; oxygen, red; sulfur, yellow.

5.3.6 Individual PIP-box modifications must be considered in the context of the entire PIP-box sequence.

All p21$_{\mu}$ peptides with greater than 20 nM affinity for PCNA display charge or hydrogen bond complementary residues on either side of the hydrophobic residue at position 4, supporting the idea that this configuration aids stabilisation of the 3$_{10}$-helical binding conformation and in turn enhances affinity. This is seen in **p21$_{\mu}$-Pogo** and **p21$_{\mu}$-XPG** that bind with higher affinity than the **p21$_{\mu}$** peptide (Figure 2), and Figure 4D clearly shows that side-chains of residues at position 2 and 6 are generally angled toward one another.

Cooperative interactions, in the form of charge-complementary pairs, may be more important in non-canonical PIP-box peptides. The RFC seven amino-acid PIP-box contains a charge complementary pair either side of the hydrophobic position 4 residue, however this is in the less common orientation where the negatively charged residue (Asp) is at position 3 and the positively charged residue (Arg/Lys) is at positions 5/6. Similarly, the divergent, non-canonical PARG PIP-box in the **p21$_{\mu}$-PARG** peptide displays relatively high PCNA affinity at 90.0 nM despite the lack of a canonical Gln^{P1}, and contains His150^{P7} rather

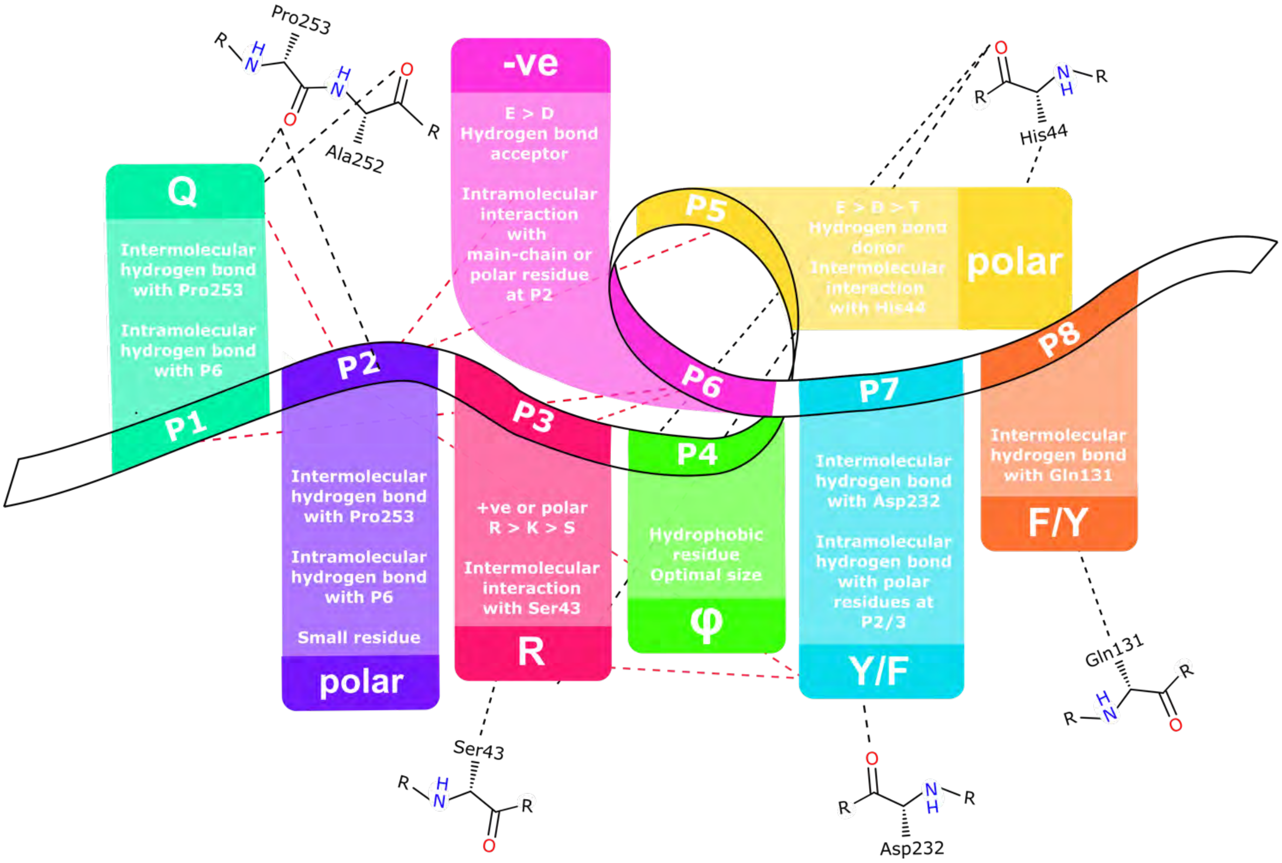

Pro253
Ala252
Q
Intermolecular hydrogen bond with Pro253
Intramolecular hydrogen bond with P6
P1
-ve
E > D
Hydrogen bond acceptor
Intramolecular interaction with main-chain or polar residue at P2
P2
Intermolecular hydrogen bond with Pro253
Intramolecular hydrogen bond with P6
Small residue
polar
P3
+ve or polar
R > K > S
Intermolecular interaction with Ser43
R
Ser43
P4
Hydrophobic residue
Optimal size
φ
P5
E > D > T
Hydrogen bond donor
Intermolecular interaction with His44
polar
His44
P6
P7
Intermolecular hydrogen bond with Asp232
Intramolecular hydrogen bond with polar residues at P2/3
Y/F
Asp232
P8
Intermolecular hydrogen bond with Gln131
F/Y
Gln131

Figure 6: Summary of the guideline to design a high affinity PIP-box and the interactions each residue participates in when bound to PCNA. PCNA residues shown as line structures. Intermolecular interactions to side-chain or main-chain shown as black dashed lines, intramolecular interactions from side-chain or main-chain shown as red dashed lines.

than phenylalanine or tyrosine. Interestingly, this peptide binds with higher affinity than **p21$_{\mu}$-F150H**, suggesting the expected affinity loss due to the absence of the Gln144^{P1} is overcome by the remaining non-conserved residues, such as the charge complementary Lys146^{P3} and Asp149^{P6}, and hydrogen bond between His150^{P7} and Ser145^{P2} (Figure 5, red). The PARG PIP-box also contains the Thr^{P5}Asp^{P6} motif observed in the high affinity PCNA partners p21 and PL. The highly divergent PIP-box of the **p21$_{\mu}$-pol ι** peptide also shows relatively high PCNA affinity (111 nM), in alignment with the higher than expected affinity displayed by the native **pol ι$_{415-437}$** peptide (0.39 μM, Table 1).[16] This PIP-box lacks a second aromatic residue at position 8, in addition, position 1 is not glutamine. However, the pol ι PIP-box does display a number of interactions that may stabilise the binding conformation and improve affinity in the absence of these traditional elements. For example, a hydrogen bond between Asp425^{P5} and His44 of PCNA (c.f. Figure 5, light green), and an intramolecular Tyr427^{P8} and Lys421 hydrogen bond (Figure S16F, yellow).

Some peptides such as **p21$_{\mu}$-pol δ$_{p66}$** and **p21$_{\mu}$-FEN1** bind with unexpectedly lower affinity than **p21$_{\mu}$**, despite strict canonical sequences. **p21$_{\mu}$-FEN1** is particularly interesting as not only does it contain charge complementarity by means of residues (Arg146^{P3} to Asp148^{P5} and Asp149^{P6}) on either side of the hydrophobic position 4 residue, but also an Arg146^{P3} which considerably enhanced PCNA binding affinity when introduced as a point modification in p21$_{\mu}$ (**p21$_{\mu}$-S146R**, 4.30 nM). The other sequence differences in **p21$_{\mu}$-FEN1** (compared to p21$_{\mu}$), are the Phe^{P7}Phe^{P8} motif (seen in **p21$_{\mu}$-Y151F**, 10.6 nM) and Leu147^{P4}

Table 1. The effect of sequence flanking the PIP-box on affinity. A comparison of p21$_{\mu}$:PIP-box hybrid peptides and the analogous native peptides from which the PIP-box derives.

Name	Sequence	Affinity K_D		Native peptides	Affinity K_D	Ref.
p21$_{\mu}$-Pogo	KRRQKKITDYFHSKR	8.82 nM	<	SAVLQKKITDYFHPKK*	100 nM	20, 24
p21$_{\mu}$-Cdt1	KRRQRRVTDFFHSKR	8.76 nM	<	^{2}MEQRRVTDFFARRR^{15}	7.20 μM	26
p21$_{\mu}$-pol δ$_{p66}$	KRRQVSITGFFHSKR	124 nM	<	^{452}KANRQVSITGFFQRK^{466}	15.6 μM	14
p21$_{\mu}$-FEN1	KRRQGRLDDFFHSKR	166 nM	<	^{331}SRQGSTQGRLDDFFKVTGSL^{350}	59.9 μM	14
p21$_{\mu}$-p15	KRRQKGIGEFFHSKR	234 nM	<	^{41}APVCVRPTPKWQKGIGEFFAA^{72}	5.56 μM	27
p21$_{\mu}$-PARG	KRRDSKITDHFHSKR	90.0 nM	<	^{402}QHGKKDSKITDHFMRLPKA^{420}	3.3 μM	23
p21$_{\mu}$-pol ι	KRRKGLIDYYLHSKR	111 nM	<	^{415}ALNTAKKGLIDYYLMPSLSTTSR^{437}	0.39 μM	16
p2$_{1\mu}$-pol η	KRRMQTLESFFHSKR	2.54 μM	>	^{693}CKRPRPEGMQTLESFFKPLTH^{713}	0.4 μM	16
p21$_{\mu}$-pol κ	KRRKHTLDIFFHSKR	3.30 μM	<	^{861}PKHTLDIFFK^{874}PLTH †	4.9 μM	16
p21$_{\mu}$-Cdt2	KRRMRKICTYFHSKR	ND*	>>	^{704}SSMRKICTYFHRKS^{717}	57 nM	26
p21$_{\mu}$-RNAseH 2B	KRRMKSIDTFFHSKR	640 nM	<	^{290}DKSGMKSIDTFFGVKNKKKIGKV^{312}	35 μM	28

* PL peptide is a mutant hybrid peptide, and not an entirely native sequence.
† PLTH is not part of native sequence and was added to sequence to improve binding affinity.

(seen in **p21$_{\mu}$-M147L**, 20.5 nM) that did not drastically impact affinity in the point modified peptides. Perhaps the lower affinity of **p21$_{\mu}$-FEN1** is owed in part to the Asp^{P5}Asp^{P6} combination, as the **p21$_{\mu}$-T148D** peptide showed a notably lower PCNA affinity relative to **p21$_{\mu}$** (96.5 nM, c.f. 4.32 nM for **p21$_{\mu}$**). This is an example of how each PIP-box modification must be considered in context of the entire PIP-box sequence, and the importance of different sequence combinations. The common interactions that appear to contribute to enhanced PCNA binding affinity have been summarised in Figure 6 to provide guidelines to predict cooperative interactions and hence affinity.

This concept, that the sequence as a whole must be considered to anticipate PCNA affinity, extends to the sequence that flanks the PIP-box. The p21$_{\mu}$ peptides considered here all comprise the same p21-derived PIP-box flanking sequence, so any changes in affinity or structure can be attributed to changes induced by the PIP-box sequence itself. The affinity of p21$_{\mu}$ peptides containing a native PIP-box was compared to the reported affinity of the analogous native peptide sequence to reveal an interesting trend: All five of the p21$_{\mu}$ native canonical PIP-box hybrid peptides display higher affinity compared to the native peptide (Table 1). The most significant difference is greater than 820-fold for the Cdt1 PIP-box peptides, where **p21$_{\mu}$-Cdt1** binds with 8.76 nM and **Cdt1$_{2-15}$** binds with 7.20 μM affinity.[26] This may be correlated to the number of positively-charged residues in the flanking region, where the native canonical sequences contain four or less positively-charged residues compared to the **p21$_{\mu}$** with four arginine residues. In contrast, four of the six non-canonical PIP-box:p21$_{\mu}$ hybrids showed lower affinity relative to the native sequence (Table 1).

5.3.7 Rationally designed PIP-box sequences: Designing the highest affinity PCNA partner to date

PIP-box sequences were designed to investigate different cooperative amino-acid combinations, and were inspired by the binding affinity results of the modified p21$_{\mu}$ peptides in conjunction with the structural observations from co-crystal structures, computationally modelled peptides and native PIP-box structures which are summarised graphically in Figures 5 and 6. Three PIP-box sequences were synthesised to trial whether amino-acid combinations not previously observed together could give rise to high PCNA affinity. For example, would a Glu148^{P5} interact with the positively charged residues at P2/P3, and the Glu149^{P6} interact with His44 simultaneously to improve PCNA affinity of **p21$_{\mu}$-RD1**.

Three peptides were synthesised, the affinity for PCNA determined by SPR and the PCNA-bound structures computationally modelled to reveal, remarkably, that all three peptides bound with higher affinity than the native **p21$_{\mu}$** (Figure 2D, Table S2). The highest affinity peptide **p21$_{\mu}$-RD2** bound PCNA with 1.12 nM affinity, which is 11-fold higher affinity than **p21$_{\mu}$** and 3-fold higher than **p21$_{139\text{-}160}$** despite being 7 amino-acids shorter. This makes **p21$_{\mu}$-RD2**, a PIP-box peptide rationally designed amino-acid by amino-acid, the highest affinity PCNA-binding peptide or protein to date, displaying a significant improvement in affinity for PCNA. The computationally modelled structures of these peptides bound to PCNA were analysed to confirm whether the hypothesised cooperative interactions were observed and contributed to this impressive affinity and allowed us to refine the observations incorporated into the secondary interaction map in Figure 6 that provides guidelines to design high affinity a PIP-box sequence.

In **p21$_{\mu}$-RD1** and **p21$_{\mu}$-RD3** the bond between Tyr150^{P7} and Pro253 is stronger than that of **p21$_{\mu}$**, at distances of 3.5 Å and 3.4 Å respectively which may contribute to the improved affinity for PCNA. The highest affinity peptide **p21$_{\mu}$-RD2** (K$_D$ 1.12 nM), has four defined intramolecular hydrogen bond interactions (Figure 7C) which all involve PIP-box residues and may stabilise the 3$_{10}$-helical binding conformation and lead to the enhanced PCNA affinity. In particular, the Glu149^{P6} forms a main-chain hydrogen bond to

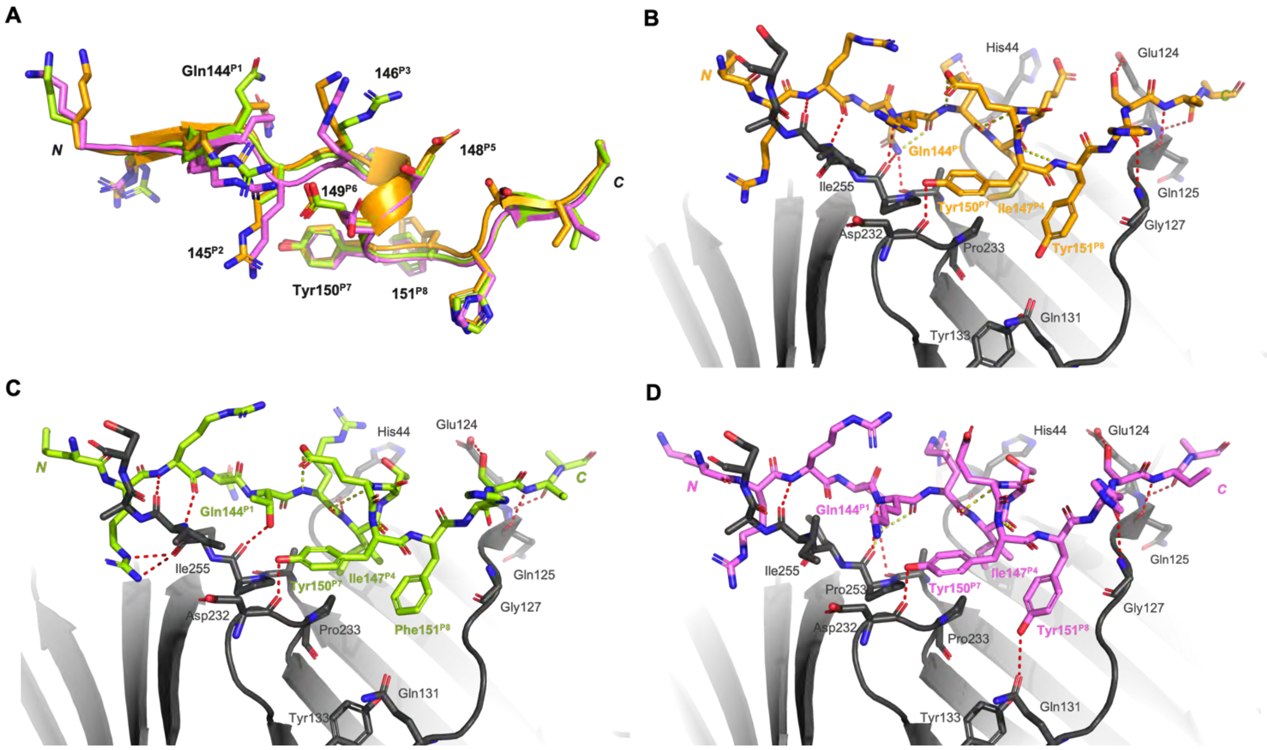

A
Gln144^{P1}
146^{P3}
N
148^{P5}
C
149^{P6}
145^{P2}
Tyr150^{P7}
151^{P8}
B
N
His44
Glu124
Gln144^{P1}
Ile255
Tyr150^{P7}
Ile147^{P4}
Gln125
Asp232
Gly127
Pro233
Tyr151^{P8}
Gln131
Tyr133
C
N
His44
Glu124
C
Gln144^{P1}
Ile255
Tyr150^{P7}
Ile147^{P4}
Gln125
Asp232
Gly127
Pro233
Phe151^{P8}
Gln131
Tyr133
D
N
His44
Glu124
C
Gln144^{P1}
Gln125
Ile255
Pro253
Tyr150^{P7}
Ile147^{P4}
Asp232
Gly127
Pro233
Tyr151^{P8}
Gln131
Tyr133

Figure 7: Computational modelling of rationally designed PIP-box peptide **p21$_{\mu}$-RD1**, **p21$_{\mu}$-RD2**, **p21$_{\mu}$-RD3** on PCNA. PCNA shown in grey as a cartoon, with interacting amino-acid shown as sticks and labelled in grey. Intermolecular polar interactions are shown as red dashed lines and intramolecular polar interactions shown as yellow dashed lines. **A** Overlay of the structure of the rationally-designed PIP-box peptides. **B p21$_{\mu}$-RD1** shown in orange and conserved residues labelled. **C p21$_{\mu}$-RD2** shown in green and conserved residues labelled. **D p21$_{\mu}$-RD3** shown in pink and conserved residues labelled.

Arg146^{P3}. **p21$_{\mu}$-RD2** contains an Arg146^{P3} which makes a hydrogen bond to Ser43 in the computationally modelled structure bound to PCNA, which is also observed for **p21$_{\mu}$-S146R** that bound PCNA with high affinity (4.30 nM, Figure 5, yellow). In addition, Thr145^{P3} side-chain of **p21$_{\mu}$-RD2** hydrogen bonds to the Pro253 carbonyl; and the Tyr150^{P7} phenol hydrogen bonds with Asp232 main-chain carbonyl (Figure 7C). The 146Arg^{P3} modification resulted in increased affinity in **p21$_{\mu}$-S146R** and was included in **p21$_{\mu}$-RD2** and **p21$_{\mu}$-RD3**; however, the affinity of **p21$_{\mu}$-RD3** was not significantly different to **p21$_{\mu}$** and suggests the other modifications may be working to oppose an increase in affinity. Interestingly, **p21$_{\mu}$-RD1** with the Tyr^{P7}Tyr^{P8} motif only makes an intermolecular hydrogen bond between Tyr150^{P7} and Asp232, as seen in **p21$_{\mu}$-F150Y**; however **p21$_{\mu}$-RD3** makes an intermolecular interaction through the phenol of both Tyr150^{P7} (to Asp232) and Tyr151^{P8} (to Gln131). This is further evidence that sequence modifications work cooperatively to alter structure and affinity.

5.4 Conclusions and outlook

This work provides a comprehensive study using a set of PCNA-binding peptides and highlights a series of interactions that synergistically contribute to high PCNA affinity, as summarised in Figure 6. The glutamine residue and hydrophobic triad are essential to enhance PCNA binding affinity and adhere a peptide to the PCNA surface. The PIP-box position 1 glutamine residue and its hydrogen bond network contribute notably to PCNA binding affinity, although glutamine is not essential to confer PCNA binding. The PIP-box position 4 hydrophobic residue, that fills the hydrophobic cleft, is essential to a high affinity PCNA interaction. A variety of combinations of tyrosine and phenylalanine (FY, YF, FF or YY) are most commonly preferred at positions 7 and 8 of the PIP-box, in order to stabilise the characteristic 3_{10}-helical turn. However, the highest binding affinity was observed for peptides containing either Phe^{P7}Tyr^{P8} or Tyr^{P7}Phe^{P8} (rather than Tyr^{P7}Tyr^{P8} or Phe^{P7}Phe^{P8}). A tyrosine residue at PIP-box position 7 makes an intramolecular interaction, commonly with the PIP-box position 2 residue, where the phenol is oriented toward the *N*-terminus of the peptide; whereas a position 8 tyrosine generally hydrogen bonds with Gln131 of PCNA within the hydrophobic cleft. The co-crystal structure of **p21$_{\mu}$-F150Y** bound to PCNA is the first Tyr^{P7}Tyr^{P8} peptide reported to interact with PCNA. The PIP-box Tyr151^{P7} forms a hydrogen bond to Thr145^{P2}, but a hydrogen bond from position 8 to Gln131 is absent which suggests these two proposed interactions cannot occur at once.

Complementary hydrogen bonding or charge-pairs either side of the hydrophobic residue at position 4 stabilise the 3_{10}-helix binding conformation, and appear particularly important to increase PCNA affinity of non-canonical PIP-box sequences. Introduction of a positively-charged residue at position 3 is more favourable than at position 2 to increase PCNA binding affinity. The non-conserved PIP-box combinations that lead to increased PCNA binding affinity include a positively charged residue at position 3 and negatively charged residue at position 6, where both are preceded by a small polar amino-acid that may hydrogen bond intramolecularly or with the PCNA surface. Negatively charged residues at position 5 more commonly hydrogen bond intermolecularly with PCNA, in particular with His44, whereas position 6 residues often form intramolecular hydrogen bonds. Lastly, an increase in positive-charge in the sequence that flanks a canonical PIP-box motif appears to further enhance binding affinity. This is the first study that has extensively looked at the effect of non-conserved PIP-box combinations on PCNA-binding affinity, and has identified key modifications that increase affinity such as the PIP-box position 3 arginine.

p21$_{\mu}$-RD2, with PIP-box QTRITEYF, is the highest affinity PCNA-binding partner reported to date (1.12 nM) and is seven amino-acids shorter than the previous title-holder in **p21$_{139\text{-}160}$** (4.32 nM). This peptide conforms explicitly to the guidelines for secondary interactions that lead to high affinity PCNA binding set out in Figure 6. It is important to note that **p21$_{\mu}$-RD2** provides just one solution to the puzzle of obtaining high affinity to PCNA, and many other combinations may also lead to high PCNA binding affinity. This study has highlighted the complexity of the secondary interaction network that gives rise to high affinity binding of short peptides to PCNA, and has made significant progress to define the rules of high affinity PCNA interaction. The importance of these interactions, in particular the intermolecular interactions with PCNA, may be further probed by mutation of key PCNA residues such as His44. Furthermore, it emphasises that to predict PCNA-binding affinity the PIP-box sequence must be considered as a whole, where small sequence changes can give rise to large changes in affinity, dependent on the rest of the PIP-box sequence. The insights gained here can be used to inform design of new PCNA-binding peptides, or tune the affinity of PCNA peptides and proteins to probe interactions important in DNA-replication and -repair. Consequently, this study provides an insight into how nature has fine-tuned affinity of the native PCNA (Figure 7C) which all involve PIP-box residues and may stabilise the 3_{10}-helical binding conformation and lead to the enhanced PCNA affinity. In particular, the Glu149^{P6} forms a main-chain hydrogen bond to binding proteins to allow tight regulation of DNA-replication and -repair processes. This knowledge may be further leveraged to design inhibitors of human PCNA interactions for therapeutic applications.

5.5 Methods

5.5.1 Peptide synthesis

All peptides were synthesised using Fmoc/t-Bu Solid-Phase Peptide Synthesis [29]. Six peptides were synthesised using Rink Amide resin on a Liberty Blue peptide synthesiser (CEM Corp., Matthews, NC, USA); thirty-one peptides were synthesised on a Prelude peptide synthesizer (PTI, Tucson, AZ, USA); five peptides were synthesised manually as detailed in the supplementary information. The remaining ten peptides were purchased from Shanghai Royobiotech at >95% purity. Synthesised peptides were purified using semi-preparatory RP-HPLC using a Phenomenex Luna C18(2) or Phenomenex Aeris Peptide C18 column (10 mm × 250 mm, 5 µm) over a linear gradient of water and acetonitrile, with 0.1% TFA, at 4 mL/min and UV detection at 220 nm. Peptide identity was confirmed by HRMS using an Agilent 6230 ESI-TOF MS. Peptide purity was characterised on an Agilent 1260 Infinity analytical RP-HPLC equipped with a Phenomenex Luna C18(2) column (4.6 mm × 250 mm, 5 µm) using a linear gradient of 0-50% acetonitrile with 0.1% TFA, in water with 0.1% TFA, over 15 min at 1.5 mL/min and visualised at 220 nm. Detailed methods are included in the Supplementary Informationand characterisation data are listed in Table S1.

5.5.2 Protein expression and purification

A glycerol stock of *E. coli* BL21 (λDE3) cells carrying a PCNA-pMCSG19 plasmid (with no purification tag) were grown in a 50 mL overnight culture. Eight 1 L baffled flasks of LB with 100 µg/mL of ampicillin were inoculated with 6.3 mL of the overnight culture. Cultures were incubated at 37°C until OD600 of 0.5 and induced with 0.5 mM IPTG. Cultures were grown overnight at 16°C with shaking at 200 rpm. Cultures were pelleted at 5000xg for 20 min. After removing the supernatant, pellets were resuspended in 30 mL of Buffer A (20 mM Tris pH 7.5, 20 mM NaCl, 2 mM DTT), then lysed by 5 rounds of cell disruption by a microfluidics cell disrupter. Lysate was pelleted at 45,000xg for 60 min and the supernatant was collected for purification.

PCNA was purified at 4°C by fast protein liquid chromatography (FPLC), using an anion exchange DEAE column (HiTrap DEAE FF 5 mL column), equilibrated in Buffer A (20 mM Tris pH 7.5, 20 mM NaCl, 2 mM DTT), and protein was eluted using Buffer B (20 mM Tris pH 7.5, 0.7 M NaCl, 2 mM DTT). Fractions were analysed by SDS-PAGE and those of interest indicating containing protein at ~28 kDa were selected and pooled and treated with ammonium sulphate to bring the concentration to 1.5 M. Protein was purified again by a hydrophobic column (HiTrap Phenyl FF [high sub] 5 mL column) equilibrated in Buffer C (20 mM Tris pH 7.5, 20 mM NaCl, 2 mM DTT, 0.5 mM EDTA, 1.5 M Ammonium Sulphate), and protein was eluted using Buffer D (20 mM Tris pH 7.5, 0.5 mM EDTA, 2 mM DTT). Fractions were analysed by SDS-PAGE and those of interest were dialysed overnight in Buffer E (20 mM Tris pH 7.5, 20 mM NaCl, 1 mM DTT).

Protein pool was concentrated using a centrifugal filter unit (30 kDa molecular mass cut off) to a volume of less than 10 mL and purified using the size exclusion column (HiPrep 26/60 Sephacryl S-200 HR 300 mL column), equilibrated in Buffer F (20 mM Tris pH 7.5, 50 mM NaCl, 2 mM DTT, 0.5 mM EDTA), and protein was eluted using the same buffer. Fractions were analysed by SDS-PAGE and those of interest were pooled and purified using an anion exchange Q Sepharose column (ENrich Q 10 × 100 mm 8 mL column), equilibrated in Buffer G (20 mM Tris pH 7.5, 20 mM NaCl, 2 mM DTT), protein was eluted using Buffer H (20 mM Tris pH 7.5, 0.7 M NaCl, 2 mM DTT). Fractions were analysed by SDS-PAGE and those of interest were pooled and dialyzed overnight against storage Buffer I (20 mM Tris pH 7.5, 10% glycerol, 2 mM DTT, 0.5 mM EDTA). Protein for crystallography was concentrated to ~10 mg/mL using a centrifugal filter unit (50 kDa molecular mass cut off) and stored at -80°C.

5.5.3 SPR assays

The running buffer used for ligand attachment and analyte binding experiments was 10 mM HEPES buffer with 150 mM NaCl, 3 mM EDTA and 0.05% Tween20, adjusted to pH 7.4 with 2M NaOH. A GE CM5 (series S) sensor chip was primed with running buffer and preconditioned per the manufacturer’s recommendation with successive injections (2 × 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, respectively. The surface was then activated with an injection of 0.2 M EDC and 50 mM NHS (600 s, 10 µL/min). PCNA (5 µL, 12 mg/mL) was diluted into running buffer (245 µL). Only once the preactivation was complete was the protein further diluted to a final concentration of 25 µg/mL in 10 mM NaAc (~pH 4.6) by addition of PCNA/HEPES (50 µL) to a solution of 100 mM NaAc (50 µL) and water (400 µL). This solution was immediately injected over only one flow cell (10 µL/min) until ~1500 RU was reached at stabilisation. Both flow cells were then blocked with 1.0 M ethanolamine pH 8.5 (600 s, 10 µL/min). The chip was left to stabilise for two hours before sample injections commenced.

Peptides (approx. 2 mg) were dissolved in milliQ H_2O and centrifuged (7800 rpm, 10 min) to remove any particulate. The peptide stock concentration was determined by 205 nm absorbance (A_{205}), where 2 µL of the stock was further diluted in water (10-50 fold) and a measurement taken in triplicate with a Nanodrop2000 and baseline referenced to 750 nm absorbance. The ε_{205} for each peptide was calculated using an online calculator (http://nickanthis.com/tools/a205.html)[30]; however an additional glycine residue was added to each peptide sequence to account for the terminal amide of the synthesised peptides. The peptide stock solution concentration was then calculated per $c = (A_{205}/\varepsilon_{205} \times l) \times DF$ where concentration is in molar, A_{205} is absorbance at 205 nm calculated as an average of three readings, l is the pathlength in cm (0.1 cm for Nanodrop), and ε_{205} is the molar absorptivity at 205 nm and DF is the dilution factor. The stock concentrations are tabulated in Table S2. The peptides were then diluted into running buffer before further dilution as necessary.

Steady state affinity experiments were conducted at a flow rate of 30 µL/min, with a starting contact time of 40 s and dissociation of 60 s, and extended if a steady state could not be reached. A 1 in 2 serial dilution, 8-times, was performed for each peptide and the resulting solutions were injected sequentially from lowest to highest concentration, preceded by a buffer only blank injection. Following each injection, the surface was regenerated with 2 M NaCl (2 × 30 s, 25 µL/min). After an optimal concentration range was found, the series of injections were repeated to ensure reproducibility. The top concentration for the final concentration range for each peptide is listed in Table S2. All data was analysed using the GE Biosystems Biacore S200 Evaluation Software. All data are summarised in Table S2 and Figure 2.

5.5.4 Protein-peptide co-crystallisation experiments

PCNA was mixed with peptide of interest at 1:1.2 molar ratio, and after incubation on ice for 30 minutes, the sample was pelleted at 16,000×*g* for 10 min to remove aggregates. The supernatant containing peptide bound protein was stored at -80°C. Crystals were grown by hanging drop vapor diffusion method in 24-well limbro plates containing 500 µL well solution, by mixing 1 µL protein and peptide with equal volume of well solution.[31-33] Initial co-crystallisation screens with all p21$_{\mu}$ modified peptides and PCNA were attempted. Diffracting crystals of PCNA bound to **p21$_{\mu}$** (residues 141-150) were formed in 8% tacsimate and 20% PEG at 16°C after 4 weeks. Diffracting crystals of PCNA bound to **p21$_{\mu}$-F150Y** were formed in 0.18 M magnesium acetate and 20% PEG at room temperature after 8 weeks. Crystals were mounted on cryoloops, cryoprotected using paratone-N, and flash cooled in liquid nitrogen.[31-33] Data was collected at 100 K using the MX1 beamline at the Australian Synchrotron (Clayton, Vic).[34] Diffraction data was indexed and integrated using XDS (X-ray Detector Software).[35] Pointless (CCP4i)[36] was used to create a mtz file for scaling. Data was scaled using Aimless (CCP4i)[37, 38] to a resolution of 3.30 Å for **p21$_{\mu}$** and 2.43 Å for **p21$_{\mu}$-F150Y**. Phasing was solved by molecular replacement using Phaser MR (CCP4i)[39] using a search model for **p21$_{\mu}$** of (PDB: 1AXC, human)[15] and for **p21$_{\mu}$-F150Y** (PDB: 4RJF, human).[22] Solutions were refined in phenix.refine[40] in iterative rounds with manual rebuilding in Coot.[41, 42] Data collection and refinement statistics for PCNA in complex with **p21$_{\mu}$** or **p21$_{\mu}$-F150Y** are summarised in Table S3. The final structures are deposited on the RCSB database under accession numbers 7KQ1 and 7KQ0, respectively.

5.5.5 Computational modelling

Models of PCNA and p21 analogues peptide structures were constructed using the solved structure of PCNA bound with **p21$_{\mu}$** peptide as a starting template (PDB: 7KQ1). The peptides analysed were: **p21$_{\mu}$-F150Y**, **p21$_{\mu}$-S146R**, **p21$_{\mu}$-M147I**, **p21$_{\mu}$-D149E**, **p21$_{\mu}$-FY150/151YF**, **p21$_{\mu}$-PARG**, **p21$_{\mu}$-Pogo**, **p21$_{\mu}$-pol δ$_{p66}$**, **p21$_{\mu}$-pol ι**, **p21$_{\mu}$-RFC**, **p21$_{\mu}$-RD1**, **p21$_{\mu}$-RD2**, **p21$_{\mu}$-RD3**. The residue(s) being investigated were mutated to the amino-acid of interest, and unresolved side-chains of residues were modelled into the computational structure. Energy minimisation/annealing (n=30) for refinement was carried out in ICM-Pro Molsoft.[43, 44] Refined models were analysed using PyMOL to validate the model by comparing against the **p21$_{\mu}$** structure, and assess side-chain interactions.[45] The resulting structures were visualised in PyMOL and are depicted in Figures S4-S15. Additional analysis was carried out using the RING server[46] and PoseView.[47]

5.6 Data availability

Atomic coordinates and structure factors for the reported crystal structures have been deposited with the RCSB Protein Data bank under accession numbers 7KQ1 and 7KQ0.

5.7 Acknowledgements

This research was undertaken in part using the MX1 beamline at the Australian Synchrotron, part of ANSTO. The facilities of the OptoFab node of the Australian National Fabrication Facility (ANFF) and associated Commonwealth and SA State Government funding are also gratefully acknowledged.

5.8 Funding

The research was supported by the Australian Research Council Centre of Excellence in Nanoscale BioPhotonics (CNBP) (CE140100003). A.J.H. and B.A.V. are supported by Australian Government Research Training Program Stipends (RTPS).

5.9 Conflicts of interest

We declare no conflicts of interest with the contents of this article.

REFERENCES

1. A. De Biasio and F. J. Blanco, Proliferating cell nuclear antigen structure and interactions: too many partners for one dancer?, *Adv. Prot. Chem. Struct. Bio.*, 2013, **91**, 1-36.
2. T. Tsurimoto, PCNA Binding Proteins, *Front. Biosci.*, 1999, **4**, d849-858.
3. G. Maga and U. Hubscher, Proliferating cell nuclear antigen (PCNA): A dancer with many partners, *J. Cell Sci.*, 2003, **116**, 3051-3060.
4. G. L. Moldovan, B. Pfander and S. Jentsch, PCNA, the maestro of the replication fork, *Cell*, 2007, **129**, 665-679.
5. E. M. Boehm, M. S. Gildenberg and M. T. Washington, The Many Roles of PCNA in Eukaryotic DNA Replication, *Enzymes*, 2016, **39**, 231-254.
6. I. Stoimenov and T. Helleday, PCNA on the crossroad of cancer, *Biochem. Soc. Trans.*, 2009, **37**, 605-613.
7. D. Zhongyun, M. Wortman, Z. Tan and K. Dillehay, WO 2012/033938 A2, *Identification of PCNA Targeting Compounds for Cancer Therapy and PCNA Function Regulation,* 2012.
8. M. De March, N. Merino, S. Barrera-Vilarmau, R. Crehuet, S. Onesti, F. J. Blanco and A. De Biasio, Structural basis of human PCNA sliding on DNA, *Nat. Comm.*, 2017, **8**, 13935.
9. J. Majka and P. M. J. Burgers, The PCNA–RFC Families of DNA Clamps and Clamp Loaders, *Prog. Nucleic Acid Res. Mol. Biol.*, 2004, **78**, 227-260.
10. A. J. Horsfall, A. D. Abell and J. B. Bruning, Targeting PCNA with Peptide Mimetics for Therapeutic Purposes, *ChemBioChem*, 2019, **21**, 442-450.
11. A. Prestel, N. Wichmann, J. M. Martins, R. Marabini, N. Kassem, S. S. Broendum, M. Otterlei, O. Nielsen, M. Willemoes, M. Ploug, W. Boomsma and B. B. Kragelund, The PCNA interaction motifs revisited: thinking outside the PIP-box, *Cell. Mol. Life Sci.*, 2019, **76**, 4923–4943.
12. A. Gonzalez-Magana and F. J. Blanco, Human PCNA Structure, Function and Interactions, *Biomolecules*, 2020, **10**, 570.
13. K. L. Wegener, A. E. McGrath, N. E. Dixon, A. J. Oakley, D. B. Scanlon, A. D. Abell and J. Bruning, Rational design of a 310-helical PIP-box mimetic targeting PCNA - the human sliding clamp, *Chem. Eur. J.*, 2018, **24**, 11325-11331.
14. J. B. Bruning and Y. Shamoo, Structural and thermodynamic analysis of human PCNA with peptides derived from DNA polymerase-delta p66 subunit and flap endonuclease-1, *Structure*, 2004, **12**, 2209-2219.
15. J. M. Gulbis, Z. Kelman, J. Hurwitz, M. O'Donnell and J. Kuriyan, Structure of the C-Terminal Region of p21 WAF1/CIP1 Complexed with Human PCNA, *Cell*, 1996, **87**, 297-306.
16. A. Hishiki, H. Hashimoto, T. Hanafusa, K. Kamei, E. Ohashi, T. Shimizu, H. Ohmori and M. Sato, Structural basis for novel interactions between human translesion synthesis polymerases and proliferating cell nuclear antigen, *J. Biol. Chem.*, 2009, **284**, 10552-10560.
17. E. M. Boehm and M. T. Washington, R.I.P. to the PIP: PCNA-binding motif no longer considered specific, *BioEssays*, 2016, **38**, 1117-1122.
18. E. Warbrick, D. P. Lane, D. M. Glover and L. S. Cox, A small peptide inhibitor of DNA replication defines the site of interaction between the cyclin-dependent kinase inhibitor p21 WAF1 and proliferating cell nuclear antigen, *Curr. Biol.*, 1995, **5**, 275-282.
19. E. Warbrick, PCNA binding through a conserved motif, *BioEssays*, 1998, **20**, 195-199.
20. D. I. Zheleva, N. Z. Zhelev, P. M. Fischer, S. V. Duff, E. Warbrick, D. G. Blake and D. P. Lane, A Quantitative Study of the in Vitro Binding of the C-Terminal Domain of p21 to PCNA: Affinity, Stoichiometry, and Thermodynamics, *Biochemistry*, 2000, **39**, 7388-7397.
21. E. Warbrick, A functional analysis of PCNA-binding peptides derived from protein sequence, interaction screening and rational design, *Oncogene*, 2006, **25**, 2850-2859.
22. A. J. Kroker and J. B. Bruning, p21 Exploits Residue Tyr151 as a Tether for High-Affinity PCNA Binding, *Biochemistry*, 2015, **54**, 3483-3493.
23. T. Kaufmann, I. Grishkovskaya, A. A. Polyansky, S. Kostrhon, E. Kukolj, K. M. Olek, S. Herbert, E. Beltzung, K. Mechtler, T. Peterbauer, J. Gotzmann, L. Zhang, M. Hartl, B. Zagrovic, K. Elsayad, K. Djinovic-Carugo and D. Slade, A novel non-canonical PIP-box mediates PARG interaction with PCNA, *Nucleic Acids Res.*, 2017, **45**, 9741-9759.
24. G. Kontopidis, S.-Y. Wu, D. I. Zheleva, P. Taylor, C. McInnes, D. P. Lane, P. M. Fischer and M. D. Walkinshaw, Structural and biochemical studies of human proliferating cell nuclear antigen complexes provide a rationale for cyclin association and inhibitor design, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 1871-1876.
25. K. N. Choe and G.-L. Moldovan, Forging Ahead through Darkness: PCNA, Still the Principal Conductor at the Replication Fork, *Mol. Cell*, 2016, **65**, 380-392.
26. A. Hayashi, N. N. Giakoumakis, T. Heidebrecht, T. Ishii, A. Panagopoulos, C. Caillat, M. Takahara, R. G. Hibbert, N. Suenaga, M. Stadnik-Spiewak, T. Takahashi, Y. Shiomi, S. Taraviras, E. von Castelmur, Z. Lygerou, A. Perrakis and H. Nishitani, Direct binding of Cdt2 to PCNA is important for targeting the CRL4Cdt2 E3 ligase activity to Cdt1, *Life Sci. Alliance*, 2018, **1**, e201800238.
27. M. De March, S. Barrera-Vilarmau, E. Crespan, E. Mentegari, N. Merino, A. Gonzalez-Magana, M. Romano-Moreno, G. Maga, R. Crehuet, S. Onesti, F. J. Blanco and A. De Biasio, p15PAF binding to PCNA modulates the DNA sliding surface, *Nucleic Acids Res.*, 2018, **46**, 9816–9828.

28. C. M. Duffy, B. J. Hilbert and B. A. Kelch, A Disease-Causing Variant in PCNA Disrupts a Promiscuous Protein Binding Site, *J. Mol. Biol.*, 2016, **428**, 1023-1040.

29. G. B. Fields and R. L. Noble, Solid phase peptide synthesis utilizing 9-fluorenylmethoxycarbonyl amino acids, *Int. J. Pept. Protein Res.*, 1990, **35**, 161-214.

30. N. J. Anthis and G. M. Clore, Sequence-specific determination of protein and peptide concentrations by absorbance at 205 nm, *Protein Sci.*, 2013, **22**, 851-858.

31. A. C. Marshall, A. J. Kroker, L. A. Murray, K. Gronthos, H. Rajapaksha, K. L. Wegener and J. B. Bruning, Structure of the sliding clamp from the fungal pathogen Aspergillus fumigatus (AfumPCNA) and interactions with Human p21, *FEBS J.*, 2017, **284**, 985-1002.

32. R. L. Frkic, B. S. Chua, Y. Shin, B. D. Pascal, S. J. Novick, T. M. Kamenecka, P. R. Griffin and J. B. Bruning, Structural and Dynamic Elucidation of a Non-acid PPARγ Partial Agonist: SR1988, *Nuclear Receptors Reseach*, 2018, **5**, 101350.

33. J. L. Pederick, A. P. Thompson, S. G. Bell and J. B. Bruning, d-Alanine-d-alanine ligase as a model for the activation of ATP-grasp enzymes by monovalent cations, *J. Biol. Chem.*, 2020, **295**, 7894-7904.

34. N. P. Cowieson, D. Aragao, M. Clift, D. J. Ericsson, C. Gee, S. J. Harrop, N. Mudie, S. Panjikar, J. R. Price, A. Riboldi-Tunnicliffe, R. Williamson and T. Caradoc-Davies, MX1: a bending-magnet crystallography beamline serving both chemical and macromolecular crystallography communities at the Australian Synchrotron, *J. Synchrotron Rad.*, 2015, **22**, 187-190.

35. W. Kabsch, XDS (X-Ray Detector Software), *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2010, **66**, 125-132.

36. P. Evans, Scaling and assessment of data quality, *Acta Crystallogr D Biol Crystallogr*, 2006, **62**, 72-82.

37. M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin and K. S. Wilson, Overview of the CCP4 suite and current developments, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2011, **67**, 235-242.

38. E. Potterton, P. J. Briggs, M. G. W. Turkenburg and E. Dodson, A graphical user interface to the CCP4 program suite., *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2003, **59**, 1131-1137.

39. A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storonia and R. J. Reada, Phaser crystallographic software, *J. Appl. Crystallogr.*, 2007, **40**, 658-674.

40. D. Liebschner, P. V. Afonine, M. L. Baker, G. Bunkóczi, V. B. Chen, T. I. Croll, B. Hintze, L. W. Hung, S. Jain, A. J. McCoy, N. W. Moriarty, R. D. Oeffner, B. K. Poon, M. G. Prisant, R. J. Read, J. S. Richardson, D. C. Richardson, M. D. Sammito, O. V. Sobolev, D. H. Stockwell, T. C. Terwilliger, A. G. Urzhumtsev, L. L. Videau, C. J. Williams and P. D. Adams, Phenix : Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2019, **75**, 861-877.

41. P. V. Afonine, R. W. Grosse-Kunstleve, N. Echols, J. J. Headd, N. W. Moriarty, M. Mustyakimov, T. C. Terwilliger, A. Urzhumtsev, P. H. Zwart and P. D. Adams, Towards automated crystallographic structure refinement with phenix.refine, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2012, **68**, 352-367.

42. P. Emsley and K. Cowtan, Coot: model-building tools for molecular graphics, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2004, **60**, 2126-2132.

43. R. Abagyan, M. Totrov and D. Kuznetsov, ICM - New Method for Protein Modeling and Design: Applications to Docking and Structure Prediction from the Distorted Native Conformation, *J. Comput. Chem.*, 1994, **15**, 488-506.

44. R. Abagyan and M. Totrov, Biased Probability Monte Carlo Conformational Searches and Electrostatic Calculations for Peptides and Proteins, *J. Mol. Biol.*, 1994, **235**, 983-1002.

45. Schrodinger, LLC, unpublished work.

46. D. Piovesan, G. Minervini and S. C. Tosatto, The RING 2.0 web server for high quality residue interaction networks, *Nucleic Acids Res.*, 2016, **44**, W367-374.

47. K. Stierand, P. C. Maass and M. Rarey, Molecular complexes at a glance: automated generation of two-dimensional complex diagrams, *Bioinformatics*, 2006, **22**, 1710-1716.

Supplemental Data.

UNLOCKING THE PIP-BOX: A PEPTIDE LIBRARY REVEALS INTERACTIONS THAT DRIVE HIGH-AFFINITY BINDING TO HUMAN PCNA

S5.1 Synthesis and characterisation of peptides

Unless otherwise indicated, all starting materials were purchased from commercial sources and used without further purification. All peptides were synthesised by the Fmoc/tBu solid-phase peptide synthesis using one of three protocols detailed below, with all L-amino-acids (unless otherwise specified) purchased from Chem-Impex Int'l.: Fmoc-Ala-OH, Fmoc-Arg(Pbf)-OH, Fmoc-Asp(tBu)-OH, Fmoc-Asn(Trt)-OH, Fmoc-Cys(Trt)-OH, Fmoc-Gln(Trt)-OH, Fmoc-Glu(tBu)-OH, Fmoc-Gly-OH, Fmoc-His(Trt)-OH, Fmoc-Ile-OH, Fmoc-Leu-OH, Fmoc-Lys(Boc)-OH, Fmoc-Met-OH, Fmoc-Thr(tBu)-OH, Fmoc-Tyr(tBu)-OH, Fmoc-Trp(Boc)-OH, Fmoc-Ser(tBu)-OH, Fmoc-Phe-OH, Fmoc-Pro-OH, Fmoc-Val-OH. Peptides were subsequently cleaved from the resin (and simultaneously globally deprotected) and purified per '*Cleavage from solid-support, isolation and purification'*. Purity of all compounds was confirmed by analytical RP-HPLC on an Agilent 1260 HPLC equipped with a Phenomenex Luna C18(2) column (250 x 4.6 mm) over a gradient of 5-50% B (15 min) and visualised at 220 nm. High-resolution mass spectra were collected using an Agilent 6230 ESI-TOF via direct injection in ACN with 0.1% formic acid as the running buffer. Characterisation data for all peptides in recorded in Table S1.

S5.1.1 Peptide synthesis

All peptides were synthesised using one of the following methods, excepting ten peptides that were purchased from Shanghai Royobiotech at >95% purity, with purity and identity confirmed on receipt by HPLC and HRMS. *Purchased peptides:* **p21$_{\mu}$-M147L**, **p21$_{\mu}$-M147A**, **p21$_{\mu}$-M147W**, **p21$_{\mu}$-M147V**, **p21$_{\mu}$-Q144D**, **p21$_{\mu}$-Q144M**, **p21$_{\mu}$-Q144S**, **p21$_{\mu}$-FY150/151YF**, **p21$_{\mu}$-p15**, and **p21$_{\mu}$-Cdt1**.

Liberty Blue synthesis

All peptides were synthesised by the Fmoc/tBu solid-phase peptide synthesis protocol on a CEM Liberty Blue Automated Microwave Peptide Synthesiser (CEM Corp., Matthews, NC, USA) using the standard manufacturer's conditions. The peptides were assembled on 0.1 mmol scale on Chem Impex Rink Amide AM resin (0.47 mmol/g) or Mimotopes Rink Amide resin (0.456 mmol/g). The resin was initially swollen in DCM (10 mL, 15 min) and then the resin washed with DMF (2 × 5 mL) and transferred to the microwave reaction vessel. The resin-bound Fmoc-groups were deprotected with a mixture of 20% piperidine and 0.1 M OxymaPure in DMF using the standard microwave deprotection method with a maximum temperature of 90°C. Couplings were performed with Fmoc-protected amino-acids (0.2 M in DMF, 5 equiv), OxymaPure (1 M in DMF, 5 equiv) and DIC (0.5 M in DMF, 5 equiv) under the 'Standard Coupling' microwave method with a maximum temperature of 90°C, except for coupling of Fmoc-His(Trt)-OH which was coupled using a maximum 50°C 10 min coupling procedure; and Fmoc-Arg(Pbf)-OH which used the default 'Arginine Double Coupling' microwave method which included two couplings steps – the first at room temperature and the second at a maximum of 75°C. Following assembly of the desired sequence the *N*-terminal protecting group was removed. The resin was then removed from the synthesiser, washed with DCM (3 × 5 mL) and then diethyl ether (3 × 5 mL) and air dried with suction. The peptides were then cleaved from the resin according to the protocol '*Cleavage from solid-support, isolation and purification'* detailed below.

Peptides synthesised by this method: **p21$_{139\text{-}160}$**‡, **p21$_{\mu}$**, **p21$_{\mu}$-F150Y**, **p21$_{\mu}$-Y151F**, **p21$_{\mu}$-F150H**, **p21$_{\mu}$-M147I**

‡ The *N*-terminal glycine was subjected to a standard double coupling cycle as initial syntheses indicated incomplete incorporation at this position. Please note, this sequence is particularly prone to aspartimide formation.

Prelude synthesis

Peptides were assembled using a Protein Technologies Prelude® Peptide synthesiser. Chem Impex Rink Amide resin (0.1 mmol, 0.47 mmol/g) was preswollen in DCM for 15 min, and washed with DMF (3 × 10 mL, 1.5 min). The Fmoc-protecting group was removed by treatment with 20% piperidine (3 × 8 mL, 7 min) and the resin washed with DMF (5 × 10 mL, 1.5 min). An amino-acid was then coupled by addition of the required Fmoc-amino-acid (7.5 equiv, 150 mM in DMF), HCTU (5 equiv, 0.5 M in DMF) and DIPEA (10 equiv, 1 M in DMF) and bubbled with nitrogen for 15 min before the solution was drained, and a fresh coupling solution added to the resin and bubbled with nitrogen for 10 min. The solution was drained and the resin washed with DMF (4 × 8 mL, 1.5 min). Fmoc deprotection and coupling steps were repeated until the desired sequence was achieved, and the final *N*-terminal Fmoc group deprotected. The resin was then removed from the synthesiser and washed with DCM (3 × 5 mL) and then diethyl ether (3 × 5 mL) and air dried with suction. The peptides were then cleaved from the resin according to the protocol '*Cleavage from solid-support, isolation and purification'* detailed below.

Peptides synthesised by this method: **p21$_\mu$-Q144K**, **p21$_\mu$-Q144N**, **p21$_\mu$-T145K**, **p21$_\mu$-T145R**, **p21$_\mu$-S146K**, **p21$_\mu$-S146R**, **p21$_\mu$-T148D**, **p21$_\mu$-T148E**, **p21$_\mu$-D149E**, **p21$_\mu$-TS145/146KK**, **p21$_\mu$-TS145/146RK**, **p21$_\mu$-TS145/145KR**, **p21$_\mu$-TS145/146RR**, **p21$_\mu$-TD148/149EE**, **p21$_\mu$-TS148/149DE**, **p21$_\mu$-RNaseH2B**, **p21$_\mu$-MCMT**, **p21$_\mu$-pol λ**, **p21$_\mu$-pol β**, **p21$_\mu$-pol δ$_{p66}$**, **p21$_\mu$-FEN1**, **p21$_\mu$-RFC**, **p21$_\mu$-RecQ5**, **p21$_\mu$-DNALig1**, **p21$_\mu$-WRN**, **p21$_\mu$-XPG**, **p21$_\mu$-Cdt2**, **p21$_\mu$-Pogo**, **p21$_\mu$-pol η**, **p21$_\mu$-pol κ**.

Manual synthesis

Rink Amide PL resin (0.1 mmol, 322 mg, 0.31 mmol/g, Agilent) was swollen in 1:1 DMF/DCM (15 mL) for 15 min. The Fmoc-protecting group was removed by treatment of the resin with a solution of 20% piperidine and 0.1 M HOBt in DMF (5 mL) for 15 min. The solution was drained and the resin washed with DMF (3 × 5 mL). Amino-acid couplings were achieved by addition of a solution of Fmoc-protected amino-acid (5 equiv), HATU (5 equiv) and DIPEA (10 equiv) in DMF (5 mL), to the resin and stirred intermittently for 1 h. The solution was drained and the resin washed with DMF (5 × 5 mL). The *N*-terminal Fmoc-protecting group was removed by treatment of the resin with a solution of 20% piperidine and 0.1 M HOBt in DMF (5 mL) for 10 min, the solution was drained and the resin washed with DMF (5 × 5 mL). A TNBS test* was used to verify each coupling (negative/colourless) and deprotection (positive/red) step, with steps repeated as necessary. Successive couplings and Fmoc-deprotections were repeated to achieve the desired sequence. The resin was washed with DCM (3 × 5 mL) and then diethyl ether (3 × 5 mL) and air dried with suction. The peptides were then cleaved from the resin according to the protocol '*Cleavage from solid-support, isolation and purification'* detailed below.

* *TNBS test:* A small spatula of swollen resin taken out and 1 drop each of TNBS (100 μL 5% w/v in H_2O added to 900 μL of DMF) and DIPEA solutions (100 μL in 900 μL of DMF) added and allowed to develop for 1 min. Clear/yellow beads indicated no free amine (negative), while red/orange beads showed free amine was present (positive).

Peptides synthesised by this method: **p21$_\mu$-pol ι**, **p21$_\mu$-PARG**, **p21$_\mu$-RD1**, **p21$_\mu$-RD2**, **p21$_\mu$-RD3**.

Cleavage from solid-support, isolation and purification

Following complete assembly of the peptide and deprotection of the final Fmoc group, the peptides were subsequently cleaved from the resin and the side-chain protecting groups simultaneously globally deprotected by treatment of the resin with 92.5/2.5/2.5/2.5 TFA:TIPS:DODT:H_2O (5 mL) for 2 hours. The cleavage mixture was pipetted from the resin and concentrated under a stream of nitrogen to 0.5-1 mL. The peptide was then precipitated by addition of diethyl ether (10 mL) and the mixture cooled to -20°C. The precipitate was pelleted by centrifugation (7600 rpm, 10 min), and the supernatant decanted. The pellet was dried under a nitrogen stream, and then dissolved in 1:1 ACN/H_2O, before being syringe filtered (0.2 μm) and lyophilised to yield the crude peptide as a fluffy white powder. The peptides were purified by semi-preparative RP-HPLC on a Gilson GX-Prep system using a Phenomenex Luna C18(2) column (10 × 250 mm), over a linear ACN/H_2O gradient optimised for each peptide sample. RP-HPLC solvents were (A) H_2O with 0.1% TFA and (B) ACN with 0.1% TFA. The product containing fractions were pooled and lyophilised. The identity of the final compounds was confirmed by High Resolution Mass Spectrometry on an Agilent 6230 ESI-TOF LCMS. Purity of the peptides was confirmed by analytical RP-HPLC on an Agilent 1260 HPLC equipped with a Phenomenex Luna C18(2) column (250 × 4.6 mm) over a gradient of 5-50% B (15 min) and visualised at 220 nm. Characterisation data for all peptides in recorded in Table S1.

Table S1: Peptide characterisation data. All peptides are *C*-terminally amidated. % Purity calculated at 220 nm.

Name	Sequence	Mw	MF	M+ Calc	[M+4]$^{4+}$ Calc	ESI+ [M+4]^{4} Found	Purity %
p21$_{\mu}$ (141-155)	^{141}KRRQTSMTDFYHSKR	1940.20	$C_{82}H_{134}N_{30}O_{23}S_1$	1938.9959	485.7570	485.7566	96.3
p21 (139-160)	^{139}GRKRRQTSMTDFYHSKRRLIFS	2770.20	$C_{120}H_{197}N_{43}O_{31}S_1$	2769.4722	693.3760	693.3793	97.4
p21$_{\mu}$-Q144D	KRRDTSMTDFYHSKR	1927.16	$C_{81}H_{131}N_{29}O_{24}S_1$	1925.9642	482.4991	482.5010	98.8
p21$_{\mu}$-Q144M	KRRMTSMTDFYHSKR	1943.27	$C_{82}H_{135}N_{29}O_{22}S_2$	1941.9778	486.5024	486.5020	98.9
p21$_{\mu}$-Q144S	KRRSTSMTDFYHSKR	1899.15	$C_{80}H_{131}N_{29}O_{23}S_1$	1897.9693	475.5003	475.4999	97.3
p21$_{\mu}$-Q144K	KRRKTSMTDFYHSKR	1940.24	$C_{83}H_{138}N_{30}O_{22}S_1$	1939.0323	485.7661	485.7658	93.4
p21$_{\mu}$-Q144N	KRRNTSMTDFYHSKR	1926.18	$C_{81}H_{132}N_{30}O_{23}S_1$	1924.9802	482.2531	482.2527	90.7
p21$_{\mu}$-T145K	KRRQKSMTDFYHSKR	1967.27	$C_{84}H_{139}N_{31}O_{22}S_1$	1966.0432	492.5188	492.5206	90.2
p21$_{\mu}$-T145R	KRRQRSMTDFYHSKR	1995.28	$C_{84}H_{139}N_{33}O_{22}S_1$	1994.0493	499.5203	499.5215	91.0
p21$_{\mu}$-S146K	KRRQTKMTDFYHSKR	1981.30	$C_{85}H_{141}N_{31}O_{22}S_1$	1980.0588	496.0227	496.0235	87.0
p21$_{\mu}$-S146R	KRRQTRMTDFYHSKR	2009.31	$C_{85}H_{141}N_{33}O_{22}S_1$	2008.0650	503.0243	503.0250	94.0
p21$_{\mu}$-TS145/146KK	KRRQKKMTDFYHSKR	2008.37	$C_{87}H_{146}N_{32}O_{21}S_1$	2007.1061	3+: 670.0492	670.3799	91.3
p21$_{\mu}$-TS145/146RK	KRRQRKMTDFYHSKR	2036.38	$C_{87}H_{146}N_{34}O_{21}S_1$	2035.1122	509.7861	509.7864	93.0
p21$_{\mu}$-TS145/146KR	KRRQKRMTDFYHSKR	2036.38	$C_{87}H_{146}N_{34}O_{21}S_1$	2035.1122	509.7861	509.7859	91.9
p21$_{\mu}$-TS145/146RR	KRRQRRMTDFYHSKR	2064.39	$C_{87}H_{146}N_{36}O_{21}S_1$	2063.1184	516.7876	516.7877	90.9
p21$_{\mu}$-M147I	KRRQTSITDFYHSK	1765.98	$C_{77}H_{124}N_{26}O_{22}$	1764.9383	442.2426	442.2419	99.1
p21$_{\mu}$-M147L	KRRQTSLTDFYHSKR	1922.16	$C_{83}H_{136}N_{30}O_{23}$	1921.0395	481.2679	481.2672	99.8
p21$_{\mu}$-M147A	KRRQTSATDFYHSKR	1880.08	$C_{80}H_{130}N_{30}O_{23}$	1878.9925	470.7561	470.7556	99.2
p21$_{\mu}$-M147W	KRRQTSWTDFYHSKR	1995.22	$C_{88}H_{135}N_{31}O_{23}$	1994.0347	499.5167	499.5161	99.5
p21$_{\mu}$-M147V	KRRQTSVTDFYHSKR	1908.14	$C_{82}H_{134}N_{30}O_{23}$	1907.0238	477.7640	477.7632	98.5
p21$_{\mu}$-T148D	KRRQTSMDDFYHSKR	1954.19	$C_{82}H_{132}N_{30}O_{24}S_1$	1952.9751	489.2518	489.2514	91.5
p21$_{\mu}$-T148E	KRRQTSMEDFYHSKR	1968.21	$C_{83}H_{134}N_{30}O_{24}S_1$	1966.9908	492.7557	492.7562	90.4
p21$_{\mu}$-D149E	KRRQTSMTEFYHSKR	1954.23	$C_{83}H_{136}N_{30}O_{23}S_1$	1953.0115	489.2609	489.2620	88.5
p21$_{\mu}$-TD148/149EE	KRRQTSMEEFYHSKR	1982.24	$C_{84}H_{136}N_{30}O_{24}S_1$	1981.0065	496.2596	496.2595	89.4
p21$_{\mu}$-TS148/149DE	KRRQTSMDEFYHSKR	1968.21	$C_{83}H_{134}N_{30}O_{24}S_1$	1966.9908	492.7557	492.7555	90.1
p21$_{\mu}$-F150H	KRRQTSMTDHYHSKR	1930.16	$C_{79}H_{132}N_{32}O_{23}S_1$	1928.9864	483.2546	483.2542	94.4
p21$_{\mu}$-F150Y	KRRQTSMTDYYHSKR	1956.20	$C_{82}H_{134}N_{30}O_{24}S_1$	1954.9908	489.7557	489.7552	93.8
p21$_{\mu}$-Y151F	KRRQTSMTDFFHSKR	1924.20	$C_{82}H_{134}N_{30}O_{22}S_1$	1923.0010	481.7582	481.7576	94.2
p21$_{\mu}$-FY150/151YF	KRRQTSMTDYFHSKR	1940.20	$C_{82}H_{134}N_{30}O_{23}S_1$	1938.9959	485.7570	485.7589	98.1
p21$_{\mu}$-DNALig1	KRRQRSIMSFFHSKR	1963.32	$C_{85}H_{143}N_{33}O_{19}S_1$	1962.0959	491.5320	491.5311	98.2
p21$_{\mu}$-FEN1	KRRQGRLDDFFHSKR	1945.20	$C_{84}H_{137}N_{33}O_{21}$	1944.0667	487.0247	487.0241	99.2
p21$_{\mu}$-Pogo	KRRQKKITDYFHSKR	1990.32	$C_{88}H_{148}N_{32}O_{21}$	1989.1497	498.2954	498.2965	98.0
p21$_{\mu}$-XPG	KRRQLRIDSFFHSKR	1973.29	$C_{87}H_{145}N_{33}O_{20}$	1972.1344	494.0416	494.0409	99.1
p21$_{\mu}$-MCMT	KRRQTTITSHFHSKR	1882.14	$C_{80}H_{136}N_{32}O_{21}$	1881.0558	471.2719	471.2714	93.9
p21$_{\mu}$-p15	KRRQKGIGEFFHSKR	1873.18	$C_{83}H_{137}N_{31}O_{19}$	1872.0707	469.0257	469.0265	96.5
p21$_{\mu}$-Cdt1	KRRQRRVTDFFHSKR	2016.33	$C_{87}H_{146}N_{36}O_{20}$	2015.1514	504.7959	504.7954	94.8
p21$_{\mu}$-pol δ$_{p66}$	KRRQVSITGFFHSKR	1846.15	$C_{82}H_{136}N_{30}O_{19}$	1845.0598	462.2729	462.2727	99.2
p21$_{\mu}$-RecQ5	KRRQNLIRHFFHSKR	2022.37	$C_{90}H_{148}N_{36}O_{18}$	2021.1772	506.3024	506.3019	99.2
p21$_{\mu}$-WRN	KRRQWKLLRDFHSKR	2053.42	$C_{92}H_{153}N_{35}O_{19}$	2052.2082	514.0600	514.0593	99.3
p21$_{\mu}$-PARG	KRRDSKITDHFHSKR	1910.16	$C_{81}H_{136}N_{32}O_{22}$	1909.0507	478.2707	478.2705	98.4
p21$_{\mu}$-pol β	KRRQLQKVHFHSKR	1847.18	$C_{81}H_{139}N_{33}O_{17}$	1846.1027	462.5337	462.5330	96.5
p21$_{\mu}$-pol λ	KRRSVPVLELFHSKR	1851.21	$C_{83}H_{143}N_{29}O_{19}$	1850.1115	463.5359	463.5353	92.4
p21$_{\mu}$-pol ι	KRRKGLIDYYLHSKR	1932.29	$C_{87}H_{146}N_{30}O_{20}$	1931.1330	483.7913	483.7919	87.6
p21$_{\mu}$-pol κ	KRRKHTLDIFFHSKR	1968.32	$C_{89}H_{146}N_{32}O_{19}$	1967.1442	492.7941	492.7939	93.5
p21$_{\mu}$-RFC$_{p14}$	KRRMDIRKFFHSKR	1904.30	$C_{84}H_{142}N_{32}O_{17}S_1$	1903.0951	476.7818	476.7813	93.4
p21$_{\mu}$-RNaseH2B	KRRMKSIDTFFHSKR	1936.29	$C_{85}H_{142}N_{30}O_{20}S_1$	1935.0737	484.7764	484.7761	97.7
p21$_{\mu}$-Cdt2	KRRMRKICTYFHSKR	2009.46	$C_{87}H_{149}N_{33}O_{18}S_2$	2008.1198	503.0380	503.0390	93.8
p21$_{\mu}$-pol n	KRRMQTLESFFHSKR	1950.28	$C_{85}H_{140}N_{30}O_{21}S_1$	1949.0530	488.2713	488.2719	96.6
p21$_{\mu}$-RD1	KRRQRKMEDYYHSKR	2080.39	$C_{88}H_{146}N_{34}O_{23}S_1$	2079.1021	520.7835	520.7836	91.0
p21$_{\mu}$-RD2	KRRQTRITEYFHSKR	2005.30	$C_{87}H_{145}N_{33}O_{22}$	2004.1242	502.0389	502.0396	97.1
p21$_{\mu}$-RD3	KRRQKRITEYYHSKR	2048.39	$C_{89}H_{150}N_{34}O_{22}$	2047.1664	512.7994	512.8001	92.3

S5.2 SPR

Table S2: Peptide SPR data against PCNA. Top Conc is the highest concentration of 8x 1 in 2 dilutions used to calculate the steady state affinity. K$_D$ is the affinity constant. SE, standard error; On/Off, indicate times for contact and dissociation phases of each run. All peptides are *C*-terminally amidated.

Name	Sequence	ε_{205}*	Top Conc (nM)	Affinity K$_D$ (nM)	K$_D$ SE (nM)	On/Off (s)
p21$_{\mu}$ (141-155)	^{141}KRRQTSMTDFYHSKR	67860	500	12.3	0.598	30/60
p21 (139-160)	^{139}GRKRRQTSMTDFYHSKRRLIFS	98620	50	4.319	1.31	60/90
p21$_{\mu}$-Q144D	KRRDTSMTDFYHSKR	67460	2000	714	30.4	40/60
p21$_{\mu}$-Q144M	KRRMTSMTDFYHSKR	69290	5000	1544	159	40/60
p21$_{\mu}$-Q144S	KRRSTSMTDFYHSKR	67460	5000	1032	116	40/60
p21$_{\mu}$-Q144K	KRRKTSMTDFYHSKR	67460	5000	1123	148	40/60
p21$_{\mu}$-Q144N	KRRNTSMTDFYHSKR	67860	5000	772	117	40/60
p21$_{\mu}$-T145K	KRRQKSMTDFYHSKR	67860	2000	98.05	10.8	40/60
p21μ-T145R	KRRQRSMTDFYHSKR	69210	1000	88.51	13.2	40/60
p21$_{\mu}$-S146K	KRRQTKMTDFYHSKR	67860	300	12.97	1.55	60/90
p21$_{\mu}$-S146R	KRRQTRMTDFYHSKR	69210	100	4.297	1.35	60/90
p21$_{\mu}$-TS145/146KK	KRRQKKMTDFYHSKR	67860	300	20.18	4.80	60/90
p21$_{\mu}$-TS145/146RK	KRRQRKMTDFYHSKR	69210	300	17.30	4.66	60/90
p21$_{\mu}$-TS145/146KR	KRRQKRMTDFYHSKR	69210	300	37.00	11.0	60/90
p21$_{\mu}$-TS145/146RR	KRRQRRMTDFYHSKR	70560	30	1.833	0.453	60/90
p21$_{\mu}$-M147I	KRRQTSITDFYHSK	66030	200	11.1	0.258	40/60
p21$_{\mu}$-M147L	KRRQTSLTDFYHSKR	66030	200	20.49	1.75	40/60
p21$_{\mu}$-M147A	KRRQTSATDFYHSKR	66030	20000	7591	626	40/60
p21$_{\mu}$-M147W	KRRQTSWTDFYHSKR	86430	5000	3566	165	40/60
p21$_{\mu}$-M147V	KRRQTSVTDFYHSKR	66030	500	29.29	1.83	40/60
p21$_{\mu}$-T148D	KRRQTSMDDFYHSKR	67860	2000	96.47	7.32	40/60
p21$_{\mu}$-T148E	KRRQTSMEDFYHSKR	67860	200	77.40	2.62	60/90
p21$_{\mu}$-D149E	KRRQTSMTEFYHSKR	67860	500	12.67	1.4	60/90
p21$_{\mu}$-TD148/149EE	KRRQTSMEEFYHSKR	67860	1000	48.52	1.16	60/90
p21$_{\mu}$-TD148/149DE	KRRQTSMDEFYHSKR	67860	2000	345.8	2.67	60/90
p21$_{\mu}$-F150H	KRRQTSMTDHYHSKR	64460	1000	159.0	62.8	40/60
p21$_{\mu}$-F150Y	KRRQTSMTDYYHSKR	65340	300	20.2	0.424	40/60
p21$_{\mu}$-Y151F	KRRQTSMTDFFHSKR	70380	300	10.61	1.48	60/90
p21$_{\mu}$-FY150/151YF	KRRQTSMTDYFHSKR	67860	30	2.175	0.531	60/90
p21$_{\mu}$-DNALig1	KRRQRSIMSFFHSKR	71730	250	40.74	9.20	60/90
p21$_{\mu}$-FEN1	KRRQGRLDDFFHSKR	69900	2000	165.6	2.07	40/60
p21$_{\mu}$-Pogo	KRRQKKITDYFHSKR	66030	100	8.816	3.30	60/90
p21$_{\mu}$-XPG	KRRQLRIDSFFHSKR	69900	200	16.39	4.81	60/90
p21$_{\mu}$-MCMT	KRRQTTITSHFHSKR	65150	2000	106.0	24.4	40/60
p21$_{\mu}$-p15	KRRQKGIGEFFHSKR	68550*	2000	234.0	42.4	40/60
p21$_{\mu}$-Cdt1	KRRQRRVTDFFHSKR	71250	150	8.760	2.22	40/60
p21$_{\mu}$-pol δ$_{p66}$	KRRQVSITGFFHSKR	68550	1000	123.7	4.24	40/60
p21$_{\mu}$-RecQ5	KRRQNLIRHFFHSKR	75500	20000	>40000**		
p21$_{\mu}$-WRN	KRRQWKLLRDFHSKR	81700	2000	1155	67.3	40/60
p21$_{\mu}$-PARG	KRRDSKITDHFHSKR	64750	2000	90.03	18.4	40/60
p21$_{\mu}$-pol β	KRRQLQKVHFHSKR	62770	20000	5732	822	40/60
p21$_{\mu}$-pol λ	KRRSVPVLELFHSKR	59550	20000	8144	766	40/60
p21$_{\mu}$-pol ι	KRRKGLIDYYLHSKR	63110	150	111.2	5.41	60/90
p21$_{\mu}$-pol κ	KRRKHTLDIFFHSKR	73350	10000	3296	353	40/60
p21$_{\mu}$-RFC$_{p14}$	KRRMDIRKFFHSKR	68550	1000	145.2	34.5	40/60
p21$_{\mu}$-Cdt2	KRRMRKICTYFHSKR	69500		NS***		
p21$_{\mu}$-pol η	KRRMQTLESFFHSKR	70380	10000	2537	745	40/60
p21$_{\mu}$-RNAseH2B	KRRMKSIDTFFHSKR	69980	2000	640.4	51.1	40/60
p21$_{\mu}$-RD1	KRRQRKMEDYYHSKR	66690	100	8.152	1.27	40/60
p21$_{\mu}$-RD2	KRRQTRITEYFHSKR	67380	50	1.118	0.330	60/90
p21$_{\mu}$-RD3	KRRQKRITEYYHSKR	64860	200	6.70	1.58	40/60

* Determined using online calculator, from Anthis 2013.[1]

** Peptide bound very non-specifically to protein where the RU significantly exceeded the expected maximum and a K$_D$ could not be determine.

*** Peptide bound significantly to the reference cell of sensor chip such that a K$_D$ could not be determined.

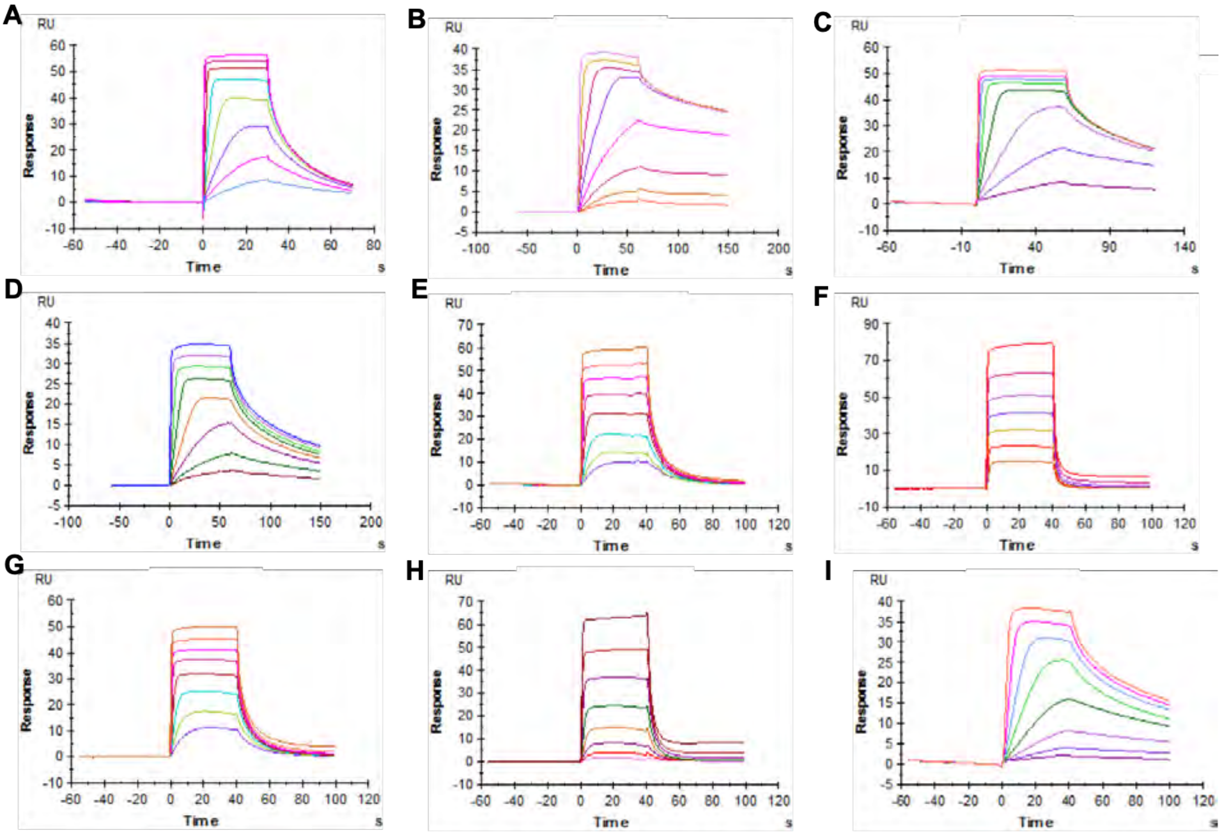


Figure S1: Representative sample of SPR sensorgrams. **A p21$_{\mu}$ B p21$_{\mu}$-S146R C p21$_{\mu}$-FY150/151YF D p21$_{\mu}$-Y151F E p21$_{\mu}$-FEN1 F p21$_{\mu}$-pol η G p21$_{\mu}$-PARG H p21$_{\mu}$-RNaseH2B I p21$_{\mu}$-RD1**

S5.3 Co-crystal experiments

Table S3: Data collection and refinement statistics of PCNA bound with **p21$_{\mu}$** (RCSB PDB ID: 7KQ1), and PCNA bound with **p21$_{\mu}$-F150Y** (RCSB PDB ID: 7KQ0). Statistics for the highest-resolution shell are shown in parentheses.

PDB ID	**7KQ1**	**7KQ0**
Wavelength	0.9537	0.9537
Resolution range	41.09 - 3.3 (3.418 - 3.3)	41.15 - 2.4 (2.486 - 2.4)
Space group	P 32 2 1	P 3
Unit cell	136.66 136.66 104.005 90 90 120	142.563 142.563 41.03 90 90 120
Unique reflections	17208 (1689)	32881 (3598)
Multiplicity	19.9 (18.2)	10.7 (11.1)
Completeness (%)	99.83 (99.64)	90.13 (99.94)
Mean I/sigma(I)	10.45 (0.91)	40.58 (13.50)
^{a} **R-merge**	0.2823 (4.453)	0.04334 (0.1841)
^{b} **Rpim**	0.065 (0.755)	0.014 (0.570)
CC1/2	0.999 (0.469)	1 (0.987)
Reflections used in refinement	17192 (1684)	32881 (3598)
Reflections used for R-free	1716 (171)	2031 (224)
^{c} **R-work**	0.2419 (0.3591)	0.1537 (0.2023)
^{d} **R-free**	0.2685 (0.3618)	0.2011 (0.2813)
Number of non-hydrogen atoms	6084	6086
macromolecules	6084	5976
Protein residues	816	807
RMS(bonds)	0.003	0.004
RMS(angles)	0.69	0.68
Ramachandran favoured (%)	94.78	98.10
Ramachandran allowed (%)	4.73	1.64
Ramachandran outliers (%)	0.50	0.25
Rotamer outliers (%)	0.00	8.80
Clashscore	8.80	3.15
Average B-factor	108.50	32.34

[a] $R_{merge} = \Sigma\,|I - <I>|/\Sigma I$.

[b] $Rpim = \Sigma h\,[1/(^{n}h - 1)]^{1/2}\,\Sigma i|<I_h> - I_{h,i}|/\Sigma_h\,\Sigma_i\,I_{h,i}\ ^{2}$

[c] $R_{work} = \Sigma|F_o - F_c|/\Sigma|F_o|$ for all data excluding data used to calculate R_{free}.

[d] $R_{free} = \Sigma|F_o - F_c|/\Sigma|F_o|$, for all data.

Table S4: Secondary Interaction Summary for co-crystal structure of **p21$_{\mu}$** with PCNA (PDB ID: 7KQ1) calculated using the RING server. Only peptide chain interactions reported (chains B, D and F). Interactions reported are an average of the number of interactions observed for the three chains. RING session ID: 5ef50f2b0e9f94078ea226bd

		Intermolecular		Intramolecular		
	Residue	**VDW**	**H-Bond**	**VDW**	**H-Bond**	**Total**
Fl	**141**	0	0	0	0	
	142	0.67	0	0	0	
	143	0	0.67	0	0	
PIP-box	*** 144**	1.67	0	0	0	
	145	2.33	1.00	0.67	0.33	
	146	0.67	0.33	0.33	1.00	
	*** 147**	6.67	1.00	0.33	1.67	
	148	0.33	0	0	0	
	149	0	0	0	0	
	*** 150**	2.67	0.00	0	0	
	*** 151**	2.00	0.67	0	0	
Fl	**152**	1.00	1.00	0	0	
	153	1.67	0.33	0	0	
	154	0.67	1.00	0	0	
	155	0	0	0	0	
	Total	20.3	6.0	1.3	3.0	30.67
	PIP-box	16	3	1	3	23.67
	Flanking (Fl)	4	3	0	0	7
	Conserved (*) PIP-box residues	11.00	1	0	2	14.00
	Non-conserved PIP-box residues	3.33	1.33	1.00	1.33	7.00
	Other: Intermolecular pi-stack between F-Tyr151 to E-Tyr133, and D-Tyr151 to C-Tyr250					

Table S5: Secondary Interaction Summary for co-crystal structure of **p21$_{\mu}$-F150Y** with PCNA (PDB ID: 7KQ0) calculated using the RING server. Only peptide chain interactions reported (chains B, D and F). Interactions reported are an average of the number of interactions observed for the three chains. RING session ID: 5f595dab0e9f94078ea22f07

		Intermolecular		Intramolecular		
	Residue	**VDW**	**H-Bond**	**VDW**	**H-Bond**	**Total**
Fl	**141**	0	0	0	0	
	142	1	0	0	0	
	143	0	1	0	0	
PIP-box	*** 144**	1	0	0	0	
	145	1	1	0	1	
	146	1	0	1	0	
	*** 147**	3	1	2	1	
	148	1	0	0	0	
	149	0	0	0	0	
	*** 150**	2	0	0	0	
	*** 151**	2	0	0	0	
Fl	**152**	0	1	0	0	
	153	0	0	0	0	
	154	0	0	0	0	
	155	1	0	0	0	
	Total	13.0	4.0	2.67	2.67	22.33
	PIP-box	11	2	3	3	18.33
	Flanking (Fl)	2	2	0	0	4
	Conserved (*) PIP-box residues	6.67	1	2	1	10.67
	Non-conserved PIP-box residues	2.00	0.67	1.00	1.33	5.00
	Other:					

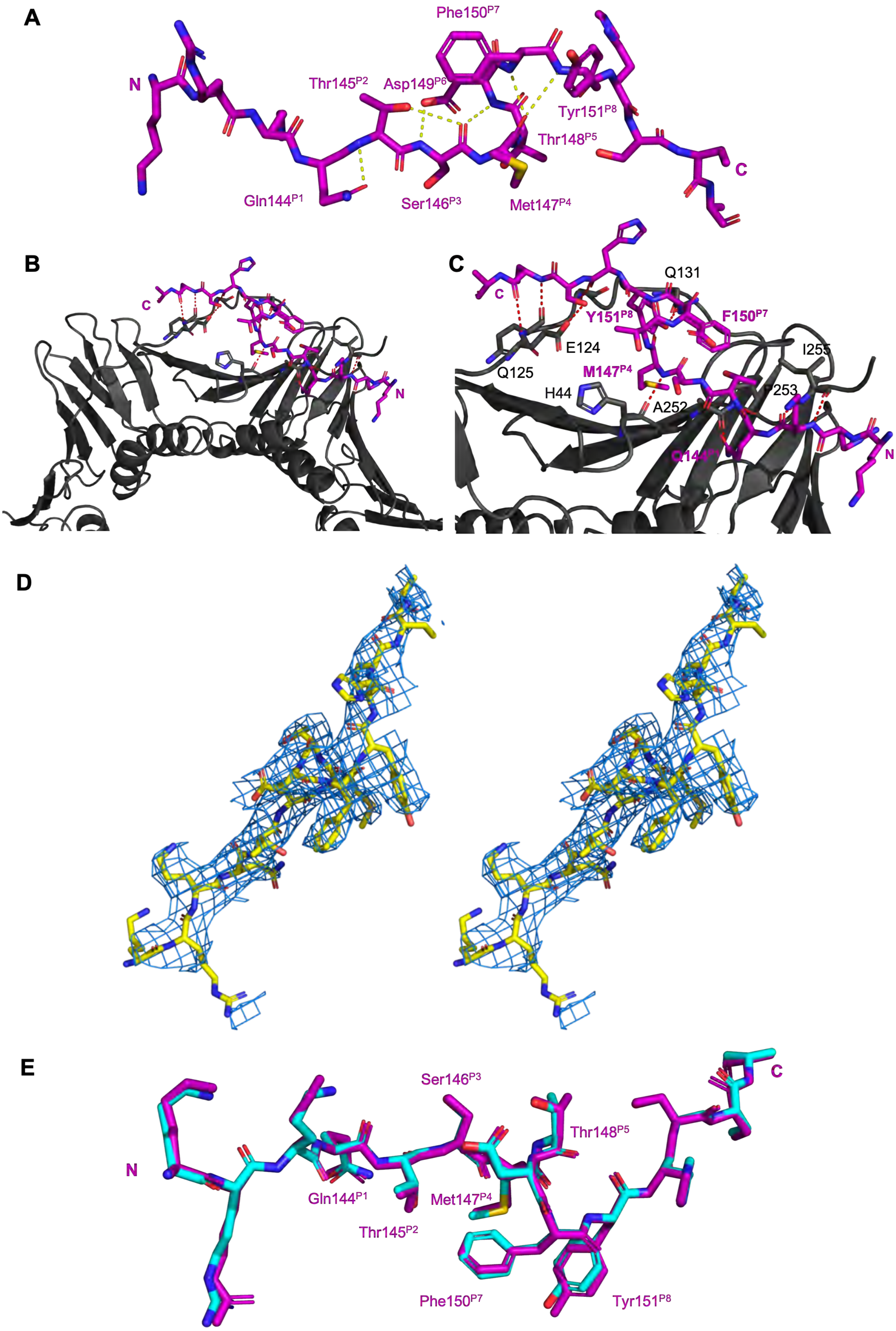

A
Phe150^{P7}
N
Thr145^{P2}
Asp149^{P6}
Tyr151^{P8}
Thr148^{P5}
C
Gln144^{P1}
Ser146^{P3}
Met147^{P4}
B
C
N
C
C
Q131
Y151^{P8}
F150^{P7}
E124
I255
Q125
M147^{P4}
H44
A252
P253
Q144^{P1}
N
D
E
Ser146^{P3}
C
Thr148^{P5}
N
Gln144^{P1}
Met147^{P4}
Thr145^{P2}
Phe150^{P7}
Tyr151^{P8}

Figure S2: Co-crystal structure of **p21**$_{\mu}$ (purple, sticks) bound to PCNA monomer (grey, cartoon). Heteroatoms indicated as blue, nitrogen; red, oxygen; yellow, sulfur. **A** Intramolecular interactions shown as yellow dashes, and PIP-box residues labelled in purple. **B** & **C** Intermolecular interactions shown as red dashes, PCNA residues labelled in grey/white and conserved PIP-box residues labelled in purple. **D** Representative electron density of **p21**$_{\mu}$ (yellow, sticks) shown as a wall-eye stereo image 2Fo-Fc composite omit map, view contoured at 1.5σ. **E** Overlay of **p21**$_{\mu}$ (purple, sticks) bound to PCNA (not shown) co-crystal structure, with **p21**$_{\mu}$ (cyan, sticks) that has been energy minimised on the PCNA surface which shows a high degree of structural similarity, and validates the computational modelling approach.

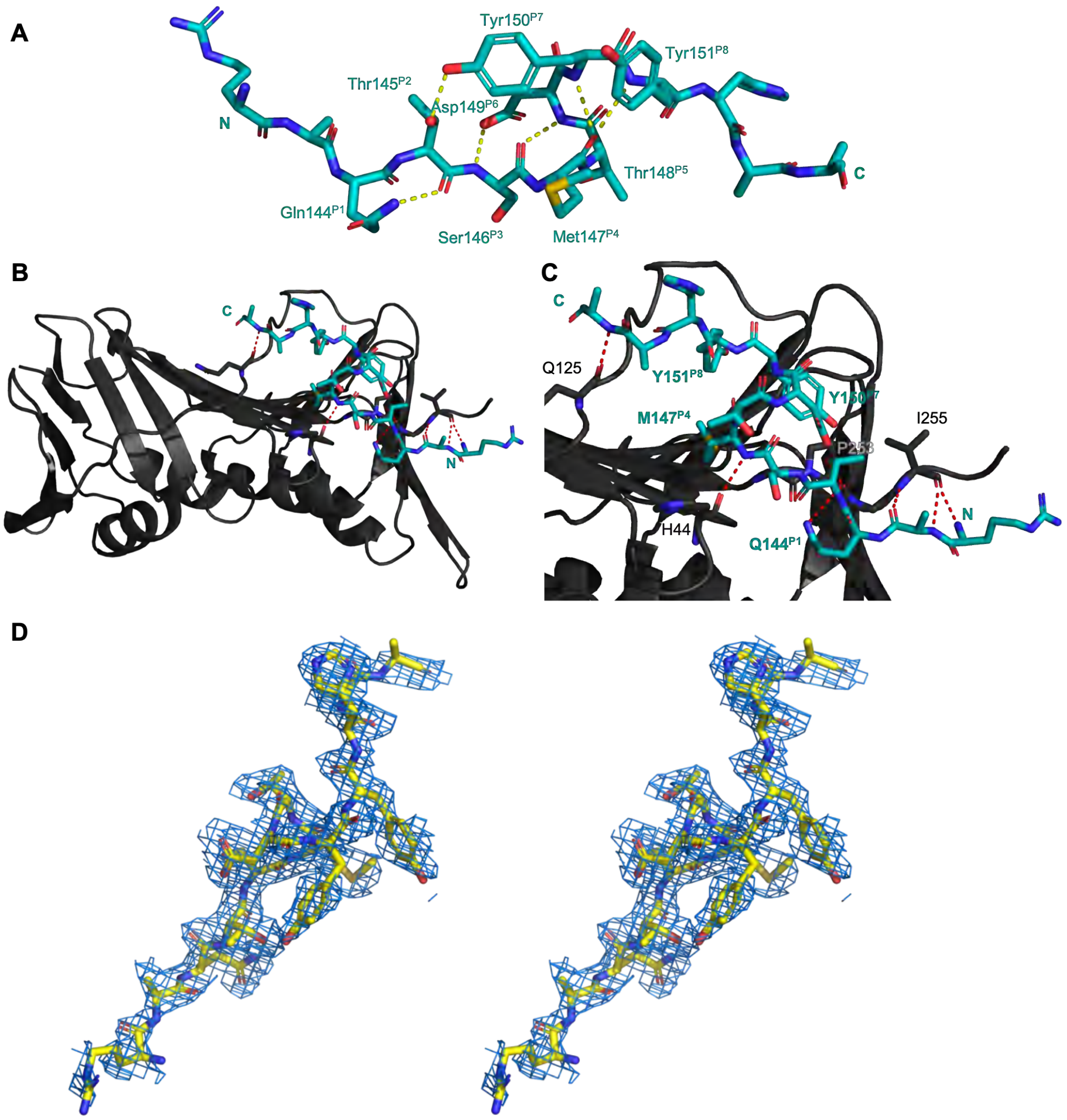

A
Tyr150P7
Thr145P2
Tyr151P8
Asp149P6
N
Thr148P5
C
Gln144P1
Ser146P3
Met147P4
B
C
N
C
C
Q125
Y151P8
Y150P7
M147P4
I255
P253
H44
Q144P1
N
D

Figure S3: Co-crystal structure of **p21$_{\mu}$-F150Y** (blue, sticks) bound to PCNA monomer (grey, cartoon). Heteroatoms indicated as blue, nitrogen; red, oxygen; yellow, sulfur **A** Intramolecular interactions shown as yellow dashes, and PIP-box residues labelled in blue. **B** & **C** Intermolecular interactions shown as red dashes, PCNA residues labelled in grey/white and conserved PIP-box residues labelled in blue. **D** Representative electron density of **p21$_{\mu}$-F150Y** (yellow, sticks) shown as a wall-eye stereo image 2Fo-Fc composite omit map, view contoured at 1.5σ.

S5.4 Computational modelling of PCNA monomers bound to p21$_{\mu}$ peptides

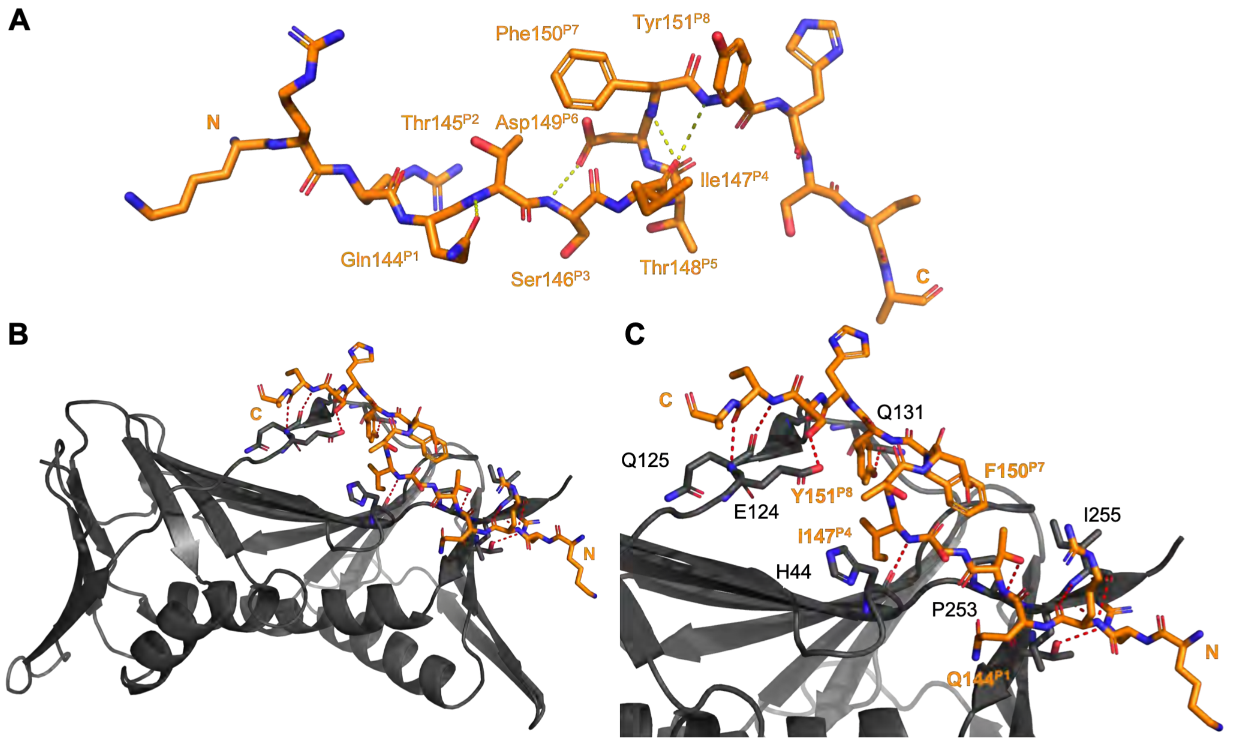

A
N
Phe150^{P7}
Tyr151^{P8}
Thr145^{P2}
Asp149^{P6}
Ile147^{P4}
Gln144^{P1}
Ser146^{P3}
Thr148^{P5}
C
B
C
N
C
C
Q131
Q125
E124
Y151^{P8}
F150^{P7}
I147^{P4}
I255
H44
P253
Q144^{P1}
N

Figure S4: Computationally modelled structure of **p21$_{\mu}$-M147I** (orange, sticks) on the PIP-box binding site of a PCNA monomer (grey, cartoon). Heteroatoms indicated as blue, nitrogen; red, oxygen; yellow, sulfur. **A** Intramolecular interactions shown as yellow dashes, and PIP-box residues labelled in orange. **B** & **C** Intermolecular interactions shown as red dashes, PCNA residues labelled in grey/white and conserved PIP-box residues labelled in orange.

Table S6: Secondary Interaction Summary for computationally modelled structured of p21$_{\mu}$-M147I with PCNA calculated using the RING server. Chain B interactions only. RING Session ID: 5f3b21280e9f94078ea22cfe

		Intermolecular		Intramolecular		
	Residue	**VDW**	**H-Bond**	**VDW**	**H-Bond**	**Total**
FI	**141**	0	0	0	0	
	142	3	1	0	0	
	143	1	2	0	0	
PIP-box	*** 144**	1	0	0	0	
	145	1	1	2	0	
	146	1	0	0	0	
	*** 147**	2	1	0	1	
	148	2	0	0	0	
	149	0	0	0	0	
	*** 150**	3	0	0	0	
	*** 151**	3	2	0	0	
FI	**152**	0	1	0	0	
	153	1	0	0	0	
	154	1	1	0	0	
	155	0	0	0	0	
	Total	19	9	2	1	31
	PIP-box	13	4	2	1	20
	Flanking (FI)	6	5	0	0	11
	Conserved (*) PIP-box residues	6	1	0	1	8
	Non-conserved PIP-box residues	4	1	2	0	7
	Other: Intermolecular pi-stack between F-Tyr151 to E-Tyr133, and D-Tyr151 to C-Tyr250					

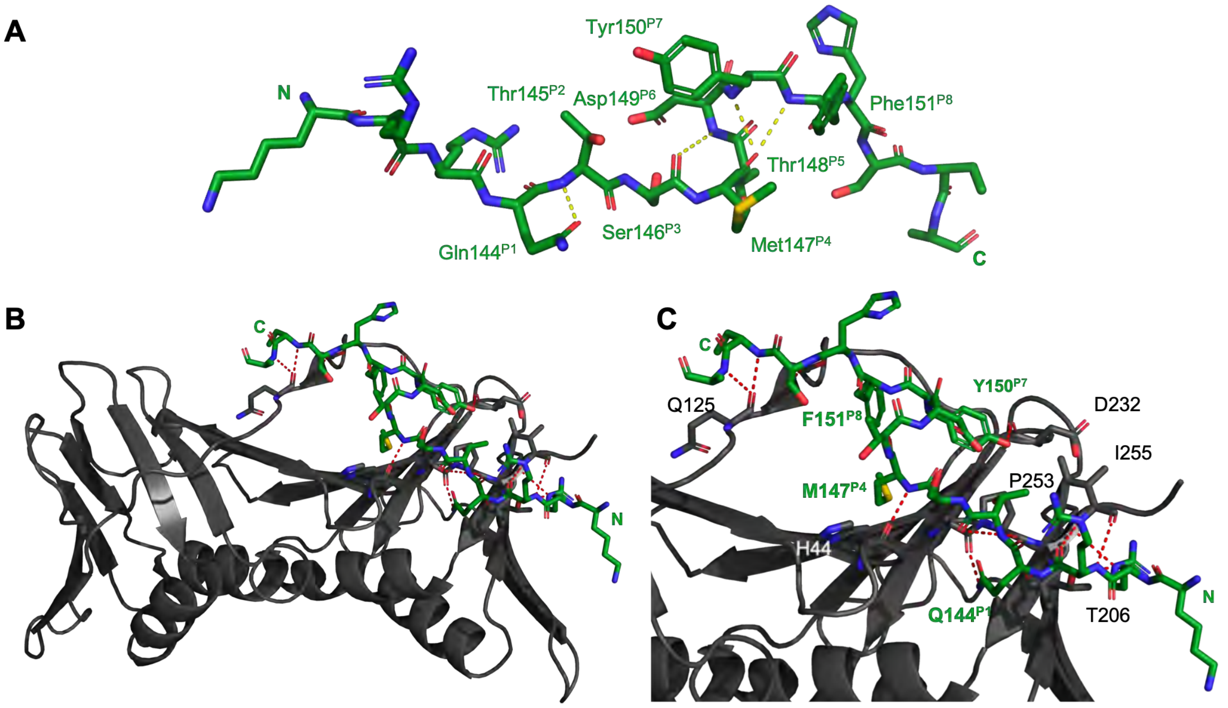

A
Tyr150P7
N
Thr145P2
Asp149P6
Phe151P8
Thr148P5
Ser146P3
Met147P4
Gln144P1
C
B
C
N
C
C
Q125
Y150P7
D232
F151P8
I255
P253
M147P4
H44
Q144P1
T206
N

Figure S5: Computationally modelled structure of **$p21_{\mu}$-FY150/151YF** (dark green, sticks) on the PIP-box binding site of a PCNA monomer (grey, cartoon). Heteroatoms indicated as blue, nitrogen; red, oxygen; yellow, sulfur. **A** Intramolecular interactions shown as yellow dashes, and PIP-box residues labelled in dark green **B** & **C** Intermolecular interactions shown as red dashes, PCNA residues labelled in grey/white and conserved PIP-box residues labelled in dark green.

Table S7: Secondary Interaction Summary for computationally modelled structured of **$p21_{\mu}$-FY150/151YF** with PCNA calculated using the RING server. Chain B interactions only. RING session ID: 5f3b22bc0e9f94078ea22d03

		Intermolecular		Intramolecular		
	Residue	**VDW**	**H-Bond**	**VDW**	**H-Bond**	**Total**
Fl	141	0	0	0	0	
	142	4	2	0	0	
	143	1	1	0	0	
PIP-box	* **144**	2	0	0	0	
	145	1	1	0	0	
	146	1	0	0	1	
	* **147**	6	1	1	2	
	148	1	0	0	0	
	149	0	0	0	0	
	* **150**	3	1	0	0	
	* **151**	2	0	0	0	
Fl	152	0	0	0	0	
	153	1	0	0	0	
	154	2	1	0	0	
	155	0	1	0	0	
	Total	24	8	1	3	36
	PIP-box	16	3	1	3	23
	Flanking (Fl)	8	5	0	0	13
	Conserved (*) PIP-box residues	11	2	1	2	16
	Non-conserved PIP-box residues	3	1	0	1	5
	Other: Intermolecular ionic interaction between B-Arg143 and A-Asp257; Intermolecular pi-stack between B-Tyr151 and A-Tyr250					

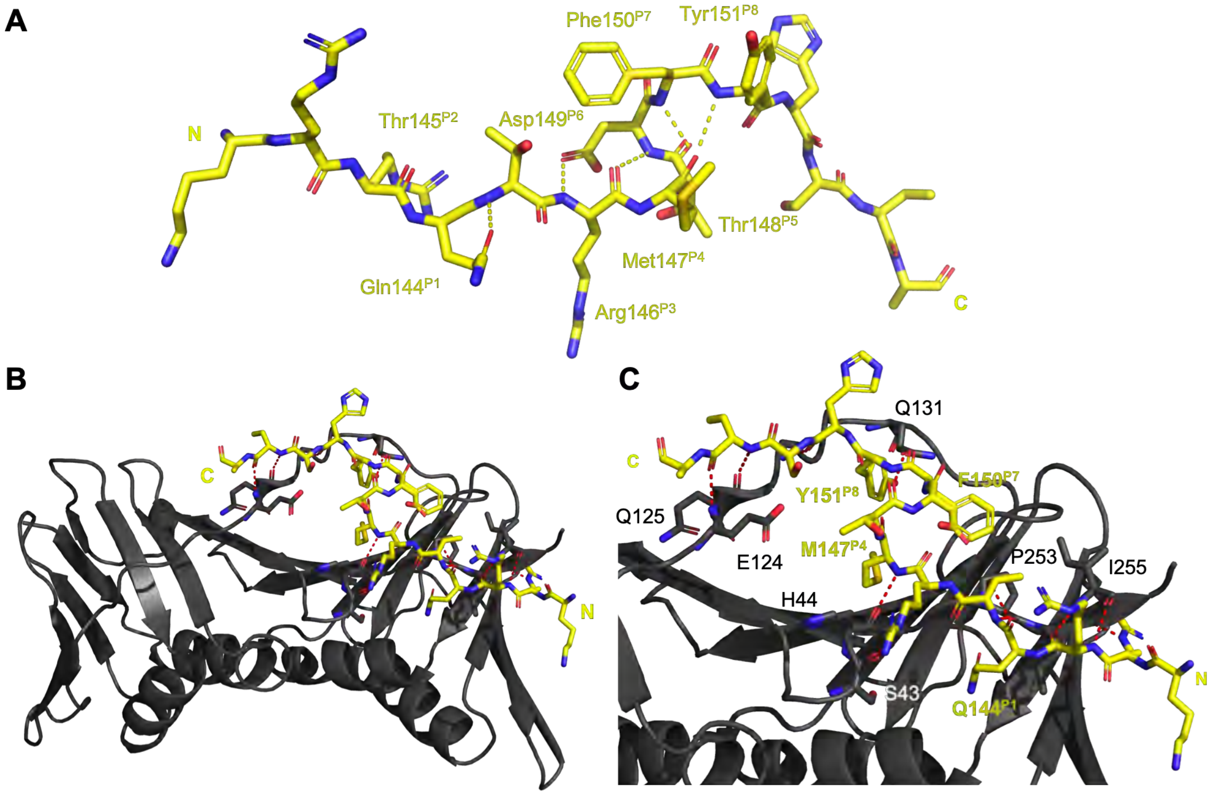

A
N
Thr145P2
Asp149P6
Phe150P7
Tyr151P8
Thr148P5
Met147P4
Gln144P1
Arg146P3
C
B
C
N
C
Q131
C
F150P7
Y151P8
Q125
E124
M147P4
P253
I255
H44
S43
Q144P1
N

Figure S6: Computationally modelled structure of **p21$_{\mu}$-S146R** (yellow, sticks) on the PIP-box binding site of a PCNA monomer (grey, cartoon). Heteroatoms indicated as blue, nitrogen; red, oxygen; yellow, sulfur. **A** Intramolecular interactions shown as yellow dashes, and PIP-box residues labelled in yellow. **B** & **C** Intermolecular interactions shown as red dashes, PCNA residues labelled in grey/white and conserved PIP-box residues labelled in yellow.

Table S8: Secondary Interaction Summary for computationally modelled structured of **p21$_{\mu}$-S146R** with PCNA calculated using the RING server. Chain B interactions only. RING Session ID: 5f3b1f6b0e9f94078ea22cf9

		Intermolecular		Intramolecular		
	Residue	**VDW**	**H-Bond**	**VDW**	**H-Bond**	**Total**
FI	141	0	0	0	0	
	142	3	1	0	0	
	143	2	2	0	0	
PIP-box	* **144**	1	0	0	0	
	145	1	1	1	0	
	146	2	1	0	1	
	* **147**	4	1	1	2	
	148	0	0	0	0	
	149	0	0	0	0	
	* **150**	3	0	0	0	
	* **151**	2	1	0	0	
FI	152	0	1	0	0	
	153	2	0	0	0	
	154	0	1	0	0	
	155	0	1	0	0	
	Total	20	10	2	3	35
	PIP-box	13	4	2	3	22
	Flanking (FI)	7	6	0	0	13
	Conserved (*) PIP-box residues	8	1	1	2	12
	Non-conserved PIP-box residues	3	2	1	1	7
	Other: Intermolecular ionic interaction between B-Arg143 and A-Asp257					

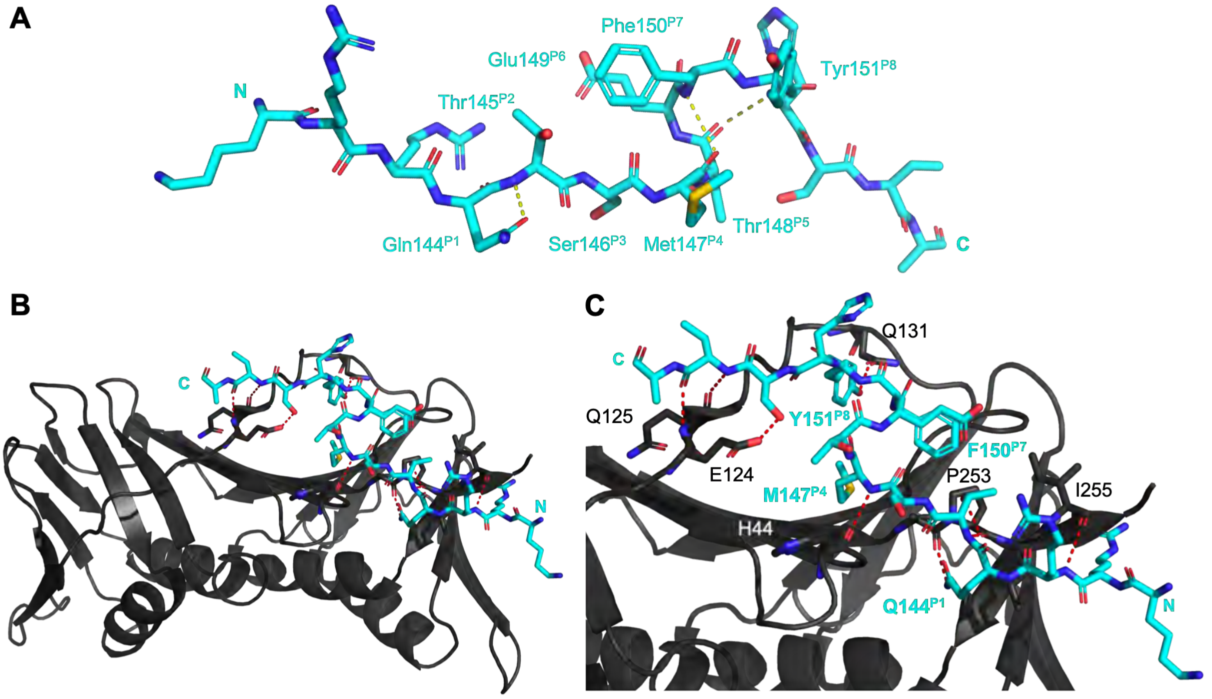

A
N
Phe150^{P7}
Glu149^{P6}
Thr145^{P2}
Tyr151^{P8}
Gln144^{P1}
Ser146^{P3}
Met147^{P4}
Thr148^{P5}
C
B
C
N
C
C
Q131
Q125
Y151^{P8}
F150^{P7}
E124
P253
I255
M147^{P4}
H44
Q144^{P1}
N

Figure S7: Computationally modelled structure of **p21$_{\mu}$-D149E** (light blue, sticks) on the PIP-box binding site of a PCNA monomer (grey, cartoon). Heteroatoms indicated as blue, nitrogen; red, oxygen; yellow, sulfur. **A** Intramolecular interactions shown as yellow dashes, and PIP-box residues labelled in light blue. **B** & **C** Intermolecular interactions shown as red dashes, PCNA residues labelled in grey/white and conserved PIP-box residues labelled in light blue.

Table S9: Secondary Interaction Summary for computationally modelled structured of **p21$_{\mu}$-D149E** with PCNA calculated using the RING server. Chain B interactions only. RING Session ID: 5f3b21fe0e9f94078ea22d00

		Intermolecular		Intramolecular		
	Residue	**VDW**	**H-Bond**	**VDW**	**H-Bond**	**Total**
Fl	141	0	0	0	0	
	142	2	2	0	0	
	143	0	2	0	0	
PIP-box	*** 144**	2	0	0	0	
	145	1	1	1	0	
	146	1	0	0	1	
	*** 147**	4	1	1	2	
	148	1	0	0	0	
	149	0	0	1	0	
	*** 150**	3	0	0	0	
	*** 151**	4	2	0	0	
Fl	152	1	0	0	0	
	153	1	0	0	0	
	154	0	1	0	0	
	155	0	0	0	0	
	Total	20	9	3	3	35
	PIP-box	16	4	3	3	26
	Flanking (Fl)	4	5	0	0	9
	Conserved (*) PIP-box residues	9	1	2	2	14
	Non-conserved PIP-box residues	3	1	2	1	7
	Other: Intermolecular ionic interaction between B-Arg143 and A-Asp257; Intermolecular pi-stack between B-Tyr151 and A-Tyr250					

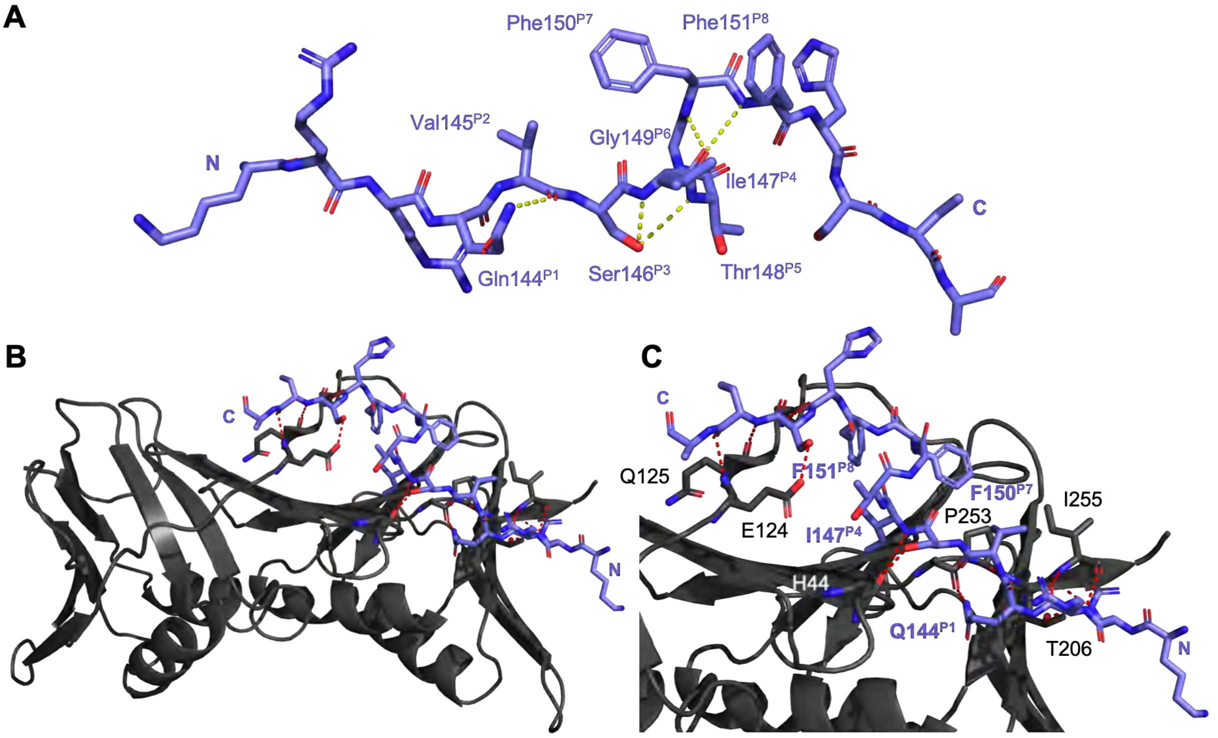

A
Phe150P7
Phe151P8
Val145P2
Gly149P6
N
Ile147P4
C
Gln144P1
Ser146P3
Thr148P5
B
C
N
C
C
Q125
F151P8
F150P7
I255
E124
I147P4
P253
H44
Q144P1
T206
N

Figure S8: Computationally modelled structure of **p21$_{\mu}$-pol δ$_{p66}$** (light purple, sticks) on the PIP-box binding site of a PCNA monomer (grey, cartoon). Heteroatoms indicated as blue, nitrogen; red, oxygen; yellow, sulfur. **A** Intramolecular interactions shown as yellow dashes, and PIP-box residues labelled in light purple. **B** & **C** Intermolecular interactions shown as red dashes, PCNA residues labelled in grey/white and conserved PIP-box residues labelled in light purple.

Table S10: Secondary Interaction Summary for computationally modelled structured of **p21$_{\mu}$-pol δ$_{p66}$** with PCNA calculated using the RING server. Chain B interactions only. RING Session ID: 5f3b202a0e9f94078ea22cfb

		Intermolecular		Intramolecular		
	Residue	**VDW**	**H-Bond**	**VDW**	**H-Bond**	**Total**
FI	141	0	0	0	0	
	142	3	2	0	0	
	143	1	2	0	0	
PIP-box	* **144**	2	0	0	0	
	145	1	1	1	0	
	146	0	0	0	0	
	* **147**	5	0	1	2	
	148	0	0	0	0	
	149	0	0	0	0	
	* **150**	1	0	0	0	
	* **151**	2	0	0	0	
FI	152	0	1	0	0	
	153	2	0	0	0	
	154	0	1	0	0	
	155	1	0	0	0	
	Total	18	7	2	2	29
	PIP-box	11	1	2	2	16
	Flanking (FI)	7	6	0	0	13
	Conserved (*) PIP-box residues	8	0	1	2	11
	Non-conserved PIP-box residues	1	1	1	0	3
	Other: Intermolecular pi-stack between B-Phe151 and A-Tyr250					

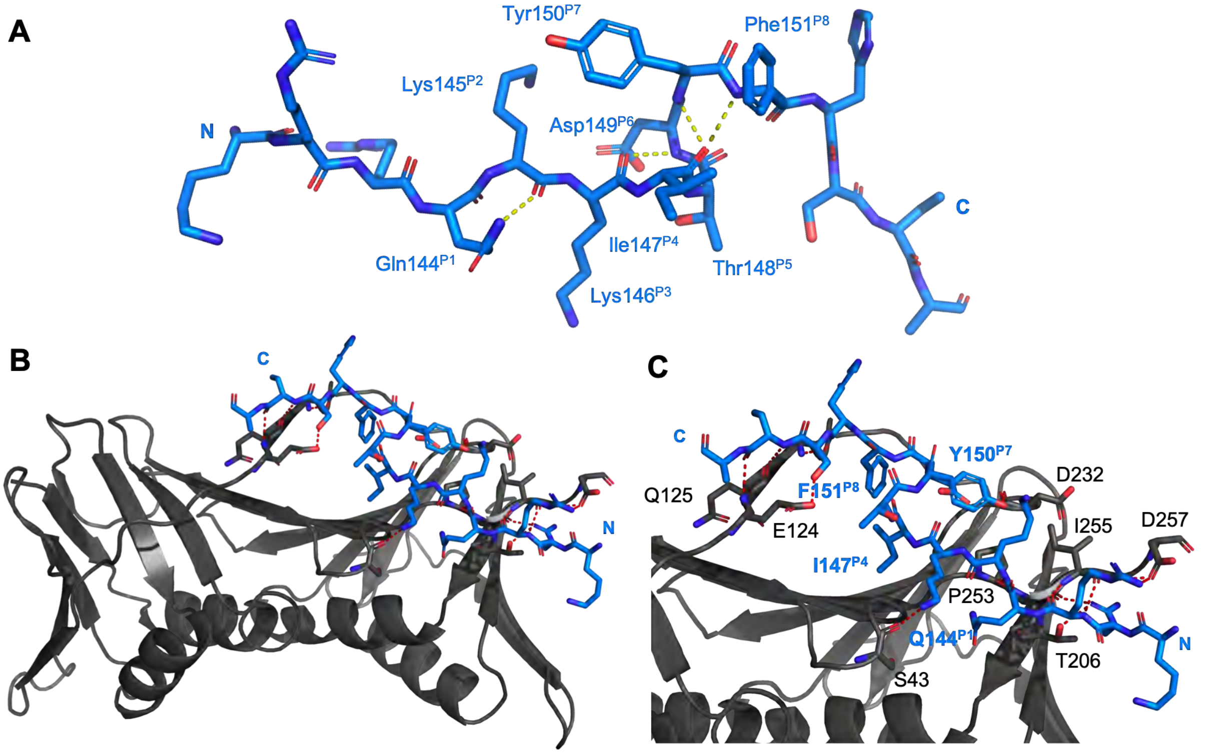

A
N
Lys145^{P2}
Tyr150^{P7}
Asp149^{P6}
Phe151^{P8}
C
Gln144^{P1}
Ile147^{P4}
Thr148^{P5}
Lys146^{P3}
B
C
N
C
C
Q125
F151^{P8}
Y150^{P7}
D232
E124
I255
D257
I147^{P4}
P253
Q144^{P1}
N
T206
S43

Figure S9: Computationally modelled structure of **p21$_{\mu}$-Pogo** (dark blue, sticks) on the PIP-box binding site of a PCNA monomer (grey, cartoon). Heteroatoms indicated as blue, nitrogen; red, oxygen; yellow, sulfur. **A** Intramolecular interactions shown as yellow dashes, and PIP-box residues labelled in blue. **B** & **C** Intermolecular interactions shown as red dashes, PCNA residues labelled in grey/white and conserved PIP-box residues labelled in blue.

Table S11: Secondary Interaction Summary for computationally modelled structured of **p21$_{\mu}$-Pogo** with PCNA calculated using the RING server. Chain B interactions only. RING Session ID: 5f3b224d0e9f94078ea22d01

		Intermolecular		**Intramolecular**		
	Residue	**VDW**	**H-Bond**	**VDW**	**H-Bond**	**Total**
FI	141	0	0	0	0	
	142	3	2	0	0	
	143	1	2	0	0	
PIP-box	*** 144**	2	0	0	0	
	145	2	0	0	0	
	146	1	0	0	1	
	*** 147**	7	0	1	2	
	148	1	0	0	0	
	149	0	0	0	0	
	*** 150**	3	1	0	0	
	*** 151**	3	0	0	0	
FI	152	0	1	0	0	
	153	0	0	0	0	
	154	0	1	0	0	
	155	0	0	0	0	
	Total	23	7	1	3	34
	PIP-box	19	1	1	3	24
	Flanking (FI)	4	6	0	0	10
	Conserved (*) PIP-box residues	12	1	1	2	16
	Non-conserved PIP-box residues	4	0	0	1	5
	Other: Intermolecular ionic interaction between B-Arg143 and A-Asp257; Intermolecular pi-stack between B-Phe151 and A-Tyr250					

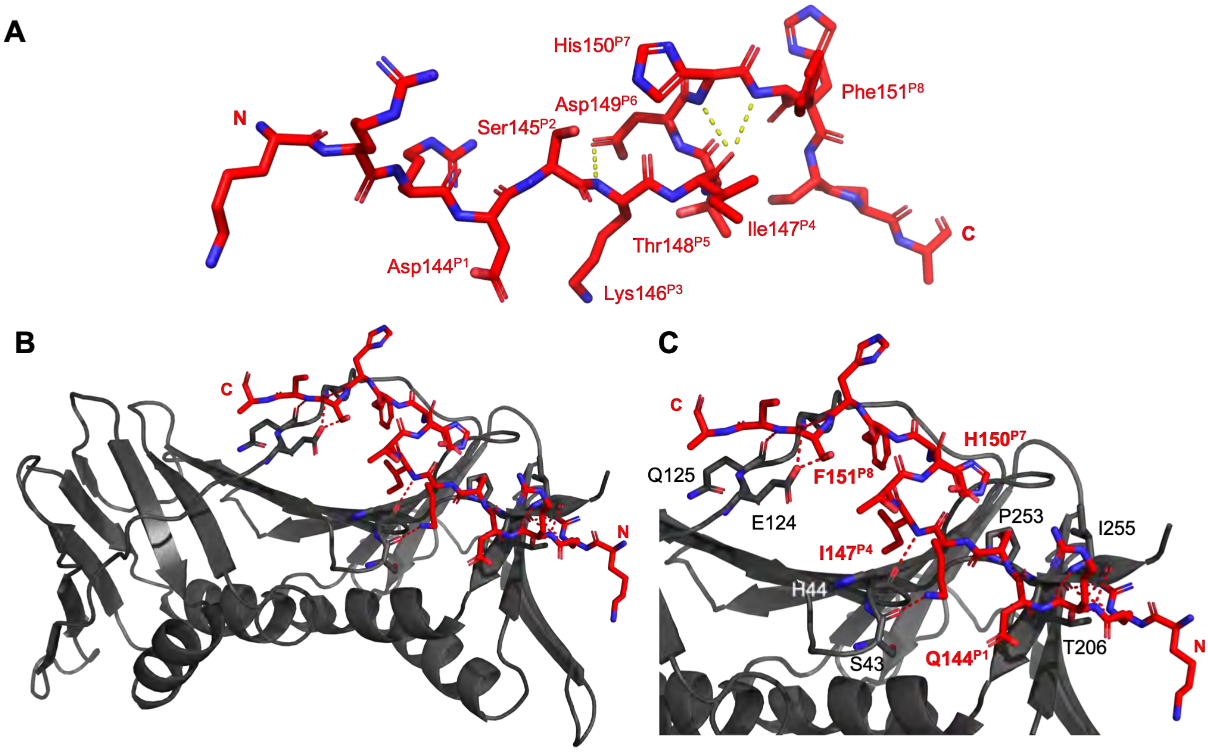

A
His150^{P7}
Phe151^{P8}
N
Asp149^{P6}
Ser145^{P2}
C
Ile147^{P4}
Thr148^{P5}
Asp144^{P1}
Lys146^{P3}
B
C
N
C
C
H150^{P7}
Q125
F151^{P8}
E124
P253
I255
I147^{P4}
H44
S43
Q144^{P1}
T206
N

Figure S10: Computationally modelled structure of **p21$_{\mu}$-PARG** (red, sticks) on the PIP-box binding site of a PCNA monomer (grey, cartoon). Heteroatoms indicated as blue, nitrogen; red, oxygen; yellow, sulfur. **A** Intramolecular interactions shown as yellow dashes, and PIP-box residues labelled in red. **B** & **C** Intermolecular interactions shown as red dashes, PCNA residues labelled in grey/white and conserved PIP-box residues labelled in red.

Table S12: Secondary Interaction Summary for computationally modelled structured of **p21$_{\mu}$-PARG** with PCNA calculated using the RING server. Chain B interactions only. RING Session ID: 5f3b1de50e9f94078ea22cf7

		Intermolecular		Intramolecular		
	Residue	**VDW**	**H-Bond**	**VDW**	**H-Bond**	**Total**
FI	141	0	0	0	0	
	142	2	2	0	0	
	143	0	2	0	0	
PIP-box	*** 144**	1	0	0	0	
	145	1	0	0	0	
	146	1	0	0	1	
	*** 147**	2	1	0	2	
	148	0	0	0	0	
	149	0	0	0	0	
	*** 150**	3	0	0	0	
	*** 151**	3	0	0	0	
FI	152	0	1	0	0	
	153	1	0	0	0	
	154	0	1	0	0	
	155	0	0	0	0	
	Total	14	7	0	3	24
	PIP-box	11	1	0	3	15
	Flanking (FI)	3	6	0	0	9
	Conserved (*) PIP-box residues	5	1	0	2	8
	Non-conserved PIP-box residues	3	0	0	1	4
	Other: Intermolecular pi-stack between B-Tyr151 and A-Tyr250					

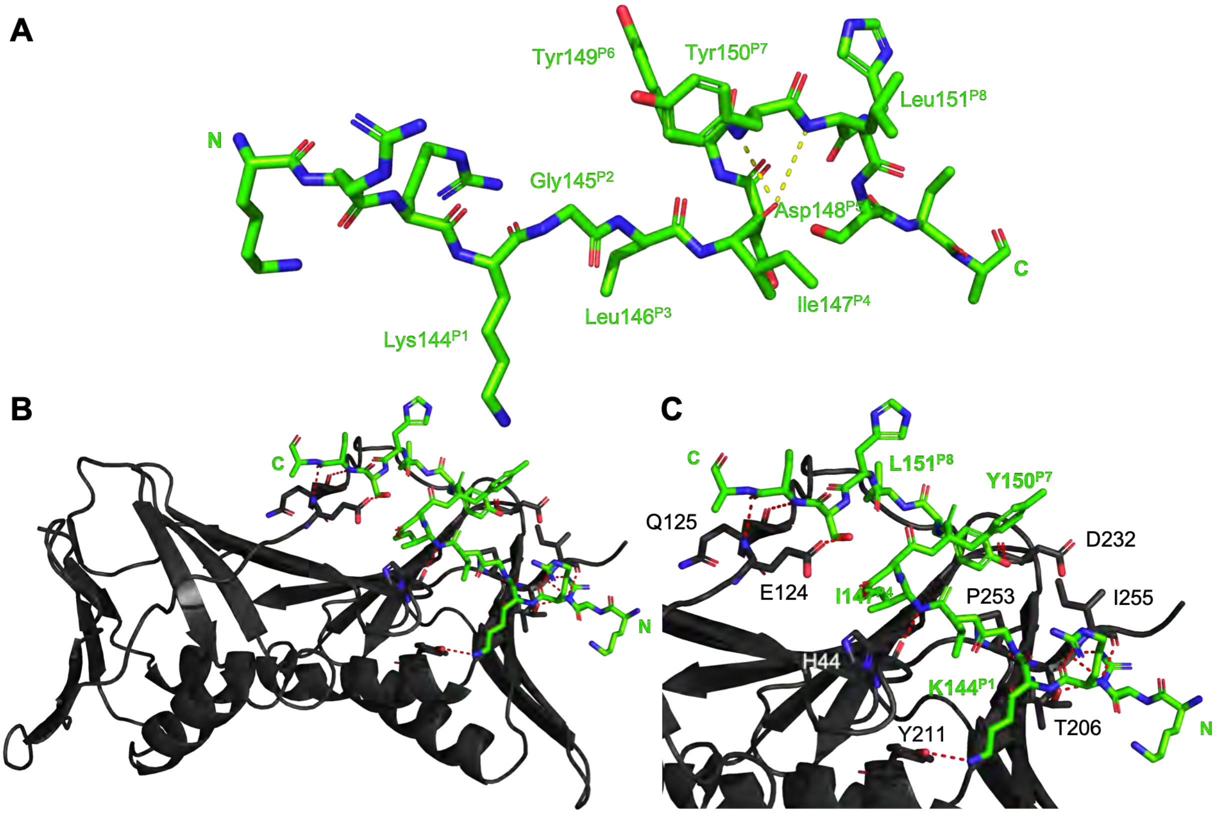

A
Tyr149^{P6}
Tyr150^{P7}
Leu151^{P8}
N
Gly145^{P2}
Asp148^{P5}
C
Ile147^{P4}
Leu146^{P3}
Lys144^{P1}
B
C
N
C
C
L151^{P8}
Y150^{P7}
Q125
D232
E124
I147^{P4}
P253
I255
H44
K144^{P1}
Y211
T206
N

Figure S11: Computationally modelled structure of **p21$_{\mu}$-pol ι** (green, sticks) on the PIP-box binding site of a PCNA monomer (grey, cartoon). Heteroatoms indicated as blue, nitrogen; red, oxygen; yellow, sulfur. **A** Intramolecular interactions shown as yellow dashes, and PIP-box residues labelled in green **B** & **C** Intermolecular interactions shown as red dashes, PCNA residues labelled in grey/white and conserved PIP-box residues labelled in green

Table S13: Secondary Interaction Summary for computationally modelled structured of **p21$_{\mu}$-pol ι** with PCNA calculated using the RING server. Chain B interactions only. RING Session ID: 5f3b1e580e9f94078ea22cf8

		Intermolecular			Intramolecular		
	Residue	**VDW**	**H-Bond**		**VDW**	**H-Bond**	**Total**
Fl	141	0	0		0	0	
	142	2	2		0	0	
	143	0	2		0	0	
PIP-box	* **144**	1	0		0	0	
	145	1	1		0	0	
	146	2	0		1	0	
	* **147**	3	1		0	1	
	148	1	0		0	0	
	149	0	0		0	0	
	* **150**	2	1		0	0	
	* **151**	1	0		0	0	
Fl	152	0	1		0	0	
	153	2	0		0	0	
	154	0	1		0	0	
	155	0	0		0	0	
	Total	15	9		1	1	26
	PIP-box	11	3		1	1	16
	Flanking (Fl)	4	6		0	0	10
	Conserved (*) PIP-box residues	6	2		0	1	9
	Non-conserved PIP-box residues	5	1		1	0	7
	Other:						

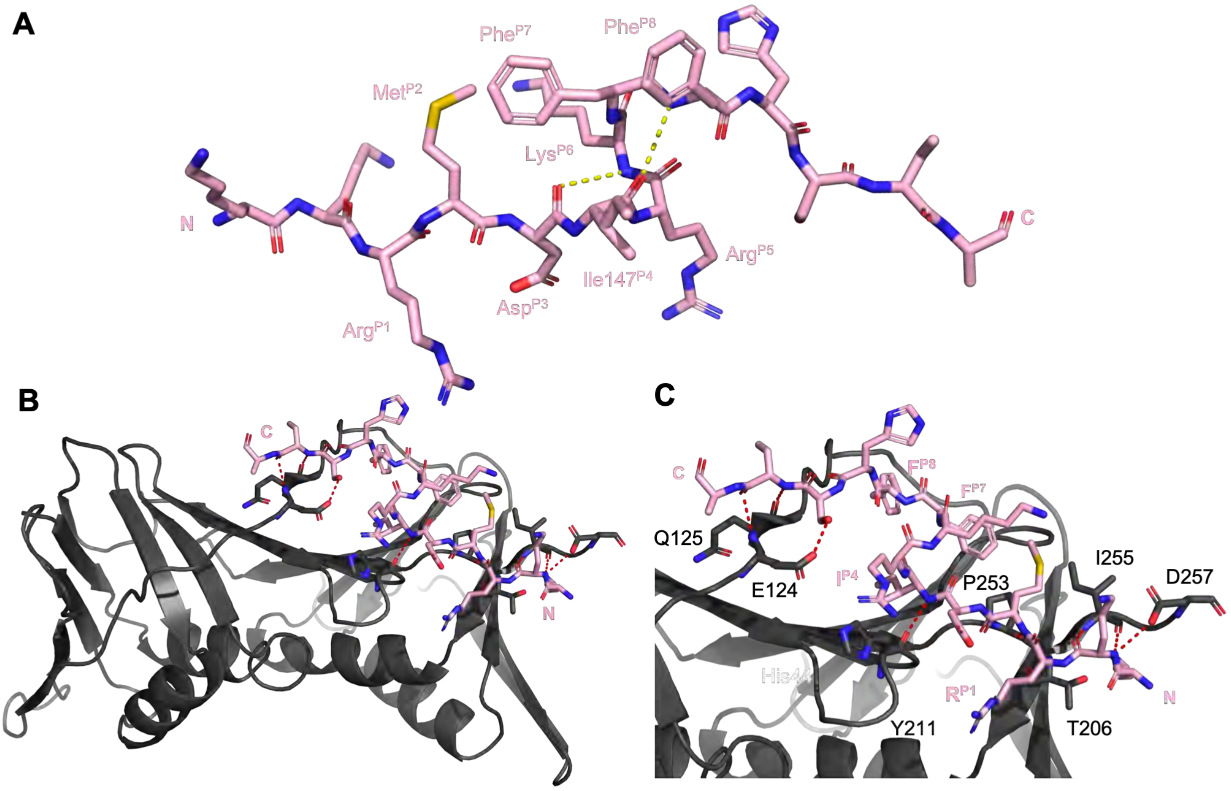


Figure S12: Computationally modelled structure of **p21$_{\mu}$-RFC** (light pink, sticks) on the PIP-box binding site of a PCNA monomer (grey, cartoon). Heteroatoms indicated as blue, nitrogen; red, oxygen; yellow, sulfur. **A** Intramolecular interactions shown as yellow dashes, and PIP-box residues labelled in light pink **B** & **C** Intermolecular interactions shown as red dashes, PCNA residues labelled in grey/white and conserved PIP-box residues labelled in light pink

Table S14: Secondary Interaction Summary for computationally modelled structured of **p21$_{\mu}$-RFC** with PCNA calculated using the RING server. Chain B interactions only. RING Session ID: 5f3b20890e9f94078ea22cfd

		Intermolecular		Intramolecular		
	Residue	**VDW**	**H-Bond**	**VDW**	**H-Bond**	**Total**
FI						
	142	3	0	0	0	
	143	1	2	0	0	
PIP-box	**144**	1	0	0	0	
	145	2	1	1	0	
	146	1	0	0	2	
	* **147**	2	1	0	2	
	148	1	0	0	0	
	149	0	0	0	0	
	* **150**	2	0	0	0	
	* **151**	2	0	0	0	
FI	152	0	0	0	0	
	153	2	0	0	0	
	154	0	1	0	0	
	155	0	0	0	0	
	Total	17	5	1	4	27
	PIP-box	11	2	1	4	18
	Flanking (FI)	6	3	0	0	9
	Conserved (*) PIP-box residues	6	1	0	2	9
	Non-conserved PIP-box residues	5	1	1	2	9
	Other: Intermolecular ionic interaction between B-Arg143 and A-Asp257; Intermolecular pi-stack between B-Tyr151 and A-Tyr250					

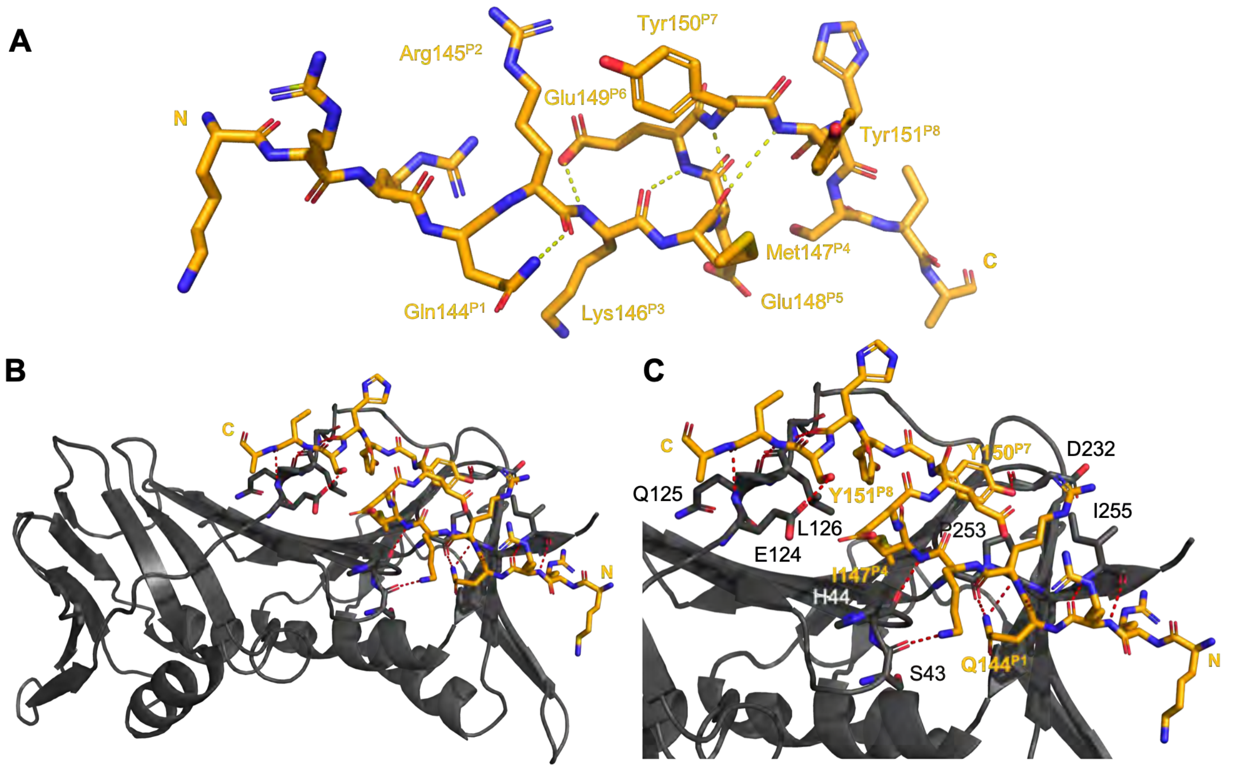

A
Tyr150^{P7}
Arg145^{P2}
Glu149^{P6}
N
Tyr151^{P8}
Met147^{P4}
C
Gln144^{P1}
Lys146^{P3}
Glu148^{P5}
B
C
N
C
C
Y150^{P7}
D232
Q125
Y151^{P8}
I255
L126
P253
E124
M147^{P4}
H44
S43
Q144^{P1}
N

Figure S13: Computationally modelled structure of **p21$_{\mu}$-RD1** (light orange, sticks) on the PIP-box binding site of a PCNA monomer (grey, cartoon). Heteroatoms indicated as blue, nitrogen; red, oxygen; yellow, sulfur. **A** Intramolecular interactions shown as yellow dashes, and PIP-box residues labelled in light orange. **B** & **C** Intermolecular interactions shown as red dashes, PCNA residues labelled in grey/white and conserved PIP-box residues labelled in light orange.

Table S15: Secondary Interaction Summary for computationally modelled structured of **p21$_{\mu}$-RD1** with PCNA calculated using the RING server. Chain B interactions only. RING Session ID: 5f3b21c10e9f94078ea22cff

		Intermolecular			Intramolecular		
	Residue	**VDW**	**H-Bond**		**VDW**	**H-Bond**	**Total**
Fl	141	0	0		0	0	
	142	3	2		0	0	
	143	1	1		0	0	
PIP-box	*** 144**	2	0		0	0	
	145	1	1		1	0	
	146	0	0		0	1	
	*** 147**	4	1		1	2	
	148	1	0		0	0	
	149	0	0		0	0	
	*** 150**	3	1		0	0	
	*** 151**	3	0		0	0	
Fl	152	0	1		0	0	
	153	1	0		0	0	
	154	0	1		0	0	
	155	0	0		0	0	
	Total	19	8		2	3	32
	PIP-box	14	3		2	3	22
	Flanking (Fl)	5	5		0	0	10
	Conserved (*) PIP-box residues	9	2		1	2	14
	Non-conserved PIP-box residues	2	1		1	1	5
	Other: Intermolecular ionic interaction between B-Arg143 and A-Asp257; Intermolecular pi-stack between B-Tyr151 and A-Tyr250						

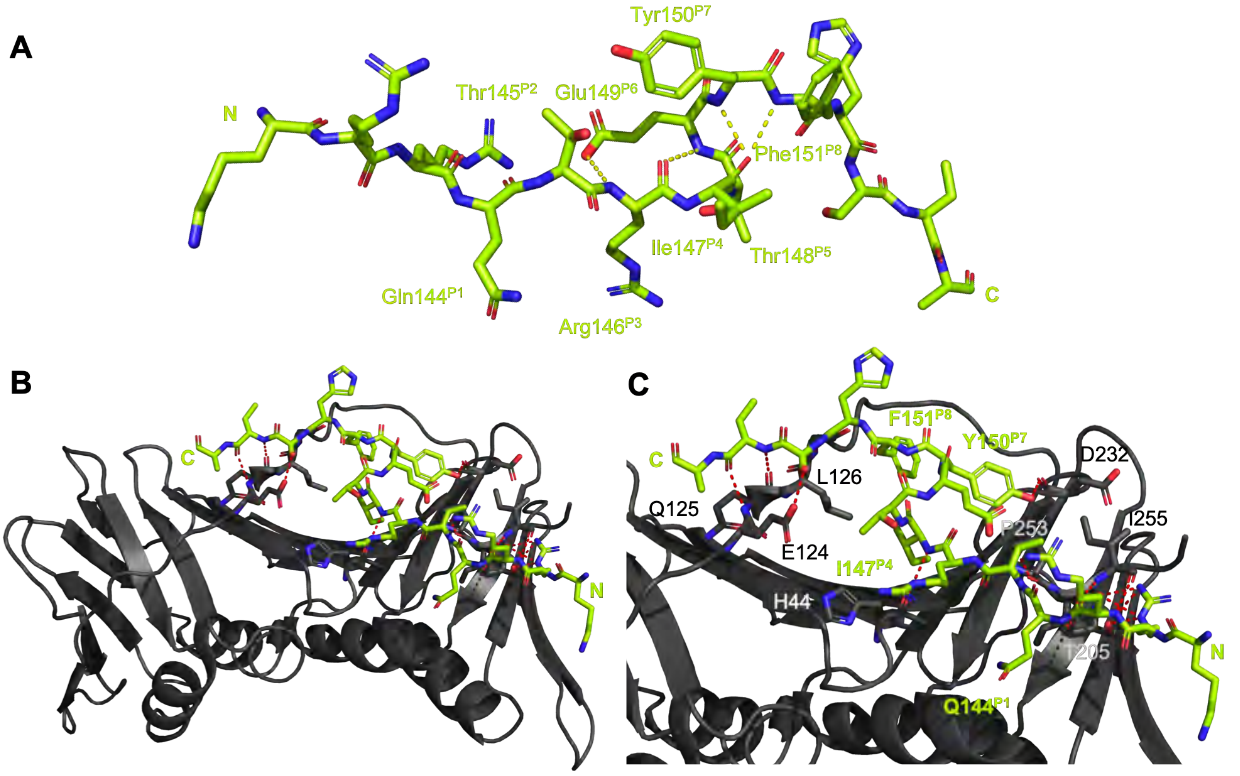

A
Tyr150^{P7}
N
Thr145^{P2}
Glu149^{P6}
Phe151^{P8}
Ile147^{P4}
Thr148^{P5}
Gln144^{P1}
Arg146^{P3}
C
B
C
N
C
C
F151^{P8}
Y150^{P7}
L126
D232
Q125
P253
I255
E124
I147^{P4}
H44
T205
N
Q144^{P1}

Figure S14: Computationally modelled structure of **p21$_{\mu}$-RD2** (green, sticks) on the PIP-box binding site of a PCNA monomer (grey, cartoon). Heteroatoms indicated as blue, nitrogen; red, oxygen; yellow, sulfur. **A** Intramolecular interactions shown as yellow dashes, and PIP-box residues labelled in green. **B** & **C** Intermolecular interactions shown as red dashes, PCNA residues labelled in grey/white and conserved PIP-box residues labelled in green.

Table S16: Secondary Interaction Summary for computationally modelled structured of **p21$_{\mu}$-RD2** with PCNA calculated using the RING server. Chain B interactions only. RING Session ID: 5f3b1ff30e9f94078ea22cfa

		Intermolecular			**Intramolecular**		
	Residue	**VDW**	**H-Bond**		**VDW**	**H-Bond**	**Total**
FI	141	0	0		0	0	
	142	2	2		0	0	
	143	0	1		0	0	
PIP-box	*** 144**	1	0		0	0	
	145	1	1		0	0	
	146	2	1		0	1	
	*** 147**	7	1		1	2	
	148	1	0		0	0	
	149	0	0		0	0	
	*** 150**	2	1		0	0	
	*** 151**	2	0		0	0	
FI	152	0	0		0	0	
	153	0	0		0	0	
	154	2	1		0	0	
	155	0	1		0	0	
	Total	20	9		1	3	33
	PIP-box	16	4		1	3	24
	Flanking (FI)	4	5		0	0	9
	Conserved (*) PIP-box residues	10	2		1	2	15
	Non-conserved PIP-box residues	4	2		0	1	7
	Other:						

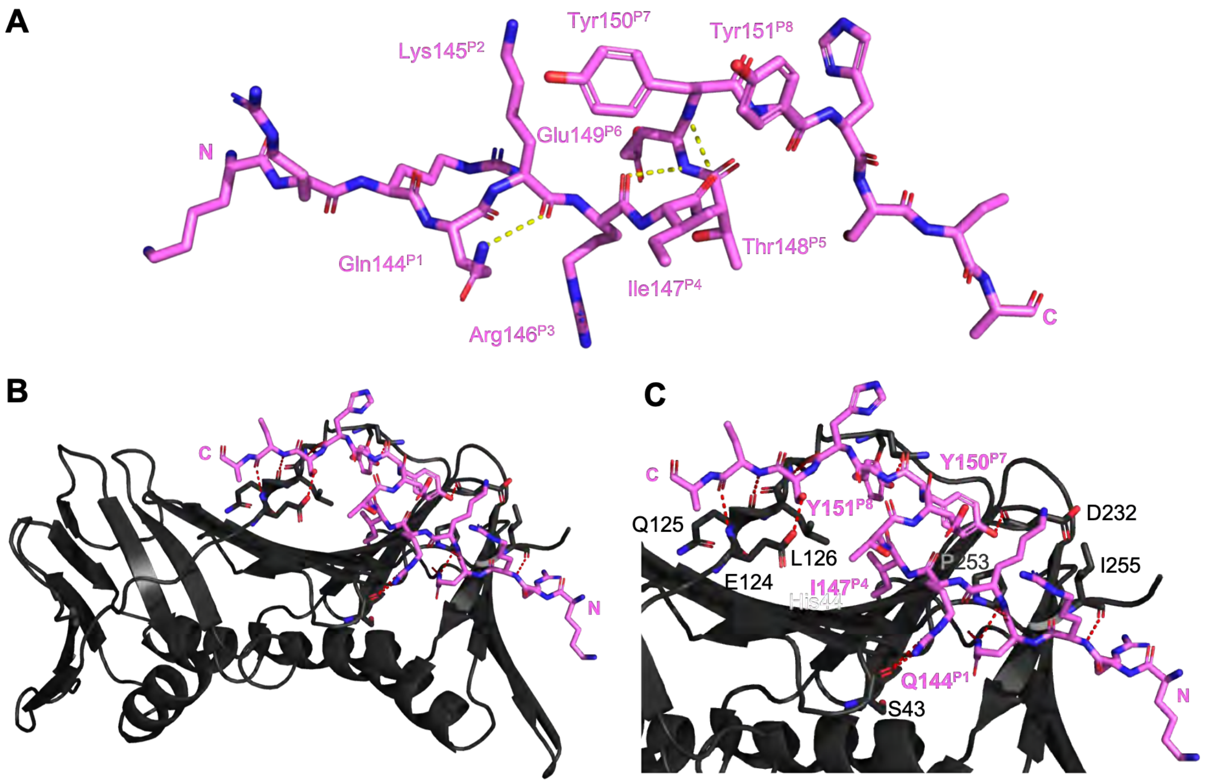

A
Tyr150^{P7}
Tyr151^{P8}
Lys145^{P2}
Glu149^{P6}
N
Gln144^{P1}
Thr148^{P5}
Ile147^{P4}
Arg146^{P3}
C
B
C
N
C
C
Y150^{P7}
D232
Q125
Y151^{P8}
L126
P253
I255
E124
I147^{P4}
His44
Q144^{P1}
S43
N

Figure S15: Computationally modelled structure of **p21$_{\mu}$-RD3** (pink, sticks) on the PIP-box binding site of a PCNA monomer (grey, cartoon). Heteroatoms indicated as blue, nitrogen; red, oxygen; yellow, sulfur. **A** Intramolecular interactions shown as yellow dashes, and PIP-box residues labelled in pink. **B** & **C** Intermolecular interactions shown as red dashes, PCNA residues labelled in grey/white and conserved PIP-box residues labelled in pink.

Table S17: Secondary Interaction Summary for computationally modelled structured of **p21$_{\mu}$-RD3** with PCNA calculated using the RING server. Chain B interactions only. RING Session ID: 5f3b22e70e9f94078ea22d04

		Intermolecular		Intramolecular		
	Residue	**VDW**	**H-Bond**	**VDW**	**H-Bond**	**Total**
FI	141	0	0	0	0	
	142	2	1	0	0	
	143	2	2	0	0	
PIP-box	* **144**	1	0	0	0	
	145	0	0	2	0	
	146	2	1	0	1	
	* **147**	5	0	0	2	
	148	1	0	0	0	
	149	0	0	0	0	
	* **150**	2	1	0	0	
	* **151**	2	1	0	0	
FI	152	0	1	0	0	
	153	2	0	0	0	
	154	0	1	0	0	
	155	0	0	0	0	
	Total	19	8	2	3	32
	PIP-box	13	3	2	3	21
	Flanking (Fl)	6	5	0	0	11
	Conserved (*) PIP-box residues	8	1	0	2	11
	Non-conserved PIP-box residues	3	1	2	1	7
	Other:					

S5.4.1 Analysis of structures

Table S18: Root-mean-squared deviation (RMSD) values of mutant peptides docked to the monomer of PCNA compared to structures of **p21$_{\mu}$** bound to PCNA and p21 bound to PCNA.

Name	Structure Type	RMSD value against Wild Type p21$_{139-160}$ (monomer of 1AXC)	RMSD value against monomer of p21$_{\mu}$ bound to PCNA (PDB ID:7KQ1)
p21$_{\mu}$	Co-crystal – 7KQ1	0.511	-
p21$_{\mu}$	Computational model	0.316	0.233
p21$_{\mu}$-F150Y	Co-crystal – 7KQ0	0.342	0.451
p21$_{\mu}$-S146R	Computational model	0.570	0.176
p21$_{\mu}$-M147I	Computational model	0.571	0.191
p21$_{\mu}$-D149E	Computational model	0.556	0.197
p21$_{\mu}$-FY150/151YF	Computational model	0.559	0.183
p21$_{\mu}$-PARG	Computational model	0.571	0.192
p21$_{\mu}$-Pogo	Computational model	0.554	0.271
p21$_{\mu}$-pol δ$_{p66}$	Computational model	0.559	0.186
p21$_{\mu}$-pol ι	Computational model	0.563	0.199
p21$_{\mu}$-RFC	Computational model	0.567	0.209
p21$_{\mu}$-RD1	Computational model	0.600	0.270
p21$_{\mu}$-RD2	Computational model	0.558	0.181
p21$_{\mu}$-RD3	Computational model	0.552	0.193

Table S19: Buried surface area (BSA) for PIP-box residues from the co-crystal structures and computationally modelled peptides calculated using the PDBePISA server v1.52 (https://www.ebi.ac.uk/pdbe/pisa/). Å² / Buried area percentage

PIP-box Residue	144^{P1}	145^{P2}	146^{P3}	147^{P4}	148^{P5}	149^{P6}	150^{P7}	151^{P8}
Peptide								
p21 (1AXC)[3]	107.86 / 70	48.31/ 70	30.39 / 60	135.10 / 100	47.60 / 60	0 / 0	59.31 / 50	132.20 / 90
p21$_{\mu}$ (7KQ1)	98.61 / 70	46.23 / 70	31.97 / 60	140.61 / 100	34.23 / 40	0 / 0	66.67 / 50	130.92 / 90
p21$_{\mu}$-F150Y (7KQ0)	105.46 / 70	45.61 / 70	32.11 / 60	134.21 / 100	42.33 / 50	0 / 0	73.82 / 60	128.04 / 100
p21$_{\mu}$-S146R	93.41 / 70	46.73 / 70	68.11 / 50	142.96 / 100	32.99 / 40	0 / 0	74.02 / 60	133.52 / 90
p21$_{\mu}$-M147I	102.47 / 60	35.07 / 60	28.13 / 50	129.11 / 100	54.96 / 60	0 / 0	81.05 / 60	142.77 / 100
p21$_{\mu}$-D149E	98.26 / 60	50.34 / 80	29.24 / 50	140.27 / 100	42.59 / 50	0 / 0	75.43 / 60	136.21 / 100
p21$_{\mu}$-FY150/151YF	105.76 / 70	56.32 / 70	25.76 / 50	139.54 / 100	43.9 / 50	0 / 0	88.7 / 60	115.46 / 90
p21$_{\mu}$-PARG	37.43 / 30	60.13 / 80	60.46 / 50	126.43 / 100	50.59 / 60	0 / 0	66.62 / 50	117.12 / 100
p21$_{\mu}$-Pogo	73.05 / 60	71.12 / 60	53.08 / 50	131.07 / 100	62.18 / 60	0 / 0	83.77 / 60	120.27 / 90
p21$_{\mu}$-pol δ$_{p66}$	98.14 / 60	50.67 / 50	26.13 / 40	123.79 / 100	57.64 / 60	0 / 0	65.75 / 50	123.84 / 90
p21$_{\mu}$-pol ι	88.44 / 50	33.75 / 70	46.18 / 50	142.62 / 100	59.19 / 70	0 / 0	97.25 / 60	74.44 / 60
p21$_{\mu}$-RFC	76.16 / 40	86.28 / 60	23.01 / 40	138.71 / 100	85.43 / 50	0 / 0	62.59 / 60	96.61 / 70
p21$_{\mu}$-RD1	99.56 / 60	56.85 / 40	55.85 / 50	143.05 / 100	77.77 / 60	0 / 0	81.89 / 60	127.73 / 90
p21$_{\mu}$-RD2	72.88 / 50	57.05 / 90	65.92 / 50	136.97 / 100	52.52 / 50	0 / 0	76.75 / 50	104.72 / 100
p21$_{\mu}$-RD3	88.02 / 80	22.38 / 20	63.49 / 50	137.11 / 100	60.31 / 80	0 / 0	75.71 / 50	123.64 / 90

S5.5 Comparison of structures to native sequences

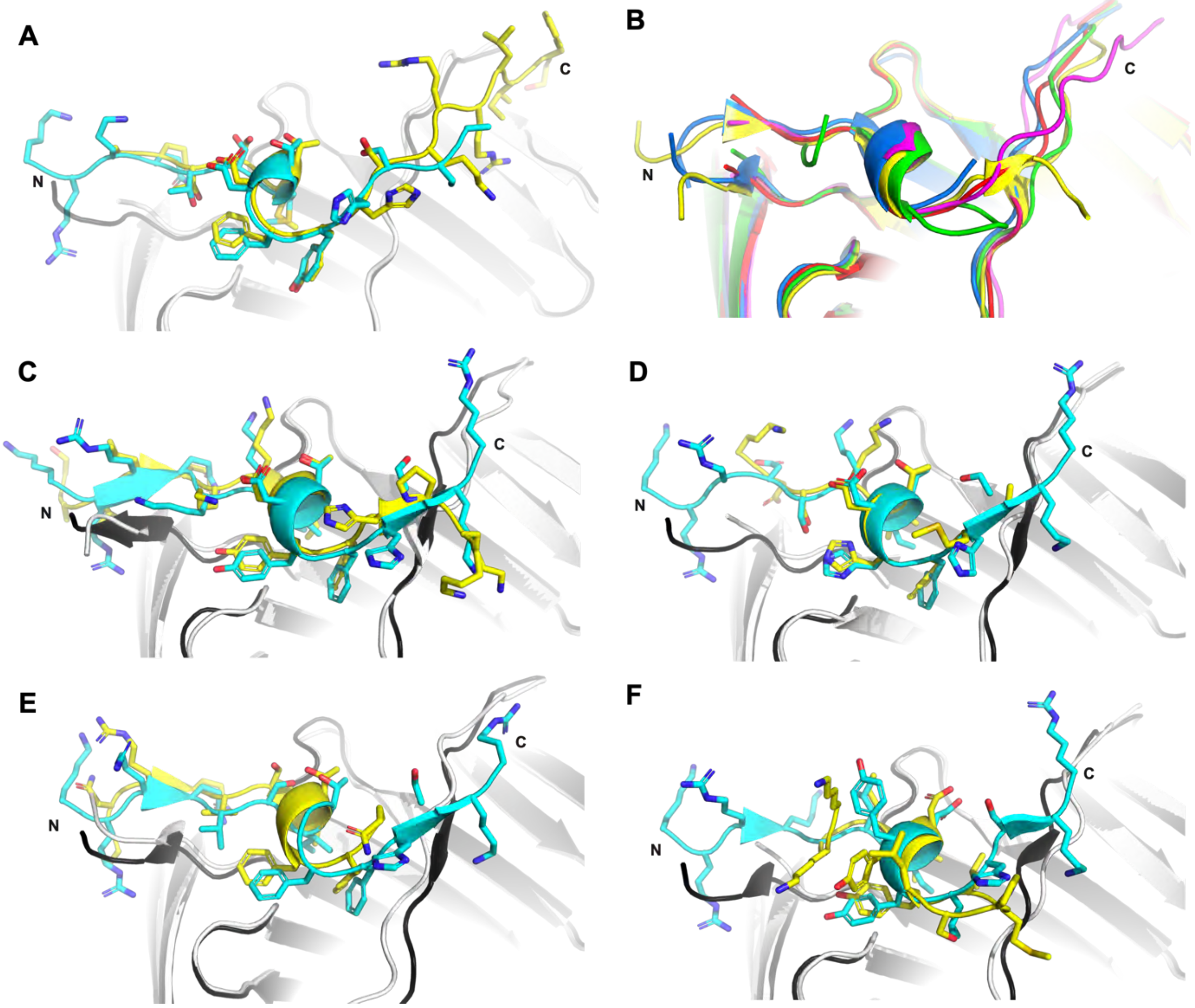


Figure S16: Overlaid structures of p21$_{\mu}$:PIP-box hybrid (PCNA, white; peptide, blue) and native PIP-box peptide (PCNA, black; peptide, yellow) shown in cartoon representation and side-chains of peptide as sticks. **A p21$_{\mu}$** (7KQ1) and **p21$_{139\text{-}160}$** (1AXC) **B** Overlay of Native PIP-box peptides (**p21$_{139\text{-}160}$** 1AXC, pink; **PL** 1VYJ, yellow; **pol δ$_{p66\ 452\text{-}466}$** 1U76, blue; **PARG$_{402\text{-}420}$** 5MAV, red; **pol ι $_{415\text{-}437}$** 2ZVM, green **C p21$_{\mu}$-Pogo** and **PL** (IVYJ) **D p21$_{\mu}$-PARG** and **PARG$_{402\text{-}420}$** (5MAV) **E p21$_{\mu}$-pol δ** and pol **δ$_{p66\ 452\text{-}466}$** (1U76) **F p21μ-pol ι** and **pol ι $_{415\text{-}437}$** (2ZVM)**.**

Table S20: Root-mean-squared deviation (RMSD) values of alternative PIP-box modified peptides docked to monomer of PCNA compared to the structures of their respective native peptides bound to PCNA

Alternative PIP-box mutant peptide docked to PCNA	Structure of native peptide bound to PCNA	RMSD value
p21$_{\mu}$-PARG	5MAV	0.496
p21$_{\mu}$-Pogo	1VYJ	0.646
p21$_{\mu}$-pol δ$_{p66}$	1U76	0.696
p21$_{\mu}$-pol ι	2ZVM	1.136

Table S21: Number of interactions for p21$_{\mu}$:hybrid peptides compared to native analogues determined from co-crystal or computationally modelled structures with RING server. For native structures the interactions are averaged over all subunit present in the pdb coordinate file.

					# Interactions					
Peptide	**PDB ID:**	**Ref**	**Affinity**	**# res**	**Total**	**PIP-box**	**Conserved**	**Non-conserved**	**Flanking**	**RING [4] session ID**
p21 $_{139-160}$	1AXC	[3]	5.96 nM	22	60.67	25	9.33	15.67	35.67	5ef515200e9f94078ea226cb
p21$_{\mu}$	7KQ1		26.1 nM	15	30.67	23.67	16.67	7	7	5ef50f2b0e9f94078ea226bd
PL $_{1-16}$ [mutant]	1VYJ	[5]	100 nM	16	22.33	17.67	13.33	4.33	4.67	5ef515730e9f94078ea226cc
p21$_{\mu}$-Pogo			9.12 nM	15	34	24	16	5	10	5f3b224d0e9f94078ea22d01
pol $\delta_{p66\ 452-466}$	1U76	[6]	15.6 μM	15	30.33	24.33	17	7.33	6	5ef516b50e9f94078ea226cf
p21$_{\mu}$-pol δ_{p66}			268 nM	15	29	16	11	3	13	5f3b202a0e9f94078ea22cfb
PARG $_{402-420}$	5MAV	[7]	3.3 μM	19	26	19	4.83	14.17	7	5ef5164c0e9f94078ea226ce
p21$_{\mu}$-PARG			401 nM	15	24	15	8	4	9	5f3b1de50e9f94078ea22cf7
pol ɩ $_{415-437}$	2ZVM	[8]	0.39 μM	23	22.33	17.67	17.67	4.67	4.67	5ef515e00e9f94078ea226cd
p21$_{\mu}$-pol ɩ			1.42 μM	15	26	16	9	7	10	5f3b1e580e9f94078ea22cf8

REFERENCES

1. N. J. Anthis and G. M. Clore, Sequence-specific determination of protein and peptide concentrations by absorbance at 205 nm, *Protein Sci.*, 2013, **22**, 851-858.

2. A. Wlodawer, W. Minor, Z. Dauter and M. Jaskolski, Protein crystallography for aspiring crystallographers or how to avoid pitfalls and traps in macromolecular structure determination, *FEBS J.*, 2013, **280**, 5705-5736.

3. J. M. Gulbis, Z. Kelman, J. Hurwitz, M. O'Donnell and J. Kuriyan, Structure of the C-Terminal Region of p21 WAF1/CIP1 Complexed with Human PCNA, *Cell*, 1996, **87**, 297-306.

4. D. Piovesan, G. Minervini and S. C. Tosatto, The RING 2.0 web server for high quality residue interaction networks, *Nucleic Acids Res.*, 2016, **44**, W367-374.

5. G. Kontopidis, S.-Y. Wu, D. I. Zheleva, P. Taylor, C. McInnes, D. P. Lane, P. M. Fischer and M. D. Walkinshaw, Structural and biochemical studies of human proliferating cell nuclear antigen complexes provide a rationale for cyclin association and inhibitor design, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 1871-1876.

6. J. B. Bruning and Y. Shamoo, Structural and thermodynamic analysis of human PCNA with peptides derived from DNA polymerase-delta p66 subunit and flap endonuclease-1, *Structure*, 2004, **12**, 2209-2219.

7. T. Kaufmann, I. Grishkovskaya, A. A. Polyansky, S. Kostrhon, E. Kukolj, K. M. Olek, S. Herbert, E. Beltzung, K. Mechtler, T. Peterbauer, J. Gotzmann, L. Zhang, M. Hartl, B. Zagrovic, K. Elsayad, K. Djinovic-Carugo and D. Slade, A novel non-canonical PIP-box mediates PARG interaction with PCNA, *Nucleic Acids Res.*, 2017, **45**, 9741-9759.

8. A. Hishiki, H. Hashimoto, T. Hanafusa, K. Kamei, E. Ohashi, T. Shimizu, H. Ohmori and M. Sato, Structural basis for novel interactions between human translesion synthesis polymerases and proliferating cell nuclear antigen, *J. Biol. Chem.*, 2009, **284**, 10552-10560.

Chapter 6.

A CELL PERMEABLE BIMANE-CONSTRAINED PCNA-INTERACTING PEPTIDE

Research Article: RSC Chemical Biology, **2021**, In press

A cell permeable bimane-constrained PCNA-interacting peptide

Aimee J. Horsfall[1-3], Beth A. Vandborg[1-4], Zoya Kikhtyak[5], Denis B. Scanlon[1,2], Wayne D. Tilley[5], Theresa E. Hickey[5], John B. Bruning[1-4] and Andrew D. Abell[1-3*]

[1] Institute of Photonics and Advanced Sensing (IPAS), The University of Adelaide, Adelaide, South Australia, 5005, Australia

[2] School of Physical Sciences, The University of Adelaide, Adelaide, South Australia 5005, Australia

[3] Australian Research Council Centre of Excellence for Nanoscale BioPhotonics (CNBP)

[4] School of Biological Sciences, The University of Adelaide, Adelaide, South Australia 5005, Australia

[5] Dame Roma Mitchell Cancer Research Laboratories, Adelaide Medical School, Faculty of Health and Medical Sciences, University of Adelaide, Adelaide, South Australia 5005, Australia

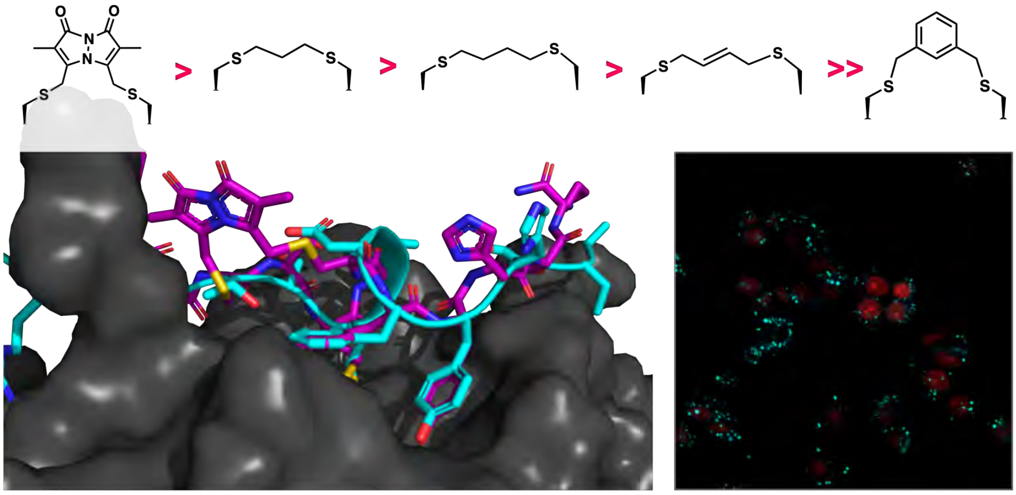
*Corresponding author

© 2021 RSC Chemical Biology, Royal Society of Chemistry

First published online: 21st July 2021

ABSTRACT

The human sliding clamp protein known as Proliferating Cell Nuclear Antigen (PCNA) orchestrates DNA-replication and -repair and as such is an ideal therapeutic target for proliferative diseases, including cancer. Peptides derived from the human p21 protein bind PCNA with high affinity via a 3_{10}-helical binding conformation and are known to shut down DNA-replication. Here, we present studies on short analogues of p21 peptides (143-151) conformationally constrained with a covalent linker between *i*, *i*+4 separated cysteine residues at positions 146 and 149 to access peptidomimetics that target PCNA. The resulting macrocycles bind PCNA with K_D values ranging from 570 nM to 3.85 μM, with the bimane-constrained peptide **7** proving the most potent. Subsequent X-ray crystallography and computational modelling studies of the macrocycle peptides bound to PCNA indicated only the high-affinity peptide **7** adopted the classical 3_{10}-helical binding conformation. This suggests the 3_{10}-helical conformation is critical to high affinity PCNA binding, however NMR secondary shift analysis of peptide **7** revealed this secondary structure was not well-defined in solution. Peptide **7** is cell permeable and localised to the cell cytosol of breast cancer cells (MDA-MB-468), revealed by confocal microscopy showing blue fluorescence of the bimane linker. The inherent fluorescence of the bimane moiety present in peptide **7** allowed it to be directly imaged in the cell uptake assay, without attachment of an auxiliary fluorescent tag. This highlights a significant benefit of using a bimane constraint to access conformationally constrained macrocyclic peptides. This study identifies a small peptidomimetic that binds PCNA with higher affinity than previous reported p21 macrocycles, and is cell permeable, providing a significant advance toward development of a PCNA inhibitor for therapeutic applications.



Five macrocyclc p21-derived peptides targeting human sliding clamp, PCNA were developed. The most potent was inherently fluorescent and cell permeable.

STATEMENT OF AUTHORSHIP

Title of Paper	A cell permeable, fluorescent peptidomimetic targeting human PCNA
Publication Status	☐ Published ☐ Accepted for Publication ■ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	**Research Article:** A. J. Horsfall, B. A. Vandborg, Z. Kikhtyak, D. B. Scanlon, T. E. Hickey, W. D. Tilley, J. B. Bruning and A. D. Abell, *RSC Chemical Biology*, 2021

Principal Author

Name of Principal Author (Candidate)	Aimee J Horsfall		
Contribution to the Paper	Designed and synthesised peptides, NMR assignment and analysis, analysed & discussed results, wrote and edited manuscript		
Overall percentage (%)	50%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	15/02/2021

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);
ii. permission is granted for the candidate in include the publication in the thesis; and
iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Beth A Vandborg		
Contribution to the Paper	Protein synthesis & purification, protein crystallography, computational modelling studies, discussed results, edited manuscript		
Signature		Date	01/03/2020

Name of Co-Author	Zoya Kikhtyak		
Contribution to the Paper	Cell uptake assays, discussed results, and edited manuscript.		
Signature	[redacted]	Date	19.02.2021

Name of Co-Author	Denis B Scanlon		
Contribution to the Paper	Peptide synthesis, discussed results, and edited manuscript.		
Signature	[redacted]	Date	01/03/2021

Name of Co-Author	Theresa E Hickey		
Contribution to the Paper	Supervised cell uptake assays, discussed results, and edited manuscript.		
Signature	[redacted]	Date	19/02/21

Name of Co-Author	Wayne D Tilley		
Contribution to the Paper	Supervised cell uptake assays, discussed results, and edited manuscript.		
Signature	[redacted]	Date	19/02/21

Name of Co-Author	John B Bruning		
Contribution to the Paper	Supervised AJH & BAV, discussed results, and edited manuscript.		
Signature	[redacted]	Date	18-3-21

6.1 Introduction

Proliferating cell nuclear antigen (PCNA) interacts with more than 200 proteins to act as an essential mediator of DNA-replication and -repair processes.[1-3] As such, it has been identified as a key therapeutic target for the treatment of diseases defined by aberrant DNA-replication, including many cancers.[4-7] The cell cycle regulator protein, p21, binds PCNA with the highest known affinity of all PCNA interacting proteins, through a motif referred to as the PCNA-Interacting Protein (PIP) box.[7-10] p21 is an intrinsically disordered protein, however on binding PCNA, it adopts a well-defined single 3_{10}-helical turn to insert three hydrophobic residues (Met147, Phe150 and Tyr151) into a cleft on the PCNA surface (Figure 1).[11] This 3_{10}-helix is observed in the binding conformation of almost all PCNA-interacting proteins or peptides, and is often flanked by two short β-strands.[7]

A p21 peptide (residues 139-160) derived from the *C*-terminus of p21 contains the PIP-box QTSMTDFY, and has been reported to bind PCNA with an affinity of ~5 nM.[7, 8, 12, 13] This peptide competitively inhibits interaction between the major processive polymerase δ and PCNA to shut down SV40 DNA replication *in vitro*.[14] Thus, p21-derived peptides provide an optimal template to design anti-cancer therapeutics that target PCNA. A conformation favourable to binding can be stabilised in such a peptide by covalently linking two appropriately spaced amino-acid side-chains, either by reaction together, or reaction with a bifunctional reagent.[15-18] To this end, we previously reported the first macrocyclic peptidomimetic of $p21_{139-152}$, where a lactam bridge was introduced between a lysine or 2,4-diaminobutyric acid residue at position 145, and a glutamic acid residue at 149 to provide two peptidomimetics referred to as ACR1 and ACR2, respectively.[19] The lactam bridge was shown by NMR to stabilise a 3_{10}-helix in solution, where the 3_{10}-helix was better defined in the smaller macrocycle ACR2, than in ACR1. Interestingly, ACR1 bound PCNA with higher affinity than ACR2, indicating the well-defined 3_{10}-helix in solution did not correlate with higher PCNA affinity. A crystal structure of ACR2 bound to PCNA suggested an unfavourable interaction of the lactam linker amide bond of ACR2 and Phe150 of the peptide, which may account for the lower affinity for PCNA. Consequently, we set out to investigate a variety of different linkers in a short p21 peptide ($p21_{143-154}$, peptide **1**) in order to study how different peptide linkers influence the secondary structure and affinity of macrocyclic peptides bound to PCNA. A range of different macrocycles can be rapidly accessed from the same parent peptide containing two cysteines (peptide **2**) using dithiol-bisalkylation chemistry.[18, 20] Here, five macrocyclic p21 peptide derivatives were synthesised (peptides **3**-**7**) and their affinity for PCNA determined by surface plasmon resonance (SPR). The conformation of the macrocycles bound to PCNA was studied by X-ray crystallography and computational modelling, to investigate how the linkers interact with the PCNA surface, and how the resulting peptide structures relate to binding affinity for PCNA. These studies reveal a bimane-constrained peptide with high affinity for PCNA that is cell permeable and is a promising lead for development of a potential therapeutic that targets PCNA.

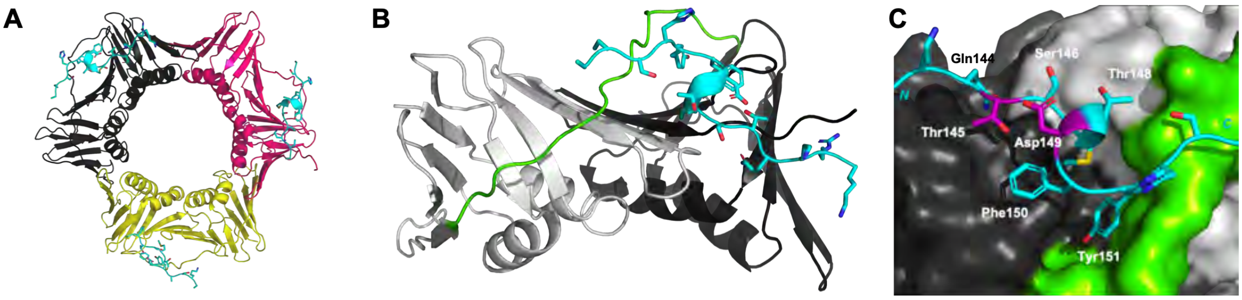


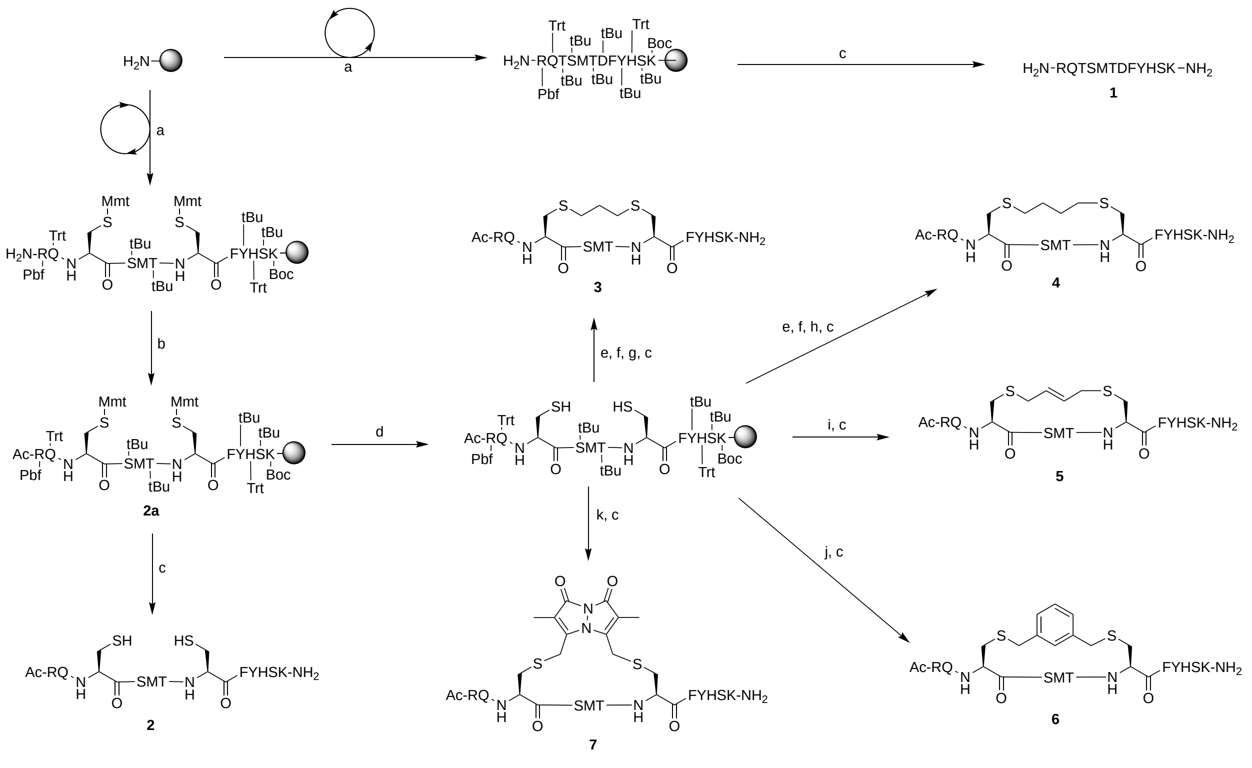
Figure 1: $p21_{141-155}$ (cyan, peptide **1**) bound to PCNA (PDB: 7KQ1). **A** Ring-shaped PCNA with three peptides (cyan) bound to the PIP-box binding site. PCNA monomers (three) shown in shades of grey, pink and yellow. **B** Single PCNA subunit with two domains shown in shades of grey and inter-domain connecting loop (IDCL) in green. $p21_{141-155}$ shown in cyan. **C** PCNA shown as surface representation (grey) with IDCL highlighted (green). $p21_{141-155}$ shown in cartoon and side-chains as sticks. PIP-box residues are labelled and the residues (145 and 149) to be modified and reacted to form a constraint are highlighted in purple.

6.2 Results and discussion

6.2.1 Design and synthesis

The peptide $p21_{139-160}$ is reported to bind PCNA with high specificity and provides an ideal template to develop PCNA inhibitors. It is known that the $p21_{139-160}$ sequence can be shortened without drastically impacting affinity for PCNA[13, 19] and consequently a short variant, $p21_{143-154}$ (**1**), was chosen as the starting scaffold for this study for ease of synthesis and to improve synthetic yields. The sequence was modified to include two cysteine residues at positions 145 and 149, in place of threonine and aspartic acid of $p21_{143-154}$ (Figure 1C, purple), to give peptide **2**. Positions 145 and 149 were chosen for modification as these side-chains are close in space when bound to PCNA (Figure 1C), and were successfully linked in a p21 lactam bridge peptide in our earlier study to stabilise the 3_{10}-helical binding conformation.[19] This structure-informed design gave rise to an *i*-*i*+4 constraint, although *i*-*i*+3 constraints are more commonly reported for 3_{10}-helices. The cysteine-containing peptide can be derivatised into a range of macrocycles, where the cysteine residues are linked by dithiol *bis*-alkylation (Scheme 1). Three alkyl linkers (propyl, butyl and *trans*-butenyl),[21, 22] the aromatic *m*-xylene linker,[20, 22] and our fluorescent bimane linker[23] were chosen as a chemically distinct range of linkers which vary in length and rigidity, in order to investigate how the different linkers impact peptide structure and PCNA binding affinity.

All peptides were synthesised by a standard Fmoc/tBu solid-phase peptide synthesis (SPPS) protocol detailed in the Experimental Section (see SI) and shown in Scheme 1. For the peptide macrocycles, the linear amino-acid sequence RQCSMTCFYHSK was assembled on a solid-support, and the *N*-terminal amine capped by reaction with acetic anhydride to give resin-bound **2a** (Scheme 1). The cysteine side-chain protecting groups, 4-methoxytrityl (Mmt), were then selectively deprotected on-resin by repetitive 1 min treatment with 2% TFA in DCM. This provided a resin-bound peptide with two free thiols, which were subsequently reacted with either 1,3-dibromopropane, 1,4-dibromobutane, *trans*-1,4-dibromobut-2-ene, *m*-dibromoxylene or dibromobimane (Scheme 1) under literature conditions (See Experimental Section in SI).[21, 23, 24] Reaction of the resin-bound peptides with 1,3-dibromopropane or 1,4-dibromobutane required addition of NaI to generate the more reactive alkyliodides *in situ*,[21] whereas reaction with *trans*-1,4-



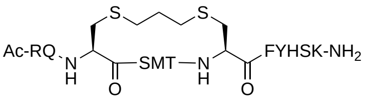
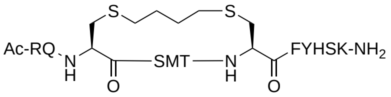
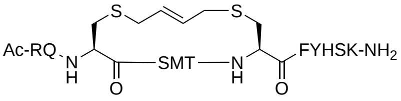
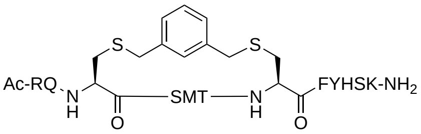
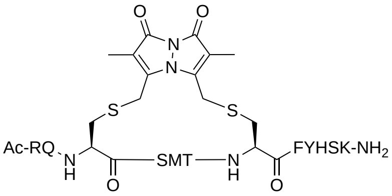
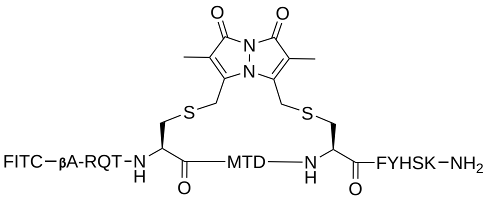
Scheme 1: Peptide synthesis scheme (**1**-**7**) and generation of multiple cyclic peptides from a single parent utilising cysteine bis-alkylation to give macrocycles **3**-**7**. **a.** Successive Fmoc-amino-acid coupling and Fmoc-deprotection steps to assemble the peptide sequence. Coupling: HATU (5 equiv), DIPEA (10 equiv), Fmoc-AA-OH (5 equiv), DMF, 1 h. Deprotection: 20% piperidine with 0.1 M HOBt, DMF, 10min. **b.** Acetylation: Ac_2O (50 equiv), DIPEA (50 equiv), DMF, 15 min. **c.** Cleavage: 92.5:2.5:2.5:2.5 TFA/TIPS/DODT/H_2O, 2 h. **d.** Mmt deprotection: 2% TFA in DCM, 1 min x 40. **e.** NaI (17.5 equiv), TCEP (0.5 equiv), DMF, N_2, 15 min. **f.** DIPEA (35 equiv), DMF, 20 min. **g.** 1,3-dibromopropane (3.5 equiv), MW 2 min, 125°C. **h.** 1,4-Dibromobutane (3.5 equiv), MW 2 min, 125°C. **i.** *trans*-1,4-Dibromo-2-butene (2 equiv), DIPEA (4 equiv), DMF, 3 h. **j.** Dibromo-*m*-xylene (2 equiv), DIPEA (4 equiv), DMF, 3 h. **k.** Dibromobimane (2 equiv), DIPEA (4 equiv), DMF, 3 h.

dibromobut-2-ene, *m*-dibromoxylene or dibromobimane proceeds under mild basic conditions (in DMF with DIPEA).[25] The resulting macrocyclic peptides were then cleaved from the resin by treatment with 92.5% TFA containing scavengers (2.5% of each TIPS, DODT and H_2O) for 2 h to give the crude macrocyclic peptides **3**-**7** (Scheme 1). The peptides were purified by semi-preparative RP-HPLC, and the identity and purity confirmed by HRMS and analytical RP-HPLC (>90 %), respectively. Full details and characterisation data are described in the Experimental Section (see SI).

6.2.2 Binding affinity

The binding affinity of the parent peptide **1**, linear cysteine-containing precursor **2** and five macrocyclic peptides (**3**-**7**), for PCNA was determined by SPR (Table 1). Peptide **1** bound PCNA with a K_D value of 102 nM, in line with reported affinities of short p21 peptides for PCNA.[8, 9, 13, 19, 26] The cysteine-modified peptide **2** displayed significantly reduced PCNA affinity, a result that is consistent with earlier reports that polar residues at positions 145 and 149 in p21 (Thr and Asp) stabilise the binding conformation via hydrogen bonding, which in turn enhances binding affinity for PCNA.[13, 27, 28] The affinity of the macrocyclic peptides **3**-**7** for PCNA ranged from 570 nM to 3.86 μM, which indicates all macrocycles were able to interact with PCNA in contrast to the precursor peptide **2**. The affinity in all cases was lower than the native peptide **1**, in line with observations made for our previously reported lactam macrocycles.[19] Macrocyclic peptide **3**,

Table 1: SPR binding affinity K_D. SE – standard error. NS represents a non-specific interaction with the sensor chip, consequently the binding affinity for PCNA could not be determined. Ac – acetyl.

Peptide	Sequence	Linker (atom length)	Affinity K_D ± SE (μM)	χ^2
1	H-RQTSMTDFYHSK-NH_2	-	0.102 ± 0.005	0.070
2	Ac-RQCSMTCFYHSK-NH_2	-	NS	-
3		Propyl (7)	0.769 ± 0.078	0.250
4		Butyl (8)	1.99 ± 0.140	0.160
5		*trans*-Butenyl (8)	2.82 ± 0.080	0.106
6		*m*-Xylene (9)	3.86 ± 0.035	0.527
7		Bimane (9)	0.570 ± 0.030	0.118
8	FITC-βA-RQTSMTDFYHSK-NH_2	-	-	-
9		Bimane (9)	25.2 ± 1.90	0.055

with its propyl linker, gave the second highest affinity for PCNA at 769 nM, whereas the larger butyl linker of peptide **4** (8 atom linker c.f. 7 atoms of **3**) resulted in reduced affinity for PCNA with a K_D value of 1.99 μM. The more rigid *trans*-butenyl-based linker in peptide **5** resulted in a K_D value for PCNA of 2.82 μM, and the binding affinity for PCNA of aromatic *m*-xylene linked peptide **6** (9 atom linker) was lower again at 3.86 μM. Considered together, these observations suggest that the longer and more rigid linkers result in a lower binding affinity for PCNA. The bimane-linked peptide **7** proved to be a particularly potent binder with a K_D value of 570 nM, which is only 5-fold lower affinity than the native peptide **1** and is the highest affinity macrocyclic peptide to bind PCNA reported to date. The bimane linker contains 9 atoms but the bimane moiety, in contrast to the xylene ring, is not rigid and is able to flex along the plane of symmetry that bisects the two nitrogen atoms, which may explain in part the observed higher PCNA affinity.

6.2.3 Structural analysis

Next, the binding conformation of macrocycles **3**-**7** bound to PCNA was investigated by X-ray crystallography and computational modelling studies to better understand the structural mechanisms responsible for the difference in PCNA binding affinities. Co-crystal structures of peptide **3** (PDB: 7M5L), **5** (PDB: 7M5M) and **6** (PDB: 7M5N) bound to PCNA were solved and are shown in Figures 2-3, as well as

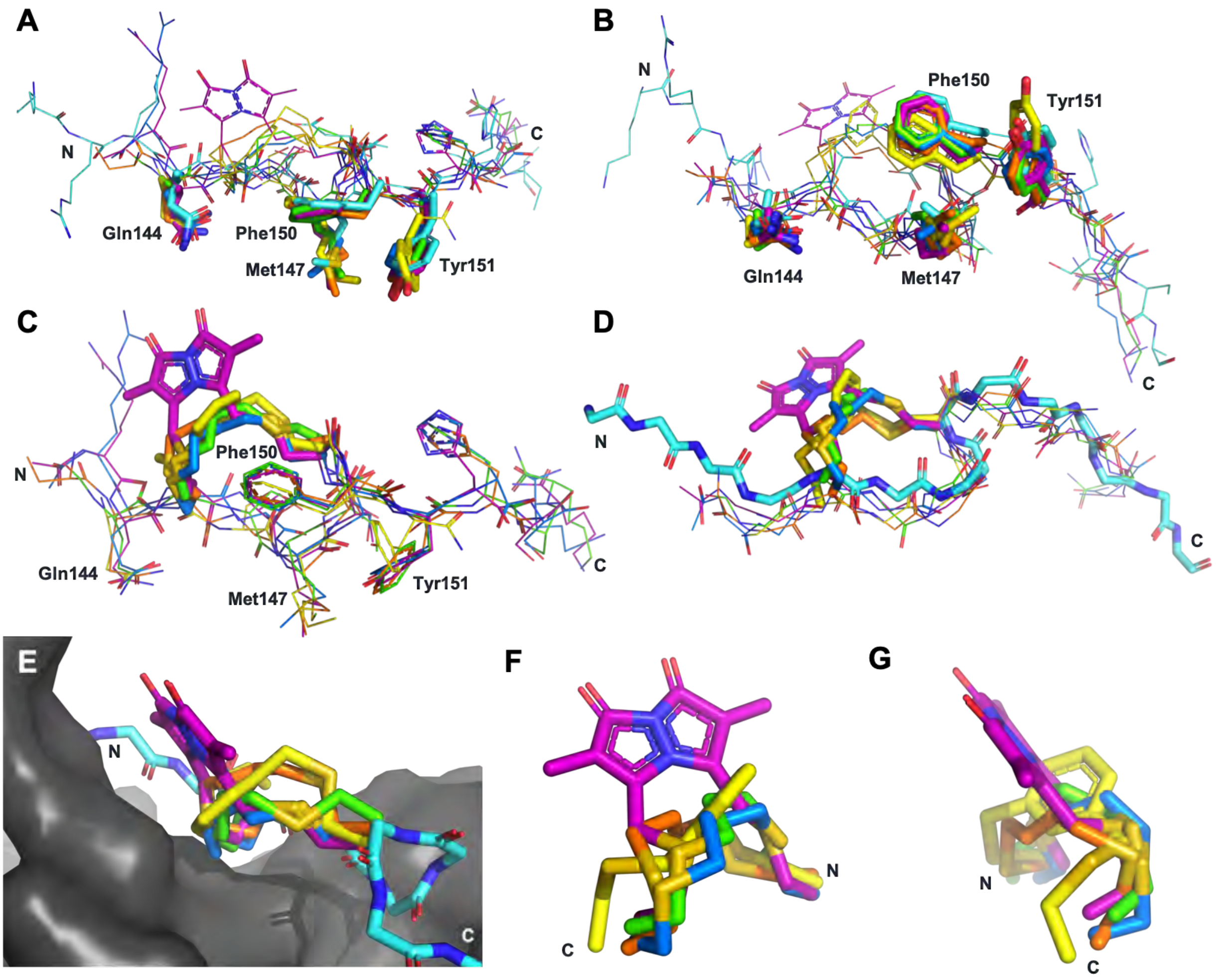


Figure 2: Differences in the linker position, and impact of the linker on backbone conformation. Peptides $p21_{141-155}$ (light blue, 7KQ1), **3** (green, co-crystal 7M5L), **4** (blue, computationally modelled), **5** (orange, co-crystal 7M5M), **6** (yellow, co-crystal 7M5N) and **7** (purple, computationally modelled). **A** & **B** The conserved residues (labelled, shown as sticks) adopt similar conformations in all macrocycles **C** The linkers (shown as sticks) induce subtle changes in the backbone of the peptide, in particular the direction of the amides. **D** The overall backbone conformation of the macrocycles is still able to mimic the $p21_{141-155}$ conformation. **E** The linker length and rigidity alters where the linker sits relative to the PCNA surface **F** & **G** Conformation of the peptide linkers relative to one another.

Figures S2, S4 and S5. Data collection and refinement statistics are summarised in Table S2. Attempts to obtain a co-crystal structure of peptides **1**, **4** and **7** bound to PCNA were unsuccessful at this time. Peptide **1** was however modelled onto the PCNA surface, to confirm the short peptide interacted with the protein surface in the same manner (see Figure S1), and was constructed from the previously published structure of $p21_{141-155}$ bound to PCNA (PDB: 7KQ1).[13] The remaining macrocycles (**4** and **7**) were computationally modelled onto the PCNA surface, where the crystal structure of **3** (PDB: 7M5L) was used as the starting structure. Peptide **3** was energy minimised on the PCNA surface in order to verify the computational method, and indicates a high degree of similarity to the crystal structure (Figure S2F-G). Macrocycles **4** and **7** were then modelled onto the PCNA surface and energy minimised in the same way and are shown in Figures 2-3, and S3 and S6.

The structures of all five macrocyclic peptides bound to PCNA show the conserved PIP-box residues Gln144, Met147, Phe150 and Tyr151 inserted onto the PCNA surface, in a similar mode to the p21 native peptide structures (7KQ1,[13] Figure 2A-D). The structures of macrocyclic peptides **3**-**7** bound to

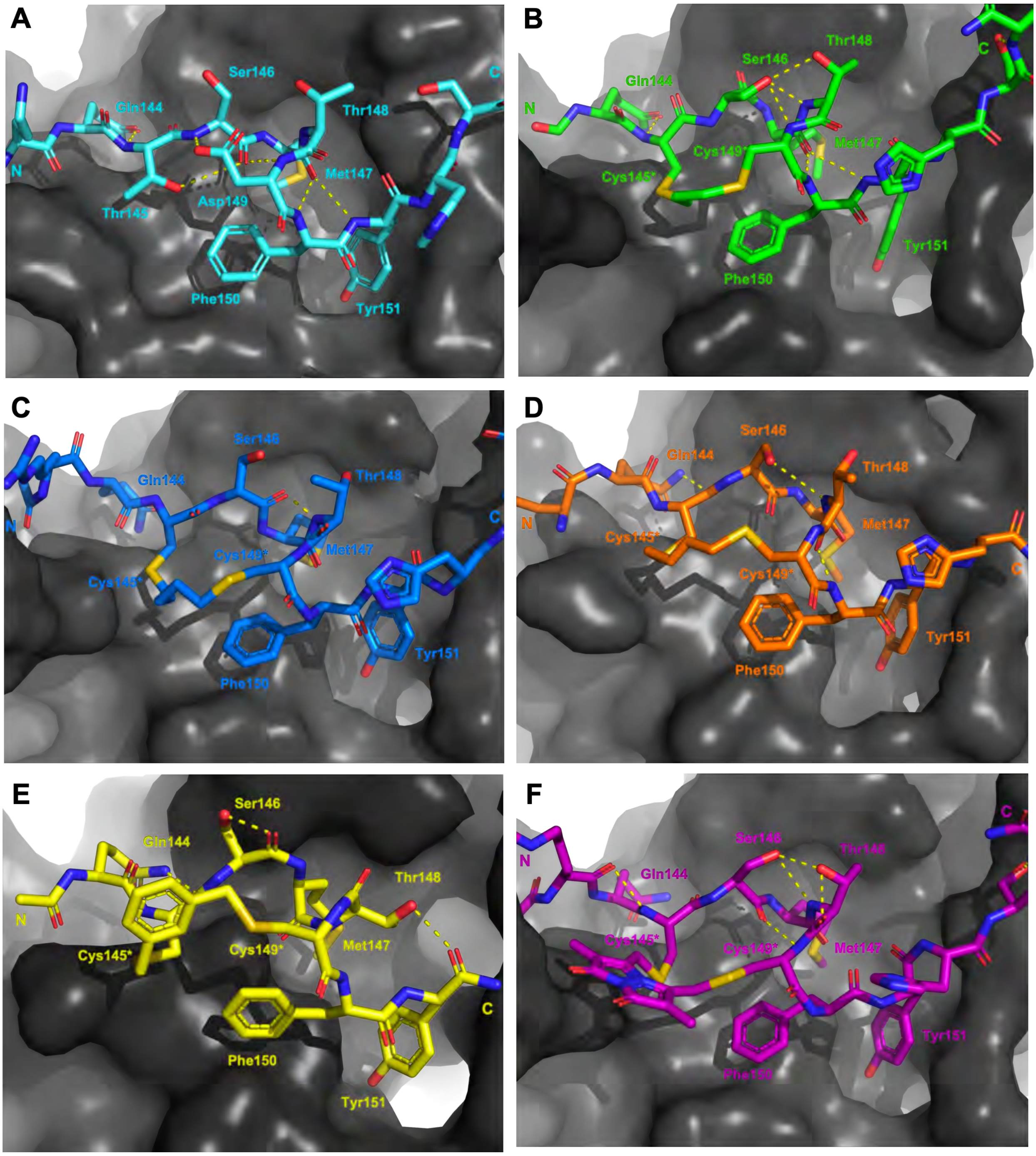

A
Ser146
Gln144
Thr148
Met147
Asp149
Thr145
Phe150
Tyr151
N
C
B
Thr148
Ser146
Gln144
Met147
Cys149*
Cys145*
Tyr151
Phe150
N
C
C
Ser146
Thr148
Gln144
Met147
Cys149*
Cys145*
Tyr151
Phe150
N
C
D
Ser146
Gln144
Thr148
Met147
Cys145*
Cys149*
Tyr151
Phe150
N
C
E
Ser146
Gln144
Thr148
Cys145*
Cys149*
Met147
Phe150
Tyr151
N
C
F
Ser146
Thr148
Gln144
Cys145*
Cys149*
Met147
Tyr151
Phe150
N
C

Figure 3: Structures of peptides (stick representation) bound to PIP-box binding site on PCNA (grey, surface). Intramolecular polar interactions are shown as yellow dashed lines. **A.** p21$_{141\text{-}155}$ (light blue), PDB:7KQ1[13]. **B.** Peptide **3** – with propyl linker (green), co-crystal structure (PDB: 7M5L). **C.** Peptide **4** – with butyl linker (blue), computationally modelled. **D.** Peptide **5** – with *trans*-butenyl linker (orange) co-crystal structure (PDB: 7M5M). **E.** Peptide **6** – with *m*-xylene linker (yellow), co-crystal structure (PDB: 7M5N). **F.** Peptide **7** – with bimane linker (purple), computationally modelled.

PCNA indicate a percentage buried surface area (%BSA) of 60% for the PIP-box residues collectively (Table S3), suggesting similar structures overall. These observations together indicate the linker, in all cases, has not compromised the ability of the macrocycles to interact with PCNA. The difference in the linker conformations is shown in Figure 2C-H. Interestingly, the main-chain hydrogen bond between residue 146 (*i*) and 149 (*i*+3), that defines the classic 3_{10}-helical binding conformation of PCNA-binding peptides, is only observed for peptide **7** which possesses the highest binding affinity of a macrocycle for PCNA, to date.

The co-crystal structure of the propyl-linked peptide **3** bound to PCNA shows the conserved Gln144 side-chain makes an intermolecular 3.2 Å hydrogen bond with Ala252 (Figure S2A-D) and a 2.9 Å intramolecular hydrogen bond to the Cys145* amide (Figure 3B, green). Additionally, the Cys145* amide makes a 2.9 Å mainchain hydrogen bond to the carbonyl of Pro253. The aromatic residues Phe150 and Tyr151 are positioned similarly to the analogous residues in $p21_{141\text{-}155}$ (Figure 2), however Tyr151 does not make a hydrogen bond with Gln131 as in $p21_{141\text{-}155}$. The Met147 carbonyl makes an intramolecular hydrogen bond to the amide NH of Tyr151 (3.4 Å) and also to Phe150 (2.4 Å), which defines an unconventional α-helical PCNA binding conformation (c.f. 3_{10}-helix from 146 – 149 for $p21_{141\text{-}155}$). The Ser146 side-chain makes a 2.4 Å and 3.1 Å intramolecular bond to the amides of Thr148 and Cys149* respectively, to further stabilise this conformation. The propyl linker sits above Pro253 of PCNA (4.4 Å) and Phe150 of the peptide (4.4 Å) to provide favourable hydrophobic interactions (Figure 3B, green) which may be in part responsible for the relatively high affinity of **3** (769 nM).

The co-crystal structure of *trans*-butenyl-linked peptide **5** bound to PCNA shows the conserved hydrophobic residues (Met147, Phe150, Tyr151) are in a similar conformation to the analogous residues in $p21_{141\text{-}155}$ (Figure 2); however, the %BSA for Met147 is notably lower at 70% than in $p21_{141\text{-}155}$, or the other macrocycles (100%, Table S3). Three intermolecular hydrogen bonds are observed between peptide **5** and PCNA to anchor the macrocycle onto the surface (Figure S4). These are between: the Met147 amide NH and the mainchain carbonyl of His44 (2.3 Å); the Gln144 side-chain to the Ala252 carbonyl (2.8 Å) and the Pro253 carbonyl and Cys145 amide NH (3.4 Å). A weak 3.1 Å interaction between the Cys145* sulfur and Pro253 carbonyl is also observed. A 2.2 Å intramolecular hydrogen bond between Thr148 carbonyl and amide of Phe150 defines a γ-turn (Figure 3D, orange). The butenyl linker of peptide **5** is raised off the PCNA surface and the only interaction with the PCNA surface is from the aforementioned intermolecular bond of the Cys145* sulfur and Pro253. The carbon-chain of the linker in macrocycle **5** does not appear to interact with the protein surface, and does not stabilise a 3_{10}-helix, which may together account for the lower affinity of peptide **5** (2.82 µM) compared to peptide **3** (769 nM).

The co-crystal structure of peptide **6** bound to PCNA shows the conserved Gln144 and Met147 positioned similar to those in $p21_{141\text{-}155}$, however the aromatic residues are notably shifted (Figure 2). Tyr151 protrudes into the hydrophobic cleft where it is 3.5 Å from Gln131 located in the bottom of the cleft and interacts with the Tyr133 phenol at 3.1 Å (Figure S5). The *m*-xylene linker sits up off the PCNA surface and does not make any clear interactions, except the Cys145* amide makes a weak 3.3 Å with Pro253 (Figure S5). A weak intramolecular 3.0 Å hydrogen bond between the Met147 carbonyl and Phe150 amide defines a 3_{10}-helical turn, however this not at the same location as in native $p21_{141\text{-}155}$ (i.e. 146 to 149). Additionally, this conformation does not appear well stabilised as there are only three other intramolecular interactions present. The first between the Ser146 side-chain and carbonyl; second between the Thr148 side-chain and carbonyl; and third between the Thr148 side-chain and the Tyr151 carbonyl (Figure 3E, yellow). The sequence that flanks the PIP-box is not resolved in this co-crystal structure which may also suggest a weakly defined structure, and may account for the lower affinity of **6** (K_D 3.86 μM), than for peptides **3**-**5** and **7**.

The computational model of peptide **4** bound to PCNA shows the conserved residues are positioned similarly to $p21_{141-155}$, except the Tyr151 %BSA is reduced to 70% compared to 90-100% in $p21_{141-155}$ and peptides **3**, **5**-**7** (Table 3). The Gln144 side-chain makes a 3.0 Å hydrogen bond with Ala252, however this is the only intermolecular hydrogen bond made by a PIP-box residue in peptide **4** (Figure S3). One intramolecular 2.0 Å interaction is observed between the Ser146 carbonyl and Thr148 amide (Figure 3C, blue). The overall lack of secondary interactions likely contributes to the lower affinity observed for **4** (1.99 μM) compared to **3** (769 nM). The butyl linker is >4 Å from the PCNA surface and does not make any clear interactions with PCNA, though a weak 3.0 Å interaction between the Cys145* sulfur and Pro253 carbonyl may be present.

The computational model of peptide **7** bound to PCNA (Figure 2, purple) shows the backbone is shifted compared to $p21_{141-155}$, indicated by the decreased distance between the mainchain amide of Met147 and the mainchain carbonyl atom of Val45 in PCNA (4.5 Å to 4.1 Å). This shift does not impact the overall %BSA of the sidechain in the hydrophobic pocket which remains at 100% (Table S3). Peptide **7** is the only macrocycle that adopts the classical 3_{10}-helix binding conformation, where the carbonyl of Ser146 makes a 2.6 Å hydrogen bond to Cys149* amide (Figure 3F, purple). This conformation is further stabilised by four additional intramolecular bonds which exist between the Ser146 side-chain and Thr148 amide (3.1 Å), the Arg143 carbonyl to Cys145* amide (2.7 Å); the Thr148 side-chain to carbonyl (2.7 Å); and the Ser153 side-chain to the Lys154 carbonyl (2.4 Å) (Figure 3F). This stabilised 3_{10}-helical binding conformation likely contributes to the high binding affinity of peptide **7** for PCNA observed (570 nM). Interestingly, there are no significant intermolecular hydrogen bonds made between peptide **7** and PCNA, except for a 2.7 Å hydrogen bond of Arg143 to Asp257 and a weak 3.2 Å interaction of the bimane carbonyl and the main-chain carbonyl of Glu256 (Figure S6).

The classical 3_{10}-helical conformation seen for PCNA-binding peptides was only observed here for PCNA-bound macrocycle **7**. Consequently, the solution-phase structure of the high PCNA affinity peptide **7** was investigated by NMR to investigate whether the bimane linker in **7** pre-organises the peptide backbone into this conformation. Secondary shifts of the bimane-linked peptide **7** were determined for the α-proton (Hα), amide proton (NH), α-carbon (Cα) and carbonyl carbon (CO) for each residue (Figure 4, Table S6). The secondary shift for each signal (Hα, NH, Cα, CO; Figure 4) was calculated as the difference between the observed resonance in macrocycle **7** (Table S5) and the observed resonance in the linear precursor peptide **2** for each residue (Table S4). The secondary shifts in the region of interest (i.e. between residues 146 and 149), were almost all shifted in the direction expected for helical structure, that is, negative proton (<-0.1 ppm) and positive carbon secondary shifts (> 1 ppm). Random coil structure is indicated by secondary shifts close to 0 ppm. Large negative NH secondary shifts for residues 148-150, in addition to negative Hα and positive Cα secondary shifts for residues 147 and 148 together suggest helical structure in this region (Figure 4). Furthermore, the CO secondary shifts of residues 145-147 are positive, which is

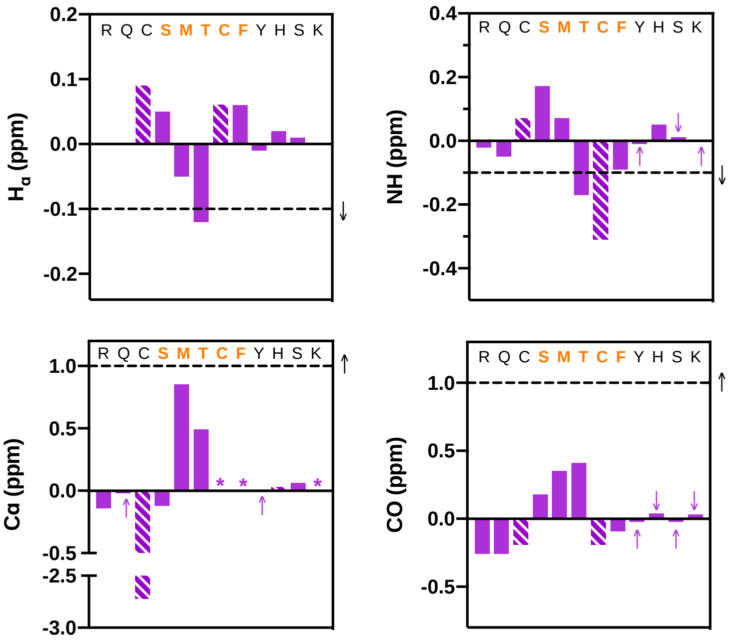


Figure 4: NMR structural data (10% aq. D_2O, pH ~5) for peptide **7**: Hα, Cα and carbonyl (CO) carbon secondary shifts are calculated relative to the corresponding resonances of peptide **2**. Each column represents a consecutive residue. Dashed columns represent bimane-modified cysteine residues. The sequence at the top of each graph shows the segment where helical structure is anticipated in orange. Purple arrows highlight values that are close to zero on the side of zero the value lies. Stars (*) represent data that could not be reliably extracted from the spectrum. The horizontal dashed lines and black arrow indicate the generally accepted threshold (and direction) to indicate helical structure. Three consecutive residues surpassing the horizontal dashed lines strongly indicates the presence of helical structure.

in the direction anticipated for helical structure, although these are small in magnitude. Overall, the secondary shifts for residues 148-149 are in the direction expected for helical structure, however these shifts are quite small, which suggests the peptide is still flexible and samples a variety of conformers. Circular dichroism of peptide **7** displays a deep minimum near 200 nm (Figure S8) which is also consistent with a flexible structure. The computationally modelled structure of **7** bound to PCNA indicates the peptide backbone adopts the classic 3_{10}-helix defining hydrogen bonds expected for a PIP-box peptide when bound to PCNA. This importantly demonstrates that the bimane linker does not restrict the peptide from adopting this key conformation, consistent with previous reports.[25]

6.2.4 Cell imaging

The bimane-linker is inherently fluorescent, which equips peptide **7** with the additional advantage that it can be directly subjected to a cell uptake assay and imaged by confocal microscopy without further derivatisation. Breast cancer cells (MDA-MB-468) were treated with 10 µM of peptide **7** in order to determine whether the high affinity bimane-constrained peptide was cell permeable. Two controls peptides were prepared: a linear $p21_{143-154}$ peptide with an *N*-terminal fluorescein appended (peptide **8**, Table 1); along with a derivative of peptide **7** that is macrocyclised with the bimane linker and also includes an *N*-terminal fluorescein (peptide **9**, Table 1). Peptides **8** and **9** were subjected to the same cell uptake experiment to compare to cell uptake of the linear and cyclised analogues, and the impact of a fluorescein-tag. Cells were seeded into culture dishes, incubated for 48 h, then treated with the peptide for

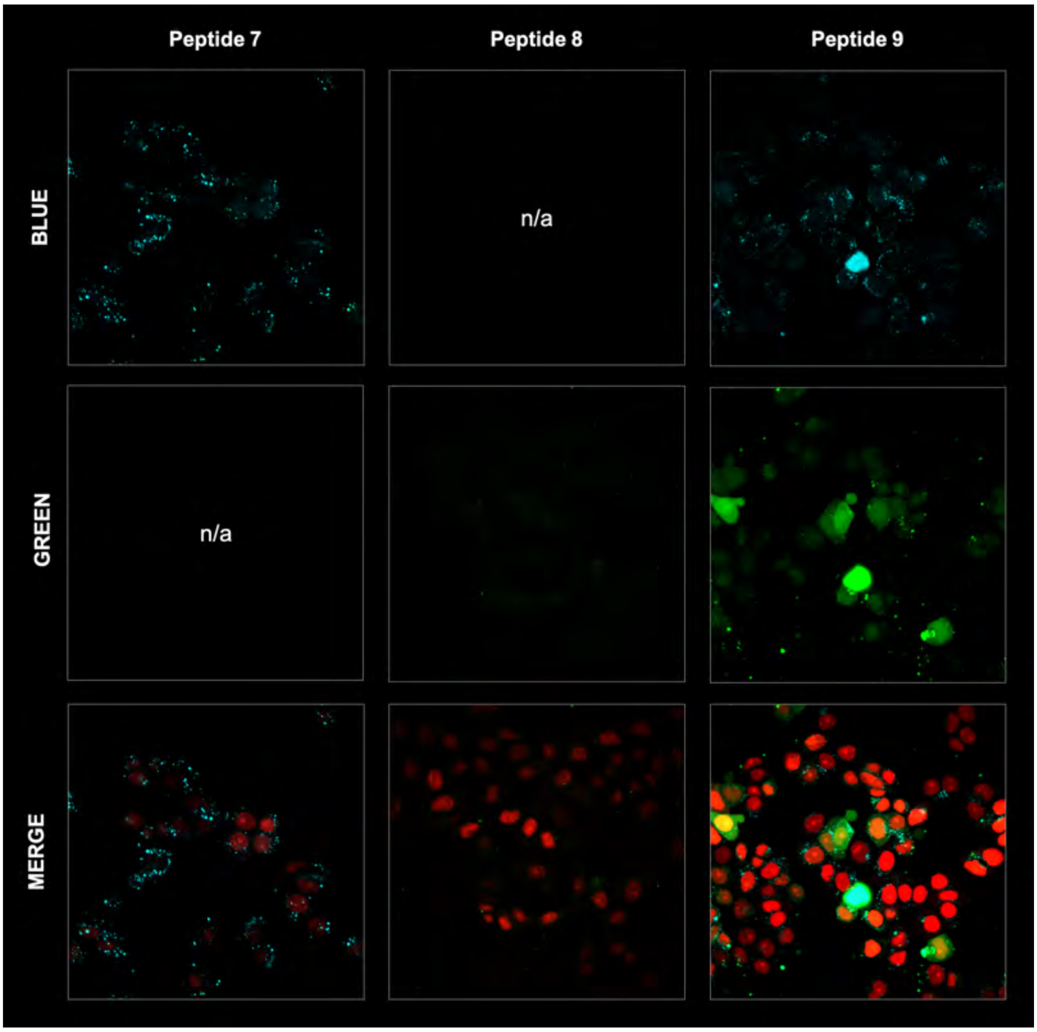


Figure 5: Breast cancer cells MDA-MB-468 treated with 10 µM of peptide **7**, peptide **8** or peptide **9**, then fixed and imaged by confocal microscopy. Peptide **7** is cell permeable as shown by punctate blue fluorescence throughout the cytoplasm. Peptide **8** is not cell permeable, where no green fluorescence is evident in the cell image. Peptide **9** is cell permeable, with both blue and green fluorescence present throughout the cell cytoplasm. The BLUE channel (ex. 405 nm, em. 410-485 nm) indicates Bimane fluorescence (peptides **7** and **9**); the GREEN channel (ex. 488 nm, em. 490-534 nm) indicates the FITC fluorophore (peptides **8** and **9**); the bottom panels show the BLUE and GREEN channel fluorescence overlaid with mKate fluorescence (mKate, ex. 594 nm, em. 600-700 nm) that marks the nucleus.

24 h. After treatment, media containing the peptide was removed and cells were washed with ice-cold PBS to remove all traces of media and peptide, fixed with 4% formaldehyde and imaged by confocal fluorescence microscopy. The nucleus, fluorescently tagged with genetically-expressed mKate, was imaged at ex. 594 nm, em. 600-700 nm and these images overlaid with the peptide fluorescence images (Figure 5). Peptide **8** treated cells were imaged at ex. 488 nm, and em. 490-534 nm (GREEN channel in Figure 5), which revealed no green fluorescence corresponding to the fluorescein-tagged peptide, indicating the linear peptide was not able to enter the cells. In contrast, peptide **7** treated cells were imaged with ex. 405 nm, and em. 410-485 nm (BLUE channel in Figure 5) to reveal blue fluorescence within the cytoplasm, corresponding to the bimane fluorophore, indicating that the bimane-cyclised peptide is cell permeable.

The fluorescence signal is mostly punctate, suggesting a significant portion of intracellular **7** is endosomally trapped.[29, 30] Macrocyclic peptides generally have increased cell permeability compared to the linear analogue.[31-34] This is the first example of a biologically relevant bimane-constrained peptide able to enter cells and be directly imaged using the bimane fluorescence. Imaging peptide **9** on both the GREEN and BLUE channels, revealed blue punctate fluorescence throughout the cytoplasm, as for peptide **7**. Green fluorescence was also observed throughout the cytoplasm for the peptide **9** treated cells, which is colocalised with the blue fluorescence. Together this indicates that peptide **9** is also cell permeable, with a similar distribution to peptide **8**. The green fluorescence was somewhat diffuse, suggesting that some free fluorescein may be present. Auxiliary fluorophores can alter the secondary structure, target binding and cell permeability of the peptide and consequently the activity in imaging assays may not accurately reflect true behaviour of the parent peptide. For this reason, there is a need to develop and utilise technologies that incorporate imaging modality directly into the therapeutic scaffold (such as the bimane) to allow direct imaging of the compound of interest. Here we demonstrate that an *N*-terminal fluorescein-tag appended to peptide **7**, as in peptide **9**, dramatically decreases the binding affinity for the protein target (PCNA, Table 1), though in this case the cellular uptake and distribution is not significantly impacted.

6.3 Conclusions

Two cysteine residues were incorporated at positions 146 and 149 of a p21 peptide (143-151, peptide **1**) known to bind PCNA to provide peptide **2**, which was then derivatised into five different macrocycles by dithiol bisalkylation. This gave rise to peptides with a propyl (**3**), butyl (**4**), *trans*-butenyl (**5**), *m*-xylene (**6**) and bimane (**7**) linker installed between the cysteine side-chains. The affinity of the resulting peptides **1**-**7** for PCNA was determined by SPR, with K_D values for the macrocycles (**3-7**) ranging from 570 nM to 3.85 μM. The bimane-constrained macrocyclic peptide **7** is the highest affinity peptidomimetic reported to bind PCNA, with a K_D value of 570 nM. Peptide **7** was the only macrocycle that adopted a classical 3_{10}-helical binding conformation upon binding PCNA, which suggests this conformation is likely responsible for the high affinity binding. However, it is interesting that the remaining four peptide macrocycles (**3**-**6**) still bind PCNA, but do not adopt this conformation upon binding. The 3_{10}-helical structure of high affinity peptide **7** was not well-defined in solution, suggesting that the binding conformation is not substantially preorganised prior to the peptide binding PCNA. However, it is significant that the bimane linker does not preclude PCNA binding, and still allows this key conformation to be adopted upon binding. This observation is in line with our previous study[19] which suggests that pre-defining a rigid peptide backbone may not improve PCNA binding affinity, and a linker that affords some flexibility is preferable to enable the peptide to adopt its ideal conformation on binding. Additionally, the inherently fluorescent peptide **7** is cell permeable and in contrast, a fluorescein-tagged linear p21 peptide of the same length (peptide **8**) was not cell permeable. This highlights how macrocyclization can improve cell permeability of short peptides. Furthermore, these results emphasise the utility of the bimane moiety as a peptide linker as it can influence peptide structure, and the resulting peptides can be directly imaged without further derivatisation to investigate cell uptake of a bimane peptidomimetic. This allows the behaviour of the molecule of therapeutic interest to be assessed, instead of a related analogue that includes a fluorescent tag; and eliminates the need for an additional synthetic step. Attachment of an auxiliary fluorophore such as fluorescein to a peptide, can impair the target binding affinity (as seen for peptide **9**) and may alter cellular permeability or intracellular distribution relative to the untagged analogue. In summary, this study identifies a short, cell permeable, high affinity PCNA-binding peptidomimetic, as a significant advance towards a pre-clinical anti-cancer therapeutic. Future work will focus on imparting nuclear permeability to macrocyclic peptides, such as peptide **7**, in order to determine the ability of such p21-peptidomimetics to inhibit DNA-replication in cancer cells.

6.4 Acknowledgements

The research was supported by the Australian Research Council Centre of Excellence for Nanoscale BioPhotonics (CNBP) (CE140100003). A.J.H. and B.A.V are supported by Australian Government Research Training Program Stipends (RTPS). This research was undertaken in part using the MX2 beamline at the Australian Synchrotron, part of ANSTO, and made use of the Australian Cancer Research Foundation (ACRF) detector. This research was undertaken in part using the MX1 beamline[35] at the Australian Synchrotron, part of ANSTO. The facilities of the OptoFab node of the Australian National Fabrication Facility (ANFF) and associated Commonwealth and SA State Government funding are also gratefully acknowledged. This work was supported by grants from the National Health and Medical Research Council of Australia (W.D.T., T.E.H., ID 1084416, ID 1130077;), the National Breast Cancer Foundation (NBCF; W.D.T.; ID PS-15-041) and a Movember & National Breast Cancer Foundation Collaboration Initiative grant (MNBCF-17-012 to W.D.T., T.E.H.). T.E.H. is currently supported by an NBCF Fellowship (IIRS-19-009).

6.5 Conflicts of interest

There are no conflicts of interest to declare

REFERENCES

1. T. Tsurimoto, PCNA Binding Proteins, *Front. Biosci.*, 1999, **4**, d849-858.
2. G. Maga and U. Hubscher, Proliferating cell nuclear antigen (PCNA): A dancer with many partners, *J. Cell Sci.*, 2003, **116**, 3051-3060.
3. G. L. Moldovan, B. Pfander and S. Jentsch, PCNA, the maestro of the replication fork, *Cell*, 2007, **129**, 665-679.
4. S. N. Naryzhny and H. Lee, Characterization of proliferating cell nuclear antigen (PCNA) isoforms in normal and cancer cells: There is no cancer-associated form of PCNA, *FEBS Lett.*, 2007, **581**, 4917-4920.
5. I. Stoimenov and T. Helleday, PCNA on the crossroad of cancer, *Biochem. Soc. Trans.*, 2009, **37**, 605-613.
6. D. Zhongyun, M. Wortman, Z. Tan and K. Dillehay, WO 2012/033938 A2, *Identification of PCNA Targeting Compounds for Cancer Therapy and PCNA Function Regulation,* 2012.
7. A. J. Horsfall, A. D. Abell and J. B. Bruning, Targeting PCNA with Peptide Mimetics for Therapeutic Purposes, *ChemBioChem*, 2019, **21**, 442-450.
8. D. I. Zheleva, N. Z. Zhelev, P. M. Fischer, S. V. Duff, E. Warbrick, D. G. Blake and D. P. Lane, A Quantitative Study of the in Vitro Binding of the C-Terminal Domain of p21 to PCNA: Affinity, Stoichiometry, and Thermodynamics, *Biochemistry*, 2000, **39**, 7388-7397.
9. J. B. Bruning and Y. Shamoo, Structural and thermodynamic analysis of human PCNA with peptides derived from DNA polymerase-delta p66 subunit and flap endonuclease-1, *Structure*, 2004, **12**, 2209-2219.
10. E. M. Boehm and M. T. Washington, R.I.P. to the PIP: PCNA-binding motif no longer considered specific, *BioEssays*, 2016, **38**, 1117-1122.
11. J. M. Gulbis, Z. Kelman, J. Hurwitz, M. O'Donnell and J. Kuriyan, Structure of the C-Terminal Region of p21 WAF1/CIP1 Complexed with Human PCNA, *Cell*, 1996, **87**, 297-306.
12. E. Gibbs, Z. Kelman, J. M. Gulbis, M. O'Donnell, J. Kuriyan, P. M. J. Burgers and J. Hurwitz, The Influence of the Proliferating Cell Nuclear Antigen-interacting Domain of p21-CIP1 on DNA Synthesis Catalyzed by the Human and Saccharomyces cerevisiae Polymerase delta Holoenzymes, *J. Biol. Chem.*, 1997, **272**, 2373-2381.
13. A. J. Horsfall, B. A. Vandborg, W. Kowalczyk, T. Chav, D. B. Scanlon, A. D. Abell and J. B. Bruning, Unlocking the PIP-box: A peptide library reveals interactions that drive high affinity binding to human PCNA, *J. Biol. Chem.*, 2021, **296** 100773.
14. E. Warbrick, D. P. Lane, D. M. Glover and L. S. Cox, A small peptide inhibitor of DNA replication defines the site of interaction between the cyclin-dependent kinase inhibitor p21 WAF1 and proliferating cell nuclear antigen, *Curr. Biol.*, 1995, **5**, 275-282.
15. T. A. Hill, N. E. Shepherd, F. Diness and D. P. Fairlie, Constraining cyclic peptides to mimic protein structure motifs, *Angew. Chem. Int. Ed. (English)*, 2014, **53**, 13020-13041.
16. A. P. Higueruelo, H. Jubb and T. L. Blundell, Protein-protein interactions as druggable targets: recent technological advances, *Curr. Opin. Pharmacol.*, 2013, **13**, 791-796.
17. Q. Chu, R. E. Moellering, G. J. Hilinski, Y.-W. Kim, T. N. Grossmann, J. T. H. Yeh and G. L. Verdine, Towards understanding cell penetration by stapled peptides, *MedChemComm*, 2015, **6**, 111-119.
18. D. P. Fairlie and A. D. de Araujo, Stapling peptides using cysteine crosslinking, *Biopolymers*, 2016, **106**, 843-852.
19. K. L. Wegener, A. E. McGrath, N. E. Dixon, A. J. Oakley, D. B. Scanlon, A. D. Abell and J. Bruning, Rational design of a 310-helical PIP-box mimetic targeting PCNA - the human sliding clamp, *Chem. Eur. J.*, 2018, **24**, 11325-11331.
20. L. Peraro, T. R. Siegert and J. A. Kritzer, Conformational Restriction of Peptides Using Dithiol Bis-Alkylation, *Methods Enzymol.*, 2016, **580**, 303-332.
21. G. Zhang, F. Barragan, K. Wilson, N. Levy, A. Herskovits, M. Sapozhnikov, Y. Rodriguez, L. Kelmendi, H. Alkasimi, H. Korsmo, M. Chowdhury and G. Gerona-Navarro, A Solid-Phase Approach to Accessing Bisthioether-Stapled Peptides Resulting in a Potent Inhibitor of PRC2 Catalytic Activity, *Angew. Chem. Int. Ed. (English)*, 2018, **57**, 17073-17078.
22. H. Jo, N. Meinhardt, Y. Wu, S. Kulkarni, X. Hu, K. E. Low, P. L. Davies, W. F. DeGrado and D. C. Greenbaum, Development of alpha-helical calpain probes by mimicking a natural protein-protein interaction, *J. Am. Chem. Soc.*, 2012, **134**, 17704-17713.
23. A. J. Horsfall, K. R. Dunning, K. L. Keeling, D. B. Scanlon, K. L. Wegener and A. D. Abell, A bimane‐based peptide staple for combined helical induction and fluorescent imaging, *ChemBioChem*, 2020, **21**, 3423-3432.
24. A. D. de Araujo, H. N. Hoang, W. M. Kok, F. Diness, P. Gupta, T. A. Hill, R. W. Driver, D. A. Price, S. Liras and D. P. Fairlie, Comparative alpha-helicity of cyclic pentapeptides in water, *Angew. Chem. Int. Ed. (English)*, 2014, **53**, 6965-6969.
25. A. J. Horsfall, D. P. McDougal, D. B. Scanlon, J. B. Bruning and A. D. Abell, Approches to introduce helical structure in cysteine-containing peptides with a bimane group, *ChemBioChem*, 2021, **In press, doi.org/10.1002/cbic.202100241**.
26. A. J. Kroker and J. B. Bruning, p21 Exploits Residue Tyr151 as a Tether for High-Affinity PCNA Binding, *Biochemistry*, 2015, **54**, 3483-3493.
27. K. N. Choe and G.-L. Moldovan, Forging Ahead through Darkness: PCNA, Still the Principal Conductor at the Replication Fork, *Mol. Cell*, 2016, **65**, 380-392.
28. G. Kontopidis, S.-Y. Wu, D. I. Zheleva, P. Taylor, C. McInnes, D. P. Lane, P. M. Fischer and M. D. Walkinshaw, Structural and biochemical studies of human proliferating cell nuclear antigen complexes provide a rationale for cyclin association and inhibitor design, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 1871-1876.

29. T. B. Potocky, A. K. Menon and S. H. Gellman, Cytoplasmic and nuclear delivery of a TAT-derived peptide and a beta-peptide after endocytic uptake into HeLa cells, *J. Biol. Chem.*, 2003, **278**, 50188-50194.

30. P. Ramoino, A. Diaspro, M. Fato and C. Usai, in *Molecular Regulation of Endocytosis*, 2012, DOI: 10.5772/46061, ch. Chapter 6.

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

32. Y. Tian, Y. Jiang, J. Li, D. Wang, H. Zhao and Z. Li, Effect of Stapling Architecture on Physiochemical Properties and Cell Permeability of Stapled alpha-Helical Peptides: a Comparative Study, *ChemBioChem*, 2017, **18**, 2087-2093.

33. S. R. Perry, T. A. Hill, A. D. de Araujo, H. N. Hoang and D. P. Fairlie, Contiguous hydrophobic and charged surface patches in short helix-constrained peptides drive cell permeability, *Org. Biomol. Chem.*, 2018, **16**, 367-371.

34. A. M. Spokoyny, Y. Zou, J. J. Ling, H. Yu, Y. S. Lin and B. L. Pentelute, A perfluoroaryl-cysteine S(N)Ar chemistry approach to unprotected peptide stapling, *J. Am. Chem. Soc.*, 2013, **135**, 5946-5949.

35. N. P. Cowieson, D. Aragao, M. Clift, D. J. Ericsson, C. Gee, S. J. Harrop, N. Mudie, S. Panjikar, J. R. Price, A. Riboldi-Tunnicliffe, R. Williamson and T. Caradoc-Davies, MX1: a bending-magnet crystallography beamline serving both chemical and macromolecular crystallography communities at the Australian Synchrotron, *J. Synchrotron Rad.*, 2015, **22**, 187-190.

Supplemental Data.

A CELL PERMEABLE BIMANE-CONSTRAINED PCNA-INTERACTING PEPTIDE

S6.1 Experimental

S6.1.1 General information

Unless otherwise indicated, all starting materials were purchased from commercial sources and used without further purification. High-resolution mass spectra were collected using an Agilent 6230 ESI-TOF LCMS. RP-HPLC solvents were (A) H_2O with 0.1% TFA and (B) ACN with 0.1% TFA. Purity of all compounds was confirmed by analytical RP-HPLC on an Agilent 1260 HPLC equipped with a Phenomenex Luna C18(2) column (250 x 4.6 mm) over a gradient of 5-50% B (15 min). Purification was carried out by semi-preparative HPLC on a Gilson GX-Prep RP-HPLC system on a Phenomenex Aeris Peptide C18 (10 x 250 mm), over a gradient as specified in the individual compound sections. All graphs were generated using GraphPad Prism 8 software.

S6.1.2 Synthesis and characterisation

All peptides were synthesised by the Fmoc solid-phase peptide synthesis protocol detailed below, with all L-amino-acids (unless otherwise specified), and then *N*-terminally acetylated before cyclisation on-resin. Peptides were subsequently cleaved from the resin (and simultaneously globally deprotected). See Scheme 1 and the following general procedures.

H-RQTSMTDFYHSK-NH$_2$

Solid-phase peptide synthesis of peptide 1: Rink Amide PL resin (0.2 mmol, 644 mg, 0.31 mmol/g, Agilent) was swollen in 1:1 DMF/DCM (15 mL) for 15 min. The Fmoc-protecting group was removed by treatment of the resin with a solution of 20% piperidine and 0.1 M HOBt in DMF (8 mL) for 15 min. The solution was drained and the resin washed with DMF (3 x 8 mL). Amino-acid couplings were achieved by addition of a solution of Fmoc-protected amino-acid (5 equiv), HATU (5 equiv) and DIPEA (10 equiv) in DMF (8 mL), to the resin and stirred intermittently for 1 h. The solution was drained and the resin washed with DMF (5 x 8 mL). The *N*-terminal Fmoc-protecting group was removed by treatment of the resin with a solution of 20% piperidine and 0.1 M HOBt in DMF (8 mL) for 10 min, the solution was drained and the resin washed with DMF (5 x 8 mL). A TNBS test* was used to verify each coupling (negative/colourless) and deprotection (positive/red) step, with steps repeated as necessary. Successive couplings and Fmoc-deprotections were repeated to achieve the desired sequence. The peptide was then cleaved from the resin as described by *General Procedure for Cleavage and Isolation*. The cyclised peptide was purified by semi-preparative RP-HPLC and Phenomenex Aeris Peptide C18 Column (10 x 250 mm) over a linear gradient of 20-25% over 15 min. Pure fractions were combined and lyophilized to give the final purified product as a white fluffy powder. HRMS (ESI+) Expected $[M+3H]^{3+}$ for $C_{64}H_{98}N_{20}O_{20}S$: 500.5740, observed: $[M+3H]^{3+}$ 500.5741. RP-HPLC purity (C18, 215 nm) 95.1%.

**TNBS Test:*[1] A small spatula of swollen resin taken out and 1 drop each of TNBS (100 µL 5% w/v picrylsulfonic/trinitrobenzenesulfonic acid in H_2O added to 900 µL of DMF) and DIPEA solutions (100 µL in 900 µL of DMF) added and allowed to develop for 1 min. Clear/yellow beads indicated no free amine (negative), while red/orange beads showed free amine was present (positive).

S6.1.3 General procedure for cleavage and isolation:

The peptide was cleaved from the resin by addition of 92.5:2.5:2.5:2.5 TFA/TIPS/DODT/H_2O (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 -20°C. The precipitate was pelleted by centrifugation (7600 rpm, 10 min), the supernatant decanted. The pellet was dried under a nitrogen stream, and then dissolved in 1:1 ACN/H_2O, before syringe filtering (0.2 µm) and lyophilised.

Solid-phase Peptide Synthesis of linear on-resin precursor peptide 2a:

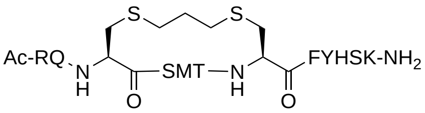
[Ac-R(Pbf)Q(Trt)C(Mmt)S(tBu)MT(tBu)C(Mmt)FY(tBu)H(Trt)S(tBu)K(Boc)-RESIN]

Rink Amide PL resin (0.2 mmol, 644 mg, 0.31 mmol/g, Agilent) was swollen in 1:1 DMF/DCM (15 mL) for 15 min. The Fmoc-protecting group was removed by treatment of the resin with a solution of 20% piperidine and 0.1 M HOBt in DMF (8 mL) for 15 min. The solution was drained and the resin washed with DMF (3 x 8 mL). Amino-acid couplings were achieved by addition of a solution of Fmoc-protected amino-acid (5 equiv), HATU (5 equiv) and DIPEA (10 equiv) in DMF (8 mL), to the resin and stirred intermittently for 1 h. The solution was drained and the resin washed with DMF (5 x 8 mL). The *N*-terminal Fmoc-protecting group was removed by treatment of the resin with a solution of 20% piperidine and 0.1 M HOBt in DMF (8 mL) for 10 min, the solution was drained and the resin washed with DMF (5 x 8 mL). A TNBS test* was used to verify each coupling (negative/colourless) and deprotection (positive/red) step, with steps repeated as necessary. Successive couplings and Fmoc-deprotections were repeated to achieve the desired sequence. After the final Fmoc-deprotection, the *N*-terminus was protected with an acetyl functionality by reaction with acetic anhydride (470 µL) and DIPEA (870 µL) in DMF (10 mL) for 15 min. The resin was washed with DMF (3 x 8 mL), and DCM (5 x 8 mL).

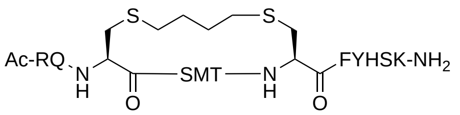
For peptides **3**-**7**, the cysteine side-chains were selectively deprotected: Mmt groups were removed by repetitive treatment of the resin with 2% TFA in DCM (8 mL) for 1 min, followed by washing with DCM (3 x 8 mL). Treatments were repeated until the solution no longer turned yellow on addition to the resin (~ 150-200 mL total). The resin was then further washed with DCM (5 x 8 mL) and DMF (5 x 8 mL).

Ac-RQTCMTCFYHSK-NH_2

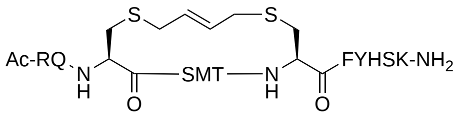
Peptide 2*:* Following peptide assembly as described in ‘Solid-phase Peptide Synthesis of linear on-resin precursor peptide to **2a**’ and N-terminal acetylation, the peptide was cleaved from the resin by General Procedure for Cleavage and Isolation to give peptide **2**. The crude peptide was purified by semi-preparative RP-HPLC and Phenomenex Aeris Peptide C18 Column (10 x 250 mm) over a linear gradient of 25-50% B (15 min). Pure fractions were combined and lyophilized to give the final purified peptide **2** as a pale yellow fluffy powder. HRMS (ESI+) Expected $[M+4H]^{4+}$ for $C_{64}H_{98}N_{20}O_{18}S_3$: 383.6778, observed: $[M+4H]^{4+}$ 383.6702. [2]RP-HPLC purity (C18, 215 nm) 88.9%.



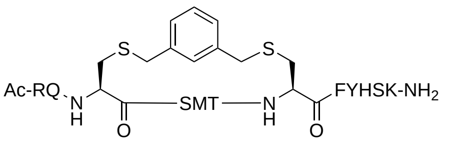
Peptide 3, propyl thioether cyclisation:[3] Following linear peptide assembly as described in ‘Solid-phase Peptide Synthesis of linear on-resin precursor peptide to **2a**’ the preswelled resin (0.1 mmol) with resin-bound N-terminally acetylated peptide and cysteines freshly deprotected, was transferred to a MW vessel with stirrer bar. NaI (17.5 equiv, 225 mg) in DMF (6 mL) was then added and the mixture stirred continuously while TCEP (0.5 equiv, 300 mg/ml, 18.9 µL) was then added and the vessel sealed and bubbled with N_2 for 15 min. Under a N_2 atmosphere, DIPEA (35 equiv, 274 µL) was added and the mixture stirred and bubbled with N_2 for a further 20 min. 1,3-Dibromopropane (3.5 equiv, 16 µL) was then added and the vessel sealed and reacted under MW for 2 min at 125°C. The vessel was then cooled, the resin removed and solution drained. The resin was washed with H_2O (5 x 5 mL), DMF (5 x 5 mL) and DCM (5 x 5 mL), then dried with diethyl ether (3 x 5 mL). The peptide was then cleaved from the resin as described by General Procedure for Cleavage and Isolation. The cyclised peptide was purified by semi-preparative RP-HPLC and Phenomenex Aeris Peptide C18 Column (10 x 250 mm) over a linear gradient of 25-50% B (15 min). Pure fractions were combined and lyophilized to give the final purified peptide **3** as a white powder. HRMS (ESI+) Expected $[M+3H]^{3+}$ for $C_{67}H_{102}N_{20}O_{18}S_3$: 524.5692, observed: $[M+3H]^{3+}$ 524.5718. RP-HPLC purity (C18, 215 nm) 91.9%.



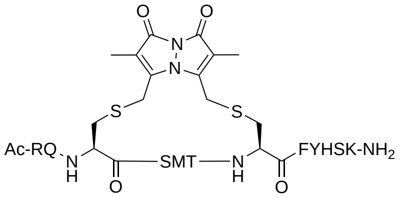
Peptide 4, butyl thioether cyclisation:[3] Following linear peptide assembly as described in 'Solid-phase Peptide Synthesis of linear on-resin precursor peptide to **2a**' the preswelled resin (0.1 mmol) with resin-bound N-terminally acetylated peptide and cysteines freshly deprotected, was transferred to a MW vessel with stirrer bar. NaI (17.5 equiv, 225 mg) in DMF (6 mL) was then added and the mixture stirred continuously while TCEP (0.5 equiv, 300 mg/ml, 18.9 µL) was then added and the vessel sealed and bubbled with N_2 for 15 min. Under a N_2 atmosphere, DIPEA (35 equiv, 274 µL) was added and the mixture stirred and bubbled with N_2 for a further 20 min. 1,4-Dibromobutane (3.5 equiv, 19 µL) was then added and the vessel sealed and reacted under MW for 2 min at 125°C. The vessel was then cooled, the resin removed and solution drained. The resin was washed with H_2O (5 x 5 mL), DMF (5 x 5 mL) and DCM (5 x 5 mL), then dried with diethyl ether (3 x 5 mL). The peptide was then cleaved from the resin as described by General Procedure for Cleavage and Isolation. The cyclised peptide was purified by semi-preparative RP-HPLC and Phenomenex Aeris Peptide C18 Column (10 x 250 mm) over a linear gradient of 25-50% B (15 min). Pure fractions were combined and lyophilized to give the final purified peptide **4** as a white powder. HRMS (ESI+) Expected $[M+3H]^{3+}$ for $C_{68}H_{104}N_{20}O_{18}S_3$: 529.2416, observed: $[M+3H]^{3+}$ 529.2410. RP-HPLC purity (C18, 215 nm) 91.8%.



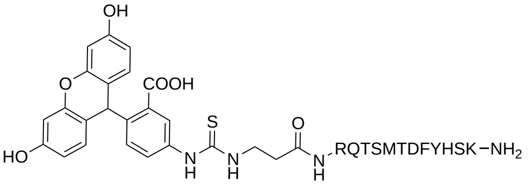
Peptide 5, trans-butene cyclisation: Following linear peptide assembly as described in '*Solid-phase Peptide Synthesis of linear on-resin precursor peptide to **2a***' the preswelled resin (0.1 mmol) with resin-bound *N*-terminally acetylated peptide and cysteines freshly deprotected, was treated with a solution of trans-1,4-dibromo-2-butene (2 equiv, 42.8 mg) and DIPEA (4 equiv, 174 µL) in DMF (4 mL) was added to the resin, and reacted for 3 h with intermittent stirring. The solution was then removed and the resin washed with DMF (5 x 5 mL) and DCM (5 x 5 mL). In the case that a small cleave of the resin revealed incomplete reaction, the procedure was repeated as above and left to react overnight. The solution was then removed and the resin washed with DMF (5 x 5 mL) and DCM (5 x 5 mL), then dried with diethyl ether (3 x 5 mL). The peptide was then cleaved from the resin as described by *General Procedure for Cleavage and Isolation*. The cyclised peptide was purified by semi-preparative RP-HPLC and Phenomenex Aeris Peptide C18 Column (10 x 250 mm) over a linear gradient of 25-50% B (15 min). Pure fractions were combined and lyophilized to give the final purified peptide **5** as a white and brown powder. HRMS (ESI+) Expected $[M+3H]^{3+}$ for $C_{68}H_{102}N_{20}O_{18}S_3$: 528.5692, observed: $[M+3H]^{3+}$ 528.5712. RP-HPLC purity (C18, 215 nm) 94.4%.



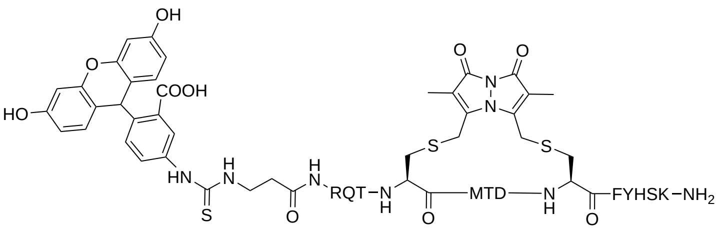
Peptide 6, xylene cyclisation: Following linear peptide assembly as described in '*Solid-phase Peptide Synthesis of linear on-resin precursor peptide to **2a***' the preswelled resin (0.1 mmol) with resin-bound *N*-terminally acetylated peptide and cysteines freshly deprotected, was treated with a solution of dibromo-m-xylene (2 equiv, 47.4 mg) and DIPEA (4 equiv, 174 µL) in DMF (4 mL) was added to the resin, immediately following the cysteine Mmt group removal and washing procedure, and reacted for 3 h with intermittent stirring. The solution was then removed and the resin washed with DMF (5 x 5 mL) and DCM (5 x 5 mL). In the case that a small cleave of the resin revealed incomplete reaction, the procedure was repeated as above and left to react overnight. The solution was then removed and the resin washed with DMF (5 x 5 mL) and DCM (5 x 5 mL), then dried with diethyl ether (3 x 5 mL). The peptide was then cleaved from the resin as described by *General Procedure for Cleavage and Isolation.* The cyclised peptide was purified by semi-preparative RP-HPLC and Phenomenex Aeris Peptide C18 Column (10 x 250 mm) over a linear gradient of 25-50% B (15 min). Pure fractions were combined and lyophilized to give the final purified peptide **6** as a white fluffy powder. HRMS (ESI+) Expected $[M+3H]^{3+}$ for $C_{72}H_{104}N_{20}O_{18}S_3$: 545.2411, observed: $[M+3H]^{3+}$ 545.2436. RP-HPLC purity (C18, 215 nm) 91.5%.



Peptide 7, bimane cyclisation:[4] Following linear peptide assembly as described in 'Solid-phase Peptide Synthesis of linear on-resin precursor peptide to 2a' the preswelled resin (0.1 mmol) with resin-bound N-terminally acetylated peptide and cysteines freshly deprotected, was treated with a solution of dibromobimane (2 equiv, 70 mg) and DIPEA (4 equiv, 174 µL) in DMF (6 mL), and reacted for 3 h with intermittent stirring. The solution was then removed and the resin washed with DMF (5 x 5 mL) and DCM (5 x 5 mL), then dried with diethyl ether (3 x 5 mL). The peptide was then cleaved from the resin as described by General Procedure for Cleavage and Isolation. The cyclised peptide was purified by semi-preparative RP-HPLC and Phenomenex Aeris Peptide C18 Column (10 x 250 mm) over a linear gradient of 25-50% B (15 min). Pure fractions were combined and lyophilized to give the final purified peptide 7 as a pale yellow fluffy powder. HRMS (ESI+) Expected $[M+3H]^{3+}$ for $C_{74}H_{106}N_{22}O_{20}S_3$: 573.9119, observed: $[M+3H]^{3+}$ 573.9140. RP-HPLC purity (C18, 215 nm) 97.3%.



Peptide 8, fluorescein attachment: Following linear peptide assembly as described in '*Solid-phase Peptide Synthesis of linear on-resin precursor peptide to **1***' the preswelled resin (0.1 mmol) with resin-bound peptide with *N*-terminal free amine, was coupled to Fmoc-β-alanine (5 equiv) with HATU (5 equiv) and DIPEA (10 equiv) in DMF for 1 h with intermittent stirring. The solution was then drained from the resin, and the Fmoc group removed by treatment of the resin with 20% piperidine in DMF (5 mL) for 10 min with intermittent stirring, and the resin then washed with DMF (3 x 5 mL), and subsequently treated with fluorescein isothiocyanate (2 equiv, 78 mg) and DIPEA (4 equiv, 70 µL) in DMF (4 mL) and stirred intermittently for 20 min. The solution was then drained and the resin thoroughly washed with DMF (5 x 5 mL) and DCM (5 x 5 mL), then dried with diethyl ether (3 x 5 mL). The peptide was then cleaved from the resin as described by *General Procedure for Cleavage and Isolation*. The cyclised peptide was purified by semi-preparative RP-HPLC and Phenomenex Luna C18(2) Column (10 x 250 mm) over a linear gradient of 25-50% B (15 min). Pure fractions were combined and lyophilized to give the final purified peptide **8** as a yellow fluffy powder. HRMS (ESI+) Expected $[M+4H]^{4+}$ for $C_{88}H_{114}N_{22}O_{26}S_2$: 490.7007, observed: $[M+3H]^{3+}$ 490.6728. RP-HPLC purity (C18, 215 nm) 93.6%.



Peptide 9, bimane cyclisation and fluorescein attachment:[4] Following linear peptide assembly as described in 'Solid-phase Peptide Synthesis of linear on-resin precursor peptide to 2a' the preswelled resin (0.1 mmol) with the exception that the N-terminal Fmoc was not removed prior to the cysteine Mmt deprotection. The resin-bound peptide, with N-terminal Fmoc and cysteines freshly deprotected, was treated with a solution of dibromobimane (2 equiv, 70 mg) and DIPEA (4 equiv, 174 µL) in DMF (6 mL), and reacted for 3 h with intermittent stirring. The solution was then removed and the resin washed with DMF (5 x 5 mL) and then treated with 20% piperidine in DMF (5 mL) for 10 min with intermittent stirring. The solution was then drained and the resin washed with DMF (3 x 5 mL) and then a solution of Fmoc-β-alanine (5 equiv), HATU (5 equiv) and DIPEA (10 equiv) in DMF added and stirred intermittently for 1 h. The solution was then drained from the resin, and the Fmoc group removed by treatment of the resin with 20% piperidine in DMF (5 mL) for 10 min with intermittent stirring, and the resin then washed with DMF (3 x 5 mL), and subsequently treated with fluorescein isothiocyanate (2 equiv, 78 mg) and DIPEA (4 equiv, 70 µL) in DMF (4 mL) and stirred intermittently for 20 min. The solution was then drained and the resin thoroughly washed with DMF (5 x 5 mL) and DCM (5 x 5 mL), then dried with diethyl ether (3 x 5 mL). The peptide was then cleaved from the resin as described by General Procedure for Cleavage and Isolation. The cyclised peptide was purified by semi-preparative RP-HPLC and Phenomenex Aeris Peptide C18 Column (10 x 250 mm) over a linear gradient of 25-50% B (15 min). Pure fractions were combined and lyophilized to give the final purified peptide 7 as a yellow powder. 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. RP-HPLC purity (C18, 215 nm) 93.4%.

S6.1.4 SPR protocol:

The running buffer used for ligand attachment and analyte binding experiments was 10 mM HEPES buffer with 150 mM NaCl, 3 mM EDTA and 0.05% Tween20, adjusted to pH 7.4 with 2 M NaOH. Experiments were performed on a GE S200 Biacore System at 25°C. A GE CM5 sensor chip was primed with running buffer and preconditioned per the manufacturer's recommendation with successive injections (2 x 50 s, 30 µL/min) of each 50 mM NaOH, 10 mM HCl, 0.1% SDS, 0.85% H_3PO_4 and glycine pH 9.5. The surface was then activated with a solution of 0.2 M EDC and 50 mM NHS (600 s, 10 µL/min).

PCNA (5 µL, 12 mg/mL) was diluted into running buffer (245 µL). Only once the preactivation was complete was the protein further diluted to a final concentration of 25 µg/mL in 10 mM NaAc (~pH 4.6) by addition of PCNA/HEPES (50 µL) to a solution of 100 mM NaAc (50 µL) and water (400 µL). This solution was immediately injected over only the second flow cell (10 µL/min) until the ~1500 RU was reached at stabilisation. Both flow cells were then blocked with 1.0 M ethanolamine pH 8.5 (600 s, 10 µL/min). The chip was left to stabilise for two hours before sample injections commenced. After stabilisation a final protein level of 1380 RU was achieved.

Peptides (approx. 2 mg by weight) were dissolved in milliQ H_2O (50 µL) and centrifuged (7800 rpm, 10 min) to remove any particulate. The peptide stock concentration was determined by UV absorbance (A_λ), where 2 µL of the stock was further diluted in water (20-40 fold) and a measurement taken in triplicate with a Nanodrop2000 and baselined to 750 nm absorbance. The peptide stock solution concentration was then calculated per $c = (A_\lambda/\varepsilon_\lambda.l).DF$ where concentration is in mol/L, A_λ is absorbance at λ nm calculated as an average of three readings, λ is the appropriate wavelength (here, 205 or 380 nm), l is the pathlength in cm (1 mm for Nanodrop), and ε_λ is the molar absorptivity at λ nm and DF is the dilution factor. The ε_{205} was estimated for peptides **1**-**6** using Anthis 2013,[5] and ε_{380} used for bimane-containing peptide **7** as reported in Shen 2011,[2] and are recorded in Table S2. The peptides were then diluted into running buffer before further dilution as necessary.

The steady state affinity experiments were all run at 30 µL/min, a contact time of 30 s and dissociation time of 40 s, followed by regeneration of 2 M NaCl (2 x 30 s). Each peptide was serially diluted (1 in 2) eight times, and run from least to most concentrated following a blank injection. After the concentration range was optimised, the experiment was repeated to ensure reproducibility. The top concentration of the final concentration range for each peptide sample is tabulated in Table S1. The data was analysed using the provided Biacore S200 Evaluation software to give a steady state K_D value and associated standard error, as shown in Table S1.

S6.1.5 NMR analysis

^{1}H and ^{13}C NMR 1D and 2D spectra were recorded on an a Varian Inova 600 MHz instrument in 10% aq. D_2O at 298K, pH ~5, and referenced to DSS at 0 ppm. ES suppression sequences were used for all ^{1}H 1D and 2D homonuclear spectra. ROESYAD and zTOCSY were obtained for all compounds. gHMBCAD and gHSQCAD experiments were collected with 256 scans. Chemical shifts are reported in ppm (δ). Full ^{1}H assignments for peptides **2** and **7** are included in Tables S4 and S5. Secondary shifts for **2** and **7** were calculated relative to the random coil shifts reported by Wishart 2011[6] and nearest neighbour corrections applied (Tables S4 and S5, Fig. S7, green and blue). The secondary shift for **7** were also calculated by subtraction from the linear precursor peptide **2** (Table S6, Fig. S7, purple).

S6.1.6 Protein expression and purification

A glycerol stock of *E. coli* BL21 (λDE3) cells carrying a PCNA-pMCSG19 plasmid (with no purification tag) were grown in a 50 mL overnight culture. Eight 1 L baffled flasks of LB with 100 µg/mL of ampicillin were inoculated with 6.3 mL of the overnight culture. Cultures were incubated at 37°C until OD600 of 0.5 and induced with 0.5 mM IPTG. Cultures were grown overnight at 16°C with shaking at 200 rpm. Cultures were pelleted at 5000xg for 20 min. After removing the supernatant, pellets were resuspended in 30 mL of Buffer A (20 mM Tris pH 7.5, 20 mM NaCl, 2 mM DTT), then lysed by 5 rounds of cell disruption by a microfluidics cell disrupter. Lysate was pelleted at 45,000xg for 60 min and the supernatant was collected for purification. PCNA was purified at 4°C by fast protein liquid chromatography (FPLC), using an anion exchange DEAE column (HiTrap DEAE FF 5 mL column), equilibrated in Buffer A (20 mM Tris pH 7.5, 20 mM NaCl, 2 mM DTT), and protein was eluted using Buffer B (20 mM Tris pH 7.5, 0.7 M NaCl, 2 mM DTT). Fractions were analysed by SDS-PAGE and those of interest indicating containing protein at ~28 kDa were selected and pooled and treated with ammonium sulphate to bring the concentration to 1.5 M. Protein was purified again by a hydrophobic column (HiTrap Phenyl FF [high sub] 5 mL column) equilibrated in Buffer C (20 mM Tris pH 7.5, 20 mM NaCl, 2 mM DTT, 0.5 mM EDTA, 1.5 M Ammonium Sulphate), and protein was eluted using Buffer D (20 mM Tris pH 7.5, 0.5 mM EDTA, 2 mM DTT). Fractions were analysed by SDS-PAGE and those of interest were dialysed overnight in Buffer E (20 mM Tris pH 7.5, 20 mM NaCl, 2 mM DTT). Equilibrated protein was purified using an anion exchange Q Sepharose column (ENrich Q 10 × 100 mm 8 mL column), equilibrated in Buffer F (20 mM Tris pH 7.5, 20 mM NaCl, 2 mM DTT), protein was eluted using Buffer G (20 mM Tris pH 7.5, 0.7 M NaCl, 2 mM DTT). Protein pool was concentrated using a centrifugal filter unit (30 kDa molecular mass cut off) to a volume of less than 10 mL and purified using the size exclusion column (HiPrep 26/60 Sephacryl S-200 HR 300 mL column), equilibrated in Buffer H (20 mM Tris pH 7.5, 50 mM NaCl, 2 mM DTT, 0.5 mM EDTA), and protein was eluted using the same buffer. Fractions were analysed by SDS-PAGE and those of interest were pooled. Fractions were analysed by SDS-PAGE and those of interest were pooled and dialyzed overnight against storage Buffer I (20 mM Tris pH 7.5, 10% glycerol, 2 mM DTT, 0.5 mM EDTA). Protein for crystallography was concentrated to ~10 mg/mL using a centrifugal filter unit (50 kDa molecular mass cut off) and stored at -80°C.

S6.1.7 Protein-peptide co-crystallisation experiments

PCNA was mixed with peptide of interest at 1:1.2 molar ratio, and after incubation on ice for 30 minutes, the sample was pelleted at 16,000×g for 10 min to remove aggregates. The supernatant containing peptide bound protein was stored at -80°C. Crystals were grown by hanging drop vapor diffusion method in 24-well limbro plates containing 500 µL well solution, by mixing 1 µL protein and peptide with equal volume of well solution.[7-9] Diffracting crystals of PCNA bound to **5** were formed in 0.18 M magnesium acetate + 19.5% (polyethylene glycol) PEG 3350 at a temperature of

16°C after one and a half weeks. Diffracting crystals of PCNA bound to **3** were formed in 0.18 M magnesium acetate 19.2% PEG 3350 at a temperature of 16°C after one and a half weeks. Diffracting crystals of PCNA bound to **6** were formed in 0.18 M magnesium acetate + 19.3% PEG 3350 at a temperature of 16°C after one and a half weeks. Crystals were mounted on cryoloops, cryoprotected using paratone-N, and flash cooled in liquid nitrogen.[7-9] Data was collected at 100 K using the MX2[10] (peptide **3**) and MX1[11] (peptides **5** and **6**) beamline at the Australian Synchrotron (Clayton, Vic). Diffraction data was indexed and integrated using XDS (X-ray Detector Software).[12] Pointless (CCP4i)[13] was used to create a mtz file for scaling. Data was scaled using Aimless (CCP4i) to a resolution of 3.30 Å for **5**, 3.30 Å for **3** and 2.80 Å for **6**. Phasing was solved by molecular replacement using Phaser MR (CCP4i)[14] using a search model (PDB: 1AXC, human).[15] Solutions were refined in phenix.refine[16] in iterative rounds with manual rebuilding in Coot.[17, 18] Data collection and refinement statistics for PCNA in complex with **3**, **5**, or **6** are summarised in Table S2. The final structures are deposited on the RCSB database under accession numbers 7M5L, 7M5M and 7M5N, respectively.

S6.1.8 Computational modelling

The model of peptide **1** bound to PCNA was constructed using the solved structure of PCNA bound with p21$_{141\text{-}155}$ peptide as a starting template (PDB ID: 7KQ1), and the necessary residues deleted and unresolved side-chains of residues were modelled into the computational structure.

Models of PCNA and peptide structures **4** and **7** were constructed by drawing the peptide structure in ChemDraw 18.0. Using SMILES, the peptide structure was made into a pdb through phenix eLBOW.[19] The .pdb and .cif files were used to dock the peptide **4** or **7**, in place of peptide **3** in the .mtz map of the PCNA-bound **3** (PDB code: 7M5L), replacing the propyl-linker peptide **3** with the computational peptide. Manual refinement of the computational linker into the .mtz map was done in Coot. Energy minimisation/annealing (n=30) for refinement was carried out in ICM-Pro Molsoft.[20, 21] Refined models were analysed using PyMOL[22] to validate the model by comparing against the p21$_{141\text{-}155}$ structure (7KQ1), and assess side-chain interactions. The resulting structures were visualised in PyMOL and are depicted in Figures S1-S6. Additional analysis was carried out using the RING server[23] and PoseView[24].

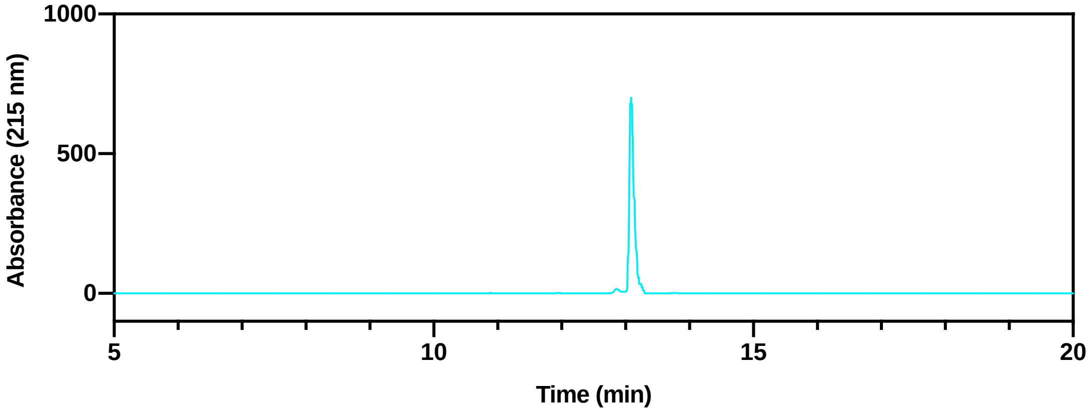
S6.1.9 Cell imaging

MDA-MB-468 mKate is a breast cancer cell line that was lenti-virally modified to stably express nuclear fluorescent (ex 588nm; em 635nm) mKate protein. Cells were maintained in DMEM (Sigma, D5671) base media supplemented with 2 mM L-Glutamine (Sigma, G7513) and 1 mM Sodium Pyruvate (Sigma, S8636) at 37°C and 5% CO_2. For experiments, cells were seeded at ~70% confluency onto glass coverslips inside wells of a 6-well culture plate containing 2 mL of media. The cells were cultured for 48 h to allow for attachment, then treated 24 h with 10 µM peptide **7** or **8** while being protected from light to sustain the florescent signal. After treatment, cells were washed with ice cold PBS (Gibco, 14190144) two times for 5 min to remove residual media and peptide. Cells were fixed with 4% PFA (10% Neutral Buffered Formalin, ChemSupply, #1258) for 10 min at rt, followed by two 5 min wash steps with PBS at rt. Coverslips were then mounted onto microscope slides using DAKO fluorescent mounting medium (S302380-2) and sealed with clear nail polish (Sally Hansen). Slides were allowed to dry overnight. The following day, samples were imaged using a Confocal Olympus FV3000 microscope (Adelaide Microscopy).

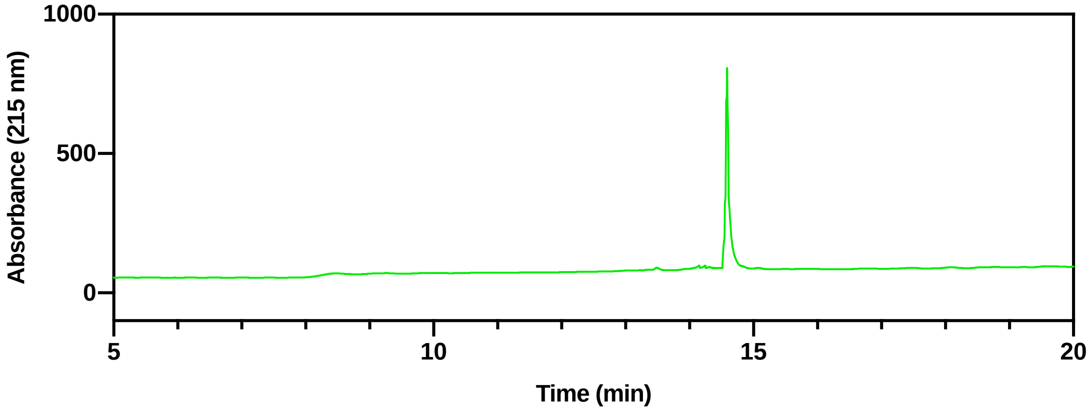
Cell nuclei were visualised using a 594 nm laser with a detection range of 600-700 nm. A 488 nm laser with a detection range of 490-534 nm was used for peptide **8**, which contains a FITC fluorophore. A 405 nm laser with a detection range of 410-485 nm was used for peptide **7**, which contains the Bimane fluorophore. Images were taken with a 30x silicon oil objective, with a 2x zoom setting with imaging software (Olympus, Cell Sens), bringing the total image magnification to 60x.

S6.2 Supplementary information

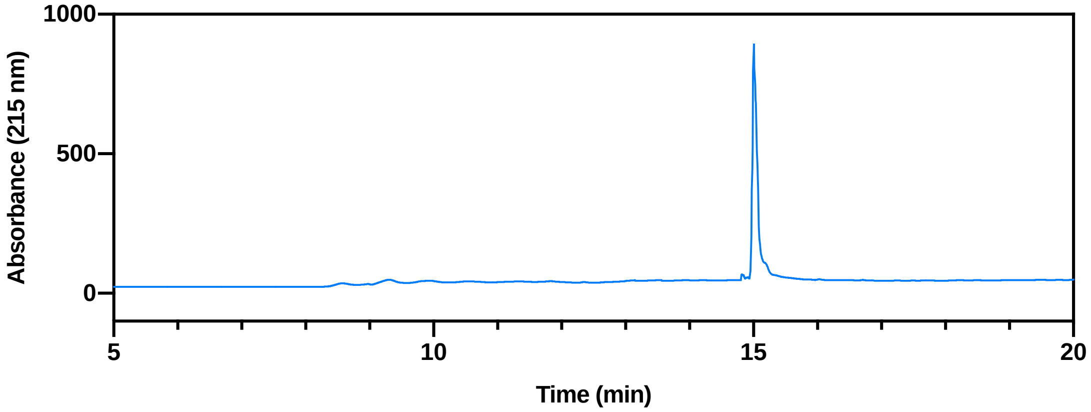
S6.2.1 HPLC spectra



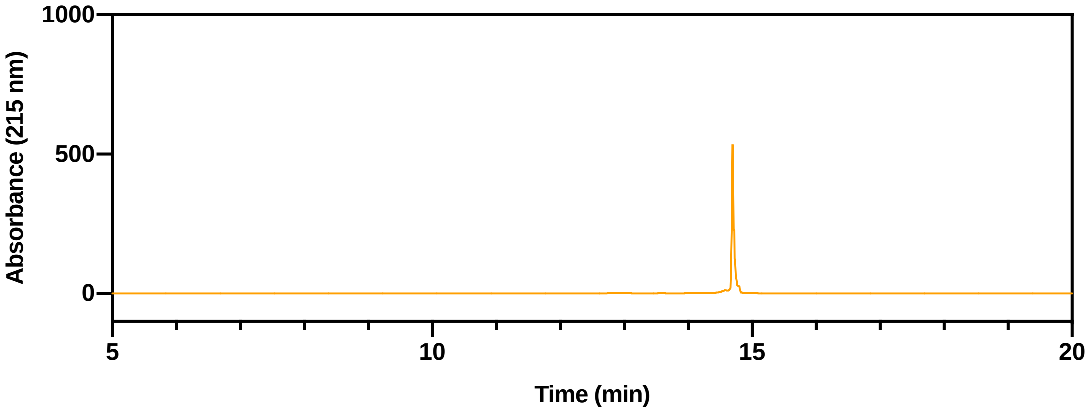
Analytical HPLC spectrum of peptide **1** at 215 nm. Collected over a gradient of 5-50% ACN with 1% TFA, in water with 1% TFA, over 15 minutes (between 5-20 min) at 1.5 mL/min.



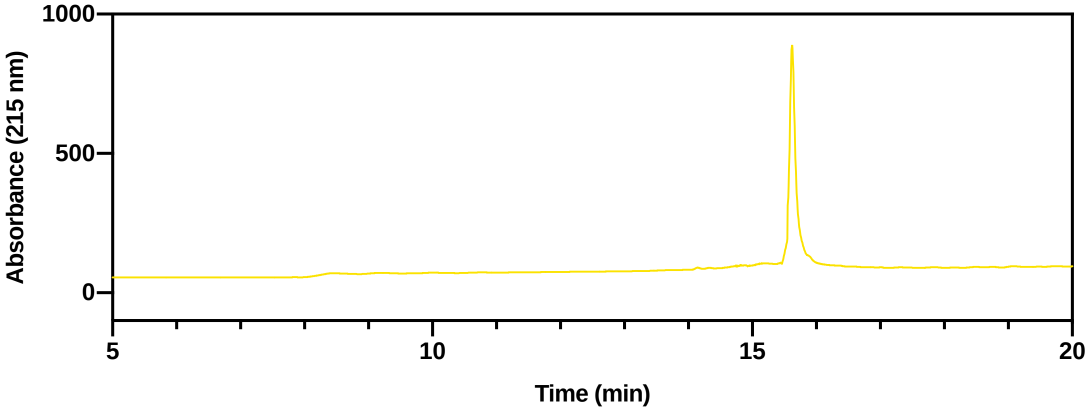
Analytical HPLC spectrum of peptide **3** at 215 nm. Collected over a gradient of 5-50% ACN with 1% TFA, in water with 1% TFA, over 15 minutes (between 5-20 min) at 1.5 mL/min.



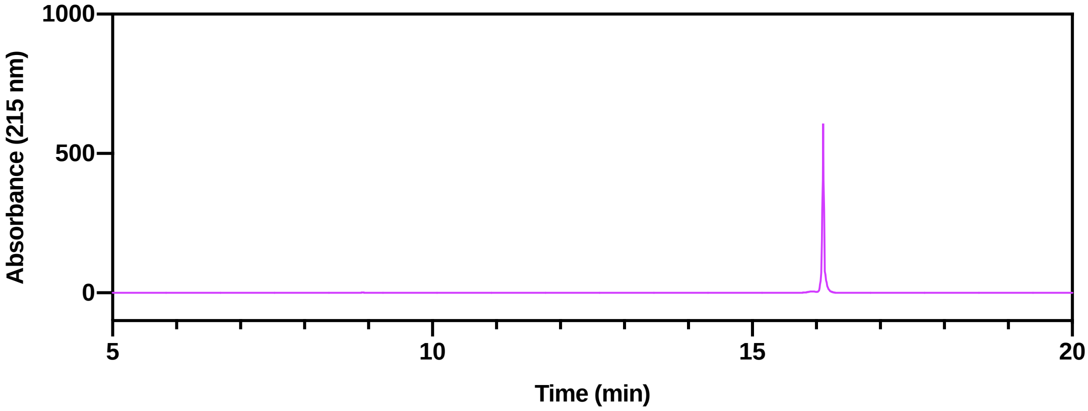
Analytical HPLC spectrum of peptide **3** at 215 nm. Collected over a gradient of 5-50% ACN with 1% TFA, in water with 1% TFA, over 15 minutes (between 5-20 min) at 1.5 mL/min.



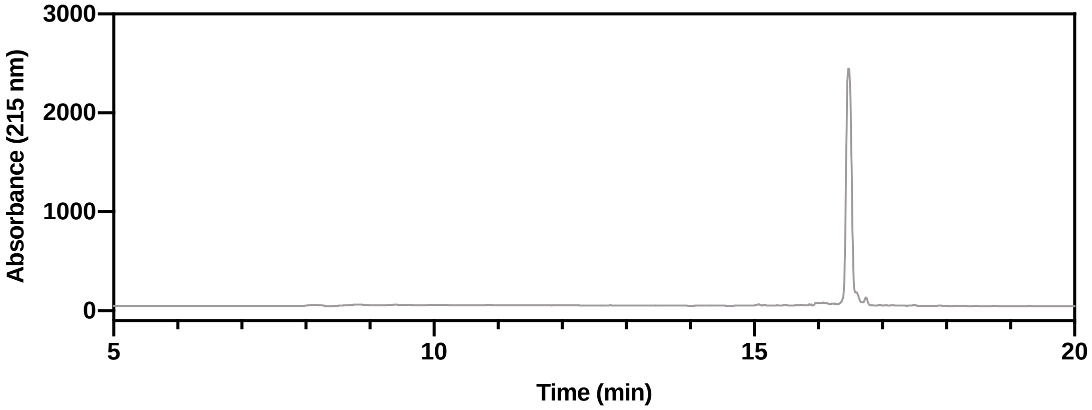
Analytical HPLC spectrum of peptide **5** at 215 nm. Collected over a gradient of 5-50% ACN with 1% TFA, in water with 1% TFA, over 15 minutes (between 5-20 min) at 1.5 mL/min.



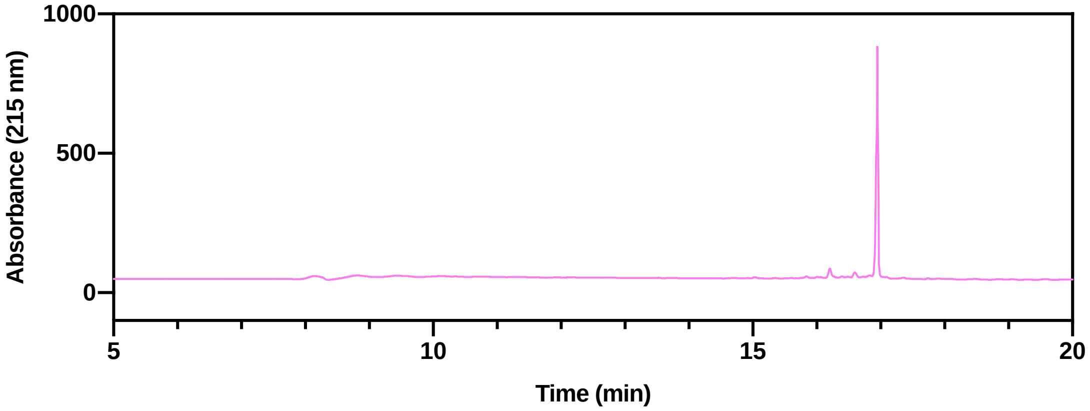
Analytical HPLC spectrum of peptide **6** at 215 nm. Collected over a gradient of 5-50% ACN with 1% TFA, in water with 1% TFA, over 15 minutes (between 5-20 min) at 1.5 mL/min.



Analytical HPLC spectrum of peptide **7** at 215 nm. Collected over a gradient of 5-50% ACN with 1% TFA, in water with 1% TFA, over 15 minutes (between 5-20 min) at 1.5 mL/min.

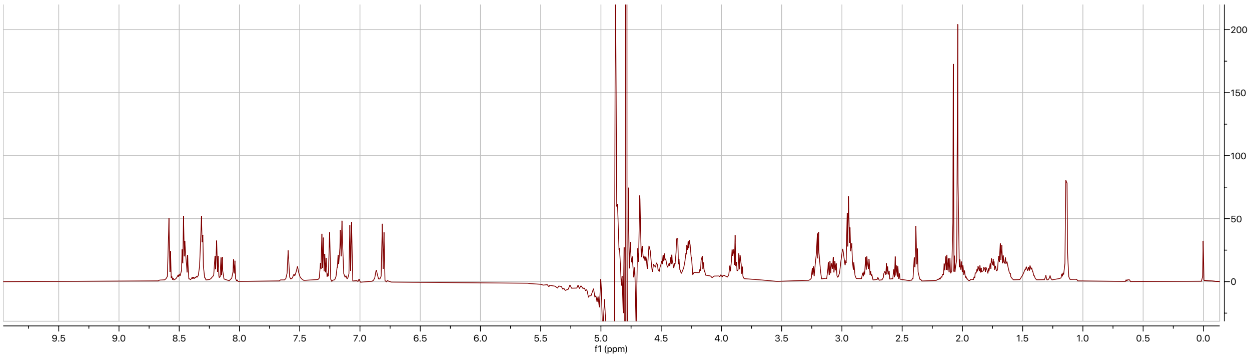


Analytical HPLC spectrum of peptide **8** at 215 nm. Collected over a gradient of 5-50% ACN with 1% TFA, in water with 1% TFA, over 15 minutes (between 5-20 min) at 1.5 mL/min.

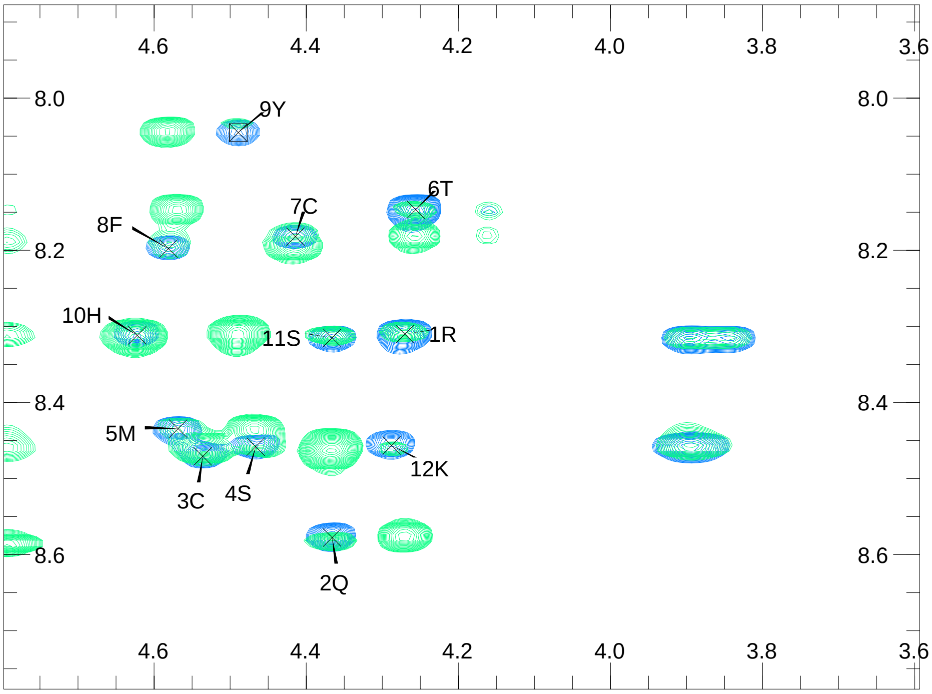


Analytical HPLC spectrum of peptide **9** at 215 nm. Collected over a gradient of 5-50% ACN with 1% TFA, in water with 1% TFA, over 15 minutes (between 5-20 min) at 1.5 mL/min.

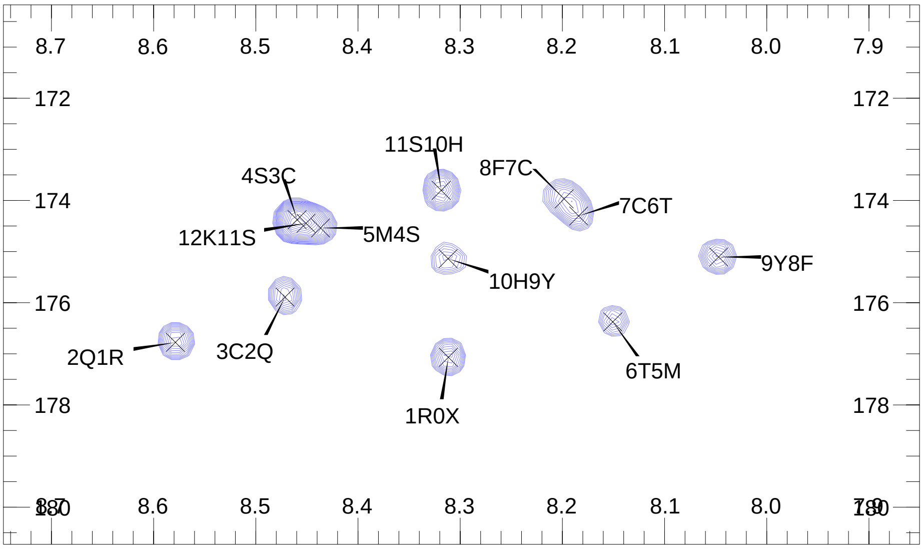
S6.2.2 NMR characterisation



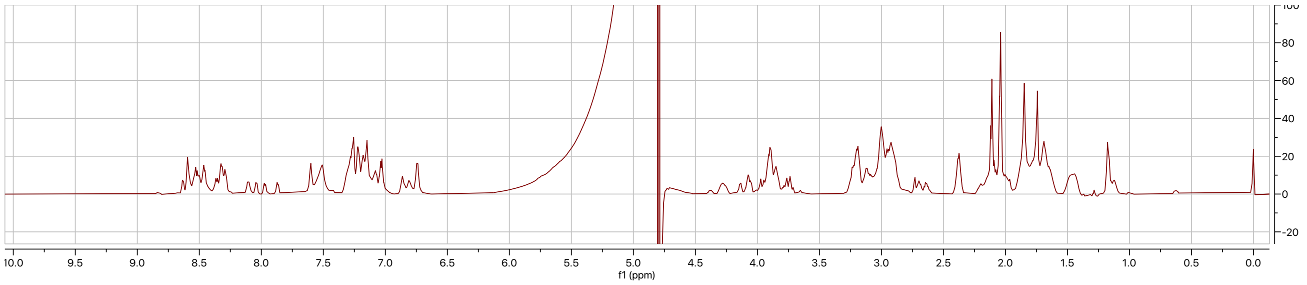
1D ^{1}H waterES spectrum for peptide **2**, referenced to DSS at 0 ppm



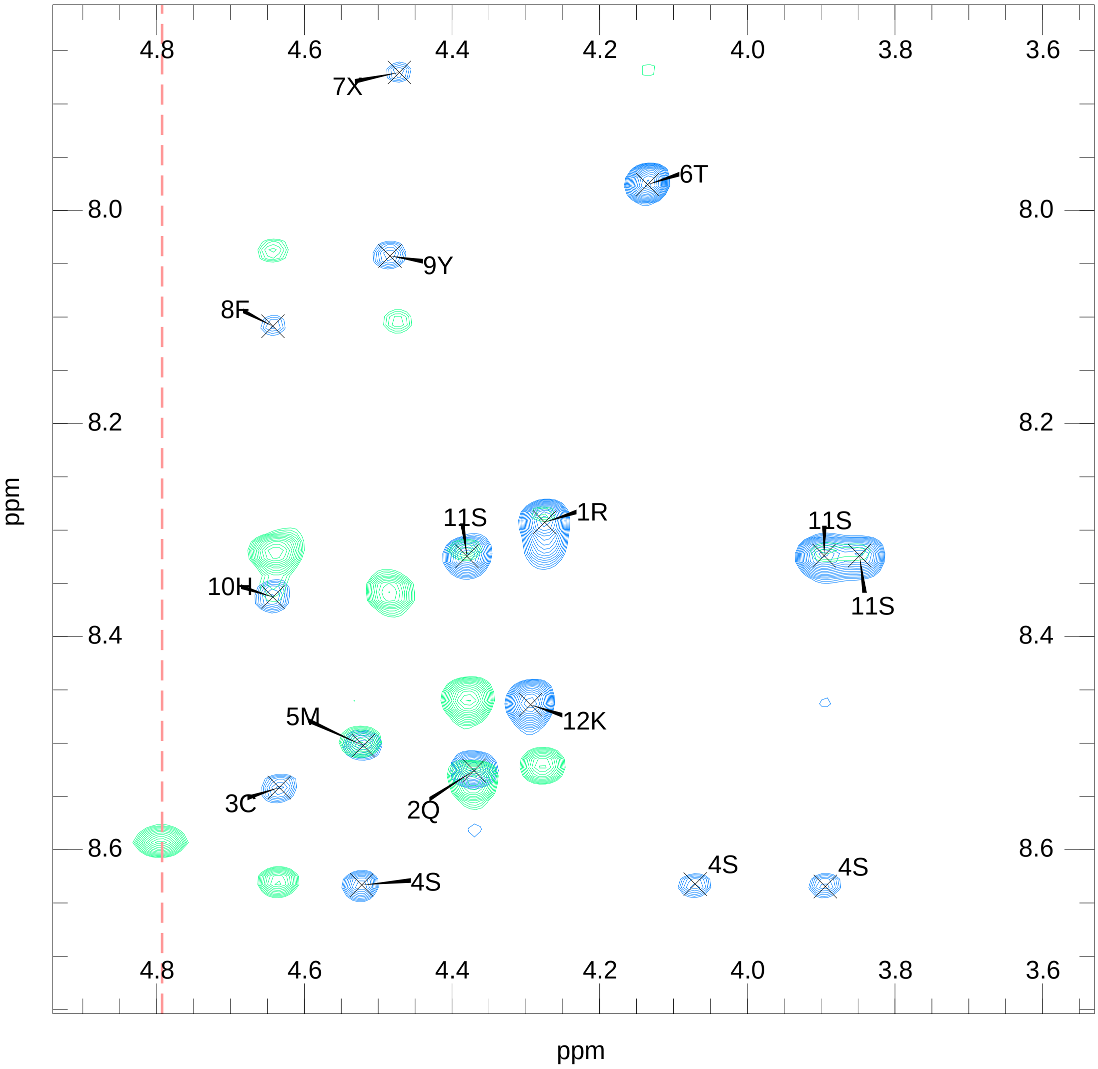
2D ^{1}H TOCSY (blue) and ROESY (green) for peptide **2** in the mainchain region. Assignments of TOCSY spectrum shown.



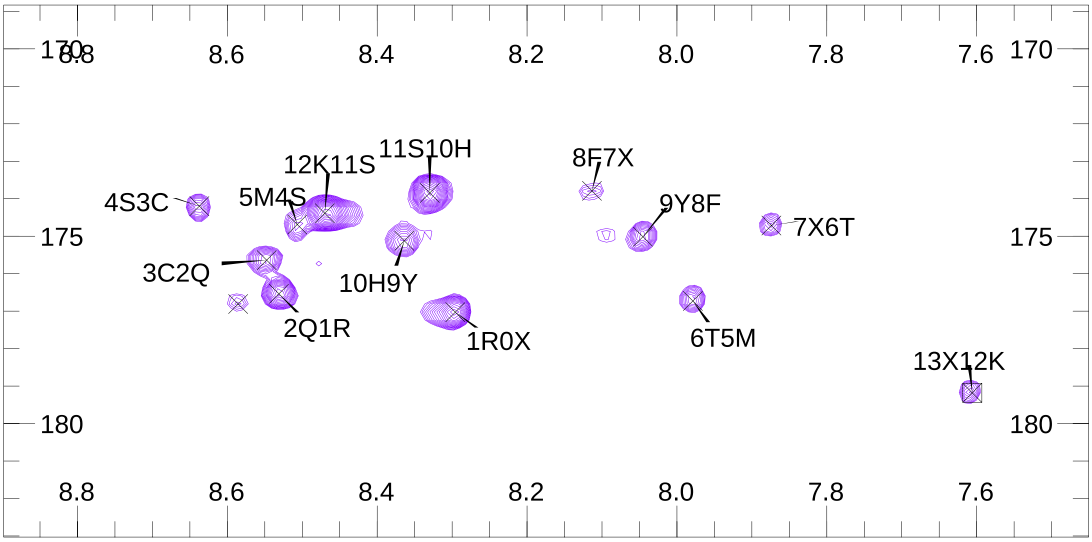
2D heteronuclear HMBC for peptide **2** in the mainchain region. Assignments are NH/CO (f2/f1).



1D ^{1}H wet1D spectrum for peptide **7**, referenced to DSS at 0 ppm



2D ^{1}H TOCSY (blue) and ROESY (green) for peptide **7** in the mainchain region. Red dotted line indicates HDO crosspeaks. Assignments of TOCSY spectrum shown.



2D heteronuclear HMBC for peptide **7** in the mainchain region. Assignments are NH/CO (f2/f1). X represents the bimane modified cysteine residue.

S6.2.3 SPR information

Table S1: SPR data. ε_λ is the extinction coefficient for the peptide at wavelength λ nm. A best estimate is used for modified peptides 3-6. ε_{205} is calculated with online calculator detailed in Anthis 2013.[5] Top Conc is the highest concentration of a 8x 1 in 2 dilution, run sequentially from lowest to highest concentration. K_D is the affinity constant, K_D SE is the standard error of the K_D, and χ^2 is a measure of the goodness of fit, all calculated using the inbuilt Biacore S200 Evaluation Software. Ass/Diss is the contact and dissociation times, respectively, in seconds for each injection cycle.

Name	ε_λ	**λ (nm)**	**Top Conc (nM)**	**Affinity K_D (nM)**	**K_D SE (nM)**	**χ^2**	**Ass/Diss (s)**
Peptide 1	56820	205	500	102.3	5.3	0.0701	40/60
Peptide 2	58200	205	20000	NS	-	-	40/60
Peptide 3	58200*	205	2000	769.1	78	0.250	40/60
Peptide 4	58200*	205	5000	1994	140	0.160	40/60
Peptide 5	58200*	205	5000	2818	80	0.106	40/60
Peptide 6	58200*	205	5000	3855	350	0.527	40/60
Peptide 7	4694[2]	380	2000	570.5	30	0.118	40/60
Peptide 9	4694[2]	380	30000	25190	1900	0.0547	40/60
*The ε_λ of **2** was used as an approximation of the peptide concentration							

S6.2.4 Crystallographic data statistics

Table S2: Data collection and refinement statistics. Data collection and refinement statistics of PCNA bound with peptide **3** (RCSB PDB ID: 7M5L), PCNA bound with peptide **5** (RCSB PDB ID: 7M5M), and PCNA bound with peptide **6** (RCSB PDB ID: 7M5N). Statistics for the highest-resolution shell are shown in parentheses.

	Peptide 3	**Peptide 5**	**Peptide 6**
PDB ID	7M5L	7M5M	7M5N
Wavelength	0.9537	0.9537	0.9537
Resolution range	48.76 - 3.0 (3.107 - 3.0)	38.17 - 3.001 (3.108 - 3.001)	38.37 - 3.11 (3.221 - 3.11)
Space group	P 21 21 21	P 31 2 1	P 31 2 1
Unit cell	70.96 84.468 134.21 90 90 90	83.579 83.579 187.494 90 90 120	83.365 83.365 181.155 90 90 120
Total reflections	33470 (3284)	312221 (29680)	27339 (2646)
Unique reflections	16736 (1643)	15794 (1532)	13672 (1195)
Multiplicity	2.0 (2.0)	19.8 (19.4)	2.0 (2.0)
Completeness (%)	99.62 (98.48)	99.58 (98.52)	94.21 (88.85)
Mean I/sigma(I)	10.20 (0.68)	13.79 (2.43)	11.31 (3.06)
R-merge	0.03674 (0.9641)	0.2337 (1.466)	0.1039 (0.585)
R-meas	0.05195 (1.363)	0.2399 (1.505)	0.1469 (0.8273)
R-pim	0.03674 (0.9641)	0.05348 (0.337)	0.1039 (0.585)
CC1/2	1 (0.496)	0.999 (0.827)	0.992 (0.557)
CC*	1 (0.814)	1 (0.952)	0.998 (0.846)
Reflections used in refinement	16684 (1619)	15761 (1530)	12911 (1195)
Reflections used for R-free	830 (72)	787 (75)	624 (56)
R-work	0.2599 (0.3938)	0.2364 (0.3278)	0.2699 (0.3649)
R-free	0.2638 (0.4255)	0.2749 (0.3665)	0.3088 (0.4653)
CC(work)	0.950 (0.577)	0.945 (0.848)	0.880 (0.659)
CC(free)	0.954 (0.715)	0.949 (0.748)	0.964 (0.674)
Number of non-hydrogen atoms	5636	5611	5324
macromolecules	5395	5443	5173
ligands	241	168	151
Protein residues	769	763	755
RMS(bonds)	0.003	0.004	0.004
RMS(angles)	0.81	0.75	0.84
Ramachandran favored (%)	95.02	92.82	94.32
Ramachandran allowed (%)	4.59	6.91	5.01
Ramachandran outliers (%)	0.39	0.27	0.68
Rotamer outliers (%)	0.00	0.19	0.22
Clashscore	12.02	27.90	13.92
Average B-factor	106.51	64.71	60.98
macromolecules	105.29	63.97	60.73
ligands	133.71	88.78	69.52
Number of TLS groups	1	1	1
Twin Law	Not applicable	Not applicable	-h, -k, l

[a] $Rmerge = \Sigma |I - \langle I \rangle| / \Sigma I$.

[b] $Rpim = \Sigma h\, [1/ (^{n}/_{h} - 1)]^{1/2}\, \Sigma i|\langle I_h \rangle - I_{h,i}| / \Sigma_h \Sigma_i I_{h,i}$ (2)

[c] $Rwork = \Sigma|F_o - F_c|/\Sigma|F_o|$ for all data excluding data used to calculate Rfree.

[d] $Rfree = \Sigma|F_o - F_c|/\Sigma|F_o|$ for all data.

S6.2.5 PCNA-peptide structures supplementary data

Table S3: Percentage buried surface area (BSA) – A^2 | % of PIP-box residues (excluding the covalently linked residues Cys145* and Cys149*) and the total buried surface area of the PIP-box residues including the covalent linker.

Peptide bound to PCNA	**p21$_{141-155}$**	**Peptide 3**	**Peptide 3**	**Peptide 4**	**Peptide 5**	**Peptide 6**	**Peptide 7**
Structure type	Co-crystal PDB: 7KQ1	Co-crystal PDB: 7M5L	Computational (confidence) model	Computational model	Co-crystal PDB: 7M5M	Co-crystal PDB: 7M5L	Computational model
Glu$_{144}$	98.61 | 70	141.15 | 50	112.31 | 80	133.02 | 80	117.83 | 60	92.99 | 40	143.85 | 80
Ser$_{146}$	31.97 | 60	47.30 | 50	40.34 | 50	35.30 | 40	38.11 | 40	35.78 | 40	72.12 | 40
Met$_{147}$	140.69 |100	147.53 | 70	113.37 | 100	124.28 | 100	142.54 | 100	130.03 | 100	138.45 | 100
Thr$_{148}$	34.23|40	58.75 | 40	17.31 | 20	23.89 | 30	20.03 | 30	29.80 | 30	27.08 | 30
Phe$_{150}$	66.67 |50	77.50 | 70	82.30 | 60	83.20 | 50	75.24 | 50	70.27 | 40	87.12 | 60
Tyr$_{151}$	132.92 | 90	133.53 | 90	129.25 | 90	140.41 | 70	142.60 | 100	165.87 | 90	145.93 | 90
Glu$_{144}$ - Tyr$_{151}$ (including linker)	252.53 | 90	568.39 | 60	533.24 | 60	561.39 | 60	589.45 | 60	552.08 | 50	652.59 | 60

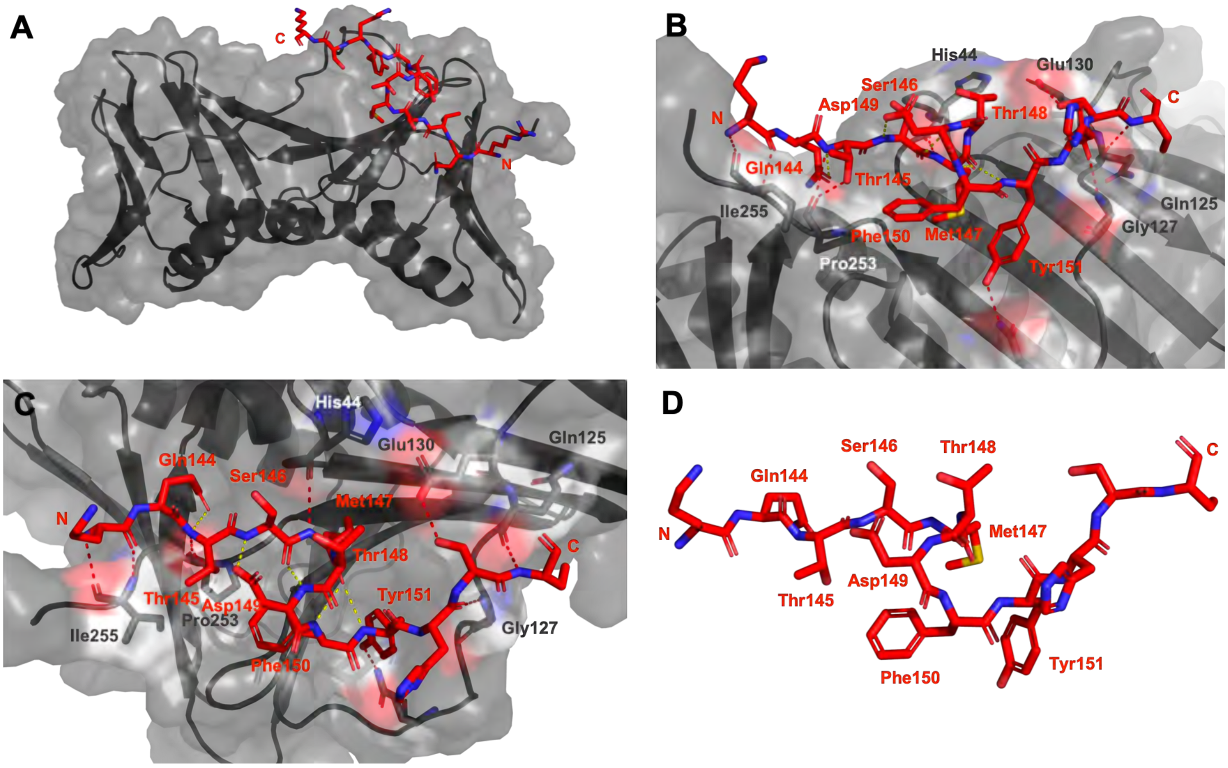

A
C
N
B
His44
Glu130
Ser146
Asp149
N
C
Thr148
Gln144
Thr145
Ile255
Phe150
Met147
Gln125
Gly127
Pro253
Tyr151
C
His44
Glu130
Gln125
Gln144
Ser146
Met147
N
Thr148
C
Thr145
Asp149
Pro253
Tyr151
Ile255
Gly127
Phe150
D
Ser146
Thr148
C
Gln144
Met147
N
Asp149
Thr145
Tyr151
Phe150

Figure S1: Computational model of peptide **1** bound to PCNA. Peptide **1** is shown in red as sticks, and PCNA in grey with cartoon and transparent surface representation. The PIP-box residues of peptide **1** are labelled in red, and PCNA residues in grey. Intermolecular interactions are shown as red dashes, and intramolecular interactions as yellow dashes.

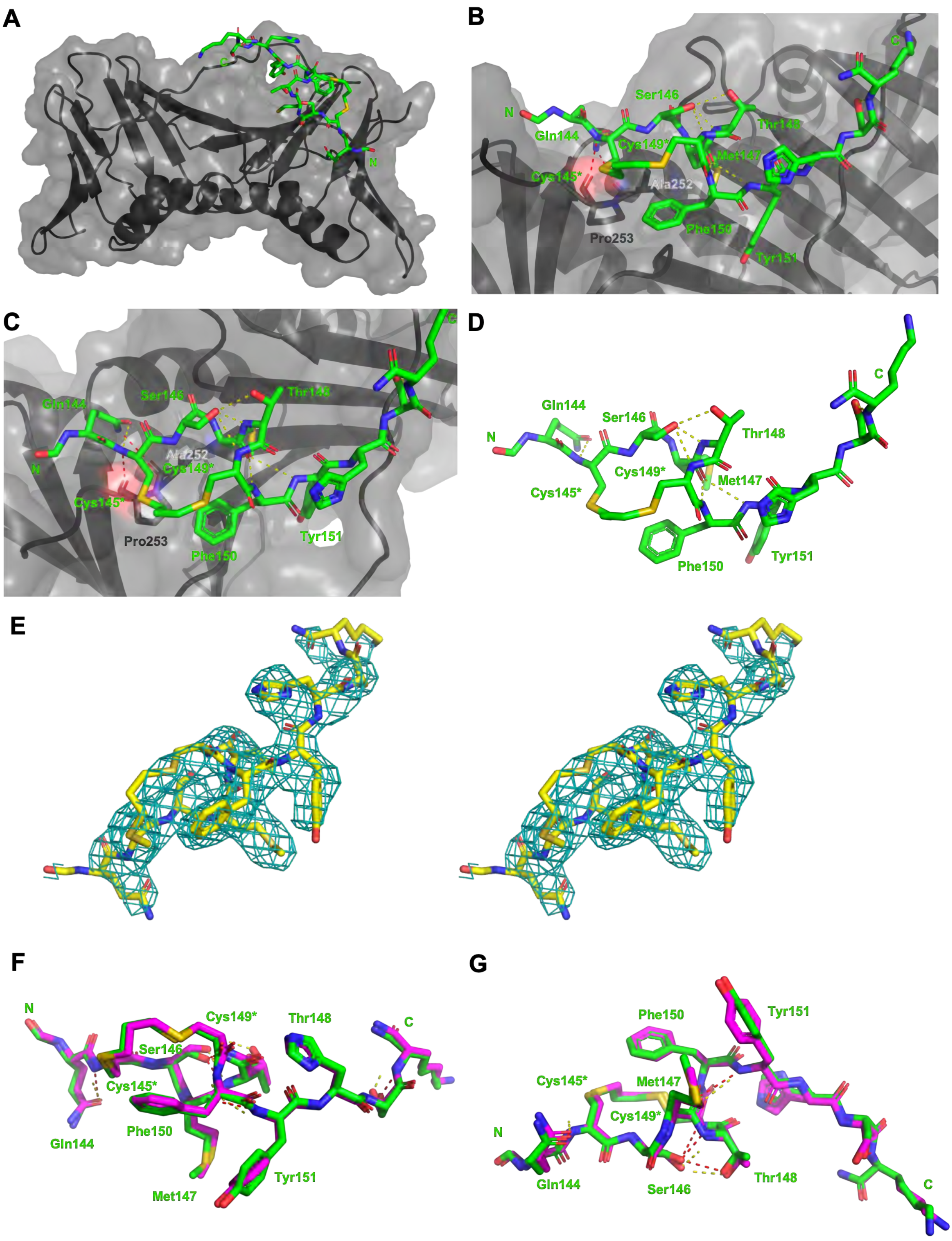

A
C
N
B
N
Gln144
Ser146
Thr148
Cys149*
Met147
Cys145*
Ala252
Pro253
Phe150
Tyr151
C
C
Gln144
Ser146
Thr148
N
Ala252
Cys149*
Cys145*
Pro253
Phe150
Tyr151
C
D
N
Gln144
Ser146
Thr148
C
Cys149*
Cys145*
Met147
Phe150
Tyr151
E
F
N
Cys149*
Thr148
C
Ser146
Cys145*
Gln144
Phe150
Met147
Tyr151
G
Phe150
Tyr151
Cys145*
Met147
Cys149*
N
Gln144
Ser146
Thr148
C

Figure S2: Co-crystal structure of peptide **3** bound to PCNA (PDB: 7M5L). Peptide **3** is shown in green as sticks, and PCNA in grey with cartoon and transparent surface representation (**A**-**C**). **B**-**D** The PIP-box residues of peptide **3** are labelled in green, and PCNA residues in grey. Intermolecular interactions are shown as red dashes, and intramolecular interactions as yellow dashes. **E** Representative electron density of peptide **3** (yellow, sticks) shown as a wall-eye stereo image of reduced model bias feature-enhanced map,[25] view contoured at 1.0σ. **F** & **G** Overlay of co-crystal structure of peptide **3** from co-crystalstructure (green) and the computational validation model of **3** (pink), where the intramolecular interactions of the crystal structure are in yellow, and in red for the model.

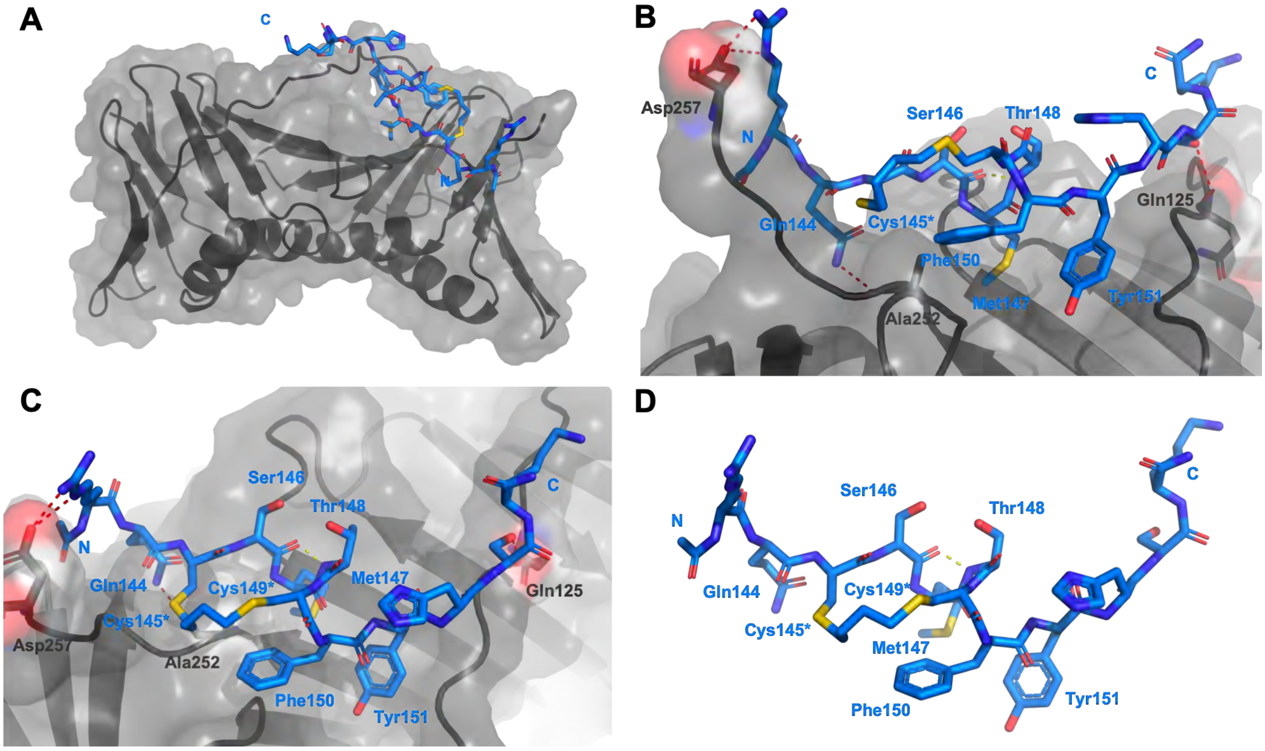

A
C
N
B
Asp257
N
Ser146
Thr148
C
Gln125
Gln144
Cys145*
Phe150
Met147
Tyr151
Ala252
C
Ser146
Thr148
C
N
Gln144
Cys149*
Met147
Gln125
Cys145*
Asp257
Ala252
Phe150
Tyr151
D
N
Ser146
Thr148
C
Gln144
Cys149*
Cys145*
Met147
Phe150
Tyr151

Figure S3: Computational model of peptide **4** bound to PCNA. Peptide **4** is shown in blue as sticks (**A**-**D**), and PCNA in grey with cartoon and transparent surface representation (**A**-**C**). The PIP-box residues of peptide **4** are labelled in blue, and PCNA residues in grey. Intermolecular interactions are shown as red dashes, and intramolecular interactions as yellow dashes.

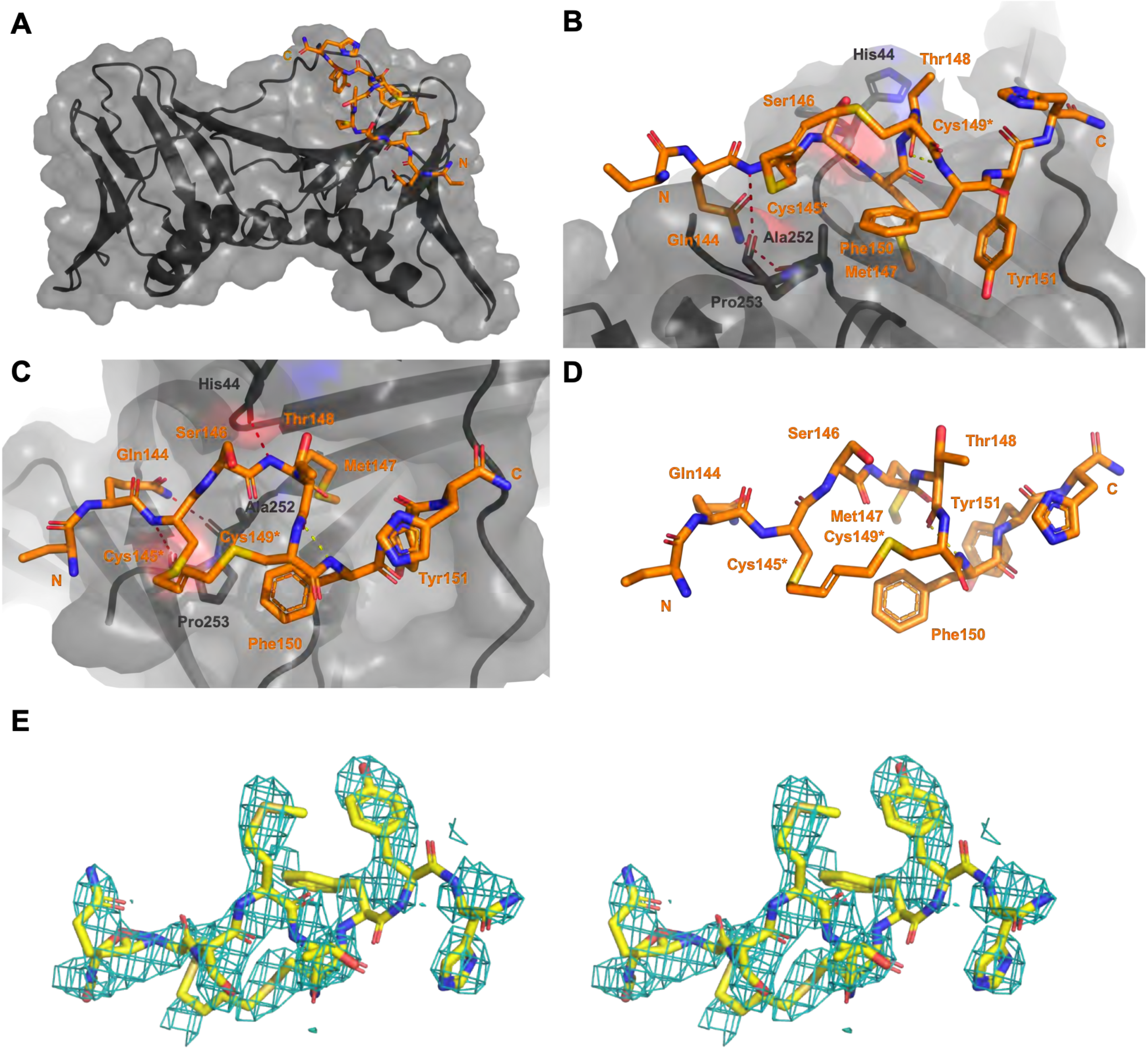

A
C
N
B
His44
Thr148
Ser146
Cys149*
C
N
Cys145*
Gln144
Ala252
Phe150
Met147
Tyr151
Pro253
C
His44
Ser146
Thr148
Gln144
Met147
C
Ala252
Cys149*
Cys145*
N
Tyr151
Pro253
Phe150
D
Ser146
Thr148
Gln144
Tyr151
C
Met147
Cys149*
Cys145*
N
Phe150
E

Figure S4: Co-crystal structure of peptide **5** bound to PCNA (PDB: 7M5M). **A**-**D** Peptide **5** is shown in orange as sticks, and PCNA in grey with cartoon and transparent surface representation (**A**-**C**). The PIP-box residues of peptide **3** are labelled in orange, and PCNA residues in grey. Intermolecular interactions are shown as red dashes, and intramolecular interactions as yellow dashes. **E** Representative electron density of peptide **5** (yellow, sticks) shown as a wall-eye stereo image of reduced model bias feature-enhanced map,[25] view contoured at 1.0σ.

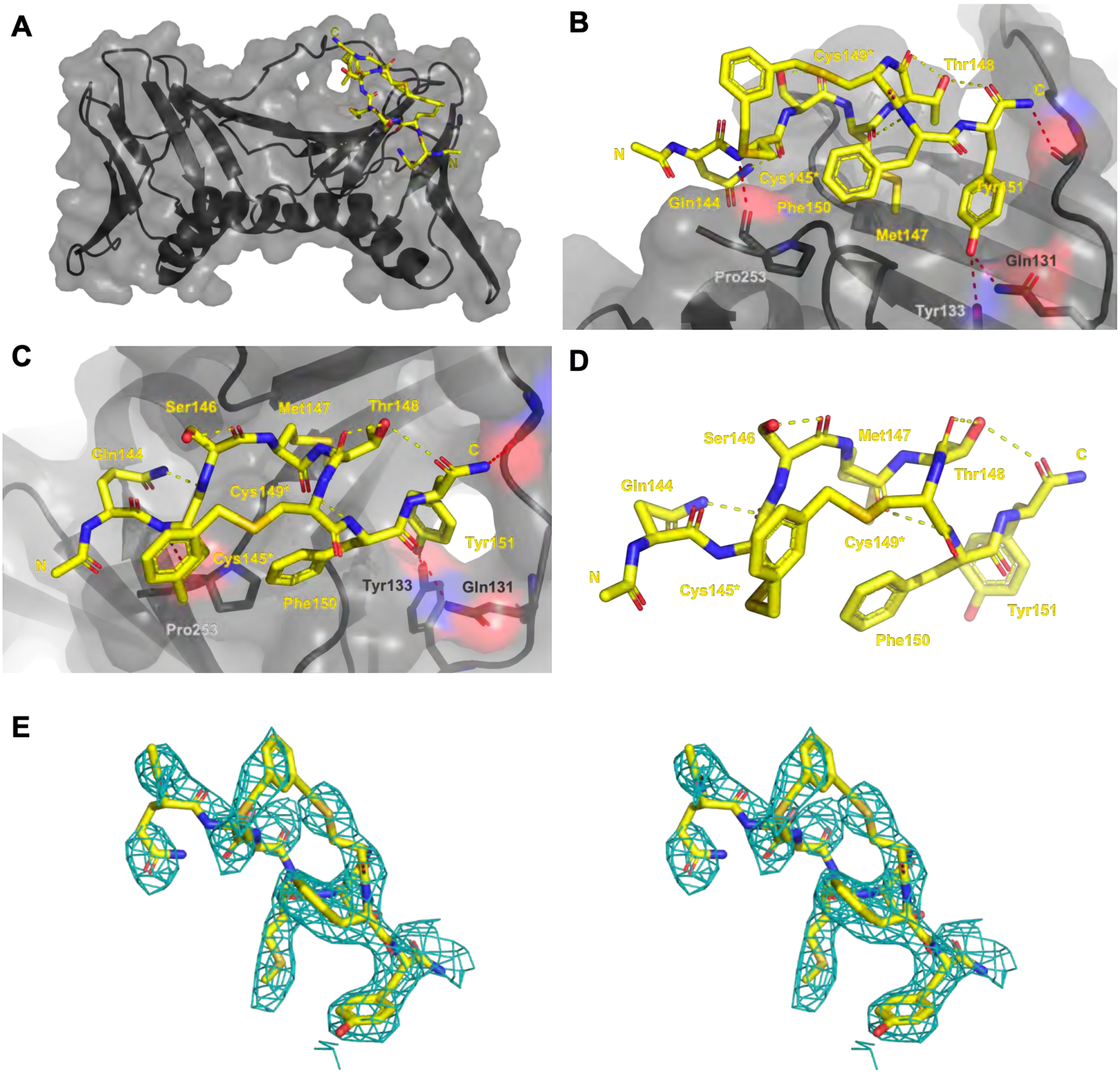

A
C
N
B
Cys149*
Thr148
N
C
Cys145*
Gln144
Phe150
Tyr151
Met147
Gln131
Pro253
Tyr133
C
Ser146
Met147
Thr148
Gln144
Cys149*
C
N
Cys145*
Tyr151
Phe150
Tyr133
Gln131
Pro253
D
Ser146
Met147
Gln144
Thr148
C
Cys149*
N
Cys145*
Tyr151
Phe150
E

Figure S5: Co-crystal structure of peptide **6** bound to PCNA (PDB: 7M5N). Peptide **6** is shown in yellow as sticks (**A**-**E**), and PCNA in grey with cartoon and transparent surface representation (**A**-**C**). The PIP-box residues of peptide **6** are labelled in yellow, and PCNA residues in grey. Intermolecular interactions are shown as red dashes, and intramolecular interactions as yellow dashes. **E** Representative electron density of peptide **6** (yellow, sticks) shown as a wall-eye stereo image of reduced model bias feature-enhanced map,[25] view contoured at 1.0σ.

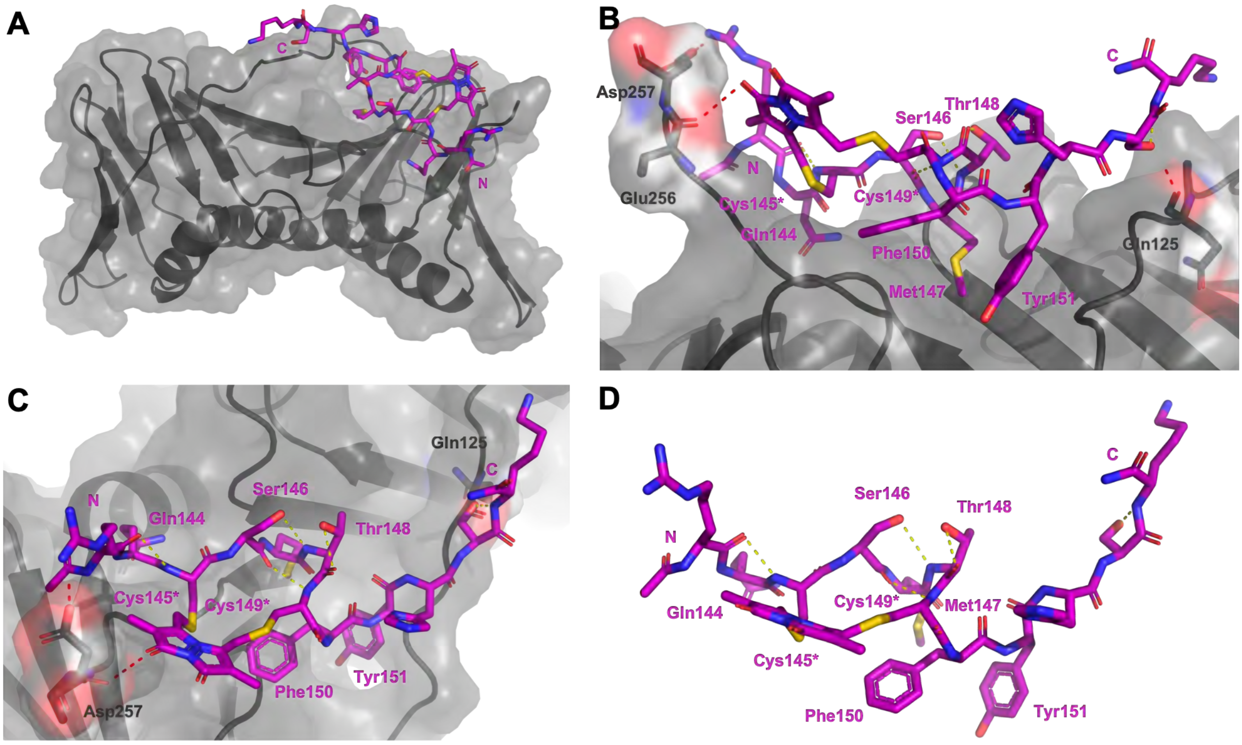

A
C
N
B
Asp257
Glu256
N
Cys145*
Gln144
Ser146
Thr148
Cys149*
Phe150
Met147
Tyr151
C
Gln125
C
N
Gln144
Ser146
Thr148
Gln125
C
Cys145*
Cys149*
Tyr151
Phe150
Asp257
D
N
Ser146
Thr148
C
Gln144
Cys149*
Met147
Cys145*
Phe150
Tyr151

Figure S6: Computational model of peptide **7** bound to PCNA. Peptide **7** is shown in purple as sticks (**A**-**D**), and PCNA in grey with cartoon and transparent surface representation (**A**-**C**). The PIP-box residues of peptide **7** are labelled in purple, and PCNA residues in grey. Intermolecular interactions are shown as red dashes, and intramolecular interactions as yellow dashes.

S6.2.6 Main-chain assignments and secondary shift analysis

Table S4: NMR resonance assignments for main-chain of peptide 2, and secondary shift calculations. All values in ppm. RC – random coil value from Wishart 2011[6]. NN – nearest neighbour corrections applied to random coil value; Hα did not require NN corrections due to sequence used here. Δδ – Secondary shift values calculated as the difference between the observed resonance (Hα, Cα, CO or NH) and the respective NN value (obs-NN), shown graphically in green on Figure S7.

2	Hα	RC	NN	Δδ	Cα	RC	NN	Δδ	CO	RC	NN	Δδ	NH	RC	NN	Δδ
1R	4.27	4.34	4.34	-0.07	56.29	56.0	56.0	0.29	176.8	176.3	175.82	0.98	8.31	8.23	8.23	0.08
2Q	4.37	4.34	4.34	0.03	55.89	55.7	55.7	0.19	175.9	176.0	175.49	0.41	8.58	8.32	8.47	0.11
3C	4.54	4.55	4.55	-0.01	58.25	58.2	58.2	0.05	174.39	174.6	174.20	0.19	8.47	8.32	8.47	0.00
4S	4.47	4.47	4.47	0.00	58.46	58.3	58.3	0.16	174.51	174.6	173.91	0.60	8.46	8.31	8.51	-0.05
5M	4.57	4.48	4.48	0.09	55.70	55.4	55.4	0.30	176.38	176.3	176.30	0.08	8.43	8.28	8.44	-0.01
6T	4.26	4.35	4.35	-0.09	62.11	61.8	61.8	0.31	174.3	174.7	174.19	0.11	8.15	8.15	8.30	-0.15
7C	4.41	4.55	4.45	-0.04	58.28	58.2	57.97	0.31	173.98	174.6	173.77	0.21	8.18	8.32	8.34	-0.16
8F	4.58	4.62	4.52	0.06	57.65	57.7	57.48	0.17	175.1	175.8	174.67	0.43	8.2	8.3	8.39	-0.19
9Y	4.49	4.55	4.55	-0.06	57.74	57.9	57.9	-0.16	175.14	175.9	175.00	0.14	8.05	8.12	8.22	-0.17
10H	4.62	4.73	4.73	-0.11	54.78	55.0	55.0	-0.22	173.8	174.1	173.46	0.34	8.31	8.42	8.42	-0.11
11S	4.37	4.47	4.47	-0.1	58.19	58.3	58.3	-0.11	174.41	174.6	173.88	0.53	8.31	8.31	8.51	-0.20
12K	4.29	4.32	4.32	-0.03	56.22	56.2	56.2	0.02	179.1	176.6	176.60	2.50	8.46	8.29	8.45	0.01

Table S5: NMR resonance assignments for main-chain of peptide 7, and secondary shift calculations. All values in ppm. RC – random coil value from Wishart 2011[6]. NN – nearest neighbour corrections applied to random coil value; Hα did not require NN corrections due to sequence used here. Δδ – Secondary shift values calculated as the difference between the observed resonance (Hα, Cα, CO or NH) and the respective NN value, shown graphically in blue on Figure S7.

7	Hα	RC	NN	Δδ	Cα	RC	NN	Δδ	CO	RC	NN	Δδ	NH	RC	NN	Δδ
1R	4.27	4.34	4.34	-0.07	56.15	56.0	56.0	0.15	176.54	176.3	175.82	0.72	8.29	8.23	8.23	0.06
2Q	4.37	4.34	4.34	0.03	55.87	55.7	55.7	0.17	175.64	176.0	175.49	-0.36	8.53	8.32	8.47	0.06
3C*	4.63	4.55	4.55	0.08	55.53	58.2	58.2	-2.67	174.20	174.6	174.20	-0.40	8.54	8.32	8.47	0.07
4S	4.52	4.47	4.47	0.05	58.34	58.3	58.3	0.04	174.69	174.6	173.91	0.09	8.63	8.31	8.51	0.12
5M	4.52	4.48	4.48	0.04	56.55	55.4	55.4	1.15	176.73	176.3	176.30	0.43	8.50	8.28	8.44	0.06
6T	4.14	4.35	4.35	-0.21	62.60	61.8	61.8	0.8	174.71	174.7	174.19	0.01	7.98	8.15	8.30	-0.32
7C*	4.47	4.55	4.45	0.02	-	58.2	57.97	-	173.79	174.6	173.77	-0.81	7.87	8.32	8.34	-0.47
8F	4.64	4.62	4.52	0.12	-	57.7	57.48	-	175.01	175.8	174.67	-0.79	8.11	8.30	8.39	-0.28
9Y	4.48	4.55	4.55	-0.07	57.73	57.9	57.9	-0.17	175.12	175.9	175.00	-0.78	8.04	8.12	8.22	-0.18
10H	4.64	4.73	4.73	-0.09	54.81	55.0	55.0	-0.19	173.84	174.1	173.46	-0.26	8.36	8.42	8.42	-0.06
11S	4.38	4.47	4.47	-0.09	58.25	58.3	58.3	-0.05	174.39	174.6	173.88	-0.21	8.32	8.31	8.51	-0.19
12K	4.29	4.32	4.32	-0.03	-	56.2	56.2	-	179.13	176.6	176.60	2.53	8.46	8.29	8.45	0.01

Table S6: NMR resonance assignments for main-chain of peptide 2 and 7, and secondary shift calculation for 7, relative to 2. All values in ppm. Δδ – Secondary shift values calculated as the difference between the observed resonance (Hα, Cα, CO or NH) of **7**, and the respective value for **2** (7-2), shown graphically in purple on Figure S7.

7-2	7Hα	2Hα	Δδ	7-2	7Cα	2Cα	Δδ	7CO	2CO	Δδ	7NH	2NH	Δδ
1R	4.27	4.27	0	1R	56.15	56.29	-0.14	176.54	176.80	-0.26	8.29	8.31	0.02
2Q	4.37	4.37	0	2Q	55.87	55.89	-0.02	175.64	175.90	-0.26	8.53	8.58	0.05
3C*	4.63	4.54	0.09	3C*	55.53	58.25	-2.72	174.20	174.39	-0.19	8.54	8.47	-0.07
4S	4.52	4.47	0.05	4S	58.34	58.46	-0.12	174.69	174.51	0.18	8.63	8.46	-0.17
5M	4.52	4.57	-0.05	5M	56.55	55.70	0.85	176.73	176.38	0.35	8.50	8.43	-0.07
6T	4.14	4.26	-0.12	6T	62.60	62.11	0.49	174.71	174.30	0.41	7.98	8.15	0.17
7C*	4.47	4.41	0.06	7C*	-	58.28	-	173.79	173.98	-0.19	7.87	8.18	0.31
8F	4.64	4.58	0.06	8F	-	57.65	-	175.01	175.10	-0.09	8.11	8.20	0.09
9Y	4.48	4.49	-0.01	9Y	57.73	57.74	-0.01	175.12	175.14	-0.02	8.04	8.05	0.01
10H	4.64	4.62	0.02	10H	54.81	54.78	0.03	173.84	173.8	0.04	8.36	8.31	-0.05
11S	4.38	4.37	0.01	11S	58.25	58.19	0.06	174.39	174.41	-0.02	8.32	8.31	-0.01
12K	4.29	4.29	0	12K	-	56.22	-	179.13	179.10	0.03	8.46	8.46	0.00

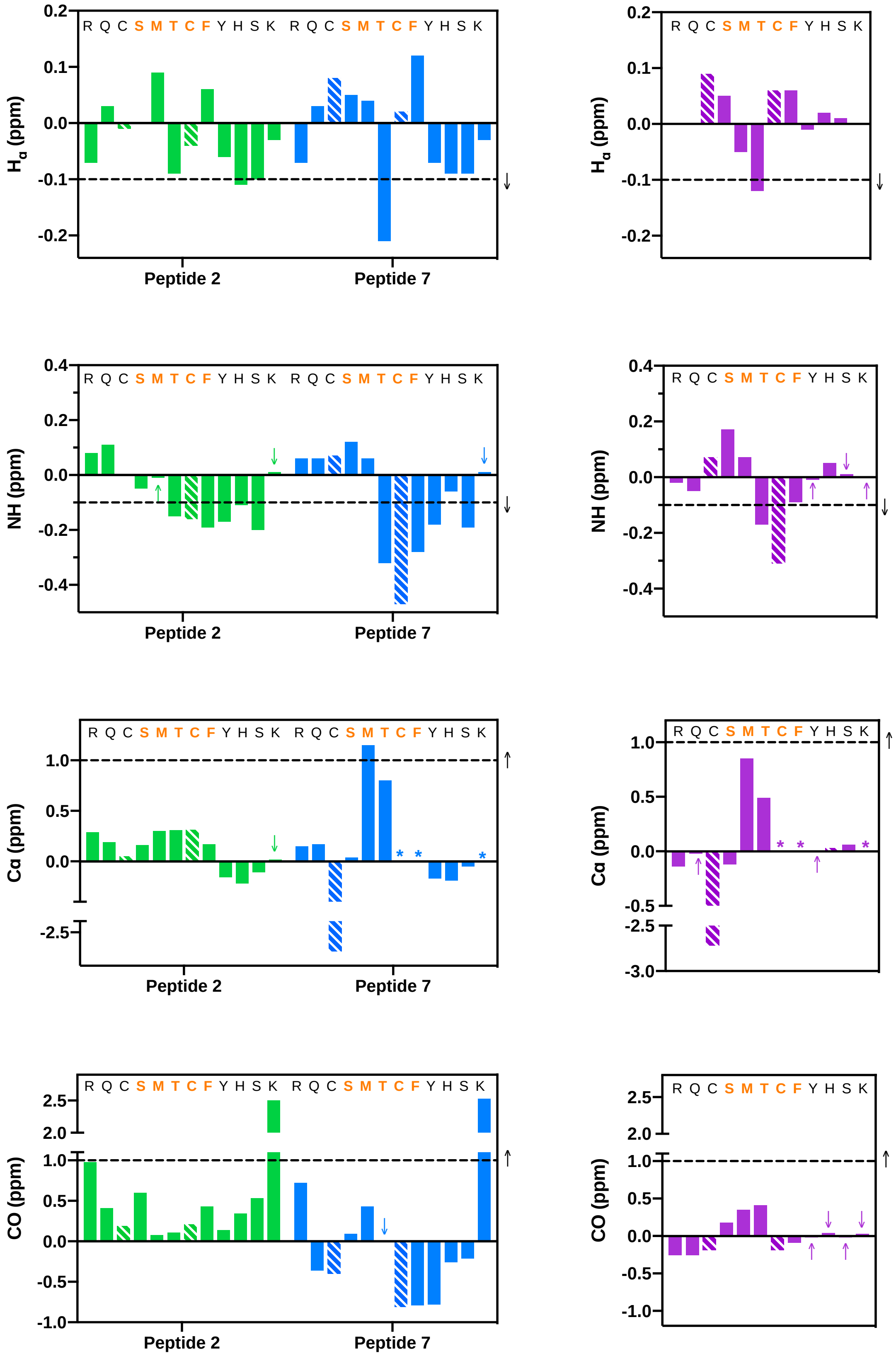


Figure S7: NMR secondary shift of mainchain resonances of peptide **2** (green) and **7** (blue) calculated relative to literature random coil values, with nearest neighbour corrections applied; and for peptide **7** relative to linear precursor peptide **2** (purple). Each bar represents the resonance for each consecutive amino-acid in the sequence, where the sequence is annotated that the top of each panel. The segment of peptide where helical structure is anticipated, is indicated in orange writing. Arrows indicate values which are too close to zero to be easily observed, stars indicate values which could not be reliably extracted from the NMR spectrum. The black dashed line, and black arrow, on each panel marks the threshold that should be surpassed by three consecutive residues to indicate the presence of a helical motif.

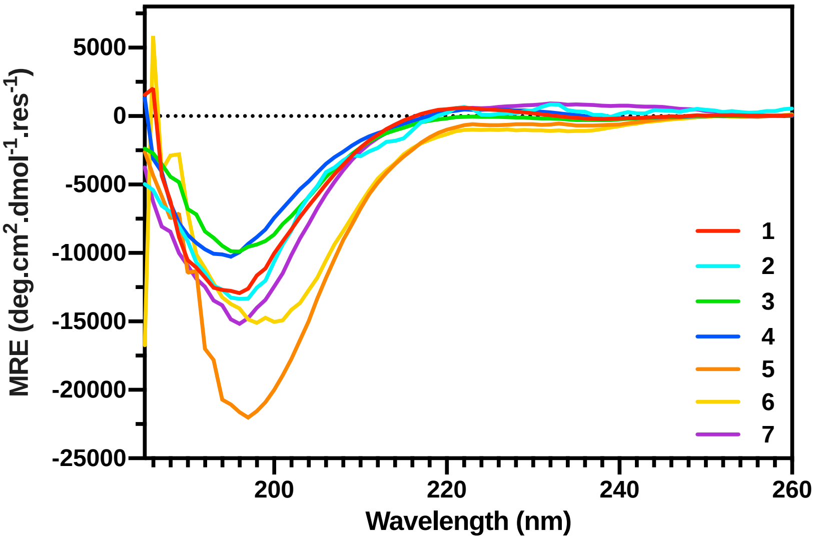


Figure S8: Circular dichroism spectra collected for peptides **1**-**7** in 10 mM phosphate buffer at pH 7.2. The spectra shown are the average of 8 scans, collected at 50 nm/min with a pitch of 1 mm in a 1 mm cuvette, using a Jasco J-810 (UniSA BioPhysical Characterisation facility); then smoothed using the in-built software with a Savitsky-Golay function and a convolution width of 7.

REFERENCES

1. W. S. Hancock and J. E. Battersvy, A new micro-test for the detection of incomplete coupling reactions in solid-phase peptide synthesis using 2,4,6-trinitrobenzene-sulphonic acid, *Anal. Biochem.*, 1976, **71**, 260-264.

2. X. Shen, C. B. Pattillo, S. Pardue, S. C. Bir, R. Wang and C. G. Kevil, Measurement of plasma hydrogen sulfide in vivo and in vitro, *Free Radical Biol. Med.*, 2011, **50**, 1021-1031.

3. G. Zhang, F. Barragan, K. Wilson, N. Levy, A. Herskovits, M. Sapozhnikov, Y. Rodriguez, L. Kelmendi, H. Alkasimi, H. Korsmo, M. Chowdhury and G. Gerona-Navarro, A Solid-Phase Approach to Accessing Bisthioether-Stapled Peptides Resulting in a Potent Inhibitor of PRC2 Catalytic Activity, *Angew. Chem. Int. Ed. (English)*, 2018, **57**, 17073-17078.

4. A. J. Horsfall, K. R. Dunning, K. L. Keeling, D. B. Scanlon, K. L. Wegener and A. D. Abell, A bimane‐based peptide staple for combined helical induction and fluorescent imaging, *ChemBioChem*, 2020, **21**, 3423-3432.

5. N. J. Anthis and G. M. Clore, Sequence-specific determination of protein and peptide concentrations by absorbance at 205 nm, *Protein Sci.*, 2013, **22**, 851-858.

6. D. S. Wishart, Interpreting protein chemical shift data, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2011, **58**, 62-87.

7. A. C. Marshall, A. J. Kroker, L. A. Murray, K. Gronthos, H. Rajapaksha, K. L. Wegener and J. B. Bruning, Structure of the sliding clamp from the fungal pathogen Aspergillus fumigatus (AfumPCNA) and interactions with Human p21, *FEBS J.*, 2017, **284**, 985-1002.

8. R. L. Frkic, B. S. Chua, Y. Shin, B. D. Pascal, S. J. Novick, T. M. Kamenecka, P. R. Griffin and J. B. Bruning, Structural and Dynamic Elucidation of a Non-acid PPARγ Partial Agonist: SR1988, *Nuclear Receptors Reseach*, 2018, **5**, 101350.

9. J. L. Pederick, A. P. Thompson, S. G. Bell and J. B. Bruning, d-Alanine-d-alanine ligase as a model for the activation of ATP-grasp enzymes by monovalent cations, *J. Biol. Chem.*, 2020, **295**, 7894-7904.

10. D. Aragao, J. Aishima, H. Cherukuvada, R. Clarken, M. Clift, N. P. Cowieson, D. J. Ericsson, C. L. Gee, S. Macedo, N. Mudie, S. Panjikar, J. R. Price, A. Riboldi-Tunnicliffe, R. Rostan, R. Williamson and T. T. Caradoc-Davies, MX2: a high-flux undulator microfocus beamline serving both the chemical and macromolecular crystallography communities at the Australian Synchrotron, *J. Synchrotron Rad.*, 2018, **25**, 885-891.

11. N. P. Cowieson, D. Aragao, M. Clift, D. J. Ericsson, C. Gee, S. J. Harrop, N. Mudie, S. Panjikar, J. R. Price, A. Riboldi-Tunnicliffe, R. Williamson and T. Caradoc-Davies, MX1: a bending-magnet crystallography beamline serving both chemical and macromolecular crystallography communities at the Australian Synchrotron, *J. Synchrotron Rad.*, 2015, **22**, 187-190.

12. W. Kabsch, XDS (X-Ray Detector Software), *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2010, **66**, 125-132.

13. P. Evans, Scaling and assessment of data quality, *Acta Crystallogr D Biol Crystallogr*, 2006, **62**, 72-82.

14. A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storonia and R. J. Reada, Phaser crystallographic software, *J. Appl. Crystallogr.*, 2007, **40**, 658-674.

15. J. M. Gulbis, Z. Kelman, J. Hurwitz, M. O'Donnell and J. Kuriyan, Structure of the C-Terminal Region of p21 WAF1/CIP1 Complexed with Human PCNA, *Cell*, 1996, **87**, 297-306.

16. D. Liebschner, P. V. Afonine, M. L. Baker, G. Bunkóczi, V. B. Chen, T. I. Croll, B. Hintze, L. W. Hung, S. Jain, A. J. McCoy, N. W. Moriarty, R. D. Oeffner, B. K. Poon, M. G. Prisant, R. J. Read, J. S. Richardson, D. C. Richardson, M. D. Sammito, O. V. Sobolev, D. H. Stockwell, T. C. Terwilliger, A. G. Urzhumtsev, L. L. Videau, C. J. Williams and P. D. Adams, Phenix : Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2019, **75**, 861-877.

17. P. V. Afonine, R. W. Grosse-Kunstleve, N. Echols, J. J. Headd, N. W. Moriarty, M. Mustyakimov, T. C. Terwilliger, A. Urzhumtsev, P. H. Zwart and P. D. Adams, Towards automated crystallographic structure refinement with phenix.refine, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2012, **68**, 352-367.

18. P. Emsley and K. Cowtan, Coot: model-building tools for molecular graphics, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2004, **60**, 2126-2132.

19. N. W. Moriarty, R. W. Grosse-Kunstleve and P. D. Adams, electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation, *Acta Crystallogr D Biol Crystallogr*, 2009, **65**, 1074-1080.

20. R. Abagyan, M. Totrov and D. Kuznetsov, ICM - New Method for Protein Modeling and Design: Applications to Docking and Structure Prediction from the Distorted Native Conformation, *J. Comput. Chem.*, 1994, **15**, 488-506.

21. R. Abagyan and M. Totrov, Biased Probability Monte Carlo Conformational Searches and Electrostatic Calculations for Peptides and Proteins, *J. Mol. Biol.*, 1994, **235**, 983-1002.

22. Schrodinger, LLC, unpublished work.

23. D. Piovesan, G. Minervini and S. C. Tosatto, The RING 2.0 web server for high quality residue interaction networks, *Nucleic Acids Res.*, 2016, **44**, W367-374.

24. K. Stierand, P. C. Maass and M. Rarey, Molecular complexes at a glance: automated generation of two-dimensional complex diagrams, *Bioinformatics*, 2006, **22**, 1710-1716.

25. P. V. Afonine, N. W. Moriarty, M. Mustyakimov, O. V. Sobolev, T. C. Terwilliger, D. Turk, A. Urzhumtsev and P. D. Adams, FEM: feature-enhanced map, *Acta Crystallogr D Biol Crystallogr*, 2015, **71**, 646-666.

Chapter 7.

A NUCLEAR PERMEABLE PEPTIDOMIMETIC TO TARGET PCNA

Research Article: In preparation for Nature Chemical Biology **2021**

A nuclear permeable peptidomimetic to target PCNA

Aimee J. Horsfall,[1-3] Theresa Chav,[1-3] Zoya Kikhtyak,[4] Jordan L. Pederick,[1,5] Wioleta Kowalcyzk,[6] Denis B. Scanlon, [1-2] Wayne D. Tilley,[4] Theresa E. Hickey,[4] Andrew D. Abell,[1-3] John B. Bruning[1,5]*

[1] Institute of Photonics and Advanced Sensing (IPAS), The University of Adelaide, Adelaide, South Australia, 5005, Australia

[2] School of Physical Sciences, The University of Adelaide, Adelaide, South Australia 5005, Australia

[3] Australian Research Council Centre of Excellence for Nanoscale BioPhotonics (CNBP)

[4] Dame Roma Mitchell Cancer Research Laboratories, Adelaide Medical School, Faculty of Health and Medical Sciences, University of Adelaide, Adelaide, South Australia 5005, Australia

[5] School of Biological Sciences, The University of Adelaide, Adelaide, South Australia 5005, Australia

[6] CSIRO Manufacturing, Clayton, Victoria 3168, Australia

*Corresponding author

ABSTRACT

Human proliferating cell nuclear antigen (PCNA) mediates DNA replication and repair, therefore inhibition of PCNA interactions presents as an ideal target to shut down DNA-replication in order to develop novel cancer treatments. Here, a short p21-derived peptide is used as a starting point and reveals a modular nuclear permeable peptidomimetic to target PCNA in breast cancer cells. We demonstrate that a fluorescent p21 macrocycle provides the smallest optimal scaffold to permit cellular entry. Additionally, we determine the optimal NLS sequence to confer nuclear uptake is derived from SV40. Lastly, we show that attaching a fluorescein-tag dramatically alters the cellular distribution of the p21 peptidomimetics. This study has identified an inherently fluorescent, nuclear permeable p21-based peptidomimetic scaffold, and thus provides a significant advance toward achieving a peptide-based therapeutic to inhibit PCNA.

STATEMENT OF AUTHORSHIP

Title of Paper	A nuclear permeable peptidomimetic to inhibit human PCNA
Publication Status	☐ Published ☐ Accepted for Publication ☐ Submitted for Publication ☒ Unpublished and Unsubmitted work written in manuscript style
Publication Details	**Research Article:** 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

Principal Author

Name of Principal Author (Candidate)	Aimee J Horsfall		
Contribution to the Paper	Designed and synthesised peptides, SPR assay, analysed & discussed results, wrote and edited manuscript		
Overall percentage (%)	50%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	15/02/2021

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);
ii. permission is granted for the candidate in include the publication in the thesis; and
iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Theresa Chav		
Contribution to the Paper	Designed and synthesised peptides, SPR assay, analysed & discussed results, and edited manuscript		
Signature		Date	

Name of Co-Author	Zoya Kikhtyak		
Contribution to the Paper	Cell uptake assays, discussed results, and edited manuscript.		
Signature		Date	19.02.2021

Name of Co-Author	Jordan L Pederick		
Contribution to the Paper	SPR assays, discussed results, and edited manuscript.		
Signature		Date	26/02/2021

Name of Co-Author	Wioleta Kowalcyzk		
Contribution to the Paper	Peptide synthesis, discussed results, and edited manuscript.		
Signature		Date	01/03/2021

Name of Co-Author	Denis B Scanlon		
Contribution to the Paper	Peptide synthesis, discussed results, and edited manuscript.		
Signature		Date	01/03/2021

Name of Co-Author	Theresa E Hickey		
Contribution to the Paper	Supervised cell uptake assays, discussed results, and edited manuscript.		
Signature		Date	19/02/21

Name of Co-Author	Wayne D Tilley		
Contribution to the Paper	Supervised cell uptake assays, discussed results, and edited manuscript.		
Signature		Date	19.02.21.

Name of Co-Author	Andrew D Abell		
Contribution to the Paper	Supervised AJH & TC, discussed results, and edited manuscript.		
Signature		Date	15/2/2021

Name of Co-Author	John B Bruning		
Contribution to the Paper	Supervised AJH, TC & JLP, discussed results, and edited manuscript.		
Signature		Date 18-2-21	

7.1 Introduction

Human proliferating cell nuclear antigen (PCNA) is a member of the sliding clamp family, an essential processivity factor, and mediator of DNA-replication, DNA-repair and cell cycle processes.[1-8] Its ring-shaped structure allows it to be loaded onto template-primer junctions to encircle double-stranded DNA, where it acts as a mobile docking platform for protein machinery, essential to DNA-replication and repair, to first associate and subsequently interact with the DNA.[2, 9-11] PCNA is consequently upregulated in a large number of cancers in response to an increased demand for DNA-replication. As such, PCNA is a key target to develop anti-cancer therapeutics.[7, 12-16]

PCNA is primarily located in the nucleus and is known to recruit over 200 proteins that carry out DNA-replication and repair, and interact with PCNA through one of three degenerate binding sites on the trimeric PCNA (Figure 1).[3, 5, 7, 11, 12, 15, 17] Each monomer is comprised of two almost symmetrical domains that are connected through an unstructured loop, termed the interdomain connecting loop (IDCL). Many proteins bind a hydrophobic cleft on the PCNA surface that is nestled between the two domains, termed the PCNA-interacting protein (PIP) binding site.[1, 18, 19] These proteins bind at the PIP binding site on PCNA through a sequence known as the PIP-box motif, defined as Qxxφxxψψ where Q is glutamine, ‘x’ is any

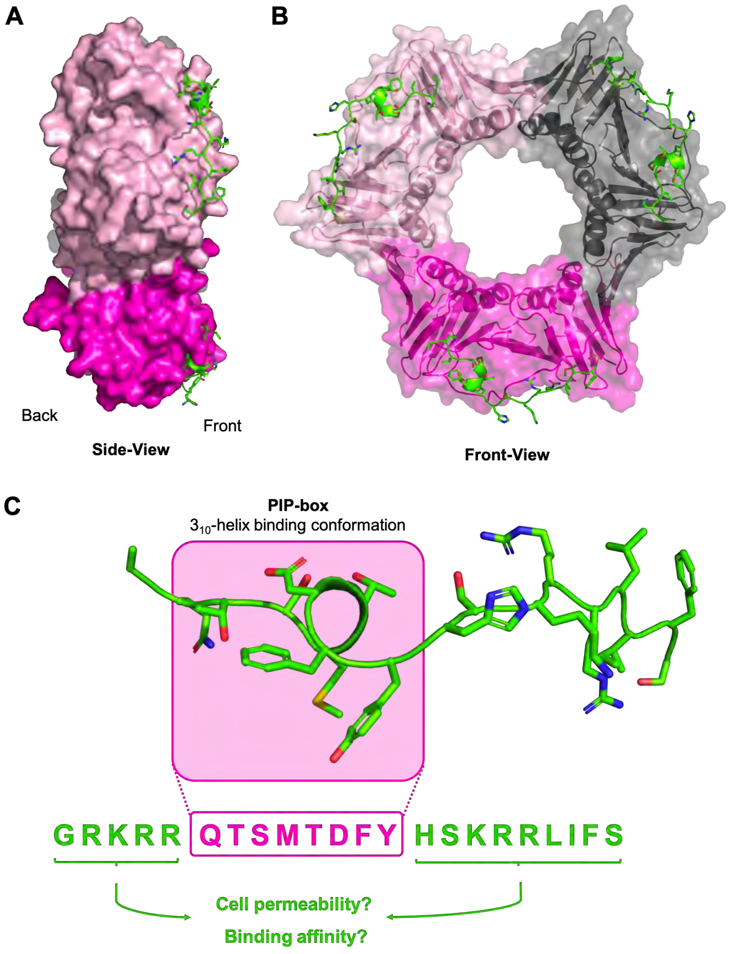


Figure 1: PCNA Structure (PDB: 1AXC) **A** Surface representation of PCNA (pink) as a side-view with $p21_{139\text{-}160}$ bound to the front face in green. **B** Front view of PCNA as a transparent surface representation, and cartoon representation that shows the three monomers (light pink, bright pink, grey) and $p21_{139\text{-}160}$ in green cartoon representation with side-chains as sticks, bound to each monomer. **C** $p21_{139\text{-}160}$ in green bound in cartoon representation with side-chains as sticks, and PIP-box highlighted in pink.

amino-acid, φ is a hydrophobic residue and ψ an aromatic residue (commonly phenylalanine or tyrosine).[19-25] One such example is the cell cycle regulator protein, p21, that contains the PIP-box QTSMTDFY and binds PCNA with the highest known affinity and competitively blocks other proteins from interacting with PCNA (Figure 1).[1, 19, 26-31] Consequently, PCNA-bound p21 stalls DNA-replication to provide a necessary checkpoint for cell cycle regulation. A 22mer peptide derived from p21 (residues 139-160) similarly inhibits polymerase δ synthesis to shut down SV40 DNA replication.[30] As such, $p21_{139\text{-}160}$ provides an ideal template to design a PCNA inhibitor for application as a cancer therapeutic.

Efforts to develop a therapeutic derived from a p21 peptide are limited by a lack of nuclear permeation. The expressed p21 protein itself is translocated to the nucleus,[1] as are similarly large expressed fragments of the protein (residues 76-164).[32, 33] It has been reported that residues 140-142 (RKR) of p21 are essential to this nuclear permeability.[33] However, smaller p21 peptides (141-160 or 139-164) that contain these residues (140-142) are not nuclear permeable, and have instead been conjugated to long nuclear locating sequence (NLS) peptides, such as Penetratin, to address the limitation of nuclear impermeability in order to investigate p21 peptide activity in cells (e.g. DNA-replication inhibition, PCNA interaction etc).[34-37] Such peptide sequences are all greater than 30 amino-acids long and here we endeavour to define a short p21-derived peptide sequence that targets PCNA and is nuclear permeable.

Here we first characterise the impact of sequentially truncating one amino-acid from either the *N*- or *C*-terminal of $p21_{139\text{-}160}$ on the binding affinity for PCNA, as determined by surface plasmon resonance (SPR), in order to determine the importance of each section of the p21 sequence on binding affinity. This then informs the design of short p21 peptides that maintain high affinity for PCNA, which is an important step towards determining the smallest functional peptide scaffold that targets PCNA. The cell and nuclear permeability of fluorescein-tagged variants of these short peptides are then assessed in a breast cancer cell line. A series of NLS-tags are then conjugated to a short linear p21 peptide, and a small macrocyclic p21 peptide, in order to achieve nuclear permeability of the conjugates in breast cancer cells. This study not only characterises the role of the sequences that flank the p21 PIP-box on PCNA binding affinity and cell permeability, but importantly identifies a small p21-derived peptidomimetic sequence that targets PCNA and is nuclear permeable.

7.2 Results

The peptide $p21_{139\text{-}160}$ is known to bind PCNA with high affinity (~5 nM)[1, 15, 19, 30] and provides a useful template to develop a PCNA inhibitor for application as a cancer therapeutic. All peptides were prepared by solid-phase peptide synthesis as detailed in the Experimental section. First, we focussed on characterising the minimum p21 sequence required to maintain high affinity binding to PCNA, as determined by SPR. Here we characterised the upper limit of 'high affinity binding' to be ~100 nM, where the mutant PL peptide binds with 100 nM and is reported to inhibit SV40 DNA-replication in HeLa cell extracts.[30, 38] Second, the role of the sequence that flanks the PIP-box on cell permeability was investigated in breast cancer cells. Lastly, nuclear permeability was imparted to short p21 peptides by conjugation of small NLS sequences, and the cellular distribution assessed in breast cancer cells by confocal microscopy.

7.2.1 Identifying the shortest sequence that maintains high affinity PCNA binding

A series of 15 p21 peptides (see Table 1) each with one additional amino-acid truncation from either the *N*- or *C*-terminus of the $p21_{139-160}$ sequence (**1-0**), up until the PIP-box (144-151), was prepared. The binding affinity of each peptide for PCNA was determined by SPR to identify the shortest sequence able to maintain high affinity binding to PCNA (Table 1). The binding affinity of **1-0** for PCNA was 5.96 nM, in good agreement with literature reports.[1, 15, 19, 30] The binding affinity of *C*-terminally truncated peptides **1-1** and **1-2** for PCNA was similar to **1-0** with K_D values of 12.4 nM and 6.70 nM. This indicates that the two most *C*-terminal residues were not important to maintain PCNA affinity. Not unexpectedly, the affinity for PCNA gradually diminished from peptide **1-1** at 12.4 nM, through to **1-6** at 37.4 nM, each which possessed one less *C*-terminal residue. There affinity of peptide **1-7** for PCNA, which lacked the seven *C*-terminal residues seen in **1-0**, was notably lower (115 nM), however the affinity remained relatively constant with any further *C*- terminal truncation (**1-8**, 98.7 nM and **1-9**, 83.0 nM). These data together suggest that at least the three residues *C*-terminal of the PIP-box (as in **1-6**) are required to maintain relatively high affinity for PCNA.

Peptides **1-10**, **1-11**, **1-12**, lacked either 1, 2, or 3 *N*-terminal residues, respectively, relative to **1-0** and maintained low nanomolar binding affinity (11.8, 9.24 and 8.05 nM, respectively). Removing the four *N*-terminal residues of **1-0**, as in **1-13**, decreased the affinity for PCNA to 31.5 nM, whereas removal of all five residues *N*-terminal to the PIP-box (**1-14**) resulted in a severe 25-fold diminishment in affinity for PCNA to 827 nM. This indicates that at least a single residue *N*-terminal to the PIP-box must be included to maintain high affinity binding to PCNA.

Consequently, a short p21 peptide which contained both a *N*-terminal and *C*-terminal truncation was prepared, where only one amino-acid *N*-terminal to the PIP-box (143, see **1-13** c.f **1-14**) was included; and three residues *C*-terminal of the PIP-box were included (154, see **1-6** c.f **1-7**). Peptide **2-0**, which spans residues 143-154 of p21, was prepared and the binding affinity for PCNA determined to be 102.3 nM. Hence, a short p21 peptide that maintains high affinity binding to PCNA was rationally designed.

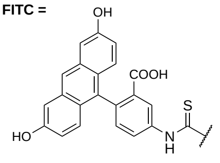
7.2.2 The influence of positive charge on cell uptake of p21 peptides

Nα-Fluorescein-tagged derivatives of **1-0** and **2-0**, designated **1-0F** and **2-0F**, were prepared and subjected to a cell uptake assay in order to determine whether these peptides are cell permeable. Breast cancer cells (T-47D), were treated with 5 μM of **2-0F** or **1-0F**, and incubated for 48 h. The cells were then fixed, stained with phalloidin and imaged by fluorescence microscopy (Figure S2). Bright green fluorescence, corresponding to the fluorescein-tagged peptide, was observed in the cytoplasm of the **1-0F** treated cells indicating this peptide was cell permeable. This fluorescence was punctate which suggests it is likely endosomally trapped.[39-41] In contrast, no green fluorescence was observed in the **2-0F** treated cells, indicating this peptide was not cell permeable. **2-0F** lacks a large number of positively charged residues relative to **1-0F**, where it is known that positively charged residues such as Arg and Lys often contribute to improved cell uptake.[42-44] Consequently, a second series of three peptides were prepared (**3-1**, **3-2** and **3-3**, Table 1) where the positively charged residues that flank the PIP-box sequence were re-introduced,

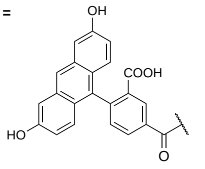
Table 1: Peptide sequences and SPR binding affinity, SE is standard error. The PIP-box binding motif is indicated in bold.

Peptide	Sequence	Affinity K_D (nM)	K_D SE (nM)
1-0	GRKRR**QTSMTDFY**HSKRRLIFS	5.96	2.07
1-1	GRKRR**QTSMTDFY**HSKRRLIF	12.4	3.40
1-2	GRKRR**QTSMTDFY**HSKRRLI	6.70	1.26
1-3	GRKRR**QTSMTDFY**HSKRRL	21.3	2.86
1-4	GRKRRQ**TSMTDFY**HSKRR	18.7	1.49
1-5	GRKRR**QTSMTDFY**HSKR	19.9	1.38
1-6	GRKRR**QTSMTDFY**HSK	37.4	2.24
1-7	GRKRR**QTSMTDFY**HS	115	9.03
1-8	GRKRR**QTSMTDFY**H	98.7	6.42
1-9	GRKRR**QTSMTDFY**	83.0	5.13
1-10	RKRR**QTSMTDFY**HSKRRLIFS	11.8	3.18
1-11	KRR**QTSMTDFY**HSKRRLIFS	9.24	2.48
1-12	RR**QTSMTDFY**HSKRRLIFS	8.05	1.76
1-13	R**QTSMTDFY**HSKRRLIFS	31.5	3.56
1-14	**QTSMTDFY**HSKRRLIFS	827	50.8
2-0	R**QTSMTDFY**HSK	102	5.3
1-0F	Fl-GRKRR**QTSMTDFY**HSKRRLIFS		
2-0F	FITC-A_βR**QTSMTDFY**HSK		
3-1	RKRR**QTSMTDFY**HSK	41.8	5.80
3-2	R**QTSMTDFY**HSKRR	45.9	4.30
3-3	RKRR**QTSMTDFY**HSKRR	38.0	10.0
3-1F	FITC-A_βRKRR**QTSMTDFY**HSK	202	18.0
3-2F	FITC-A_βR**QTSMTDFY**HSKRR	190	8.7
3-2F	FITC-A_βRKRR**QTSMTDFY**HSKRR	61.0	5.6
2-0a	mal-R**QTSMTDFY**HSK		
4-1F	FITC-A_βPKKKRKVC		
4-2F	FITC-A_βPAAKRVKLDC		
4-3F	FITC-A_βGRKKRRQRRRC		
4-4F	FITC-A_βRRWWRRWRRC		
4-1a	H-PKKKRKVC		
5-1F	FITC-A_βPKKKRKVC(suc-R**QTSMTDFY**HSK)	33.1	4.20
5-2F	FITC-A_βPAAKRVKLDC(suc-R**QTSMTDFY**HSK)	130	15.0
5-3F	FITC-A_βGRKKRRQRRRC(suc-R**QTSMTDFY**HSK)	101	9.50
5-4F	FITC-A_βRRWWRRWRRC(suc-R**QTSMTDFY**HSK)	373	30.0
6-0	[cyclo-5,9]-Ac-R**QC(-)SMTC(Bim-)FY**HSK	570	30.0
6-0a	[cyclo-5,9]-mal-R**QC(-)SMTC(Bim-)FY**HSK		
6-0F	[cyclo-6,10]-FITC-A_βR**QC(-)SMTC(Bim-)FY**HSK	25190	1900
6-1F	FITC-A_βPKKKRKVC([cyclo-5,9] suc-R**QC(-)SMTC(Bim)FY**HSK)	56	73.0
6-2F	FITC-A_βPAAKRVKLDC([cyclo-5,9] suc-R**QC(-)SMTC(Bim)FY**HSK)		
6-3F	FITC-A_βGRKKRRQRRRC([cyclo-5,9] suc-R**QC(-)SMTC(Bim)FY**HSK)		
6-4F	FITC-A_βRRWWRRWRRC([cyclo-5,9] suc-R**QC(-)SMTC(Bim)FY**HSK)		
6-1	H-PKKKRKVC([cyclo-5,9]suc-R**QC(-)SMTC(Bim)FY**HSK)	176	27.0

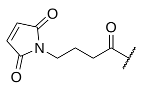
FITC =



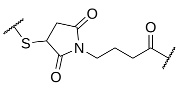
Fl =



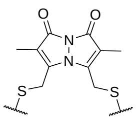
mal =



suc =



Bim =



to determine if this positive-charge could reinstate cell permeable character seen for **1-0F**. The binding affinity for PCNA of these peptides was determined by SPR and revealed a K_D value of ~40 nM for all three peptides (Table 1). Nα-Fluorescein-tagged analogues of these peptides were prepared (**3-1F**, **3-2F** and **3-3F**) and subjected to the same cell permeability assay. Fluorescence microscopy of the resulting **3-1F** or **3-2F** treated cells indicated there was no discernible green fluorescence present within the cells corresponding to the fluorescein-tagged peptides, indicating these peptides were not cell permeable (Figure S2). A small amount of green fluorescence was observed in the cell cytoplasm of the **3-3F** treated cells, indicating some of that this peptide, which contains five additional positively charged residues relative to **2-0F**, was able to enter the cells (Figure S2). This cell permeability is modest, especially in comparison to the uptake of **1-0F**, and suggests that the cell uptake of **1-0F** is not solely reliant on the positively charged residues and is improved in the presence of the hydrophobic LIFS (157-160) tail of **1-0F**.

7.2.3 Introducing an NLS to increase cell uptake

Next, it was investigated whether appending an NLS-tag to the short p21 peptide **2-0** could confer both cell and nuclear permeability. The shorter **2-0** was chosen, as the longer **1-0F** was not nuclear permeable, and our goal was to identify the smallest nuclear permeable scaffold. Four peptide sequences derived from SV40, cMyc and TAT, and a synthetic sequence termed R6W6, are reported to be capable of permitting cell and nuclear entry to a covalently tethered cargo. These sequences were chosen as they vary in length, origin and overall character. Peptides SV40, cMyc, TAT and R6W3, each with an additional cysteine included at the *C*-terminus of the sequence, and an Nα-fluorescein, were prepared to give **4-1F**, **4-2F**, **4-3F**, **4-4F** respectively (Table 1). An analogue of **2-0** was also prepared, where an *N*-alkyl maleimide was covalently attached to the *N*-terminus of the peptide to give **2-0a**. Peptide **2-0a** was then reacted in water with **4-1F**, **4-2F**, **4-3F**, or **4-4F**, to give NLS-p21 conjugates **5-1F**, **5-2F**, **5-3F** and **5-4F**, respectively. Modified MDA-MB-468 breast cancer cells that express a nuclear mKate fluorophore were treated with 10 µM of **1-0F**, **2-0F**, **5-1F**, **5-2F**, **5-3F** or **5-4F**, to determine if these peptides were cell and nuclear permeable. The treated cells were then fixed and imaged by confocal microscopy (Figure 2). As before, green punctate fluorescence corresponding to the fluorescein-tagged peptide was evident in the **1-0F** treated cells, confirming **1-0F** is cell permeable. No green fluorescence was observed for the **2-0F** treated cells, indicating the short p21 peptide was not cell permeable, consistent with the T47D cell line. The **5-1F** treated cells displayed bright green fluorescence throughout the cytoplasm, which also appears to colocalise with the red mKate fluorescence indicating the nuclei (Figure 2). This suggests that **5-1F** is both cell and nuclear permeable. **5-2F** and **5-3F** were modestly cell permeable, with a small amount of green fluorescence evident in the **5-2F** and **5-3F** treated cells (Figure 2). The fluorescence images of the **5-4F** treated cells indicated bright green fluorescence, however the lack of red fluorescence corresponding to mKate expression in the nucleus is indicative of lysed or dead cells, suggesting **5-4F** at 10 µM is cytotoxic.

7.2.4 Macrocyclic p21 peptides appended to NLS sequences

We recently reported a short macrocyclic p21 peptide, **6-0**, that is inherently fluorescent and cell permeable in a breast cancer cell line.[45] Given this peptide is already cell permeable, we envisaged that tethering it to an NLS should further enhance cell and nuclear uptake of the conjugate, compared to the **2-0** conjugates (**5-1F** – **5-4F**). An analogue of **6-0** with an *N*-alkyl maleimide covalently attached to the peptide *N*-terminus (**6-0a**) was prepared and conjugated to the four NLS-tags (**4-1F**, **4-2F**, **4-3F**, or **4-4F**) as before, to give **6-1F**, **6-2F**, **6-3F** and **6-4F**. MDA-MB-468 cells were treated with 10 μM of **6-0**, **6-1F**, **6-2F**, **6-3F** or **6-4F**, incubated for 24 h, then fixed and imaged by confocal microscopy (Figure 2). The **6-0** treated cells displayed blue punctate fluorescence within the cell cytoplasm owing to the bimane fluorophore embedded in the macrocycle (Figure S3), consistent with our previous report that **6-0** is cell permeable. An Nα fluorescein-tagged derivative of the p21 macrocycle, without an NLS, **6-0F** was also prepared and imaged (Figure 2).

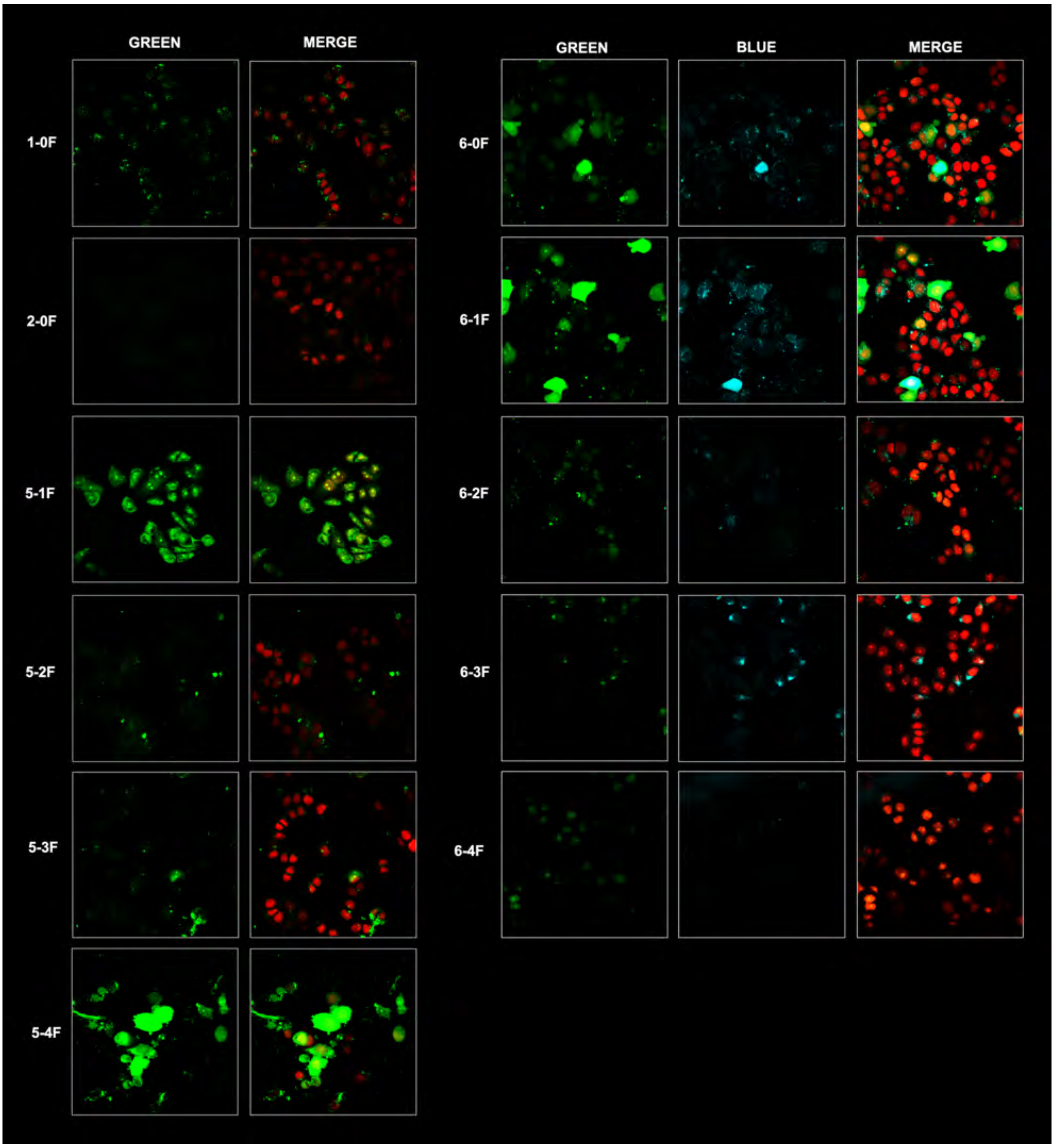


Figure 2: MDA-MB-468 breast cancer cells treated with 10 μM of peptide for 24 h. Green represents fluorescein, blue indicates the bimane fluorophore, red indicates the mKate fluorophore expressed at the cell nucleus. Merged images: all relevant channels overlaid.

Again, blue punctate fluorescence was apparent in the cell cytosol, which largely colocalises with the green fluorescein fluorescence. The green fluorescence is however more diffuse which may suggest some of the fluorescein has been liberated from the peptide. Peptide **6-0F** thus is also cell permeable, but not nuclear permeable. The **6-1F** treated cells, and to a lesser extent the **6-2F** treated cells, display blue fluorescence within the cell cytoplasm that colocalises with the green fluorescence of the fluorescein appended to the NLS-tag. The blue and green in the **6-1F** treated cells is a combination of diffuse and punctate, suggesting some of the peptide is endosomally trapped. Additionally, some of this signal overlaps with the red fluorescence of the nuclear mKate indicating some of the peptide has translocated to the nucleus. Small pockets of well colocalised green and blue fluorescence is apparent in the **6-3F** treated cells. The **6-4F** treated cells were administered a lower concentration of **6-4F** owing to the apparent cytotoxicity observed for the **5-4F** treated cells, and only faint green fluorescence is notable indicating these peptides are not significantly cell permeable.

An analogue of the cell and nuclear permeable peptide **6-1F** was prepared to investigate if the fluorescein-tag influences the cell and nuclear permeability profile, where such effects have been reported for a range of auxiliary tags. **6-1** contains the bimane-constrained p21 peptide appended to an SV40 NLS-sequence but no fluorescein. The **6-1** treated cells were imaged by confocal microscopy to reveal blue fluorescence corresponding to the **6-1** peptide, in the cell cytoplasm, but not the cell nucleus (Figure S3). This indicates that **6-1** is not nuclear permeable, in contrast to **6-1F**, and suggests the fluorescein-tag may influence cell uptake and distribution.

7.3 Discussion

A viable pre-clinical peptide-based cancer therapeutic, that inhibits PCNA and hence DNA-replication and -repair, must not only have high affinity for PCNA but must also be cell and nuclear permeable. Here we define the minimum p21 peptide scaffold required to enter the nucleus and target PCNA. We demonstrate, through a series of 15 peptides with one amino-acid sequentially truncated from either the *N*- or *C*-terminus of the p21$_{139\text{-}160}$ (**1-0**) peptide, that truncation of the 6 *C*-terminal and 4 *N*-terminal residues did not significantly impact PCNA binding affinity. This is further supported by the peptide **2-0**, which consists of the p21 sequence 143-154, that binds PCNA with high affinity (102.3 nM). The slightly longer PL peptide is reported to bind PCNA with 100 nM affinity and inhibits SV40 replication in cell extracts.[30] Thus, we expect that the nuclear permeable peptides presented here, that bind with improved affinity for PCNA (relative to the PL peptide) will be able to inhibit DNA-replication in cells. These studies are currently being undertaken.

It is somewhat interesting that though **1-0F** is cell permeable, and the shorter variant **2-0F** is not, that reintroduction of the positively charged residues (Arg/Lys in **3-1F** – **3-3F**) did not recover the cellular uptake of the peptides. It is known that peptides with large positive charge character, in particular those with poly-Arg combinations, display better cell uptake. Additionally, residues 140-142 have been shown important for the nuclear localisation of the expressed p21 protein.[32] However, from this combination of peptides it is clear that the cell permeability of **1-0F** cannot be attributed solely to the presence of the positively charged amino-acids. The cell uptake of **1-0F** peptide may be owed, in part, to the hydrophobic tail (^{157}LIFS) where a number of amphipathic peptides display good cell uptake.[43, 46, 47] Further, hydrophobic groups may associate with the cell membrane and help facilitate passive diffusion.

The SV40-derived tag improved cell and nuclear uptake of both the linear (**5-1F**) and macrocyclic (**6-1F**) p21 peptides, which suggests it is the optimal choice for further development of a p21-derived peptidomimetic to target PCNA. The cMyc (**5-2F** and **6-2F**) and TAT (**5-3F** and **6-3F**) tagged peptides did not display significant uptake, whereas the R6W3 tagged peptides **5-4F** and **6-4F** showed a clear concentration-dependent cytotoxicity. The origin of this cytotoxicity (e.g. DNA-replication inhibition or cell membrane disruption) is under investigation.

The SV40-tagged **6-1F** was revealed to be nuclear permeable, however an analogue which lacks the fluorescein-tag (**6-1**) was not cell permeable, suggesting that the fluorescein-tag has contributed to the nuclear permeability of **6-1F**. This observation highlights the impact that auxiliary fluorescent tags can have on the biophysical properties and behaviour of a peptide,[48-50] and consequently the need for bifunctional peptide modifications, or label free technologies, that allow direct analysis a compound of interest. This emphasises the benefit of introducing imaging tags that are inherently part of the peptide structure (such as the bimane modification), and not auxiliary tags (such as fluorescein), such that the compound subjected to assays (such as a cell permeability assay), is reflective of the true function of the compound that will go on to further testing. Future work will determine the ability of peptides **6-0F** and **6-1F** to inhibit DNA-replication in cancer cell lines, and additionally investigate whether the affinity of these short macrocyclic p21 peptides can be enhanced through sequence mutations.

7.4 Acknowledgements

The research was supported by the Australian Research Council Centre of Excellence in Nanoscale BioPhotonics (CNBP) (CE140100003). A.J.H. and J.L.P. are supported by Australian Government Research Training Program Stipends (RTPS). The facilities of the OptoFab node of the Australian National Fabrication Facility (ANFF) and associated Commonwealth and SA State Government funding are also gratefully acknowledged. This work was supported by grants from the National Health and Medical Research Council of Australia (W.D.T., T.E.H., ID 1084416, ID 1130077;), the National Breast Cancer Foundation (NBCF; W.D.T.; ID PS-15-041) and a Movember & National Breast Cancer Foundation Collaboration Initiative grant (MNBCF-17-012 to W.D.T., T.E.H.). T.E.H. is currently supported by an NBCF Fellowship (IIRS-19-009).

7.5 Competing interests

The authors declare there are no conflicts of interest relating to the contents of this manuscript.

REFERENCES

1. J. M. Gulbis, Z. Kelman, J. Hurwitz, M. O'Donnell and J. Kuriyan, Structure of the C-Terminal Region of p21 WAF1/CIP1 Complexed with Human PCNA, *Cell*, 1996, **87**, 297-306.
2. T. Oku, S. Ikeda, H. Sasaki, K. Fukuda, H. Morioka, E. Ohtsuka, H. Yoshikawa and T. Tsurimoto, Functional sites of human PCNA which interact with p21 (Cip1/Waf1), DNA polymerase delta and replication factor C, *Genes Cells*, 1998, **3**, 357-369.
3. G. Maga and U. Hubscher, Proliferating cell nuclear antigen (PCNA): A dancer with many partners, *J. Cell Sci.*, 2003, **116**, 3051-3060.
4. K. M. Gilljam, E. Feyzi, P. A. Aas, M. M. Sousa, R. Muller, C. B. Vågbø, T. C. Catterall, N. B. Liabakk, G. Slupphaug, F. Drabløs, H. E. Krokan and M. Otterlei, Identification of a novel, widespread, and functionally important PCNA-binding motif, *J. Cell Biol.*, 2009, **186**, 645-654.
5. E. M. Boehm, M. S. Gildenberg and M. T. Washington, The Many Roles of PCNA in Eukaryotic DNA Replication, *Enzymes*, 2016, **39**, 231-254.
6. K. T. Powers and M. T. Washington, Eukaryotic translesion synthesis: Choosing the right tool for the job, *DNA Repair*, 2018, **71**, 127-134.
7. A. Gonzalez-Magana and F. J. Blanco, Human PCNA Structure, Function and Interactions, *Biomolecules*, 2020, **10**, 570.
8. J. R. Pohler, M. Otterlei and E. Warbrick, An in vivo analysis of the localisation and interactions of human p66 DNA polymerase delta subunit, *BMC Molecular Biology*, 2005, **6**, 17.
9. J. Majka and P. M. J. Burgers, The PCNA–RFC Families of DNA Clamps and Clamp Loaders, *Prog. Nucleic Acid Res. Mol. Biol.*, 2004, **78**, 227-260.
10. M. De March, N. Merino, S. Barrera-Vilarmau, R. Crehuet, S. Onesti, F. J. Blanco and A. De Biasio, Structural basis of human PCNA sliding on DNA, *Nat. Comm.*, 2017, **8**, 13935.
11. G. L. Moldovan, B. Pfander and S. Jentsch, PCNA, the maestro of the replication fork, *Cell*, 2007, **129**, 665-679.
12. I. Stoimenov and T. Helleday, PCNA on the crossroad of cancer, *Biochem. Soc. Trans.*, 2009, **37**, 605-613.
13. D. Zhongyun, M. Wortman, Z. Tan and K. Dillehay, WO 2012/033938 A2, *Identification of PCNA Targeting Compounds for Cancer Therapy and PCNA Function Regulation,* 2012.
14. S. Y. Park, M. S. Jeong, C. W. Han, H. S. Yu and S. B. Jang, Structural and Functional Insight into Proliferating Cell Nuclear Antigen, *J. Microbiol. Biotechnol.*, 2016, **26**, 637-647.
15. A. J. Horsfall, A. D. Abell and J. B. Bruning, Targeting PCNA with Peptide Mimetics for Therapeutic Purposes, *ChemBioChem*, 2019, **21**, 442-450.
16. C. K. Søgaard, A. Nepal, V. Petrovic, A. Sharma, N.-B. Liabakk, T. S. Steigedal and M. Otterlei, Targeting the non-canonical roles of PCNA modifies and increases the response to targeted anti-cancer therapy, *Oncotarget*, 2019, **10**, 7185-7197.
17. A. De Biasio and F. J. Blanco, Proliferating cell nuclear antigen structure and interactions: too many partners for one dancer?, *Adv. Prot. Chem. Struct. Bio.*, 2013, **91**, 1-36.
18. T. S. R. Krishna, X.-P. Kong, S. Gary, P. M. Burgers and J. Kuriyan, Crystal Structure of the Eukaryotic DNA Polymerase Processtivity Factor PCNA, *Cell*, 1994, **79**, 1233-1243.
19. J. B. Bruning and Y. Shamoo, Structural and thermodynamic analysis of human PCNA with peptides derived from DNA polymerase-delta p66 subunit and flap endonuclease-1, *Structure*, 2004, **12**, 2209-2219.
20. E. Warbrick, PCNA binding through a conserved motif, *BioEssays*, 1998, **20**, 195-199.
21. A. Hishiki, H. Hashimoto, T. Hanafusa, K. Kamei, E. Ohashi, T. Shimizu, H. Ohmori and M. Sato, Structural basis for novel interactions between human translesion synthesis polymerases and proliferating cell nuclear antigen, *J. Biol. Chem.*, 2009, **284**, 10552-10560.
22. C. Punchihewa, A. Inoue, A. Hishiki, Y. Fujikawa, M. Connelly, B. Evison, Y. Shao, R. Heath, I. Kuraoka, P. Rodrigues, H. Hashimoto, M. Kawanishi, M. Sato, T. Yagi and N. Fujii, Identification of small molecule proliferating cell nuclear antigen (PCNA) inhibitor that disrupts interactions with PIP-box proteins and inhibits DNA replication, *J. Biol. Chem.*, 2012, **287**, 14289-14300.
23. E. M. Boehm and M. T. Washington, R.I.P. to the PIP: PCNA-binding motif no longer considered specific, *BioEssays*, 2016, **38**, 1117-1122.
24. A. Prestel, N. Wichmann, J. M. Martins, R. Marabini, N. Kassem, S. S. Broendum, M. Otterlei, O. Nielsen, M. Willemoes, M. Ploug, W. Boomsma and B. B. Kragelund, The PCNA interaction motifs revisited: thinking outside the PIP-box, *Cell. Mol. Life Sci.*, 2019, **76**, 4923–4943.
25. A. J. Horsfall, B. A. Vandborg, W. Kowalczyk, T. Chav, D. B. Scanlon, A. D. Abell and J. B. Bruning, Unlocking the PIP-box: A peptide library reveals interactions that drive high affinity binding to human PCNA, *J. Biol. Chem.*, 2021, **296** 100773.
26. A. J. Kroker and J. B. Bruning, p21 Exploits Residue Tyr151 as a Tether for High-Affinity PCNA Binding, *Biochemistry*, 2015, **54**, 3483-3493.
27. A. C. Marshall, A. J. Kroker, L. A. Murray, K. Gronthos, H. Rajapaksha, K. L. Wegener and J. B. Bruning, Structure of the sliding clamp from the fungal pathogen Aspergillus fumigatus (AfumPCNA) and interactions with Human p21, *FEBS J.*, 2017, **284**, 985-1002.
28. S. Waga, G. J. Hannon, D. Breach and B. Stillman, The p21 inhibitor of cyclin-dependent kinases controls DNA replication by interaction with PCNA, *Nature*, 1994, **369**, 574-578.

29. J. Chen, P. K. Jackson, M. W. Kirschner and A. Dutta, Separate domains of p21 involved in the inhibition of Cdk kinase and PCNA, *Nature*, 1995, **374**, 386-388.

30. D. I. Zheleva, N. Z. Zhelev, P. M. Fischer, S. V. Duff, E. Warbrick, D. G. Blake and D. P. Lane, A Quantitative Study of the in Vitro Binding of the C-Terminal Domain of p21 to PCNA: Affinity, Stoichiometry, and Thermodynamics, *Biochemistry*, 2000, **39**, 7388-7397.

31. C. Prives and V. Gottifredi, The p21 and PCNA partnership: a new twist for an old plot, *Cell Cycle*, 2008, **7**, 3840-3846.

32. Y. Luo, J. Hurwitz and J. Massague, Cell-cycle inhibition by independent CDK and PCNA binding domains in p21Cip1, *Nature*, 1995, **375**, 159-161.

33. A. Rodr|guez-Vilarrupla, C. D|az, N. Canela, H.-P. Rahn, O. Bachs and N. Agell, Indentification of the nuclear localization signal of p21(cip1) and consequences of its mutation on cell proliferation, *FEBS J.*, 2002, **531**, 319-323.

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

35. I. Massodi, G. L. Bidwell and D. Raucher, Evaluation of cell penetrating peptides fused to elastin-like polypeptide for drug delivery, *J. Controlled Release*, 2005, **108**, 396-408.

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

37. K. L. Ball, S. Lain, R. Fahraeus, C. Smythe and D. P. Lane, Cell-cycle arrest and inhibitirion of Cdk4 activity by small peptides based on the carboxy-terminal domain of p21, *Current Biology*, 1996, **7**, 71-80.

38. G. Kontopidis, S.-Y. Wu, D. I. Zheleva, P. Taylor, C. McInnes, D. P. Lane, P. M. Fischer and M. D. Walkinshaw, Structural and biochemical studies of human proliferating cell nuclear antigen complexes provide a rationale for cyclin association and inhibitor design, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 1871-1876.

39. T. B. Potocky, A. K. Menon and S. H. Gellman, Cytoplasmic and nuclear delivery of a TAT-derived peptide and a beta-peptide after endocytic uptake into HeLa cells, *J. Biol. Chem.*, 2003, **278**, 50188-50194.

40. P. Ramoino, A. Diaspro, M. Fato and C. Usai, in *Molecular Regulation of Endocytosis*, 2012, DOI: 10.5772/46061, ch. Chapter 6.

41. A. J. Horsfall, K. R. Dunning, K. L. Keeling, D. B. Scanlon, K. L. Wegener and A. D. Abell, A bimane ‐ based peptide staple for combined helical induction and fluorescent imaging, *ChemBioChem*, 2020, **21**, 3423-3432.

42. S. Futaki, T. Suzuki, W. Ohashi, T. Yagami, S. Tanaka, K. Ueda and Y. Sugiura, Arginine-rich peptides. An abundant source of membrane-permeable peptides having potential as carriers for intracellular protein delivery, *J Biol Chem*, 2001, **276**, 5836-5840.

43. A. T. Jones and E. J. Sayers, Cell entry of cell penetrating peptides: tales of tails wagging dogs, *J Control Release*, 2012, **161**, 582-591.

44. N. Schmidt, A. Mishra, G. H. Lai and G. C. Wong, Arginine-rich cell-penetrating peptides, *FEBS Lett.*, 2010, **584**, 1806-1813.

45. A. J. Horsfall, B. A. Vandborg, Z. Kikhtyak, D. B. Scanlon, W. D. Tilley, T. E. Hickey, J. B. Bruning and A. D. Abell, A short, cell permeable bimane-constrained PCNA-interacting peptide *Advance Article, doi.org/10.1039/D1CB00113B*, 2021.

46. S. R. Perry, T. A. Hill, A. D. de Araujo, H. N. Hoang and D. P. Fairlie, Contiguous hydrophobic and charged surface patches in short helix-constrained peptides drive cell permeability, *Org. Biomol. Chem.*, 2018, **16**, 367-371.

47. D. Kalafatovic and E. Giralt, Cell-Penetrating Peptides: Design Strategies beyond Primary Structure and Amphipathicity, *Molecules*, 2017, **22**.

48. D. Birch, M. V. Christensen, D. Staerk, H. Franzyk and H. M. Nielsen, Fluorophore labeling of a cell-penetrating peptide induces differential effects on its cellular distribution and affects cell viability, *Biochim. Biophys. Acta. Biomembr.*, 2017, **1859**, 2483-2494.

49. H. H. Szeto, P. W. Schiller, K. Zhao and G. Luo, Fluorescent dyes alter intracellular targeting and function of cell-penetrating tetrapeptides, *The FASEB Journal*, 2005, **19**, 118-120.

50. M. P. Luitz, A. Barth, A. H. Crevenna, R. Bomblies, D. C. Lamb and M. Zacharias, Covalent dye attachment influences the dynamics and conformational properties of flexible peptides, *PLoS One*, 2017, **12**, e0177139.

Supplemental Data.

A NUCLEAR PERMEABLE PEPTIDOMIMETIC TO TARGET PCNA

S7.1 Methods

S7.1.1 Synthesis and characterisation of peptides

Unless otherwise indicated, all starting materials were purchased from commercial sources and used without further purification. All peptides were synthesised by the Fmoc/tBu solid-phase peptide synthesis protocol detailed below, with all L-amino-acids (unless otherwise specified), and then *N*-terminally acetylated before cyclisation on-resin. Peptides were subsequently cleaved from the resin (and simultaneously globally deprotected). Purification was carried out by semi-preparative HPLC on a Gilson GX-Prep RP-HPLC system on a Phenomenex Aeris Peptide C18 (10 × 250 mm) column. RP-HPLC solvents were A: H_2O with 0.1% TFA, and buffer B: ACN with 0.1% TFA; 0.2 µm filtered. Elution was done with linear gradients of solvent B into A over a gradient as specified in the individual compound sections, at 4 mL/min flow rate, with UV detection at 220/254 nm. Purity of all compounds was confirmed by analytical RP-HPLC on an Agilent 1260 HPLC equipped with a Phenomenex Luna C18(2) column (250 × 4.6 mm) over a linear gradient of 5-50% B over 15 min at 1.5 mL/min flow rate, with UV detection at 220 nm. High-resolution mass spectra were collected using an Agilent 6230 ESI-TOF via direct injection in ACN with 0.1% formic acid as the running buffer. All graphs were generated using GraphPad Prism 8 software.

The following standard Fmoc-protecting amino-acids with orthogonal protecting groups were used for all peptides unless otherwise specified. All amino-acids were L- and purchased from Chem-Impex International: 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 or Boc-Arg(Pbf)-OH.

Peptides were synthesised according to the following sequence of protocols in Table S1, and characterisation data is included in Table S2.

Stage 1 Protocols: Synthesis

1A. Liberty Blue automated microwave-assisted solid-phase peptide synthesis

Peptides were synthesised by the Fmoc/tBu solid-phase peptide synthesis protocol on a CEM Liberty Blue automated microwave peptide synthesiser using the standard manufacturer's conditions. The peptides were assembled on Chem Impex Rink Amide AM resin (0.1 mmol, 213 mg, 0.47 mmol/g) or Mimetopes Rink Amide AM resin (0.1 mmol, 219 mg, 0.456 mmol/g). The resin was initially swollen in DCM (10 mL, 15 min), washed with DMF (2 × 5 mL) and transferred to the microwave reaction vessel. The resin-bound Fmoc-groups were deprotected with a mixture of 20% piperidine and 0.1 M OxymaPure in DMF using the standard microwave deprotection method with a maximum temperature of 90°C. Couplings were performed with Fmoc-protected amino-acids (0.2 M in DMF, 5 equiv), OxymaPure® (1 M in DMF, 5 equiv) and DIC (0.5 M in DMF, 5 equiv) under the 'Standard Coupling' microwave method with a maximum temperature of 90°C, except for coupling of Fmoc-L-His(Trt)-OH which was coupled using a 'maximum 50°C 10 min coupling' procedure; and Fmoc-L-Arg(Pbf)-OH which used the default 'Arginine Double Coupling' microwave method which included two couplings steps – the first at room temperature 25 min and the second at a maximum of 75°C 2 min. Additionally, in the synthesis of **1-0** the *N*-terminal glycine was double coupled. Between each coupling and deprotection step the resin was washed with DMF.

1B. Manual solid-phase peptide synthesis

Rink Amide PL resin (0.1 mmol, 322 mg, 0.31 mmol/g, Agilent) was swollen in 1:1 DMF/DCM (15 mL) for 15 min. The Fmoc-protecting group was removed by treatment of the resin with a solution of 20% piperidine and 0.1 M HOBt in DMF (5 mL) for 15 min. The solution was drained and the resin washed with DMF (3 × 5 mL). Amino-acid couplings were achieved by addition of a solution of Fmoc-protected amino-acid (5 equiv), HATU (5 equiv) and DIPEA (10 equiv) in DMF (5 mL), to the resin and stirred intermittently for 1 h. The solution was drained and the resin washed with DMF (5 × 5 mL). The *N*-terminal Fmoc-protecting group was removed by treatment of the resin with a solution of 20%

piperidine and 0.1 M HOBt in DMF (5 mL) for 10 min, the solution was drained and the resin washed with DMF (5 x 5 mL). A TNBS test* was used to verify each coupling with steps repeated as necessary. Successive couplings and Fmoc-deprotections were repeated to achieve the desired sequence.

**TNBS Test:* A small spatula of swollen resin taken out and 1 drop each of TNBS (100 µL 5% w/v in H_2O added to 900 µL of DMF) and DIPEA solutions (100 µL in 900 µL of DMF) added and allowed to develop for 1 min. Clear/yellow beads indicated no free amine (negative), while red/orange beads showed free amine was present (positive).

Stage 2 Protocols: Cysteine reaction

2A. Bimane cyclisation

Following linear peptide assembly, the cysteine side-chains were selectively deprotected: Mmt groups were removed by repetitive treatment of the resin with 2% TFA in DCM (5 mL) for 1 min, and the resin then washed with DCM (3 × 5 mL). Treatments were repeated until the solution no longer turned yellow on addition to the resin (~ 150-200 mL total). The resin was then further washed with DCM (5 × 5 mL) and DMF (5 × 5 mL) and treated with a solution of dibromobimane (2 equiv) and DIPEA (4 equiv) in DMF (6 mL), and reacted for 3 h with intermittent stirring. The solution was then removed and the resin washed with DMF (5 × 5 mL) and DCM (5 × 5 mL), then dried with diethyl ether (3 × 5 mL).

2B. 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 HPLC and MS. The solution was then lyophilised to give the crude peptide conjugate. See Figure S1 for the combination of reagent peptides.

Stage 3 Protocols: N-terminal modifications

3A. Fluorescein attachment 1

Following linear peptide assembly, and *N*-terminal deprotection a solution of fluorescein-5-isothiocyanate (5 equiv) and DIPEA (10 equiv) in DMF (5 mL) was added to the resin-bound peptide and stirred intermittently for 1 h. The solution was then removed and the resin washed with DMF (5 x 5 mL) and DCM (5 x 5 mL), then dried with diethyl ether (3 x 5 mL).

3B. Fluorescein attachment 2

Following linear peptide assembly, and *N*-terminal deprotection a solution of 5(6)-carboxyfluorescein (5 equiv), HATU (5 equiv) and DIPEA (10 equiv) in DMF (5 mL), was added to the resin-bound peptide and stirred intermittently for 1 h. The solution was drained and the resin washed with DMF (5 × 5 mL) and DCM (5 × 5 mL), then dried with diethyl ether (3 × 5 mL).

3C. Maleimide attachment

Following linear peptide assembly, a solution of 4-malemidobutyric acid (3 equiv), HATU (3 equiv, 0.5 M) and DIPEA (10 equiv, 348 µL) in DMF (5 mL) was added to the *N*-terminal deprotected peptide on resin and stirred intermittently for 2 h. The solution was removed and the resin washed with DMF (5 × 5 mL), DCM (3 × 5 mL) and DMF (5 × 5 mL). A small amount (micro spatula-full) of resin was transferred and tested for the presence of a free amine by a TNBS test* and the coupling repeated if necessary. The solution was then removed and the resin washed with DMF (5 × 5 mL) and DCM (5 × 5 mL), diethyl ether (3 × 5 mL) and air dried.

3D. Acetylation

Following linear peptides assembly, and *N*-terminal deprotection, the resin-bound peptide was treated with a solution of Ac_2O (470 µL) and DIPEA (870 µL) in DMF (4 mL) for 15 min with intermittent stirring. The solution was then drained, and the resin washed with DMF (3 × 5 mL) and DCM (3 x 5 mL) then dried with diethyl ether (3 × 5 mL).

Stage 4 Protocols: Isolation and purification

4A. Cleavage from solid-support, isolation and purification

Following complete assembly of the peptide and deprotection of the final Fmoc group, cyclisation and/or *N*-terminal modification, the peptides were subsequently cleaved from the resin and the side-chain protecting groups simultaneously globally deprotected by treatment of the resin with 92.5/2.5/2.5/2.5 TFA:TIPS:DODT:H_2O** (5 mL) for 2 h***. The TFA mixture was pipetted from the resin and concentrated under a stream of nitrogen to 0.5-1 mL. The peptide was then precipitated by addition of diethyl ether (10 mL) and the mixture cooled at -20°C for 1 h or overnight. The precipitate was pelleted by centrifugation (7600 rpm, 10 min), and the supernatant decanted. The pellet was dried under a nitrogen stream, and then dissolved in 1:1 ACN/H_2O, before being syringe filtered (0.2 µm) and lyophilised to yield the crude peptide as a fluffy white powder. The peptides were purified by semi-preparative RP-HPLC on a Gilson GX-Prep system using a Phenomenex Luna C18(2) column (10 × 250 mm), over a linear ACN/H_2O gradient. RP-HPLC solvents were (A) H_2O with 0.1% TFA and (B) ACN with 0.1% TFA. The pure product containing fractions were pooled and lyophilised. The identity of the final compounds was confirmed by High Resolution Mass Spectrometry on an Agilent 6230 ESI-TOF LCMS. Purity of the peptides was confirmed by analytical RP-HPLC on an Agilent 1260 HPLC equipped with a Phenomenex Luna C18(2) column (250 × 4.6 mm) over a gradient of 5-50% B (15 min).

For maleimide-containing peptides **2-0a and **6-0a**, DODT was omitted from the cleavage mixture – 95/2.5/2.5 TFA:TIPS:H_2O (5 mL)

*** For arginine-rich peptides **1-0** – **1-14**, **3-3**, **4-3**, **4-3F**, **4-4**, and **4-4F** the cleavage time was extended to 5 h to ensure complete deprotection of the Pbf groups.

4B. Purification

The crude peptides were purified by semi-preparative RP-HPLC using a Phenomenex Aeris Peptide C18 column (10 × 250 mm) over a linear gradient of 0.1% TFA in H_2O (A) and 0.1% TFA in ACN (B). The pure product-containing fractions were combined and lyophilised to give the final purified peptide. Purity of the peptides was confirmed by analytical RP-HPLC on an Agilent 1260 HPLC equipped with a Phenomenex Luna C18(2) column (250 × 4.6 mm) over a gradient of 5-50% B (15 min).

S7.1.2 General SPR protocol:

The running buffer used for ligand attachment and analyte binding experiments was 10 mM HEPES buffer with 150 or 300 mM NaCl, 3 mM EDTA and 0.05% Tween20, adjusted to pH 7.4 with 2M NaOH, as specified in Table S4. Experiments were performed on a S200 GE Biacore System at 25°C. A GE CM5 Series S sensor chip was primed with running buffer and preconditioned per the manufacturer's recommendation with successive injections (2 × 50 s, 30 µL/min) of each 50 mM NaOH, 10 mM HCl, 0.1% SDS, 0.85% H_3PO_4 and glycine pH 9.5. The surface was then activated with a solution of 0.2 M EDC and 50 mM NHS (600 s, 10 µL/min).

PCNA (5 µL, 12 mg/ml) was diluted into running buffer (245 µL). Only once the preactivation was complete was the protein further diluted to a final concentration of 25 µg/ml in 10 mM NaAc (~pH 4.6) by addition of PCNA/HEPES (50 µL) to a solution of 100 mM NaAc (50 µL) and water (400 µL). This solution was immediately injected over only the second flow cell (10 µL/min) until ~2500 RU was reached at stabilisation. Both flow cells were then blocked with 1.0 M ethanolamine pH 8.5 (600 s, 10 µL/min). The chip was left to stabilise for two h before sample injections commenced. After stabilisation a final protein level of 2600 RU was achieved.

Peptides (approx. 2 mg by weight) were dissolved in milliQ H_2O (50 µL) and centrifuged (7800 rpm, 10 min) to remove any particulate. The peptide stock concentration was determined by UV absorbance (A_λ), where 2 µL of the stock was further diluted in water (20-40 fold) and a measurement taken in triplicate with a Nanodrop2000 and baselined to 750 nm absorbance. The ε_λ and λ used for each peptide is specified in Table S5. The peptide stock solution concentration was then calculated per $c = (A_\lambda/\varepsilon_\lambda.l).DF$ where concentration is in molar, A_λ is absorbance at λ nm calculated as an average of three readings, λ is the appropriate wavelength, l is the pathlength in cm (1 mm for Nanodrop), and ε_λ is the molar absorptivity at λ nm and DF is the dilution factor. The peptides were then diluted into running buffer before further dilution as necessary.

The steady state affinity experiments were all run at 30 µL/min, a contact time of 40 s and dissociation time of 60 s, followed by regeneration of 2 M NaCl (2 × 30 s). Each peptide was serially diluted eight times, and run from least to most concentrated following a blank injection. The data was analysed using the provided Biacore S200 Evaluation software.

S7.1.3 Cell uptake and imaging experiments

T47D cell line with fluorescence microscopy

T47D cells were cultured at 37°C and 5% CO_2, on glass coverslips in a 6-well plate in 2mL media per well consisting of RPMI (Sigma, R0883) supplemented with 2 mM L-Glutamine (Sigma, G7513) and 10% FBS. The cells were allowed to reach 70% of confluency prior to treatment with 1 or 5 µM of p21-peptidomimetics and incubated for 4, 24 or 48 h. During and after treatment, samples were protected from light as much as possible. After treatment the cells were washed with ice cold PBS (Gibco, 14190144) two times for 5 minutes. The cells were fixed with 4% PFA (10% Neutral Buffered Formalin, ChemSupply, #1258) for 10 min at rt. Samples were washed twice for 5 min with PBS at rt and then permeabilised in 5% TritonX solution for 1 h at rt, followed by two 5 min wash steps with PBS at rt. Samples were stained with Phalloidin (Alexa Fluor568 Phalloidin, Invitrogen A12380; 1/40 in 5% BSA in PBS) for 20 minutes. Followed by two 5 minutes wash steps with PBS at room temperature. The nuclei in the samples were stained with DAPI (4',6-diamidino-2-phenylindole, dihydrochloride, Invitrogen D1306; 1/2000 in PBS) for 1 min. After 2 wash steps with PBS the coverslips were mounted onto slides using DAKO fluorescent mounting medium (S302380-2) and sealed with clear nail polish (Sally Hansen). The slides were allowed to dry o/n. The following day, the samples were imaged with an Olympus IX73 Inverted Fluorescence Microscope (Adelaide Microscopy). The nuclei of the cells were visualised by using a standard Ultra-Violet long-pass filter. Peptides that were tagged with a FITC fluorophore (**1-0F**, **2-0F**, **3-1F** – **3-3F**) a standard intermediate-blue long-pass filter was used. Visualisation of cellular F-actin (phalloidin staining) was done using a standard Cy5 long-pass filter. The images were taken with a 40x and 100x (oil) objectives and are shown in Figure S2.

MDA-MB-468 cell line with confocal microscopy

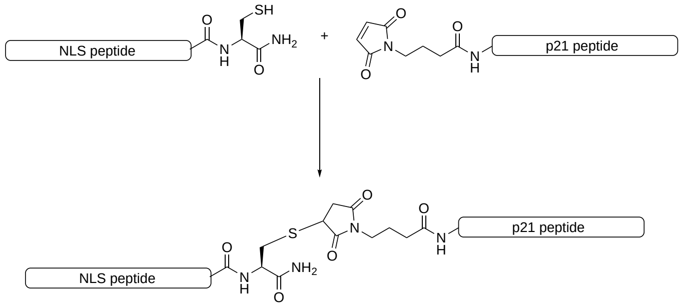
MDA-MB-468 mKate is a breast cancer cell line that was lenti-virally modified to stably express nuclear fluorescent (ex 588nm; em 635nm) mKate protein. Cells were maintained in DMEM (Sigma, D5671) base media supplemented with 2 mM L-Glutamine (Sigma, G7513) and 1 mM Sodium Pyruvate (Sigma, S8636) at 37°C and 5% CO_2. For experiments, cells were seeded at ~70% confluency onto glass coverslips inside wells of a 6-well culture plate containing 2 mL of media. The cells were cultured for 48 h to allow for attachment, then treated 24 h with 10 µM of peptide (**1-0F**, **2-0F**, **5-1F** – **5-4F**, **6-0**, **6-0F**, **6-1**, **6-1F** – **6-4F**) while being protected from light to sustain the florescent signal. After treatment, cells were washed with ice cold PBS (Gibco, 14190144) two times for 5 min to remove residual media and peptide. Cells were fixed with 4% PFA (10% Neutral Buffered Formalin, ChemSupply, #1258) for 10 min at room temperature, followed by two 5 min wash steps with PBS at rt. Coverslips were then mounted onto microscope slides using DAKO fluorescent mounting medium (S302380-2) and sealed with clear nail polish (Sally Hansen). Slides were allowed to dry overnight. The following day, samples were imaged using a Confocal Olympus FV3000 microscope (Adelaide Microscopy).

Cell nuclei were visualised using a 594 nm laser with a detection range of 600-700 nm. A 488 nm laser with a detection range of 490-534 nm was used for peptides **1-0F**, **2-0F**, **5-1F** – **5-4F**, **6-0F**, **6-1F** – **6-4F**, which contain a FITC fluorophore. A 405 nm laser with a detection range of 410-485 nm was used for peptides **6-0**, **6-0F**, **6-1**, **6-1F** – **6-4F**, which contains the Bimane fluorophore. Images were taken with a 30x silicon oil objective, with a 2x zoom setting with imaging software (Olympus, Cell Sens), bringing the total magnification of images to 60x.

S7.2 Peptide synthesis supplementary data

Table S1: Order of peptide synthesis protocols

Peptides	Stage 1 Synthesis	Stage 2 Modification #1- Cysteine reaction	Stage 3 Modification #2 *N*-terminal modification	Stage 4 Isolation & purification
1-0 – 1-14	1A. Liberty Blue Synthesis			4A. Cleavage, Isolation & Purification
1-0F	1A. Liberty Blue Synthesis		3B.Fluorescein attachment 2	4A. Cleavage, Isolation & Purification
2-0 **3-1 – 3-3** **4-2a**	1B. Manual peptide synthesis			4A. Cleavage, Isolation & Purification
2-0a	1B. Manual peptide synthesis		3C. Maleimide attachment	4A. Cleavage, Isolation & Purification
2-0F **3-1F – 3-3F** **4-1F – 4-4F**	1B. Manual peptide synthesis		3A. Fluorescein attachment 1	4A. Cleavage, Isolation & Purification
5-1F – 5-4F		2B. Conjugation of 4-1F – 4-4F with 2-0a, respectively		4B. Purification
6-0	1B. Manual peptide synthesis	2A. Bimane cyclisation	3D. Acetylation	4A. Cleavage, Isolation & Purification
6-0a	1B. Manual peptide synthesis	2A. Bimane cyclisation	3C. Maleimide attachment	4A. Cleavage, Isolation & Purification
6-0F	1B. Manual peptide synthesis	2A. Bimane cyclisation	3A. Fluorescein attachment 1	4A. Cleavage, Isolation & Purification
6-1F – 6-4F, 6-1		2B. Conjugation of 4-1F – 4-4F, 4-2a with 6-0a, respectively		4B. Purification



Product	NLS peptide	p21 peptide
5-1F	**4-1F**	**2-0a**
5-2F	**4-2F**	**2-0a**
5-3F	**4-3F**	**2-0a**
5-4F	**4-4F**	**2-0a**
6-1F	**4-1F**	**6-0a**
6-2F	**4-2F**	**6-0a**
6-3F	**4-3F**	**6-0a**
6-4F	**4-4F**	**6-0a**
6-1	**4-1a**	**6-0a**

Figure S1: Thioether conjugation of NLS peptide and p21 peptides

Table S2: Peptide characterisation data. All peptides contain a *C*-terminal amide. ESI+ data collected on an Agilent HRMS. Purity is calculated from area under the curve of analytical RP-HPLC trace.

Peptide	Sequence	Mw	MF	M+ Calc	ESI+ $[M+X]^{X+}$ Calc.	ESI+ $[M+X]^{X+}$ Found	Purity % (220 nm)
1-0	GRKRRQTSMTDFYHSKRRLIFS	2770.20	$C_{120}H_{197}N_{43}O_{31}S_1$	2769.4722	4^{+}: 693.3760	693.3793	97.4
1-1	GRKRRQTSMTDFYHSKRRLIF	2683.11	$C_{117}H_{192}N_{42}O_{29}S_1$	2681.4561	4^{+}: 671.3720	671.3718	93.2
1-2	GRKRRQTSMTDFYHSKRRLI	2535.94	$C_{108}H_{183}N_{41}O_{28}S_1$	2534.3877	4^{+}: 634.6049	634.6042	80.3
1-3	GRKRRQTSMTDFYHSKRRL	2422.78	$C_{102}H_{172}N_{40}O_{27}S_1$	2421.3036	4^{+}: 606.3339	606.3338	88.2
1-4	GRKRRQTSMTDFYHSKRR	2309.62	$C_{96}H_{161}N_{39}O_{26}S_1$	2308.2196	4^{+}: 578.0629	578.0604	87.0
1-5	GRKRRQTSMTDFYHSKR	2153.43	$C_{90}H_{149}N_{35}O_{25}S_1$	2152.1185	5^{+}: 431.4317	431.4316	91.9
1-6	GRKRRQTSMTDFYHSK	1997.25	$C_{84}H_{137}N_{31}O_{24}S_1$	1996.0174	4^{+}: 500.0123	500.0123	90.3
1-7	GRKRRQTSMTDFYHS	1869.08	$C_{78}H_{125}N_{29}O_{23}S_1$	1867.9224	4^{+}: 467.9886	467.9874	71.7
1-8	GRKRRQTSMTDFYH	1782.00	$C_{75}H_{120}N_{28}O_{21}S_1$	1780.8904	4^{+}: 446.2306	446.2305	82.0
1-9	GRKRRQTSMTDFY	1644.86	$C_{69}H_{113}N_{25}O_{20}S_1$	1643.8314	4^{+}: 411.9659	411.9653	93.2
1-10	RKRRQTSMTDFYHSKRRLIFS	2713.14	$C_{118}H_{194}N_{42}O_{30}S_1$	2711.4667	4^{+}: 678.8747	678.8744	91.2
1-11	KRRQTSMTDFYHSKRRLIFS	2556.96	$C_{112}H_{182}N_{38}O_{29}S_1$	2555.3656	5^{+}: 512.0811	512.0798	86.4
1-12	RRQTSMTDFYHSKRRLIFS	2428.79	$C_{106}H_{170}N_{36}O_{28}S_1$	2427.2706	4^{+}: 607.8257	607.8257	91.3
1-13	RQTSMTDFYHSKRRLIFS	2272.60	$C_{100}H_{158}N_{32}O_{27}S_1$	2271.1695	4^{+}: 568.8004	568.7994	79.8
1-14	QTSMTDFYHSKRRLIFS	2116.41	$C_{94}H_{146}N_{28}O_{26}S_1$	2115.0684	4^{+}: 529.7751	529.7752	75.7
2-0	RQTSMTDFYHSK	1499.67	$C_{64}H_{98}N_{20}O_{20}S_1$	1498.6987	4+: 500.5740	500.5774	80.9
1-0F	Fl- GRKRRQTSMTDFYHSKRRLIFS-OH	3128.53	$C_{141}H_{206}N_{42}O_{38}S_1$	3127.5199	4+: 782.8878	782.9110	83.0
2-0F	FITC-A$_{\beta}$RQTSMTDFYHSK	1960.13	$C_{88}H_{114}N_{22}O_{26}S_2$	1958.7716	4+: 490.7007	490.6728	98.2
3-1	RKRRQTSMTDFYHSK	1940.20	$C_{82}H_{134}N_{30}O_{23}S_1$	1938.9959	2^{+}: 970.5057	970.5058	95.2
3-2	RQTSMTDFYHSKRR	1812.05	$C_{76}H_{122}N_{28}O_{22}S_1$	1810.9009	4^{+}: 453.7330	453.7330	89.3
3-3	RKRRQTSMTDFYHSKRR	2252.60	$C_{94}H_{158}N_{38}O_{25}S_1$	2251.1981	4^{+}: 563.8073	563.8068	90.6
3-1F	FITC-A$_{\beta}$RKRRQTSMTDFYHSK	2402.70	$C_{106}H_{152}N_{32}O_{29}S_2$	2401.0844	4^{+}: 600.7750	600.7727	94.1
3-2F	FITC-A$_{\beta}$RQTSMTDFYHSKRR	2272.51	$C_{100}H_{140}N_{30}O_{28}S_2$	2270.9738	6^{+}: 370.5040	370.5032	83.9
3-2F	FITC-A$_{\beta}$RKRRQTSMTDFYHSKRR	2713.06	$C_{118}H_{174}N_{40}O_{31}S_2$	2711.2710	3^{+}: 904.7683	904.7681	91.2
2-0a	mal-RQTSMTDFYHSK	1664.82	$C_{72}H_{105}N_{21}O_{23}S$	1663.7413	4^{+}: 416.9430	416.9403	73.0
4-1F	FITC-A$_{\beta}$PKKKRKVC	1445.76	$C_{67}H_{100}N_{18}O_{14}S_2$	1444.7108	4^{+}: 362.1855	362.1915	91.2
4-2F	FITC-A$_{\beta}$PAAKRVKLDC	1559.82	$C_{71}H_{102}N_{18}O_{18}S_2$	1558.7061	4^{+}: 390.6843	390.6866	75.9
4-3F	FITC-A$_{\beta}$GRKKRRQRRRC	1959.29	$C_{82}H_{131}N_{35}O_{18}S_2$	1957.9853	3^{+}: 653.6696	653.6609	90.4
4-4F	FITC-A$_{\beta}$RRWWRRWRRC	2076.40	$C_{96}H_{126}N_{34}O_{16}S_2$	2074.9532	3^{+}: 692.6589	692.6472	93.7
4-1a	PKKKRKVC	985.30	$C_{43}H_{84}N_{16}O_8S$	984.6379	2^{+}:493.32675	493.3270	85.6
5-1F	FITC-A$_{\beta}$PKKKRKVC(suc-RQTSMTDFYHSK)	3110.58	$C_{139}H_{205}N_{39}O_{37}S_3$	3108.4521	5^{+}: 622.6982	622.6968	79.4
5-2F	FITC-A$_{\beta}$PAAKRVKLDC(suc-RQTSMTDFYHSK)	3224.64	$C_{143}H_{207}N_{39}O_{41}S_3$	3222.4474	5^{+}: 645.4973	645.4969	73.0
5-3F	FITC-A$_{\beta}$GRKKRRQRRRC(suc-RQTSMTDFYHSK)	3642.11	$C_{154}H_{236}N_{56}O_{41}S_3$	3621.7266	6^{+}: 604.6289	604.6287	77.8
5-4F	FITC-A$_{\beta}$RRWWRRWRRC(suc-RQTSMTDFYHSK)	3741.22	$C_{168}H_{231}N_{55}O_{39}S_3$	3738.6945	6^{+}: 624.1236	624.1239	84.2
6-0	[cyclo-3,7]-Ac-RQC(-)SMTC(Bim-)FYHSK		$C_{74}H_{106}N_{22}O_{20}S_3$		3+: 573.9119	573.9140	81.4
6-0a	[cyclo-3,7]-mal-RQC(-)SMTC(Bim-)FYHSK	1843.08	$C_{80}H_{111}N_{23}O_{22}S_3$	1841.7436	3^{+}: 614.9223	614.9252	70.7
6-0F	[cyclo-4,8]-FITC-A$_{\beta}$RQC(-)SMTC(Bim-)FYHSK	2138.39	$C_{97}H_{122}N_{24}O_{24}S_4$	2134.7942	3^{+}: 712.6059	712.5945	74.2
6-1F	FITC-A$_{\beta}$PKKKRKVC[mal-RQC(-)SMTC(Bim)FYHSK]	3288.89	$C_{147}H_{211}N_{41}O_{36}S_5$	3286.4544	6^{+}: 548.7502	548.7506	91.3
6-2F	FITC-A$_{\beta}$PAAKRVKLDC[mal-RQC(-)SMTC(Bim)FYHSK]	3402.91	$C_{151}H_{213}N_{41}O_{40}S_5$	3400.4497	6^{+}: 567.7494	567.7497	85.0
6-3F	FITC-A$_{\beta}$GRKKRRQRRRC[mal-RQC(-)SMTC(Bim)FYHSK]	3802.38	$C_{162}H_{242}N_{58}O_{40}S_5$	3799.7289	5^{+}:760.95358	760.9534	82.1
6-4F	FITC-A$_{\beta}$RRWWRRWRRC[mal-RQC(-)SMTC(Bim)FYHSK]	3919.49	$C_{176}H_{237}N_{57}O_{38}S_5$	3916.6969	6^{+}: 653.7906	653.7907	76.2
6-1	H-PKKKRKVC[mal- RQC(-)SMTC(Bim)FYHSK]	2828.39	$C_{123}H_{195}N_{39}O_{30}S_4$	2826.3815	4^{+}: 707.6032	707.6026	85.1

Table S4: SPR protocol conditions

SPR Run	SPR Chip loading	Buffer system
1	1500 RU	10 mM HEPES buffer with 300 mM NaCl, 3 mM EDTA and 0.05% Tween20, adjusted to pH 7.4 with 2M NaOH
2	1380 RU	10 mM HEPES buffer with 150 mM NaCl, 3 mM EDTA and 0.05% Tween20, adjusted to pH 7.4 with 2M NaOH

Table S5: SPR characterisation data. ε_λ is the absorptivity factor calculated for the wavelength λ. Top Conc is the highest concentration run of an 8 times 1 in 2 dilution. SE is the standard error. χ^2 is the fitting error. Ass/Diss are the association and dissociation times (in seconds) for the injection. Round 1 or 2 specifies the conditions of the experiment specified in Table S4.

Peptide	λ	ε_λ	Top Conc (nM)	Affinity K_D	K_D SE	χ^2	Ass/Diss time	Round no.
1-0	205	98620	300	5.96	2.07	14.6	60/90	1
			50	4.32	1.30	13.8	60/90	2
1-1	205	95840	300	12.4	3.40	15.7	60/40	1
1-2	205	84460	200	6.70	1.26	3.34	80/40	1
1-3	205	81680	500	21.3	2.86	2.32	60/40	1
1-4	205	78900	500	18.7	1.49	0.415	60/40	1
1-5	205	74770	500	19.9	1.38	0.338	60/40	1
1-6	205	70640	500	37.4	2.24	0.205	60/40	1
1-7	205	67860	500	115	9.03	0.219	60/40	1
1-8	205	65080	500	98.7	6.42	0.146	60/40	1
1-9	205	57100	500	83.0	5.13	0.115	60/40	1
1-10	205	95840	200	11.8	3.18	11.4	80/40	1
1-11	205	91710	200	9.24	2.48	10.8	80/40	1
1-12	205	88930	200	8.05	1.76	4.07	80/40	1
1-13	205	84800	500	31.5	3.56	1.03	80/40	1
1-14	205	80670	3000	827	50.8	0.136	80/40	1
2-0	205	56820	500	102.3	5.3	0.0701	40/60	2
3-1	205	67867	500	41.85	5.80	1.01	40/60	2
3-2	205	65080	1000	45.90	4.30	0.301	40/60	2
3-3	205	76120	500	37.99	10.0	7.06	60/90	2
3-1F	490	86983	1000	202.1	18.0	0.556	40/60	2
3-2F	490	86983	2000	190.0	8.7	0.100	40/60	2
3-2F	490	86983	500	61.00	5.6	1.09	40/60	2
5-1F	490	86983	250	33.11	4.20	2.62	40/60	2
5-2F	490	86983	2000	130.1	15.0	0.77	40/60	2
5-3F	490	86983	1000	100.7	9.50	3.60	40/60	2
5-4F	490	86983	1000	372.9	30.0	3.70	40/60	2
6-0	380	4694[1]	10000	887	15.4	0.880	40/60	2
6-1	380	4694[1]	1000	175.8	27	5.58	40/40	2
6-0F	380	4694[1]	30000	25190	1900	0.0547	40/60	2
6-1F	380	4694[1]	3000	535.9	73.0	4.92	40/60	2

S7.2.1 Supplementary cell images

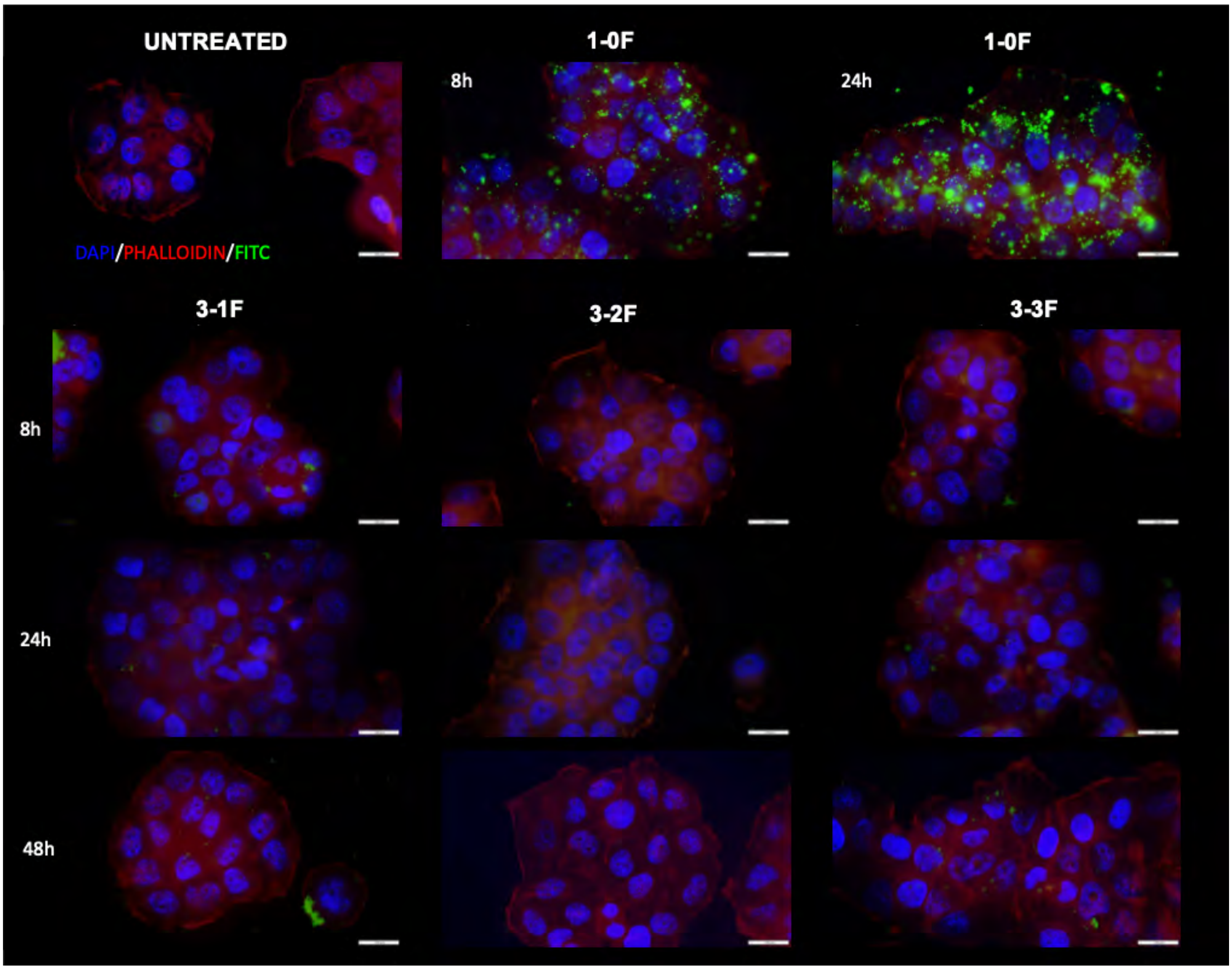

UNTREATED
1-0F
1-0F
8h
24h
DAPI/PHALLOIDIN/FITC
3-1F
3-2F
3-3F
8h
24h
48h

Figure S1: Cell images in T47D breast cancer cells, where blue DAPI stain marks the cell nucleus, red phallodin marks F-actin indicating the cell cytosol, and green Fluorescein marks the peptide. The cells were incubated for 8, 24 or 28 h with 5 μM of peptide. The scale bar indicates 50 μM.

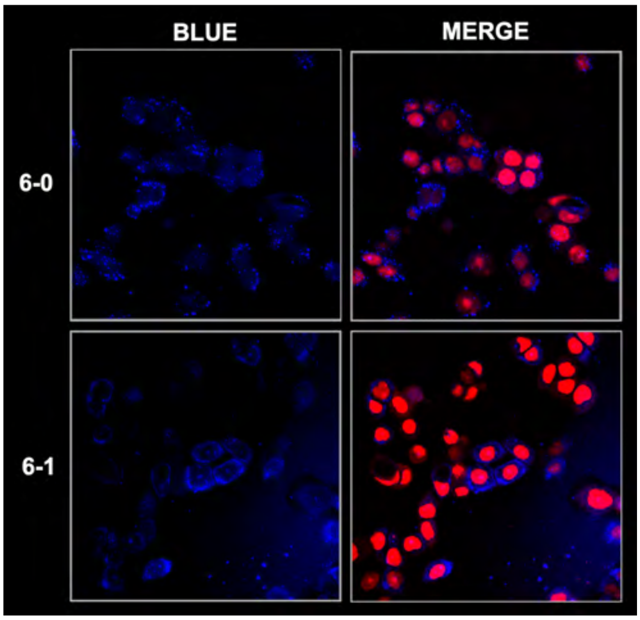

BLUE
MERGE
6-0
6-1

Figure S2: MDA-MB-468 breast cancer cells treated with 10 μM of macrocyclic bimane peptides for 24 h

REFERENCES

1. X. Shen, C. B. Pattillo, S. Pardue, S. C. Bir, R. Wang and C. G. Kevil, Measurement of plasma hydrogen sulfide in vivo and in vitro, *Free Radical Biol. Med.*, 2011, **50**, 1021-1031.

Chapter 8.

A TURN-ON FLUORESCENT PCNA SENSOR

Research Article: Bioorganic and Medicinal Chemistry Letters **2021**, 41, 128031

A turn-on fluorescent PCNA sensor

Aimee J. Horsfall,[1] Theresa Chav,[1] John B. Bruning[2] and Andrew D. Abell[1*]

[1] ARC Centre of Excellence for Nanoscale BioPhotonics, Institute of Photonics and Advanced Sensing, School of Physical Sciences, The University of Adelaide, Adelaide, South Australia, 5005, Australia

[2] Institute of Photonics and Advanced Sensing, School of Biological Sciences, The University of Adelaide, Adelaide, South Australia, 5005, Australia

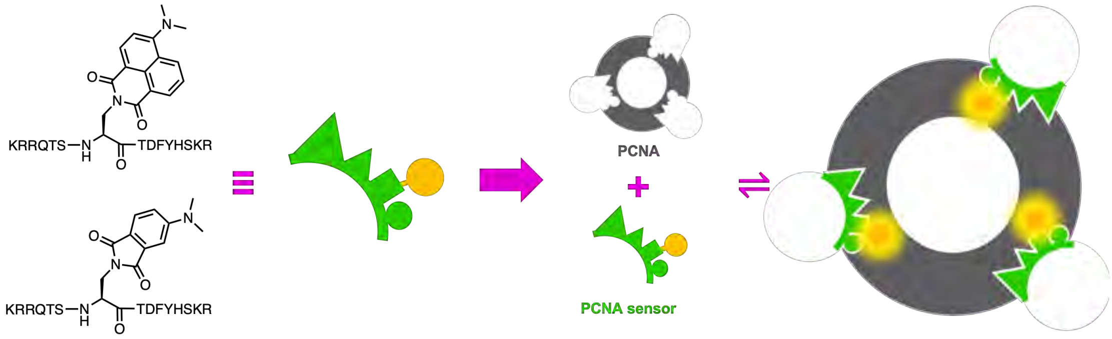
*Corresponding author

© 2021 Bioorganic and Medicinal Chemistry Letters, Elsevier

First published online: 8th April 2021

ABSTRACT

The solvatochromic amino-acids 4-DMNA or 4-DAPA, were separately introduced at position 147, 150 or 151 of a short p21 peptide (141-155) known to bind sliding clamp protein PCNA. The ability of these peptides, **1a**-**3a** and **1b**-**3b**, to act as a turn-on fluorescent sensor for PCNA was then investigated. The 4-DMNA-containing peptides (**1a**-**3a**) displayed up to a 40-fold difference in fluorescence between a polar (Tris buffer) and a hydrophobic solvent (dioxane with 5 mM 18-crown-6), while the 4-DAPA-containing peptides (**1b-3b**) displayed a significantly enhanced (300-fold) increase in fluorescence from Tris buffer to dioxane with 18-crown-6. SPR analysis of the peptides against PCNA revealed that the 151-substituted peptides **3a** and **3b** interacted selectively with PCNA, with K_D values of 921 nM and 1.28 μM, respectively. Analysis of the fluorescence of these peptides in the presence of increasing concentrations of PCNA revealed a 10-fold change in fluorescence for **3a** at 2.5 equivalents of PCNA, compared to only a 3.5-fold change in fluorescence for **3b**. Peptide **3a** is an important lead for development of a PCNA-selective turn-on fluorescent sensor for application as a cell proliferation sensor to investigate diseases such as cancer.



STATEMENT OF AUTHORSHIP

Title of Paper	A turn-on fluorescent sensor for PCNA
Publication Status	☐ Published ☐ Accepted for Publication ☒ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	**Research Article:** A. J. Horsfall, T. Chav, J. B. Bruning, and A. D. Abell, *Journal of Bioorganic & Medicinal Chemistry Letters,* 2021

Principal Author

Name of Principal Author (Candidate)	Aimee J Horsfall		
Contribution to the Paper	Designed and synthesised peptides, SPR assay, fluorescence assays, discussed results, wrote and edited manuscript		
Overall percentage (%)	60%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	15/02/2021

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Theresa Chav		
Contribution to the Paper	Designed and synthesised peptides, SPR assay, fluorescence assays, discussed results, edited manuscript		
Signature		Date	

Name of Co-Author	John B Bruning		
Contribution to the Paper	Supervised AJH & TC, discussed results, and edited manuscript.		
Signature	[redacted]	Date	18/2/21

Name of Co-Author	Andrew D Abell		
Contribution to the Paper	Supervised AJH & TC, discussed results, and edited manuscript.		
Signature		Date	15/2/2021

8.1 Introduction

A number of solvatochromic amino-acids have been reported and incorporated into short peptides because of their ability to respond to the local microenvironment.[1-3] The fluorescence emission and intensity of such amino-acids changes in response to an environmental stimulus, such as a change in solvent polarity. Consequently, the judicious incorporation of a solvatochromic amino-acid into a peptide that is known to selectively bind a protein, can provide the basis of sensor for that protein.[2-4] Binding of a solvatochromic peptide to a protein exposes the solvatochromic amino-acid to a hydrophobic environment, thus shielding it from the polar surroundings, to give a measurable change in fluorescence.[5] Such fluorescent sensors provide a number of advantages over other protein detection methods, such as antibody-based methods (e.g. histology, ELISA).[6, 7] In particular, solvatochromic peptides produce a reversible fluorescent response, while also being small enough to be cell permeable and thus enable real-time monitoring of dynamic processes.[2, 3] Such peptides can also be readily synthesised on a large scale and stored as a lyophilised powder. In addition, peptides can selectively target a surface of a non-enzymatic protein of interest,[8-10] whereas many small molecule protein sensors generally require enzyme activation to liberate a fluorophore and 'sense' the enzyme.[11] Solvatochromic peptides therefore provide an opportunity to target a wide range of proteins that can otherwise only be detected with antibody stains, or via indirect methods such as reporter assays. One such protein target for study is Proliferating Cell Nuclear Antigen (PCNA).

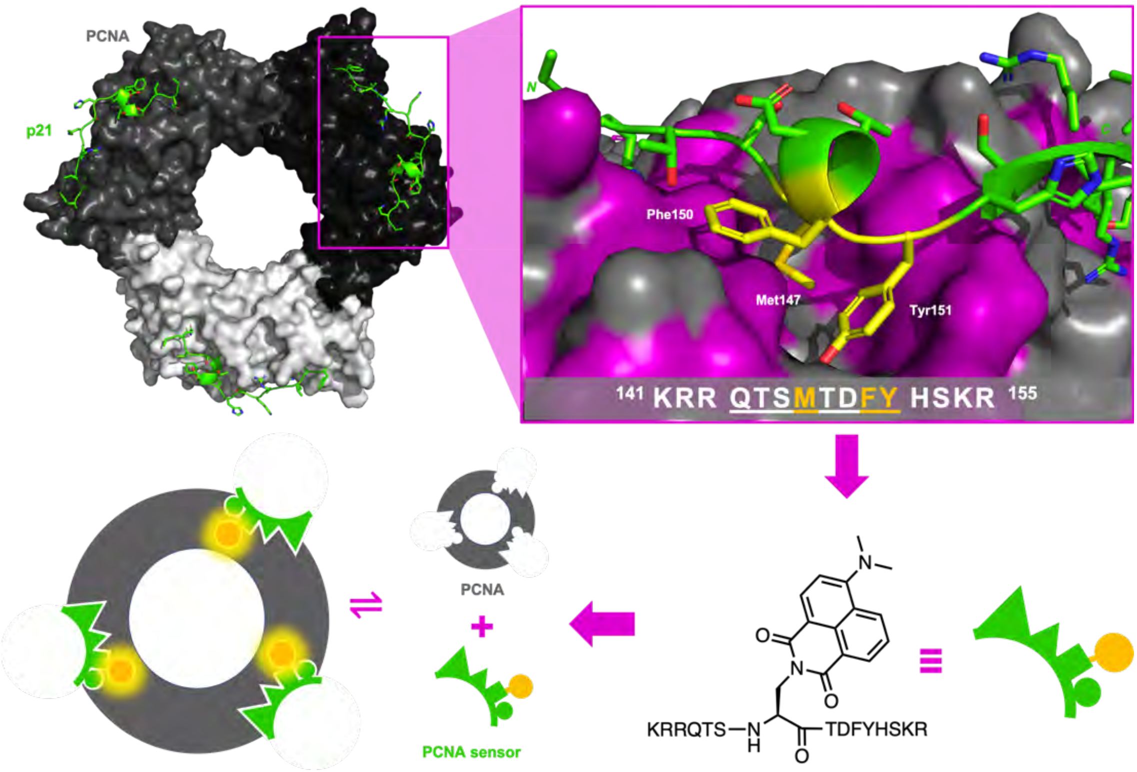
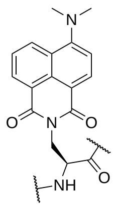
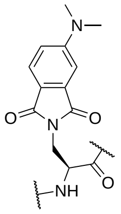


Figure 1: Design of a peptide-based sensor for PCNA. PCNA (grey, surface representation, co-crystal structure PDB:1AXC) interacts with a p21 peptide (139-160, green, main-chain shown in cartoon representation, side-chains shown as sticks) (top left), where three hydrophobic residues (yellow) are embedded onto the hydrophobic (purple, aa: I, V, L, A, G, F, M, C) surface (top right). A solvatochromic amino-acid (DMNA) within a p21 peptide sequence yields a PCNA sensor (bottom right) which upon binding PCNA can give a fluorescent signal (bottom left). Co-crystal structure was visualised with PyMOL 2.0. Atoms elemental colouring in stick representation: Oxygen, red; Nitrogen, blue; and Sulfur, gold.

Table 1. Structure of peptides and SPR Binding affinity. SE – standard error. All peptides possess a *C*-terminal amide. B$_1$ – 4-DMNA, B$_2$ – 4-DMAP

Peptide	Sequence	Fluorophore	Affinity K$_D$ (nM)	K$_D$ SE (nM)
p21$_{141-155}$	^{141}KRRQTSMTDFYHSKR^{155}	N/A	16.3	0.516
1a	KRRQTS^{147}B$_1$TDFYHSKR	B$_1$ = 	NS >25000	
2a	KRRQTSMTD^{150}B$_1$YHSKR		NS > 20000	
3a	KRRQTSMTDF^{151}B$_1$HSKR		921	120
1b	KRRQTS^{147}B$_2$TDFYHSKR	B$_2$ = 	NS > 25000	
2b	KRRQTSMTD^{150}B$_2$YHSKR		4610	570
3b	KRRQTSMTDF^{151}B$_2$HSKR		1280	260

NS – non-specific binding to protein. See supplementary Figure S2.

PCNA is a member of the sliding clamp family and is an essential mediator of DNA-replication and repair processes.[12-16] It is a toroidal shaped homotrimer that interacts with other proteins through a shallow hydrophobic groove on the clamp surface that presents as an ideal target for solvatochromic-based detection.[13, 14, 17, 18] PCNA encircles the double-stranded DNA and acts as a mobile docking platform for over 200 enzymes required for DNA-replication and repair to bind, in order to gain access to the DNA.[14, 16, 19, 20] Additionally, PCNA has been shown to be a key processivity factor for a number of proteins, such as the major processive polymerase, pol δ.[18, 21] PCNA is upregulated in many cancers in order to meet the high demand for DNA-replication required by proliferative cancers cells.[12-14, 20]

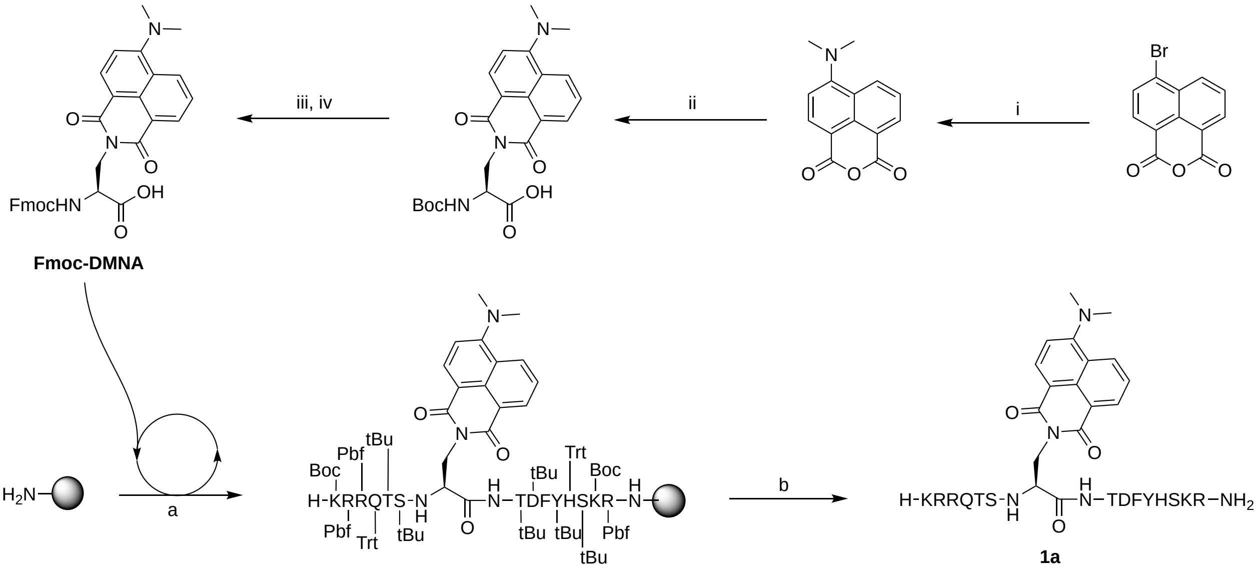
Peptides derived from the p21 protein (e.g. p21$_{139-160}$) are reported to bind selectively to PCNA and thus provide a template for developing a peptide-based fluorescent PCNA sensor.[17, 22-24] Here we report the synthesis of six p21 peptide derivatives (**1a**-**3a**, and **1b**-**3b**) which include a solvatochromic amino-acid, 4-*N,N*-dimethylamino-1,8-naphthalimidoalanine (4-DMNA) or 4-*N,N*-dimethylaminophthalimidoalanine (4-DAPA). The position of the solvatochromic acid was chosen so as not to perturb binding of the peptide to PCNA, but ensure it embeds on the PCNA surface to give the maximum possible fluorescent response. A solvatochromic peptide of this type would provide a basis for a turn-on fluorescent sensor for PCNA to investigate disease-states where PCNA is upregulated, such as cancers.[12, 20, 25]

A p21-derived peptide (residues 139-160) reported to bind PCNA with high affinity was chosen as the starting scaffold to design a fluorescent PCNA sensor.[17, 18] p21$_{139-160}$ binds PCNA with a well-defined defined 3$_{10}$-helical turn that is anchored onto the protein surface by Met147, Phe150 and Tyr151 which insert into a hydrophobic cleft on the protein surface (PDB 1AXC, Figure 1).[17, 18] Thus, following visual inspection of a co-crystal structure of a p21 peptide bound to PCNA (1AXC, Figure 1), we hypothesised that substituting solvatochromic amino-acids at positions 147, 150 or 151 would give rise to a solvatochromic peptide that targets PCNA. These three residues (Met147, Phe150 and Tyr151) largely make hydrophobic contacts with the protein surface, however a hydrogen bond from a Tyr151 of p21 to Gln131 of PCNA is commonly observed.[17, 20] A shorter p21 variant (141-155) was used for ease of synthesis and is known to bind to PCNA.[26] Six peptides were designed with either 4-*N*,*N*-dimethylamino-1,8-naphthalimidoalanine (4-DMNA) or 4-(*N*,*N*-dimethylamino)phthalimidoalanine (4-DAPA) introduced at one of position 147, 150 or 151, to give peptides **1a**-**3a** and **1b**-**3b** (Table 1).

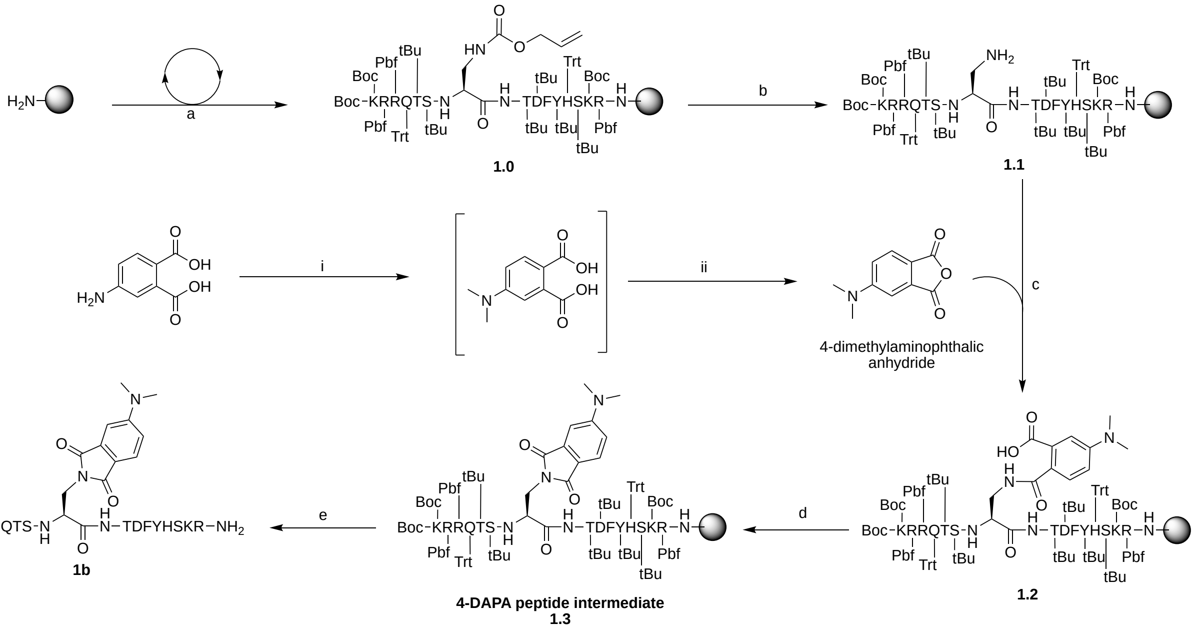
8.2 Results and discussion

All peptides were synthesised by solid-phase peptide synthesis using a standard Fmoc/t-Bu protecting group strategy, as detailed in the Experimental section (see SI) using literature protocols. Fmoc-DMNA[1] was inserted during the synthesis (Scheme 1), to give peptides **1a**-**3a** respectively. Peptides **1b**-**3b** were synthesised via an alternate route where the 4-(*N*,*N*-dimethylamino)phthalimide (4-DMAP) fluorophore was installed following peptide assembly to avoid piperidine-catalysed hydrolysis of 4-DMAP.[1] An Alloc-protected diaminopropionic acid (Dap) was introduced at the appropriate position (147, 150 or 151) to allow subsequent introduction of the 4-DMAP fluorophore. The Alloc group was deprotected on treatment with palladium triphenyl phosphene (Scheme 2, **1.0** to **1.1**) and the amine then reacted with 4-(*N*,*N*-dimethylamino)phthalic anhydride under basic conditions (Scheme 2, **1.1** to **1.2**) followed by ring-closure with HBTU/HOBt gave the 4-DMAP-functionalised amino-acid (4-DAPA, Scheme 2 **1.3**).[1, 27] The peptides were cleaved from the resin on treatment with TFA (Scheme 2), to give peptides **1b**-**3b.**

The resulting peptides were all purified by semipreparative RP-HPLC and characterised by analytical RP-HPLC and MS, however a notable M+16 peak was evident in the MS of peptides **2a** and **3a**, post-purification. This product was assigned as a methionine sulfoxide, a methionine oxidation product. Fractions of **2a** or **3a** left on the 'bench' for an extended period (e.g. overnight) with exposure to ambient



Scheme 1: General synthesis scheme for DMNA-containing peptides (a, b)[1], where **1a** is a representative example and synthesis of Fmoc-DMNA (*i-iv*). Reagents: i) 3-dimethylaminopropionitrile, 3-methylbutanol, N_2, reflux; ii) Boc-Dap-OH, $NaHCO_3$, 1,4-dioxane/water, reflux; iii) 1:1 DCM/TFA; iv) Fmoc-OSu, $NaHCO_3$, dioxane/water; a) Sequential cycles of Fmoc-AminoAcid-OH (5 equiv) coupling with HCTU (5 equiv), DIPEA (10 equiv) in DMF, 1h, followed by Fmoc-deprotection with 20% piperidine in DMF, 10 min; b) 95:2.5:2.5 TFA/TIPS/H_2O, 1 h.


a
1.0
b
1.1
i
ii
c
4-dimethylaminophthalic anhydride
1.2
d
4-DAPA peptide intermediate
1.3
e
1b
H-KRRQTS-N
N-TDFYHSKR-NH2

Scheme 2: General synthesis scheme for DMAP-containing peptides (a-e)[27] where **1b** is a representative example, and synthesis of 4-(*N*,*N*-dimethylamino)phthalic anhydride (i, ii). Reagents: i) Formaldehyde, Pd/C, H_2, MeOH, 3 h; ii) Acetic anhydride, 50°C, N_2, 12 h; a) Sequential cycles of Fmoc-AminoAcid-OH (5 equiv) coupling with HCTU (5 equiv), DIPEA (10 equiv) in DMF, 1h, followed by Fmoc-deprotection with 20% piperidine in DMF, 10 min; b) Pd(PPh$_3$) (0.8 equiv), phenylsilane (25 equiv), N_2; c) 4-(*N*,*N*-dimethylamino)phthalic anhydride (2 equiv), DIPEA (4 equiv), NMP, N_2, o/n; d) HBTU (6 equiv), HOBt (6 equiv), DIPEA (12 equiv), NMP, 2 h; e) 95:2.5:2.5 TFA/TIPS/H_2O, 1 h.

light, gave an enhanced M+16 peak compared to fractions analysed immediately after purification. This suggests that Met147 oxidation is likely the result of photo-catalysed singlet-oxygen production promoted by the fluorophore.[28, 29] A fresh sample of each crude peptide **2a** and **3a** was purified in the dark to give samples with >90% purity as determined by analytical RP-HPLC. All subsequent experiments were carried out while minimising exposure to ambient light in order to limit methionine oxidation. Separate solutions of peptides **1a**-**3a** and **1b**-**3b** in Tris buffer (pH 7.4), and dioxane containing 18-crown-6, to provide polar and hydrophobic environments respectively, were prepared. Each sample was plated in triplicate in a 96-well plate and an absorbance spectrum (300-700 nm) was collected, to reveal 390 and 440 nm as the maximum absorbance wavelengths for DAPA peptides (**1b**-**3b**) and DMNA peptides (**1a**-**3a**) respectively, in good agreement with literature.[1] Fluorescence emission spectra for peptides **1a**-**3a** (460-700 nm) and **1b**-**3b** (410-700 nm) were collected, with an excitation wavelength of 440 and 390 nm respectively. This revealed an emission maximum at 540 nm for DMNA peptides **1a**-**3a**, and at 520 nm for DAPA peptides **1b**-**3b** in dioxane with 18-crown-6. Emission maxima in Tris buffer were red-shifted (relative to the dioxane spectra), with a maximum at ~560 nm observed for **1a**-**3a**, and ~585 nm for **1b**-**3b** (see Table S1). These emission maxima are consistent with literature values for these fluorophores.[1] Fluorescence spectra in Tris buffer were of significantly lower intensity for all peptides compared to those taken in dioxane with 18-crown-6, indicating that introducing the fluorophores to these peptides does not impact solvatochromic behaviour (Figure 2, dashed v. solid line). An excitation spectrum was also collected for each peptide (**1a**-**3a**, em. 540 nm; **1b**-**3b**, em. 520 nm) which is displayed in Figure 2. The change in fluorescence emission intensity for each peptide was calculated by dividing the maximum fluorescence intensity in dioxane with 18-crown-6,

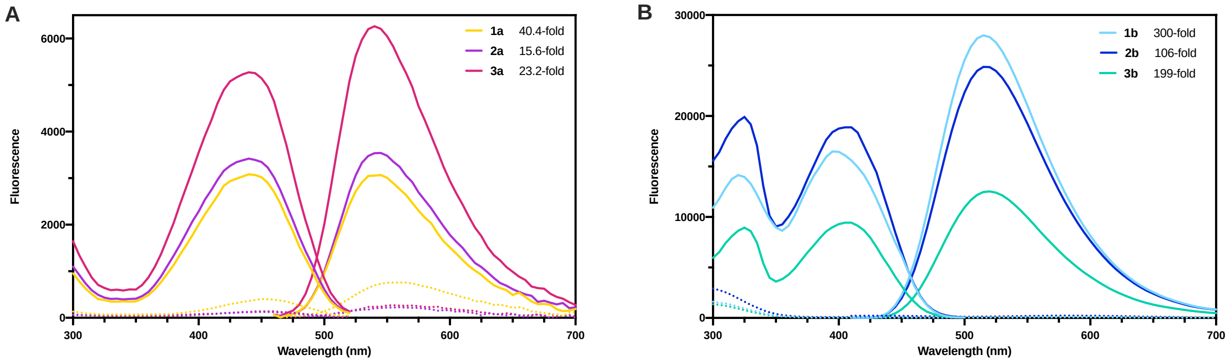


Figure 2: Fluorescence excitation and emission spectra of peptides **1a**-**3a** (**A**) and **1b**-**3b** (**B**) at 5 µM in Tris buffer (dashed line) or dioxane with 5 mM 18-crown-6 (solid line). **A** DMNA Excitation curve (300-520 nm) with emission fixed at 540 nm, emission curve (460-700 nm) with excitation fixed at 440 nm. **B** DMAP Excitation curve (300-500 nm) with emission fixed at 520 nm, emission curve (410-700 nm) with excitation fixed at 390 nm. All curves were smoothed in GraphPad Prism 9 with 10 neighbours, and all emission spectra scaled down by a factor of 5.

by the maximum fluorescence intensity in Tris buffer (Table S1). Overall, the DAPA-containing peptides **1b**-**3b** gave a greater increase in fluorescence in a hydrophobic vs polar solvent system (106 – 300-fold, Figure 2B), compared to the DMNA peptides **1a**-**3a** (15.6 – 40-fold, Figure 2A). The greatest difference in fluorescence was observed for peptide **1b**, which gave a 300-fold increase in fluorescence between Tris buffer and 18-crown-6 in dioxane. The 151-substituted peptide **3a** gave the largest fluorescence increase for the DMNA-containing peptides between the two solvent systems at 40-fold. This indicates that peptides containing the same fluorophore (e.g. compare 1a to 2a and 3a) exhibit a different magnitude increase in fluorescence intensity. This difference is likely related to the position of the fluorophore within the peptide scaffold which may influence the conformation of the peptide in solution, and give rise to a different microenvironment (more or less hydrophobic) surrounding the fluorophore. It is known that different solvents can drive peptide secondary structure folding.[30-32]

Next, peptides **1a**-**3a**, and **1b**-**3b** (800 nM) were separately added to PCNA (600 nM) in Tris buffer, to identify any change in fluorescence that would likely reflect an interaction with PCNA. Solutions of each of the peptides alone, PCNA, and Tris buffer alone were prepared as a set of controls. Each sample was plated in triplicate in a 96-well plate and fluorescence spectra collected for peptides **1a**-**3a** (ex. 440 nm, em. 460-700 nm) and for **1b**-**3b** (ex. 390 nm, em. 420-700 nm), with the results shown in Figure 3. The fluorescence of the peptides in buffer alone (Figure 3, dashed line) were different to one another, which is likely an influence of the fluorophore position in the backbone. All six peptides exhibited significantly greater fluorescence in the presence of PCNA (Figure 3, solid coloured line), compared to in the absence of PCNA (Figure 3, dashed coloured line), consistent with a more hydrophobic environment. The change in fluorescence observed on exposure to PCNA was calculated by dividing the maximum fluorescence intensity of the emission spectrum of each peptide with PCNA, by the maximum intensity of the respective peptide alone (Table S1). Peptide **2a** gave the largest increase (14-fold change) in fluorescence on exposure to PCNA (see Figure 3B). In comparison, the maximum fluorescence increase observed for a DAPA-containing peptide was 7.9-fold for **3b** (Figure 3F). Overall, the DMNA-containing peptides (**1a**-**3a**,

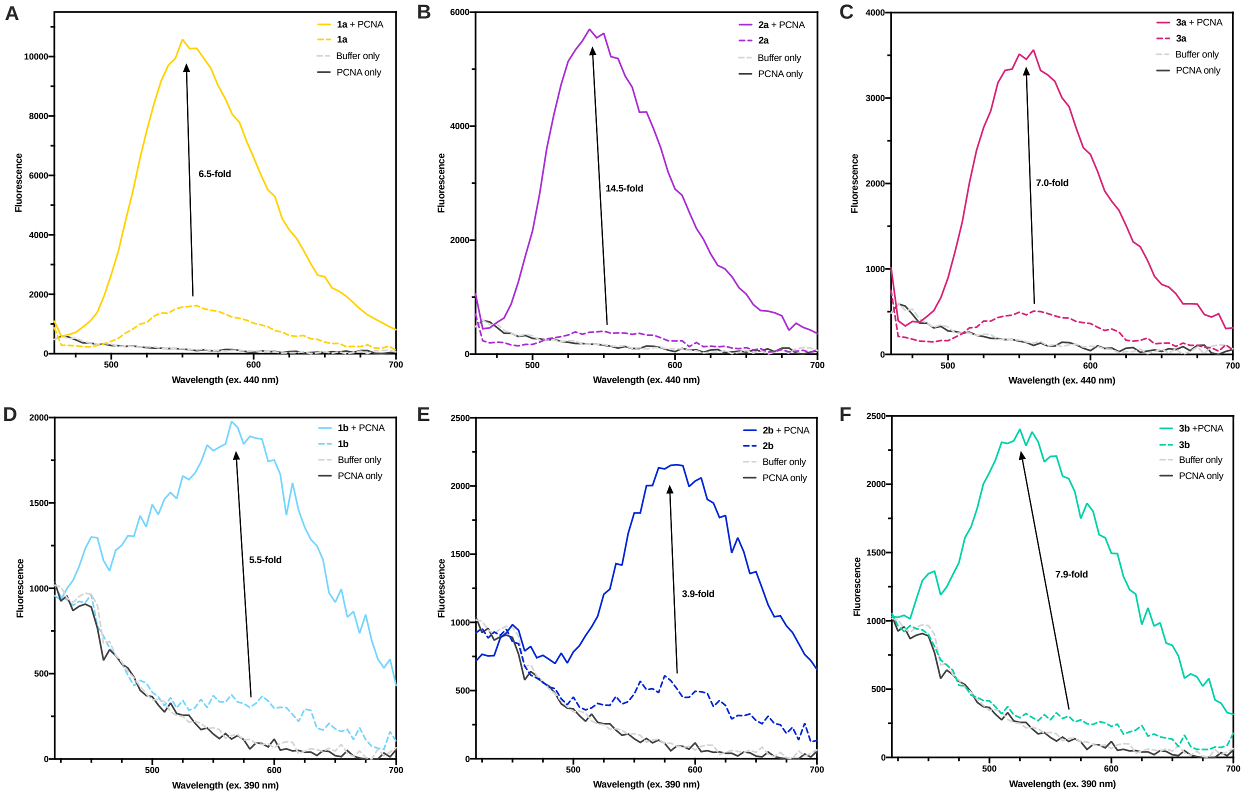


Figure 3: Fluorescence emission spectra of peptides **1a**-**3a** and **1b**-**3b** in the presence or absence of PCNA. All solutions were in Tris buffer, with peptide at 800 nM (dashed line), and peptide (800 nM) mixed with PCNA (600 nM) (solid line). **A**, **B** & **C** Fluorescence emission spectra for peptides **1a**-**3a** with excitation fixed at 440 nm. **D**, **E** & **F** Fluorescence emission spectra for peptides **1b**-**3b** with excitation fixed at 390 nm. Raw spectra are presented; however, all curves were smoothed in GraphPad Prism 9 with 10 neighbours prior to calculation of the change in fluorescence intensity (Table S1).

Figure 3A-C) gave a larger fluorescence response in the presence of PCNA, than the DAPA-containing peptides (**1b**-**3b**, Figure 3D-F). This suggests a stronger association between **1a**-**3a** and PCNA, where the DMNA-peptides spend longer associated to PCNA or DMNA and are better shielded from the polar environment, that in turn gives rise to a greater fluorescence emission.

The affinities of peptides **1a**-**3a** and **1b**-**3b** for PCNA were determined by SPR. Peptide **p21$_{141\text{-}155}$**, without a solvatochromic amino-acid, bound to PCNA with a K_D value of 16.3 nM, in good agreement with reported values for other p21 peptides.[17, 18, 24] The 147 DMNA or DAPA-substituted peptides **1a** and **1b**, and 150-DMNA substituted peptide **2a**, displayed significantly weaker affinity for PCNA (Table 1). The maximum sensorgram response for peptides **1a**, **1b** and **2a** exceeded the theoretical maximum response calculated for specific binding to PCNA, relative to the amount of protein loaded onto the sensor chip (see SI, Equation 1). This suggests that binding of **1a**, **1b** or **2a** to PCNA was largely non-specific (Figure S1) with the predicted K_Ds laying outside of the dose-response curve (Table S3). Non-specific binding implies that the fluorophores are too large to be readily accommodated within the hydrophobic cleft that houses the p21 147-residue in the case of **1a** and **1b**, or the 150-residue in the case of **2a** (Figure 1). Peptide **2b** with a smaller 4-DMAP fluorophore at position 150 bound PCNA with a K_D value of 4.61 μM. However, the sensorgram response was also high and the defined plateau in the binding response was not reached, which also suggested a non-specific binding component (Figure S1). In comparison, the SPR response of

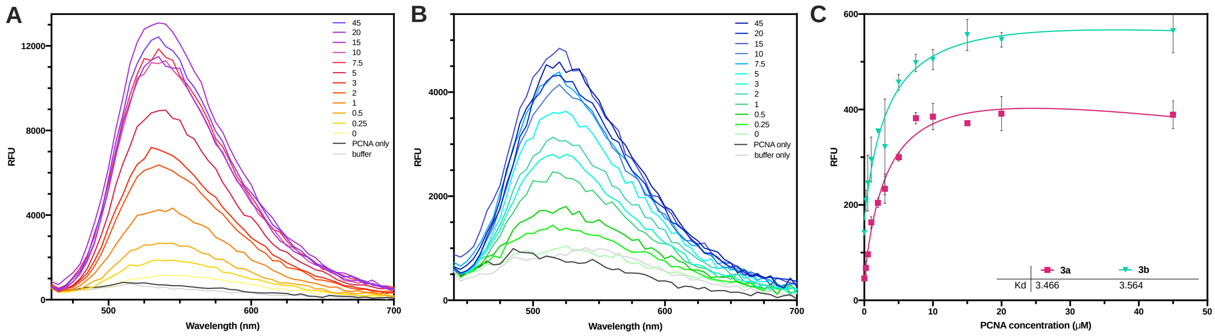


Figure 4: Dose-response experiment: Fluorescence response of 3 µM of peptide **3a** (**A**) or **3b** (**B**) in the presence of increasing concentrations of PCNA (in µM). **A** DMNA Emission curve of **3a** (460-700 nm) with excitation fixed at 440 nm. **B** DMAP Emission curve of **3b** (440-700 nm) with excitation fixed at 390 nm. Raw spectra are presented; however, all curves were smoothed in GraphPad Prism 9 with 10 neighbours prior to fluorescence change calculation (Table S1). **C** Fluorescence maxima from each spectrum in **A** and **B** plotted against PCNA concentration, and dissociation constant (K_d) calculated in GraphPad Prism (Table S2).

151 DMNA and 151 DAPA-substituted peptides **3a** and **3b** respectively, reached a plateau with increasing peptide concentration that suggests the peptides were able to selectively target a PCNA-binding site and interact specifically with PCNA. A K_D of 921.5 nM and 1.28 µM was determined for **3a** and **3b** respectively (Table 1). This suggests that both 4-DMN and 4-DMAP are well accommodated by the hydrophobic cleft usually occupied by Tyr151 residue of p21. Peptides **3a** and **3b** were consequently chosen for further study as fluorescent PCNA sensor.

Next peptides **3a** and **3b** were combined with increasing concentrations of PCNA in HEPES buffer and the fluorescence emission spectrum recorded in a dose response experiment, to determine the ratio of peptide:PCNA required to generate the maximum fluorescence response (Figure 4). The peptide:PCNA ratio was varied from 12:1 to 1:15, with the peptide concentration held constant (3 µM), where the ratio is calculated relative to a PCNA monomer. Control solutions of the peptides **3a** and **3b**, PCNA, and buffer only were also prepared. All these solutions were plated in triplicate in a 96-well plate and fluorescence emission spectra (em. 460-700 nm) were collected (**3a** ex. 390 nm, **3b** ex. 420 nm, Figure 4A and B). The maximum fluorescence intensity of the peptide-only control was compared to that of the peptide with PCNA at each peptide:PCNA ratio. This revealed a maximum 10-fold greater fluorescence intensity for peptide **3a**, at a ratio of 1:2.5 peptide:PCNA and emission maximum at 535 nm, compared to the peptide alone. Peptide **3b** revealed a maximum 3.5-fold increase in fluorescence, also achieved at a ratio of 1:2.5 peptide:PCNA and emission maximum at 520 nm. The peak fluorescence intensity for each peptide was plotted against the concentration of PCNA (Figure 4C), to reveal a dissociation constant (K_d) of 3.4 and 3.5 µM for **3a** and **3b** respectively.

8.3 Conclusions

In summary, six p21 peptides (141-155) **1a**-**3a** and **1b**-**3b** containing a solvatochromic amino-acid 4-DMNA or 4-DAPA substituted at position 147, 150 or 151, are reported. The fluorescence response of these peptides in Tris buffer, and 18-crown-6 in dioxane was measured to reveal an expected and significantly higher fluorescence response for all peptides in the more hydrophobic dioxane with 18-crown-6. The 150 DAPA-substituted peptide **1b** resulted in the largest increase from the polar to hydrophobic environment with a 300-fold change in fluorescence intensity. The DAPA-containing peptides **1b**-**3b** all gave a greater fluorescence response compared to the DMNA-containing peptides **1a**-**3a**.

The DMNA-containing peptide **2a** gave the largest fluorescence increase of 14.5-fold in the presence of PCNA, compared to the peptide alone. The largest increase in fluorescence intensity upon exposure to PCNA for a DAPA peptide was 7.9-fold for **3b**. The size of these fluorescence changes are similar to other peptides targeting shallow protein clefts.[6] Overall, the DMNA-peptides resulted in a greater fluorescence change than the DAPA-peptides, when exposed to PCNA.

Only the 151-substituted peptides, **3a** and **3b**, bound specifically to PCNA with K_Ds of 921 nM and 1.28 μM, respectively. The remaining peptides mainly interacted with PCNA in a non-specific manner indicated by the SPR sensorgrams, which suggests the inclusion of DMNA or DAPA at position 147 or 150 has perturbed the binding to PCNA.

The fluorescence response of DMNA-containing peptide **3a** reached saturation at 2.5 equivalents of PCNA to peptide, and was 10-fold greater than the fluorescence obtained from the peptide alone. The fluorescence response of peptide **3b** also saturated at 2.5 equivalents of PCNA, however this only resulted in a 3.5-fold fluorescence increase compared to the peptide alone. This indicates that the maximum fluorescence intensity is achieved at a 1:2.5 ratio of peptide to PCNA monomer, where **3a** gave a greater fluorescence increase. Therefore, **3a** provides an interesting template to develop a PCNA-selective fluorescence probe to monitor cell proliferation. Future work will investigate optimisation of the cell and nuclear permeability of the probe to enable in-cell experiments.

8.4 Acknowledgments

The research was supported by the Australian Research Council Centre of Excellence for Nanoscale BioPhotonics (CNBP) (CE140100003). A.J.H. is supported by Australian Government Research Training Program Stipends (RTPS). The facilities of the OptoFab node of the Australian National Fabrication Facility (ANFF) and associated Commonwealth and SA State Government funding are also gratefully acknowledged.

8.5 Conflicts of interest

The authors have no conflicts of interest to declare.

REFERENCES

1. G. Loving and B. Imperiali, A Versatile Amino Acid Analogue of the Solvatochromic Fluorophore 4-N,N-Dimethylamino-1,8,naphthalimide: A Powerful Tool for the Stufy of Dynamic Protein Interactions, *J. Am. Chem. Soc.*, 2008, **130**, 13630-13638.
2. L. Choulier and K. Enander, Environmentally sensitive fluorescent sensors based on synthetic peptides, *Sensors*, 2010, **10**, 3126-3144.
3. Q. Liu, J. Wang and B. J. Boyd, Peptide-based biosensors, *Talanta*, 2015, **136**, 114-127.
4. E. Pazos, O. Vazquez, J. L. Mascarenas and M. E. Vazquez, Peptide-based fluorescent biosensors, *Chem. Soc. Rev.*, 2009, **38**, 3348-3359.
5. G. S. Loving, M. Sainlos and B. Imperiali, Monitoring protein interactions and dynamics with solvatochromic fluorophores, *Trends Biotechnol.*, 2010, **28**, 73-83.
6. K. E. Sapsford, J. B. Blanco-Canosa, P. E. Dawson and I. L. Medintz, Detection of HIV-1 Specific Monoclonal Antibodies Using Enhancement of Dye-Labeled Antigenic Peptides, *Bioconjugate Chem.*, 2010, **21**, 393-398.
7. H. D. Herce, D. Schumacher, A. F. L. Schneider, A. K. Ludwig, F. A. Mann, M. Fillies, M. A. Kasper, S. Reinke, E. Krause, H. Leonhardt, M. C. Cardoso and C. P. R. Hackenberger, Cell-permeable nanobodies for targeted immunolabelling and antigen manipulation in living cells, *Nat Chem*, 2017, **9**, 762-771.
8. K. M. Marks, M. Rosinov and G. P. Nolan, In vivo targeting of organic calcium sensors via genetically selected peptides, *Chem. Biol.*, 2004, **11**, 347-356.
9. R. J. Radford, W. Chyan and S. J. Lippard, Peptide-based Targeting of Fluorescent Zinc Sensors to the Plasma Membrane of Live Cells, *Chem. Sci.*, 2013, **4**, 3080-3084.
10. R. Radford, WenChyan and S. Lippard, Peptide Targeting of Fluorescein-Based Sensors to Discrete Intracellular Locales, *Chem. Sci.*, 2014, **4**, 4512-4516.
11. K. Singh, A. M. Rotaru and A. A. Beharry, Fluorescent Chemosensors as Future Tools for Cancer Biology, *ACS Chem. Biol.*, 2018, **13**, 1785-1798.
12. I. Stoimenov and T. Helleday, PCNA on the crossroad of cancer, *Biochem. Soc. Trans.*, 2009, **37**, 605-613.
13. S. Y. Park, M. S. Jeong, C. W. Han, H. S. Yu and S. B. Jang, Structural and Functional Insight into Proliferating Cell Nuclear Antigen, *J. Microbiol. Biotechnol.*, 2016, **26**, 637-647.
14. A. Gonzalez-Magana and F. J. Blanco, Human PCNA Structure, Function and Interactions, *Biomolecules*, 2020, **10**, 570.
15. W. Strzalka and A. Ziemienowicz, Proliferating cell nuclear antigen (PCNA): a key factor in DNA replication and cell cycle regulation, *Annals of Botany*, 2011, **107**, 1127-1140.
16. E. M. Boehm, M. S. Gildenberg and M. T. Washington, The Many Roles of PCNA in Eukaryotic DNA Replication, *Enzymes*, 2016, **39**, 231-254.
17. J. M. Gulbis, Z. Kelman, J. Hurwitz, M. O'Donnell and J. Kuriyan, Structure of the C-Terminal Region of p21 WAF1/CIP1 Complexed with Human PCNA, *Cell*, 1996, **87**, 297-306.
18. J. B. Bruning and Y. Shamoo, Structural and thermodynamic analysis of human PCNA with peptides derived from DNA polymerase-delta p66 subunit and flap endonuclease-1, *Structure*, 2004, **12**, 2209-2219.
19. K. N. Choe and G.-L. Moldovan, Forging Ahead through Darkness: PCNA, Still the Principal Conductor at the Replication Fork, *Mol. Cell*, 2016, **65**, 380-392.
20. A. J. Horsfall, A. D. Abell and J. B. Bruning, Targeting PCNA with Peptide Mimetics for Therapeutic Purposes, *ChemBioChem*, 2019, **21**, 442-450.
21. X. Lu, C. K. Tan, J. Q. Zhou, M. You, L. M. Carastro, K. M. Downey and A. G. So, Direct interaction of proliferating cell nuclear antigen with the small subunit of DNA polymerase delta, *J. Biol. Chem.*, 2002, **277**, 24340-24345.
22. D. I. Zheleva, N. Z. Zhelev, P. M. Fischer, S. V. Duff, E. Warbrick, D. G. Blake and D. P. Lane, A Quantitative Study of the in Vitro Binding of the C-Terminal Domain of p21 to PCNA: Affinity, Stoichiometry, and Thermodynamics, *Biochemistry*, 2000, **39**, 7388-7397.
23. C. Prives and V. Gottifredi, The p21 and PCNA partnership: a new twist for an old plot, *Cell Cycle*, 2008, **7**, 3840-3846.
24. A. J. Kroker and J. B. Bruning, p21 Exploits Residue Tyr151 as a Tether for High-Affinity PCNA Binding, *Biochemistry*, 2015, **54**, 3483-3493.
25. N. Barboule, V. Baldin, S. Jozan, S. Vidal and A. Valette, Increase level of p21in human ovarian tumors is associated with increased expression of CDK2, Cyclin A and PCNA., *Int. J. Cancer*, 1998, **76**, 891-896.
26. K. L. Wegener, A. E. McGrath, N. E. Dixon, A. J. Oakley, D. B. Scanlon, A. D. Abell and J. Bruning, Rational design of a 310-helical PIP-box mimetic targeting PCNA - the human sliding clamp, *Chem. Eur. J.*, 2018, **24**, 11325-11331.
27. M. Sainlos and B. Imperiali, Synthesis of anhydride precursors of the environment-sensitive fluorophores 4-DMAP and 6-DMN, *Nat. Protoc.*, 2007, **2**, 3219-3225.
28. Ping Yan, Yijia Xiong, Baowei Chen, Sewite Negash, T. C. Squier and M. U. Mayer, Fluorophore-Assisted Light Inactivation of Calmodulin Involves Singlet-Oxygen Mediated Cross-Linking and Methionine Oxidation, *Biochemistry*, 2006, **45**, 4736-4748.

29. I. Pibiri, S. Buscemi, A. Palumbo Piccionello and A. Pace, Photochemically Produced Singlet Oxygen: Applications and Perspectives, *ChemPhotoChem*, 2018, **2**, 535-547.

30. A. Kentsis and T. R. Sosnick, Trifluoroethanol Promotes Helix Formation by Destabilizing Backbone Exposure: Desolvation Rather than Native Hydrogen Bonding Defines the Kinetic Pathway of Dimeric Coiled Coil Folding, *Biochemistry*, 1998, **37**, 14613-14622.

31. N. Prabhu and K. Sharp, Protein-solvent interactions, *Chem. Rev.*, 2006, **106**, 1616-1623.

32. Y. Yu, J. Wang, Q. Shao, J. Shi and W. Zhu, The effects of organic solvents on the folding pathway and associated thermodynamics of proteins: a microscopic view, *Sci. Rep.*, 2016, **6**, 19500.

Supplemental Data.

A TURN-ON FLUORESCENT PCNA SENSOR

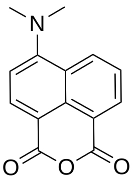
S8.1 Experimental

S8.1.1 General information

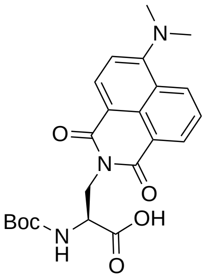
Unless otherwise indicated, all starting materials were purchased from commercial sources and used without further purification. All amino-acid building blocks used were purchased from Chem Impex Int'l Ltd, unless otherwise specified: Fmoc-Asp(tBu)-OH Fmoc-Lys(Boc)-OH, Fmoc-Arg(Pbf)-OH, Fmoc-Gln(Trt)-OH, Fmoc-Thr(tBu)-OH, Fmoc-Glu(tBu)-OH, Fmoc-Ile-OH, Fmoc-Met-OH, Fmoc-Tyr(tBu)-OH, Fmoc-Phe-OH, Fmoc-His(Trt)-OH, Fmoc-Ser(tBu)-OH, and Fmoc-Cys(Trt)-OH.

High-resolution mass spectra were collected using an Agilent 6230 ESI-TOF LCMS. RP-HPLC solvents were (A) H_2O with 0.1% TFA and (B) ACN with 0.1% TFA. Purity of all compounds was confirmed by analytical RP-HPLC on an Agilent 1260 HPLC equipped with a Phenomenex Luna C18(2) column (250 x 4.6 mm) over a gradient of 5-50% B (15 min). Purification was carried out by semi-preparative HPLC on a Gilson GX-Prep RP-HPLC system using a Phenomenex Aeris Peptide C18 column (10 x 250 mm), over a gradient as specified in the individual compound sections. All graphs were generated using GraphPad Prism 9 software.

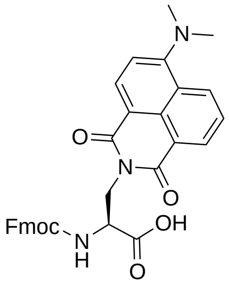
S8.1.2 Fluorophore synthesis



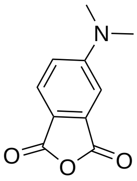
4-N,N'-Dimethylamino-1,8-nathphthalic anhydride. A suspension of 4-bromo-1,8-naphthalic anhydride (0.870 g, 3.14 mmol) in 3-methylbutanol (30 mL) was placed under a N_2 atmosphere and heated to reflux. 3-(dimethylamino)propionitrile (1.42 mL, 4 equiv, 12.6 mmol) was then added gradually to give a dark brown solution. The solution was heated at a vigorous reflux for 12 h. The solution was then allowed to cool to rt, without stirring, and the crystalline precipitate then isolated by vacuum filtration. The solid was washed with water (3 x 20 mL), then minimal cold hexane (2 x 10 mL) and dried under vacuum to give 4-*N,N'*-dimethylamino-1,8-nathphthalic anhydride as fine yellow crystals (0.669 g, 88%). NMR was consistent with literature:[1] ^{1}H NMR (500 MHz, $CDCl_3$) δ = 8.56 (dd, *J*=7.3, 0.8, 1H), 8.48 (dd, *J*=8.4, 0.7, 1H), 8.45 (d, *J*=8.3, 1H), 7.68 (dd, *J*=8.4, 7.4, 1H), 7.11 (d, *J*=8.3, 1H), 3.18 (s, 6H) ppm.



N-α-Boc-(4-N,N'-dimethylamino-1,8-nathpthalamido)alanine (Boc-DMNA). 4-(*N,N'*-Dimethylamino)-1,8-nathphthalic anhydride (1.23 g, 5.1 mmol) was dissolved in 1,4-dioxane (40 mL) under a N_2 atmosphere and brought to reflux. 3-amino-2-(Boc-amino)-propionic acid (Boc-Dap-OH, 1.04 g, 1 equiv, 5.1 mmol) was suspended in a solution of sodium bicarbonate (2.14 g, 5 equiv, 25.5 mmol) in water (20 mL) and then added slowly to the naphthalic anhydride solution such that reflux was maintained. The solution was refluxed for a further 1 h, then cooled to rt and concentrated in vacuo to ~20 mL. The residue was taken up in water (80 mL) and acidified to pH 1 with 4 M HCl. The cloudy orange solution was extracted with DCM (6 x 100 mL) and the combined organic layers were dried over $MgSO_4$ and the solvent removed *in vacuo* to give a dark orange residue of *N*-α-Boc-(4-N,N'-dimethylamino-1,8-nathpthalamido)alanine (1.66 g, 76 %). NMR was consistent with literature:[1] H NMR (500 MHz, $CDCl_3$) δ = 8.60 (dd, *J*=7.3, 1.2, 1H), 8.50 (d, *J*=8.2, 1H), 8.46 (dd, *J*=8.5, 1.2, 1H), 7.67 (dd, *J*=8.5, 7.3, 1H), 7.12 (d, *J*=8.3, 1H), 4.74 (br s, 1H), 4.68 – 4.60 (m, 1H), 4.61 – 4.52 (m, 1H), 3.13 (s, 6H), 1.30 (s, 9H) ppm.



N-α-Fmoc-(4-N,N'-Dimethylamino-1,8-nathpthalamido)alanine (Fmoc-DMNA). *N*-α-Boc-(4-N,N'-Dimethylamino-1,8-nathpthalamido)alanine (1.91 g, 4.5 mmol) was dissolved in a 1:1 solution of DCM/TFA (40 mL) and stirred at rt for 3 h. The reaction mixture was concentrated *in vacuo* and then redissolved in DCM (3 x 30 mL) and solvent removed *in vacuo*. The residue was dried under vacuum for 30 min. Fmoc-OSu (1.67 g, 1.1 equiv, 4.95 mmol) was added to the orange residue and dissolved in 1,4-dioxane (40 mL). Sodium bicarbonate (1.89 g, 5 equiv, 22.5 mmol) was suspended in water (20 mL) and added gradually to the Fmoc/DMNA solution. The reaction mixture was stirred at rt for 16 h then concentrated to ~20 mL *in vacuo*. The residue was taken up in water (50 mL) and washed with diethyl ether (3 x 40 mL). The aqueous layer was acidifed to pH 1 with 4 M HCl then extracted with DCM (4 x 60 mL). The combined organic layers were dried over $MgSO_4$ and concentrated *in vacuo*. The crude product was purified by column chromatography on silica using 1% AcOH in EtOAc, then pure fractions were combined and the solvent removed *in vacuo*. The residue was azeotroped with toluene to remove excess AcOH to give the final product as an orange solid (0.94 g, 38% over two steps). NMR was consistent with literature:[1] ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.59 (d, *J* = 7.2 Hz, 1H), 8.47 (d, *J* = 8.2 Hz, 1H), 8.44 – 8.38 (m, 1H), 7.69 (t, *J* = 7.9 Hz, 2H), 7.66 – 7.60 (m, 1H), 7.51 (d, *J* = 7.5 Hz, 1H), 7.47 (d, *J* = 7.4 Hz, 1H), 7.33 (t, *J* = 7.5 Hz, 2H), 7.26 – 7.21 (m, 2H), 7.20 – 7.13 (m, 2H), 7.05 (d, *J* = 8.1 Hz, 1H), 6.15 (d, *J* = 7.3 Hz, 1H), 4.93 – 4.81 (m, 1H), 4.79 – 4.61 (m, 2H), 4.30 – 4.13 (m, 2H), 4.03 (t, *J* = 7.7 Hz, 1H), 3.09 (s, 6H) ppm; HRMS (ESI+) Expected $[M+H]^+$ for $C_{32}H_{27}N_3O_6$ (549.1900): 500.1978, observed: $[M+H]^+$ 500.1965.



4-dimethylaminophthalimide (4-DMAP).[2] 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. in water, 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), then 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 reaction mixture was concentrated *in vacuo* before azeotroping with toluene (2 x 100 mL). The crude product was purified by column chromatagraphy on silica using DCM to give the final product as yellow crystals (0.35 g, 68% over two steps). NMR was consistent with literature:[2] ^{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.

S8.1.3 Peptide synthesis

General protocol for solid-phase peptide synthesis

Rink Amide PL resin (0.1 mmol, 322 mg, 0.31 mmol/g, Agilent) was swollen in 1:1 DMF/DCM (10 mL) for 15 min. The Fmoc-protecting group was removed by treatment of the resin with a solution of 20% piperidine in DMF (5 mL) for 15 min. The solution was drained and the resin washed with DMF (3 x 5 mL). Amino-acid couplings were achieved by addition of a solution of Fmoc-protected amino-acid (5 equiv), HCTU (5 equiv) and DIPEA (10 equiv) in DMF (4 mL) to the resin and stirred intermittently for 1 h. The solution was drained and the resin washed with DMF (5 x 5 mL). Deprotection and coupling steps were repeated until the desired sequence was achieved. A TNBS test was used to verify each coupling (negative/colourless) and deprotection (positive/red) step, with steps repeated as necessary. The final *N*-terminal Fmoc-protecting group was removed by treatment of the resin with a solution of 20% piperidine in DMF (5 mL) for 10 min, and the solution was drained and the resin washed with DMF (3 x 4 mL). The peptide-loaded resin was then washed with DCM (3 x 5 mL) and dried by washing with diethyl ether (3 x 5 mL), and cleaved from the resin as described by *General Procedure for Cleavage and Isolation*. The peptides were purified by semi-preparative RP-HPLC using a Phenomenex Aeris Peptide C18 column (10 x 250 mm) over a gradient as specified in the individual compound section.

TNBS Test:[3] A small spatula of swollen resin was taken out and 1 drop each of TNBS (100 µL 5% w/v picrylsulfonic/trinitrobenzenesulfonic acid in H_2O added to 900 µL of DMF) and DIPEA solutions (100 µL in 900 µL of DMF) added and allowed to develop for 1 min. Clear/yellow beads indicated no free amine (negative), while red/orange beads indicated free amine was present (positive).

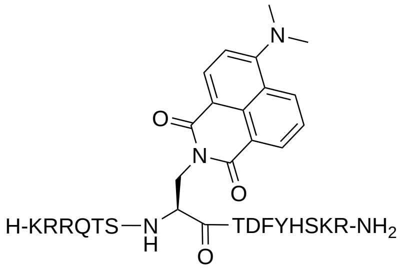
General procedure for cleavage and isolation

The peptide was cleaved from the resin by addition of 95:2.5:2.5 TFA/TIPS/H_2O (8 mL) to the resin and the suspension rocked for 1 h. The TFA solution was pipetted away 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 -20 °C. The precipitate was pelleted by centrifugation (7600 rpm, 10 min), and the supernatant was decanted. The pellet was dried under a nitrogen stream, then dissolved in 1:1 ACN/H_2O, syringe filtered (0.2 µm) and lyophilised.

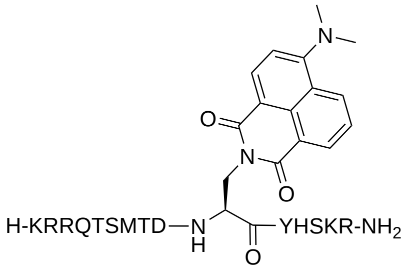
On-resin DMAP derivatisation[1]

The peptide sequence was assembled per '*General protocol for solid-phase peptide synthesis',* where the Nα of the final amino-acid 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). Dry DCM (5 mL) was then added to the resin and the mixture transferred to a sealed flask and purged with N_2. The alloc-protecting group was then selectively removed by treatment of the resin with $Pd(PPh_3)_4$ (0.8 equiv) and phenylsilane (25 equiv) 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 the Pd treatment repeated as before, twice more. The resin was then suspended in a solution of 4-DMAP (2 equiv) and DIPEA (4 equiv) in NMP (5 mL) and the mixture transferred to a sealed vessel purged with N_2, and bubbled with nitrogen overnight. The resin was then transferred to a fritted syringe and 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), and the HBTU/HOBt/DIPEA 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 diethyl ether (3 x 5 mL), and cleaved from the resin as described by *General Procedure for Cleavage and Isolation*. The peptides were purified by semi-preparative RP-HPLC as specified in the individual compound section.

Peptide ***p21**$_{141\text{-}155}$, H-KRRQTSMTDFYHSKR-NH$_2$.* The peptide was assembled per the '*General protocol for solid-phase peptide synthesis'* using commercial Fmoc-L-amino-acid building blocks. Following the final *N*-terminal Fmoc deprotection the peptide was cleaved from the resin per the '*General Protocol for Cleavage and Isolation'*. The peptide was then purified by semi-preparative RP-HPLC using a Phenomenex Luna C18(2) 250 x 10 mm column over a linear gradient of 10-30% (15 min) to give p21$_\mu$ as a white fluffy powder. HRMS (ESI+) Expected $[M+4H]^{4+}$ for $C_{82}H_{134}N_{30}O_{23}S$ (1938.9959): 485.7568, observed: $[M+4H]^{4+}$ 485.7579.

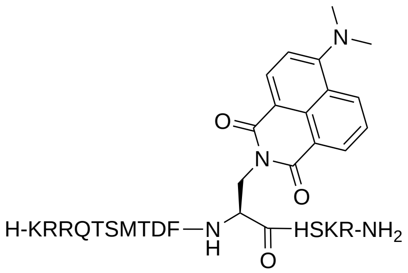


Peptide **1a**, *H-KRRQTSXTDFYHSKR-NH_2*. The peptide was assembled per the '*General protocol for solid-phase peptide synthesis*' using commercial Fmoc-L-amino-acid building blocks, and X was incorporated with use of Fmoc-DMNA (synthesised in-house) with 3 equiv, and DIPEA (6 equiv) in DMF (5 mL) for 3h. Following the final *N*-terminal Fmoc deprotection the peptide was cleaved from the resin per the '*General Protocol for Cleavage and Isolation'*. The peptide was then purified by semi-preparative RP-HPLC using a Phenomenex Luna C18(2) 250 x 10 mm column over a linear gradient of 10-40% over 20 min to give $p21_{\mu}$-147DMNA as a yellow fluffy powder. Purity (220 nm), 93%. HRMS (ESI+) Expected $[M+4H]^{4+}$ for $C_{94}H_{140}N_{32}O_{25}$ (2117.0667): 530.2745, observed: $[M+4H]^{4+}$ 530.2758.



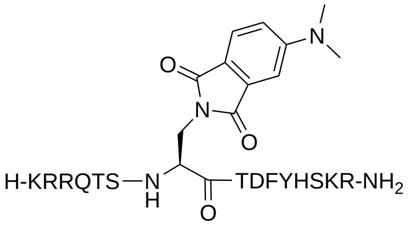
Peptide **2a**, *H-KRRQTSMTDXYHSKR-NH_2*. The peptide was assembled per the '*General protocol for solid-phase peptide synthesis*' using commercial Fmoc-L-amino-acid building blocks, and X was incorporated with use of Fmoc-DMNA, synthesised in-house. Following the final *N*-terminal Fmoc deprotection the peptide was cleaved from the resin per the '*General Protocol for Cleavage and Isolation*'. The peptide was then purified by semi-preparative RP-HPLC using a Phenomenex Luna C18(2) 250 x 10 mm column over a linear gradient of 10-40% over 20 min to give $p21_{\mu}$-150DMNA as a yellow fluffy powder. Purity (220 nm), 91%. HRMS (ESI+) Expected $[M+4H]^{4+}$ for $C_{90}H_{140}N_{32}O_{25}S$ (2102.0388): 536.5175, observed: $[M+4H]^{4+}$ 526.5211.

*Note: Keep out of light where possible, significant methionine oxidation was noticed in crude mixture.

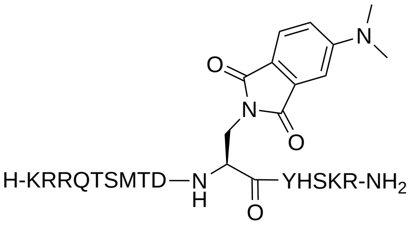


Peptide **3a**, *H-KRRQTSMTDFXHSKR-NH_2*. The peptide was assembled per the '*General protocol for solid-phase peptide synthesis*' using commercial Fmoc-L-amino-acid building blocks, and X was incorporated with use of Fmoc-DMNA, synthesised in-house. Following the final *N*-terminal Fmoc deprotection the peptide was cleaved from the resin per the '*General Protocol for Cleavage and Isolation*'. The peptide was then purified by semi-preparative RP-HPLC using a Phenomenex Luna C18(2) 250 x 10 mm column over a linear gradient of 25-50% over 25 min to give $p21_{\mu}$-151DMNA as a yellow fluffy powder. Purity (220 nm), 93%. HRMS (ESI+) Expected $[M+4H]^{4+}$ for $C_{90}H_{140}N_{32}O_{24}S$ (2085.0439): 522.2688, observed: $[M+4H]^{4+}$ 522.2707.

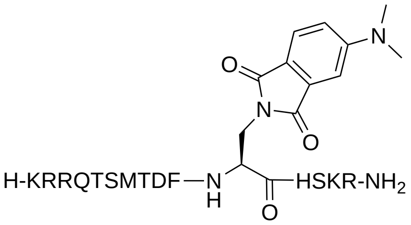
*Note: Keep out of light where possible, significant methionine oxidation was noticed in crude mixture.



Peptide **1b**, *H-KRRQTSBTDFYHSKR-NH*$_2$. The peptide was assembled per the '*General protocol for solid-phase peptide synthesis*' using commercial Fmoc-L-amino-acid building blocks, and 4-DMAP was incorporated using the '*On resin derivatisation method*'. The peptide was cleaved from the resin per the '*General Protocol for Cleavage and Isolation'*. The peptide was then purified by semi-preparative RP-HPLC using a Phenomenex Luna C18(2) 250 x 10 mm column over a linear gradient of 25-50% (15 min) to give p21$_μ$-147DMAP as a yellow fluffy powder. Purity (220 nm), 90%. HRMS (ESI+) Expected $[M+4H]^{4+}$ for $C_{90}H_{138}N_{32}O_{25}$ (2067.0511): 517.7706, observed: $[M+4H]^{4+}$ 517.7707.



Peptide **2b**, *H-KRRQTSMTDBYHSKR-NH*$_2$. The peptide was assembled per the '*General protocol for solid-phase peptide synthesis'* using commercial Fmoc-L-amino-acid building blocks, and 4-DMAP was incorporated using the '*On resin derivatisation*'. The peptide was cleaved from the resin per the '*General Protocol for Cleavage and Isolation'*. The peptide was then purified by semi-preparative RP-HPLC using a Phenomenex Luna C18(2) 250 x 10 mm column over a linear gradient of 25-50% (15 min) to give p21$_μ$-150DMAP as a yellow fluffy powder. Purity (220 nm), 95%. HRMS (ESI+) Expected $[M+4H]^{4+}$ for $C_{86}H_{138}N_{32}O_{25}S$ (2051.0232): 513.7636, observed: $[M+4H]^{4+}$ 513.7640.



Peptide **3b**, *H-KRRQTSMTDFBHSKR-NH*$_2$. The peptide was assembled per the '*General protocol for solid-phase peptide synthesis'* using commercial Fmoc-L-amino-acid building blocks, and 4-DMAP was incorporated using the '*On resin derivatisation*'. The peptide was cleaved from the resin per the '*General Protocol for Cleavage and Isolation'*. The peptide was then purified by semi-preparative RP-HPLC using a Phenomenex Luna C18(2) 250 x 10 mm column over a linear gradient of 25-50% (15 min) to give p21$_μ$-151DMAP as a yellow fluffy powder. Purity (220 nm), 90%. HRMS (ESI+) Expected $[M+4H]^{4+}$ for $C_{86}H_{138}N_{32}O_{24}S$ (2035.0282): 509.7649, observed: $[M+4H]^{4+}$ 509.7654.

S8.1.4 Plate reader experiments

Fluorescence characterisation was carried out on a BioTeK H4 Synergy Plate Reader with a Xenon light source.

Fluorescence characterisation, solvent comparison

The buffer for this experiment was 20 mM Tris.HCl pH 7.4 with 100 mM NaCl, 1 mM EDTA, 2 mM DTT and 0.05% Tween20. Peptides **1a**-**3a** and **1b**-**3b** (5 μM) were each dissolved in buffer, or dioxane with 5mM 18-crown-6. The resulting 12 solutions were then each plated in triplicate (60 μL/well) in a 96-well black clear-bottomed plate (CoStar 3905) and the fluorescence spectra were then immediately collected. An absorbance spectrum (300-700 nm) was collected for all samples. An emission spectrum for the DMNA-containing peptides (**1a**-**3a**) was collected (460-700 nm) with ex. 440 nm, and an excitation spectrum was collected (300-520 nm) with em. 540 nm. An emission spectrum for the DMAP-containing peptides (**1b**-**3b**) was collected (410-700 nm) with ex. 390 nm, and a excitation spectrum was collected (300-500 nm) with em. 520 nm. The measurements were collected from the top of the plate, and a step of 5 nm used.

*Fluorescence response of **1a**-**3a**, **1b**-**3b** in the presence of PCNA*

The buffer for this experiment was 20 mM Tris.HCl pH 7.4 with 100 mM NaCl, 1 mM EDTA, 2 mM DTT and 0.05% Tween20. Peptides **1a**-**3a** and **1b**-**3b** (1.6 µM) were each dissolved in buffer, and mixed in a 1:2 ratio with PCNA in buffer (1.2 µM), The resulting 6 solutions were then each plated in triplicate (60 µL/well) in a 96-well black plate (CoStar 3915), additionally a sample of each peptide (**1a**-**3a** and **1b**-**3b**) in buffer (800 nM) and PCNA in buffer (600 nM), and buffer alone were also plated in triplicate and incubated at 4 °C overnight. An emission spectrum for the DMNA-containing peptides (**1a**-**3a**), also the PCNA-only and buffer-only wells, was collected (460-700 nm) with ex. 440 nm. An emission spectrum for the DMAP-containing peptides (**1b**-**3b**), also the PCNA-only and buffer-only wells, was collected (410-700 nm) with ex. 390 nm. The measurements were collected from the top of the plate, and a step of 10 nm used.

PCNA titration

The buffer used for this experiment was 10 mM HEPES, 150 mM NaCl, 3 mM EDTA, 0.05% Tween20 at pH 7.4, adjusted with 2 M NaOH. Peptides **3a** and **3b** were prepared at 6 µM in buffer (600 µL). PCNA in buffer (120 µL) was prepared at each of the following concentrations: 0.5, 1, 2, 4, 6, 10, 15, 20, 30, 40 and 90 µM (120 µL each). Aliquots of either peptide **3a** or **3b** (50µL) were then combined with PCNA at each concentration (50 µL) and incubated at rt for 15 min, to give 11 samples of **3a** with PCNA, and 11 samples of **3b** with PCNA. The final concentration of peptide was 3 µM and the PCNA concentration was 0.25, 0.5, 1, 2, 3, 5, 7.5, 10, 15, 20 or 45 µM. The peptide/PCNA solution was plated in triplicate (30 µL/well) in a 384-well black, clear bottom plate (Grenier Bio-One, 781209). Additionally, PCNA (20 µM), **3a** (3 µM), **3b** (3µM) and buffer alone were plated in triplicate (30 µL/well). A full fluorescence spectrum with a 5 nm step was then collected for all samples 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 reads were collected from the top of the plate with a slit width of 9.0.

S8.1.5 SPR protocol

The running buffer used for ligand attachment and analyte binding experiments was 10 mM HEPES buffer with 150 mM NaCl, 3 mM EDTA and 0.05% Tween20, adjusted to pH 7.4 with 2M NaOH. A GE CM5 (series S) sensor chip was primed with running buffer and preconditioned per the manufacturer's recommendation with successive 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, respectively. The surface was then activated with an injection of 0.2 M EDC and 50 mM NHS (600 s, 10 µL/min). PCNA (5 µL, 12 mg/mL) was diluted into running buffer (245 µL). Only once the preactivation was complete was the protein further diluted to a final concentration of 25 µg/mL in 10 mM NaAc (~pH 4.6) by addition of PCNA/HEPES (50 µL) to a solution of 100 mM NaAc (50 µL) and water (400 µL). This solution was immediately injected over only one flow cell (10 µL/min) until ~1500 RU was reached at stabilisation. Both flow cells were then blocked with 1.0 M ethanolamine pH 8.5 (600 s, 10 µL/min). The chip was left to stabilise for two hours before sample injections commenced and a final chip loading of 1390 RU of protein was achieved.

Peptides (approx. 2 mg by weight) were dissolved in milliQ H2O (50 µL) and centrifuged (7800 rpm, 10 min) to remove any particulate. The peptide stock concentration was determined by UV absorbance (Aλ), where 2 µL of the stock was further diluted in water (20-40 fold) and a measurement taken in triplicate with a Nanodrop2000 and baselined to 750 nm absorbance. The ε_{440} used for DMNA containing peptides and ε_{420} used for DMAP containing peptides was reported by Loving[1]. The peptide stock solution concentration was then calculated per $c = (A_\lambda/\varepsilon_\lambda.l).DF$ where concentration is in molar, Aλ is absorbance at λ nm calculated as an average of three readings, λ is the appropriate wavelength (here, 440 or 420 nm), l is the pathlength in cm (0.1 cm for Nanodrop), and ε_λ is the molar absorptivity at λ nm and DF is the dilution factor. The peptides were then diluted into running buffer before further dilution as necessary.

Steady state affinity experiments were conducted at a flow rate of 30 µL/min, with a starting contact time of 40 s and dissociation of 60 s, and extended if a steady state could not be reached. All peptides were serially diluted 1 in 2, 8-times, and injected sequentially from lowest to highest concentration, preceded by a buffer only blank injection. The top concentration for each sample is listed in Table S3. Following each injection, the surface was regenerated with 2 M NaCl (2 x 30 s, 30 µL/min). After an optimal concentration range was found the series of injections were repeated to ensure reproducibility. All data was analysed using the GE Biosystems software Biacore S200 Evaluation software. All data are summarised in Table S3.

S8.2 Supplementary Information

S8.2.1 Fluorescence Data

Table S1: Fluorescence Data summary. These data are an average of three replicates. The fluorescence spectra were first smoothed in GraphPad Prism with 10 nearest neighbours and then the maximum emission wavelength determined.

Peptide	Experiment	Condition	Ex. Wavelength (nm)	Max Em. Wavelength (nm)	Fluorescence (arb. units)	Fold-change	Shift in Em Wavelength (nm)
1a	Solvent	Tris	440	565	758	40.4	-20
		Dioxane + crown	440	545	3066		
2a		Tris	440	560	227	15.6	-15
		Dioxane + crown	440	545	3541		
3a		Tris	440	555	270	23.2	-15
		Dioxane + crown	440	540	6264		
1b		Tris	390	585	93	300	-70
		Dioxane + crown	390	515	27987		
2b		Tris	390	580	234	106	-65
		Dioxane + crown	390	515	24866		
3b		Tris	390	585	63	199	-65
		Dioxane + crown	390	520	12529		
buffer	PCNA	-	440	425	1392	-	-
PCNA		-	440	425	1435	-	-
1a		peptide only	440	560	1585	6.54	-5
		+PCNA	440	555	10370		
2a		peptide only	440	545	387	14.5	-5
		+PCNA	440	540	5616		
3a		peptide only	440	560	494	7.04	-10
		+PCNA	440	550	3485		
buffer		-	390	410	1666	-	-
PCNA		-	390	410	1651	-	-
1b		peptide only	390	555	349	5.50	+20
		+PCNA	390	575	1919		
2b		peptide only	390	570	550	3.91	+5
		+PCNA	390	575	2153		
3b		peptide only	390	560	299	7.91	-35
		+PCNA	390	525	2365		
3a	PCNA Titration	0 PCNA	440	545	1149	10.4	-10
		Sat – avg >7.5eq	440	535	11950		
3b		0 PCNA	420	525	1006	3.45	-5
		Sat – avg > 7.5 eq	420	520	3475		

Table S2: Fit of fluorescence data from PCNA titration, calculated in GraphPad Prism 9 using a One site Total binding model

One site -- Total	**3a**	**3b**
Best-fit values		
Bmax	468.4	479.8
Kd	3.466	3.564
NS	-2.069	-1.066
Background	42.78	168.6
95% CI (profile likelihood)		
Bmax	411.7 to 540.7	374.0 to 671.1
Kd	2.430 to 4.992	1.668 to 7.797
NS	-3.726 to -0.6150	-4.915 to 1.748
Background	22.36 to 62.37	127.0 to 206.0
Goodness of Fit		
Degrees of Freedom	32	32
R squared	0.9711	0.9204
Sum of Squares	17574	61098
Sy.x	23.43	43.70
Number of points		
# of X values	36	36
# Y values analyzed	36	36

S8.2.2 SPR data

Table S3: SPR data. ε_λ and λ are used to calculate the peptide stock concentration. Affinity, K_D SE and Chi^2 are determined using the Biacore S200 Evaluation software. SE is the standard error. The Highest Conc is the highest peptide concentration used, where 8 peptide concentrations were prepared as a 1 in 2 serial dilution from the highest peptide concentration. The contact/dissociation is the contact and dissociation times for each injection cycle in seconds.

Peptide	**λ**	**ε_λ**	**Affinity (K_D)**	**K_D SE**	**Chi^2**	**Highest Conc (nM)**	**Contact/Dissociation (s)**	**Conclusion**
p21$_{141-155}$		67860[4]	16.3 nM	0.516 nM	0.0855	500	40/60	
1a	450	8800[1]	>25 µM			20000	40/60	Non specific
2a	450	8800	> 20 µM			10000	40/60	Non specific
3a	450	8800	921.5 nM	120 nM	1.21	2000	40/60	
1b	420	6500	> 25 µM			20000	40/60	Non specific
2b	420	6500	4.610 µM	570 nM	3.02	10000	40/60	
3b	420	6500	1.282 µM	260 nM	0.0298	2500	40/60	
Non-specific interaction was deemed on the basis of the sensorgrams continuing to increase with contact time, and as the maximum response is exceeded, where the maximum response is 103 RU for this PCNA-loaded chip.								

Sensorgrams

Maximum response expected:

RUmax = RUprotein*(Mpeptide/Mprotein) [Equation 1]

RUmax = 1390*(2000/27000)
= 103

The valency factor has been omitted as it is a 1:1 ratio for PCNAmonomer/peptide.

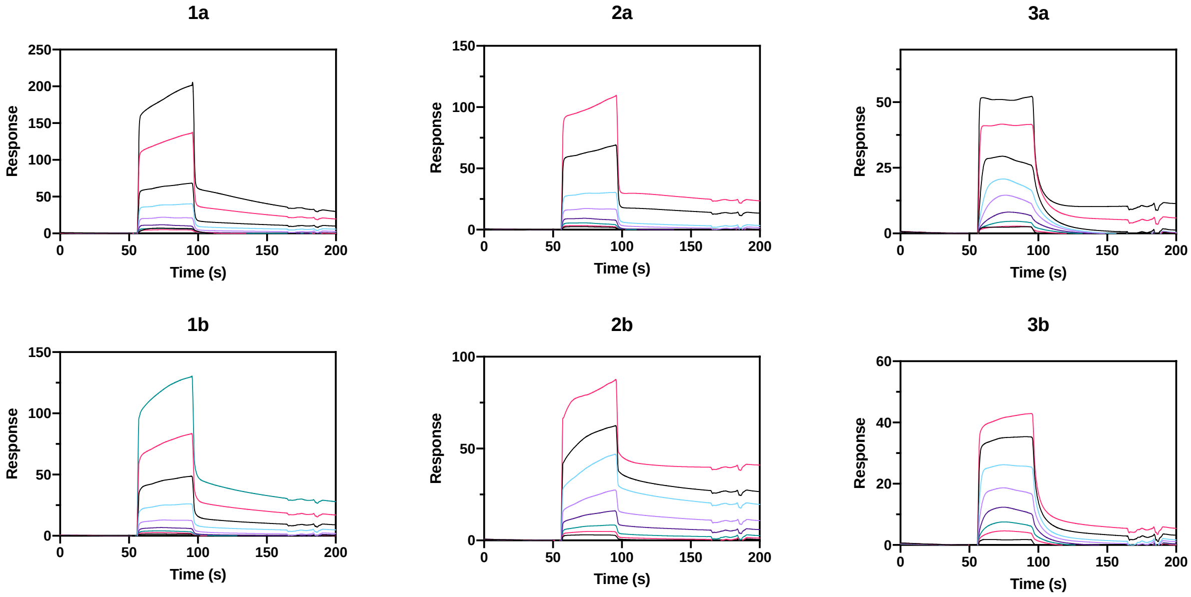


Figure S1: SPR sensorgrams of peptides **1a**-**3a** and **1b**-**3b** where the highest concentration (tallest sensorgram) for each peptide is listed in Table S2 and the subsequent curves are a 1 in 2 serial dilution. The maximum expected response is calculated using Equation 1.

REFERENCES

1. G. Loving and B. Imperiali, A Versatile Amino Acid Analogue of the Solvatochromic Fluorophore 4-N,N-Dimethylamino-1,8,naphthalimide: A Powerful Tool for the Stufy of Dynamic Protein Interactions, *J. Am. Chem. Soc.*, 2008, **130**, 13630-13638.
2. M. Sainlos and B. Imperiali, Synthesis of anhydride precursors of the environment-sensitive fluorophores 4-DMAP and 6-DMN, *Nat. Protoc.*, 2007, **2**, 3219-3225.
3. W. S. Hancock and J. E. Battersvy, A new micro-test for the detection of incomplete coupling reactions in solid-phase peptide synthesis using 2,4,6-trinitrobenzene-sulphonic acid, *Anal. Biochem.*, 1976, **71**, 260-264.
4. N. J. Anthis and G. M. Clore, Sequence-specific determination of protein and peptide concentrations by absorbance at 205 nm, *Protein Sci.*, 2013, **22**, 851-858.

Chapter 9.

SUMMARY, FUTURE DIRECTIONS AND OUTLOOK

This thesis describes the development of methodology to introduce a new inherently fluorescent bimane-based peptide modification and has defined its scope as both a covalent constraint, or an unconstrained side-chain decoration, to stabilise secondary structure in peptides as a basis for protein-protein interaction inhibitors (Chapters 2 and 3). Secondly, this thesis has detailed significant work toward optimising the primary sequence and structure of peptides and peptidomimetic macrocycles that target the human sliding clamp protein, PCNA (Chapters 4-8).

In **Chapter 2**, a bimane-based peptide modification is presented that can stabilise secondary structure and is inherently fluorescent. This bifunctional peptide modification introduces an imaging tag early in the therapeutic optimisation process, such that the exact compound of interest can be analysed directly in a biological setting (e.g. in cells). In chapter 2, dibromobimane was reacted with a range of short cysteine-containing peptides to create macrocycles that range from 16 to 31 atoms in size. The bimane was introduced while the peptide was still attached to the resin, or in phosphate buffered solution. The secondary structure of the resulting peptides was studied by CD and NMR secondary shift analysis to reveal a range of helical structures, dependent on the size of the macrocycle. In particular, the *i*-*i*+4 bimane constrained pentapeptide stabilised an α-helical conformation. Three macrocyclic peptides were shown to be cell permeable, where the inherent blue fluorescence of the bimane was used to visualise the peptides in the cells by confocal microscopy. This demonstrates that additional modification, such as appending a fluorescein-tag, is not required to assess the bimane-modified peptides in a biological setting. Such modifications can alter the biophysical properties of a peptidomimetic which may impact the cell permeability, secondary structure or ability to bind a target protein.

Next, in **Chapter 3**, this methodology was applied to a 12 amino-acid peptide sequence known to target estrogen receptor alpha (ERα), in order to investigate how the bimane influences secondary structure in a longer peptide sequence. The chosen sequence adopts an α-helical conformation in order to embed three conserved leucine residues on the ERα surface. Dibromobimane or monobromobimane were shown to react selectively with a cysteine thiol/s using an Ellman's reagent test in combination with HPLC and MS. Seven bimane modified peptides were prepared and the resulting secondary structures defined by CD and NMR secondary shift analysis. The *i*-*i*+4 bimane constrained peptide stabilised a 3_{10}-helical structure in solution, and adopted an α-helix in the presence of 20% TFE, or when computationally modelled onto the ERα surface to mimic the biologically important binding conformation. This indicates that the *i*-*i*+4 bimane-constrained macrocycle maintains sufficient flexibility in order to adopt the requisite α-helical conformation for interaction with ERα. Chapter 3 also presents a second method to stabilise helical structure with the bimane group, where a bimane appended to a cysteine side-chain in an acyclic peptide stabilises helical structure in a sequence dependent manner. We hypothesise this helical stabilisation results from a non-covalent interaction between the bimane group (at position 7), and a tryptophan residue six residues away (at position 1), as the analogous peptide with a W1A modification is not helical. Helicity is also observed in a peptide with a tyrosine at position 1, or to a lesser extent a phenylalanine. Tryptophan, tyrosine and phenylalanine have all been reported to interact with the bimane and quench the bimane fluorescence, to provide a basis for distance mapping experiments in proteins. This fluorescence quenching effect was not observed for these peptides, however the helical peptides did display an increase in fluorescence when 20% TFE was added. This might suggest that the bimane and aromatic residue are positioned further apart when the peptide adopts a classical α-helix following addition of TFE, which is known to promote α-helical formation. This change in fluorescence is an interesting observation that requires further study, but may provide a useful fluorescence-based method to detect conformational changes of bimane-modified peptides. For example, bimane-modified peptides could be designed such that a conformational change, related to the peptide binding a target protein, gives rise to a change in fluorescence and in this way, detects the binding event. Future studies will investigate if the helical stabilisation between a bimane and aromatic residue, in particular a tryptophan, can be replicated other sequences where these residues are closer together (e.g. *i*, *i*+4 c.f. *i*, *i*+6) and if the helical stabilisation is dependent on the order of the bimane and aromatic residue in the peptide (e.g. *i*, *i*-6).

The interaction of the Chapter 3 peptides with ERα and the related biological activity will be the subject of future studies. Such peptides that inhibit ERα coactivator recruitment may have therapeutic utility in the treatment of ER-positive cancers. Appendix 1 summarises methodology to introduce the bimane-modification in peptides, and highlights the bimane's ability to influence secondary structure and act as a fluorescent tag for imaging. The bimane methodology was further utilised in Chapters 6 and 8, to develop of a peptidomimetic inhibitor of PCNA, to address the second aim in this thesis.

PCNA is the human sliding clamp protein that interacts with over 200 proteins to mediate DNA-replication and -repair, where it acts as a mobile docking platform for protein machinery to interact with the DNA. Consequently, PCNA is upregulated in a large number of cancers and presents as a target to develop a cancer therapeutic. Chapters 4-8 present studies to understand the secondary interactions and structure-activity relationship of peptides that bind PCNA with high affinity, toward developing a therapeutic inhibitor of PCNA.

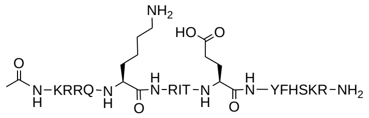
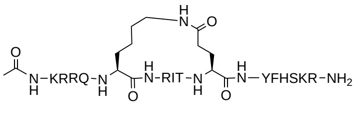
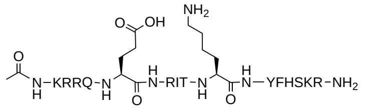
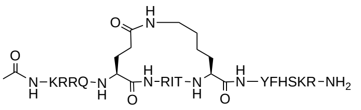
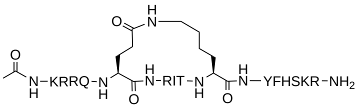
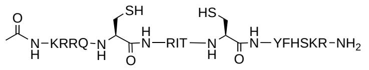
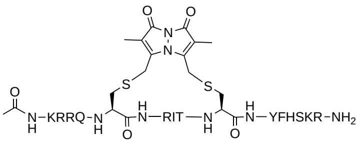
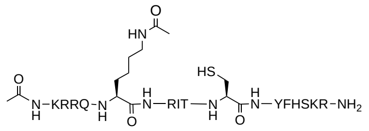
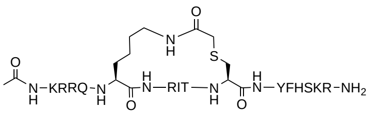
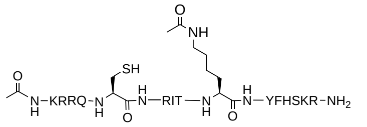
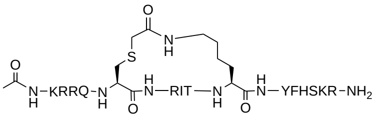
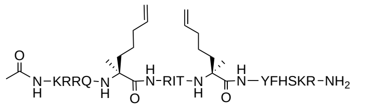
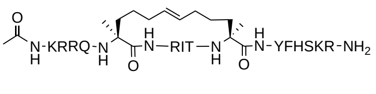
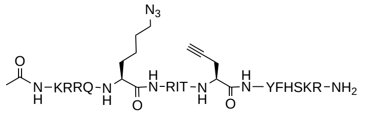
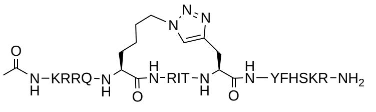
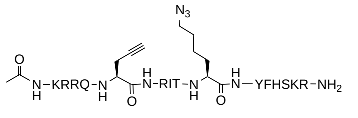
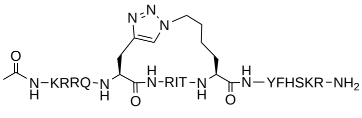
Chapter 4 provides a short literature review on the current knowledge of peptides, peptidomimetics and small molecules that target PCNA. This review highlights that research into PCNA as a therapeutic target is still quite young, but some key insights, particularly in relation to the structural elements necessary for binding PCNA are well characterised. It is known that the vast majority of proteins and peptides bind PCNA through a well-defined 3_{10}-helical turn, followed by a short β-strand. This 3_{10}-helical turn is the central element in the consensus sequence known as the PCNA-interacting protein (PIP) box, defined as Qxxφxxψψ where φ is a hydrophobic residue and ψ is an aromatic residue. Despite this, the PIP-box sequence of proteins known to bind PCNA is diverse, and we have endeavoured to unpack the subtleties in PIP-box sequence that fine tune the binding affinity for PCNA. One such observation is the complementary charged, or hydrogen bonding, pairs that are positioned either side of the hydrophobic residue at position 4 of the PIP-box, that appear in many high affinity PCNA partners. **Chapter 5** first seeks to understand the subtle differences between the binding-motif (PIP-box) of the many proteins that interact with PCNA, that give rise to large differences in affinity for PCNA. A peptide derived from the cell-cycle regulator protein, p21, was used as the starting template for this work as it binds PCNA with the highest known affinity. A series of 27 p21 peptides, with either one or two point modifications introduced in the PIP-box sequence (residues 144-151, QTSMTDFY), were prepared and the binding affinity for PCNA determined by SPR, to reveal 5 peptides with improved binding affinity over the native p21 peptide. A second series of 19 p21 peptides was prepared which incorporated PIP-box sequences of other proteins known to bind PCNA in order to sample a more diverse set of PIP-box sequence combinations, and the binding affinity for PCNA determined by SPR. Only 2 of these peptides bound PCNA with higher affinity than the native p21 sequence, reaffirming that nature has tuned the p21 PIP-box sequence for high PCNA binding affinity. Further analysis of the sequences of these two series of peptides revealed PIP-box sequence combinations that appear to correlate to high binding affinity for PCNA, such as an arginine at position 3 or a glutamic acid at position 6 of the PIP-box. A representative sample of 11 modified p21 peptides were selected and their structure when bound to PCNA was studied by either X-ray crystallography or computational modelling. This allowed identification of a series of hydrogen bonding networks that correlate to high PCNA binding affinity. In particular, an arginine at position 3 of the PIP-box hydrogen bonds to Ser43 on the PCNA surface, and a small polar residue at position 2 of the PIP-box interacts with Pro253 and Ala252 on the edge of the Q-pocket. A glutamic acid at position 5 of the PIP-box participates in a side-chain to main-chain intramolecular hydrogen bond to the PIP-box residues at positions 2 or 3, which appears more favourable when these are positive positively charged amino-acids. Conversely, a glutamic acid at position 6 of the PIP-box makes an intermolecular hydrogen bond to His44 on the PCNA surface. Though the number of hydrogen bonds (intra or intermolecular) does not strongly correlate with PCNA binding affinity, it would be interesting to rank the importance of each interaction to develop a scoring matrix that may predict PCNA binding affinity from PIP-box sequence. Here, our observations informed the design of three new PIP-box sequences that incorporated one or more of the interactions we predicted to be important for PCNA binding affinity. All three of these peptides were more potent than the native p21 sequence, and revealed a peptide with 1.21 nM affinity, the highest affinity PCNA-binding peptide reported to date that contains the PIP-box sequence QTRITEYF. This feat is particularly impressive given this peptide is seven amino-acids shorter than the previous title-holder, the native p21 sequence from residues 139-160.

Next, a series of constrained p21 peptides were prepared in **Chapter 6**, in order to stabilise a 3_{10}-helical binding conformation. Almost all peptides that bind PCNA adopt a single 3_{10}-helical turn, in p21 this is from residue 146 to 149. A p21 peptide (residues 141-154) was modified to include two cysteine residues at positions 145 and 149, and reacted with five different dihalo-reagents (1,3-dibromopropane, 1,4-dibromobutane, *trans*-1,4-dibromo-2-butene, dibromo-*m*-xylene and dibromobimane) to give five dicysteine *bis*-alkylated macrocycles. The binding affinity of these peptides for PCNA was determined by SPR, and spanned 570 nM – 3.86 μM. The most potent macrocycle was the bimane constrained p21 peptide which bound PCNA with a K_D value of 570 nM, in contrast to the cysteine-containing parent peptide which did not appear to interact with PCNA. The structure of the macrocyclic peptides bound to PCNA was studied by X-ray crystallography and computational modelling in order to determine if the 3_{10}-helical turn was stabilised by the linkers. These studies revealed that of the five macrocycles, only the bimane-constrained macrocycle was able to adopt the 3_{10}-helical binding conformation from residues 146-149. Two other peptides, constrained with a propyl or butyl linker, stabilised an unconventional 3_{10}-helical turn from residues 147-150. This indicates that the 3_{10}-helical binding conformation is integral to high affinity binding to PCNA, where the bimane-constrained macrocycle was the most potent binder, followed by the propyl-constrained macrocycle and then the butyl-contained macrocycle. The *m*-xylene constrained peptide, which did not have any main-chain hydrogen bonds, displayed much weaker affinity for PCNA at 3.84 µM. The conformation of the bimane constrained peptide was also studied in solution by NMR, which indicated there was only minor preorganisation of the peptide backbone into the 3_{10}-helix prior to binding PCNA, and suggests that the macrocycle is still flexible. This flexibility is likely key to the bimane-constrained peptide's high affinity for PCNA, where it allows the peptide sufficient leeway to adopt its preferred geometry on binding. Additionally, the bimane-constrained macrocycle was shown to be cell permeable in a breast cancer cell line, where the blue fluorescence of the bimane linker was visible in the cell cytoplasm by confocal microscopy. In contrast, a fluorescein-tagged linear p21 peptide was not cell permeable. This is a neat demonstration of how macrocyclic peptides often have improved cell uptake relative to their linear counterparts, and of the benefit of incorporating a fluorophore into the peptide structure, to enable imaging, during early structure-activity optimisation. The later addition of auxiliary fluorophores may negatively alter the biophysical properties of the peptidomimetic in development.

The next step in developing a potent peptidomimetic to target PCNA, for application as a cancer therapeutic, would be to combine the optimised PIP-box sequence (QTRITEYF) outlined in Chapter 5, and the benefits of peptide macrocycles using the information gained from the initial exploration of constrained p21 peptides in Chapter 6. Initial efforts toward this goal are briefly presented here. Nine p21 peptidomimetics were prepared, incorporating the sequence mutations shown to improve PCNA binding affinity. These macrocyclic peptidomimetics were constrained with a bimane linker, a metathesis linker, as well as the asymmetric lactam (Lys/Glu), 1,2,3-triazole (Lys-ε-N_3/Pra) and acyl-thioether (Lys-ε-Ac/Cys) linkers in two orientations (Table 1). The p21 sequence 141-155 was chosen, with the modified PIP-box (QTRITEYF) presented in Chapter 5, as it has been shown to bind PCNA with higher affinity than the short native sequence used in Chapter 6. The linear precursors were prepared by solid-phase peptide synthesis, and then cyclised on-resin using literature procedures, as detailed in the Experimental section. Two products were identified, of the desired product mass but different RP-HPLC retention times, following the cyclisation of the 5Glu/9Lys lactam peptide. The early eluting macrocycle was designated as **3.20** and the later eluting macrocycle as **3.21** (Table 1). The precise (structural) difference between peptides **3.20** and **3.21** remains under investigation, though initial 1D and 2D NMR analysis suggests both peptides are of the expected connectivity, which suggests these may be two stable conformers.

The binding affinity of the linear precursors, the macrocycles and the parent sequence (**1.00**) for PCNA was determined by SPR and the results are shown in Table 1. The affinity of **1.00** for PCNA was determined to be 2.48 nM, in line with the report in Chapter 5 where this sequence was identified. The lactam-cyclised

Table 1: Second generation p21 peptidomimetics with optimised PIP-box sequence and the SPR binding affinity for PCNA.

	Name	Affinity K_D (nM)	K_D SE (nM)		Name	Affinity K_D (nM)	K_D SE (nM)
1.00	Ac-KRRQTRITEYFHSKR-NH_2.	2.488	0.790				
2.10		115.8	21/0	2.20		8.627	2.09
3.10		19.03	4.40	3.20		217.0	46.0
				3.21		38.68	7.00
4.10		NS		4.20		91.24	14.0
5.10		NS		5.30		24.10	7.90
6.10		NS		6.30		71.12	14.0
7.10		1031	160	7.20		45.70	7.10
8.10		31.62	6.40	8.20		15.67	1.40
9.10		NS		9.20		94.26	2.20

peptide **2.20** was the most potent peptidomimetic at 8.63 nM, closely followed by the 1,2,3-triazole constrained peptide **8.20** (15.7 nM) and acyl-thioether constrained peptide **5.30** (24.1 nM). All of the linear peptides, expect the 5Glu/9Lys lactam precursor **3.10**, had lower affinity for PCNA than the corresponding cyclised analogue. This suggests that stabilisation of the secondary structure of the modified p21 peptides, by constraining two side-chains, generally leads to an increase in affinity for PCNA. Interestingly, in the macrocycles with an asymmetric bridge (lactam, triazole, acyl thioether), the analogue where the lysine residue is located on the *N*-terminal side of the constraint (**2.20**, **5.30** and **8.20**) binds PCNA with higher affinity than the reverse orientation (**3.20**/**3.21**, **6.30** and **9.20**). This suggests that the position of the polar atoms (e.g. carbonyl) within the linker is important for PCNA binding, presumably by providing additional contacts with the protein surface, or through additional intramolecular interactions. It is also known that the position of a rigid moiety within a peptide constraint (e.g. double bond, triazole etc) can influence the secondary structure of a macrocyclic peptide, and consequently alter target binding affinity. Future work will involve investigating the secondary structure of these peptides in solution, and their structure bound to PCNA in order to determine whether the constraints pre-define the key 3_{10}-helical binding conformation. Additionally, the proteolytic stability of the peptides will be characterised to determine whether the linkers are able to extend the half-life of macrocycles relative to the linear counterparts.

This initial investigation provides an important foundation for future work, and serves as a proof-of-concept that combining the individually optimised components of sequence (Chapter 5), length (Chapter 7), macrocyclization (Chapter 6) can cooperate synergistically to provide a rationally designed high affinity peptidomimetic to target PCNA. The optimal candidates will then be conjugated to an NLS sequence, informed by the knowledge from the study presented in Chapter 7.

Chapter 7 identified the shortest length of p21 peptide required to maintain high affinity binding for PCNA as residues 143-154 (102 nM), however the fluorescein-tagged derivative of this peptide was not cell permeable, in contrast to fluorescein-tagged $p21_{139-160}$. Truncated p21 peptides that included the positively charged residues (Arg/Lys) that flank the PIP-box (140-142 and 155-156) did not display significantly improved cell permeability, compared to $p21_{143-154}$ without these charged residues. This indicates that the hydrophobic tail (157-160, LIFS) likely contributes to the cell permeability of $p21_{139-160}$. Amphipathic peptides often display enhanced cell uptake which may explain this observation, or the hydrophobic tail may help the peptide associate with the cell membrane. Four different NLS-tags (from SV40, cMyc, and TAT, and the synthetic R6W3) were appended to the $p21_{143-154}$ peptide using a thiol-maleimide conjugation strategy, in an attempt to confer both cell and nuclear permeability to this short p21 peptide. Excitingly, the SV40-tagged peptide was evident in both the cell cytosol and nucleus. The NLS-tags were also appended to the cell permeable p21 macrocycle, constrained with a bimane group, that was identified in Chapter 6, to determine whether this could further enhance cell uptake. This series has the added benefit that both the p21 peptidomimetic, and the NLS-tag, are labelled with a fluorophore and thus it can be determined whether the two halves of the conjugate remain together. The SV40-tagged p21 bimane macrocycle was cell and nuclear permeable, where green and blue fluorescence of the fluorescein-tag was observed around and in the cell nucleus, though this was not as diffuse as for the linear p21-SV40 conjugate. In both the bimane macrocycle with a fluorescein, and bimane macrocycle tagged with a SV40-tag treated cells, there is notable more green fluorescence than blue fluorescence which may suggest some degradation of the peptide or cleavage of the maleimide linker in the case of SV40-tagged bimane macrocycle. However, this may also be an artefact of the imaging where the fluorescein fluorophore has notably higher quantum yield (i.e. is ‘brighter’) than the bimane. Interestingly, an analogue of the SV40-tagged p21 bimane macrocycle, that lacked the fluorescein-tag, did not appear to be cell permeable, as the blue bimane fluorescence was only observed in the cell cytoplasm. This suggests that the fluorescein-tag influences the cellular distribution of the peptidomimetic. This study provides a significant advance with the identification of the first nuclear permeable p21-derived peptidomimetics to target PCNA. Future work, which is currently underway at the time of writing, will characterise whether these peptides colocalise with PCNA, successfully inhibit DNA-replication and if there is an associated dose-response.

The third generation of peptidomimetics to target PCNA for application as a cancer therapeutic will move to incorporate the optimal peptidomimetic scaffolds that have been presented here in Chapter 9 (e.g. lactam **2.20**, acyl-thioether **8.20** or bimane **4.20**) and conjugate these structures to an SV40-tag (as done in Chapter 7) to develop nuclear permeable peptidomimetics with high affinity for PCNA. Additionally, the stability of the maleimide-thioether bond that links the p21 peptidomimetic to the NLS-tag will also be characterised, to ensure the conjugate remains intact until the p21 peptidomimetic is delivered to the nucleus. Alternate linkers will also be investigated, such as a lactam, or photocleavable groups such as the *ortho*-nitrobenzyl linkers to allow selective release of the peptidomimetic from the NLS-tag. Further modification of this scaffold will likely be required to improve the proteolytic stability of the peptidomimetic, especially due to the large number of arginine residues present which are readily targeted by proteases. Some such modifications include backbone *N*-methylation and introducing bioisoteres in place of the arginine side-chains. After a further round of optimisation and demonstration that these compounds are able to successfully interfere with DNA-replication processes in cells, the optimal candidates will be screened for selectivity between different cancers (e.g. breast vs. prostate) and with healthy cells; and then also tested in patient-derived xenograph (PDX) models to identify key clinical candidates in a clinically relevant model. The studies presented in this thesis (Chapters 4-7) have provided a solid foundation and significant advance in development of a peptide-based compound that targets PCNA, for use as a pre-clinical cancer therapeutic.

Lastly, **Chapter 8** presents studies on developing a peptide-based fluorescent probe for the detection of PCNA. Two solvatochromic amino-acids, 4-DMNA and 4-DAPA were incorporated at one of three positions (147, 150 and 151) within the $p21_{141\text{-}155}$ sequence, which are known to insert into a hydrophobic cleft on binding PCNA. The resulting six peptides gave enhanced fluorescence in the hydrophobic solvent dioxane with 5 mM 18-crown-6 when compared to fluorescence in Tris buffer, which demonstrates the solvatochromic action of the peptides. Additionally, all peptides showed a large increase in fluorescence when in solution with PCNA, compared to buffer alone, which indicates that the peptides gave a significant fluorescence response due to the presence of PCNA. This suggests an interaction is occurring between PCNA and the peptides. The binding affinity of the six peptides for PCNA was determined by SPR and revealed that only the 151-substituted peptides interact specifically with PCNA at 921 nM and 1.28 µM. The 151-substituted DMNA peptide gave the largest fluorescence response of 10-fold, in the presence of 2.5 equivalents of PCNA, relative to the 151-substituted DAPA peptide at only 3-fold. These peptides provide a useful first step toward developing a turn-on fluorescent probe to monitor PCNA in a biological setting, such as to further characterise the role of PCNA in cancer. Future work will involve characterising the cell and nuclear permeability of these peptides, such that they can be deployed in cell-based assays to monitor PCNA levels.

Supplemental Data.

SUMMARY, FUTURE DIRECTIONS AND OUTLOOK

S9.1 Experimental

S9.1.1 General information

Unless otherwise indicated, all starting materials were purchased from commercial sources and used without further purification. All amino-acid building blocks used, unless otherwise specified: Fmoc-Lys(Boc)-OH, Fmoc-Arg(Pbf)-OH, Fmoc-Gln(Trt)-OH, Fmoc-Thr(tBu)-OH, Fmoc-Glu(tBu)-OH, Fmoc-Ile-OH, Fmoc-Tyr(tBu)-OH, Fmoc-Phe-OH, Fmoc-His(Trt)-OH, Fmoc-Ser(tBu)-OH, and Fmoc-Cys(Trt)-OH were purchased from Chem Impex Int'l Ltd.

High-resolution mass spectra were collected using an Agilent 6230 ESI-TOF LCMS. RP-HPLC solvents were (A) H_2O with 0.1% TFA and (B) ACN with 0.1% TFA. Purification was carried out by semi-preparative HPLC on a Gilson GX-Prep RP-HPLC system on a Phenomenex Aeris Peptide C18 (10 x 250 mm), over a gradient as specified in the individual compound sections at 4 mL/min and visualised at 220 and 254 nm. Purity of all compounds was confirmed by analytical RP-HPLC on an Agilent 1260 HPLC equipped with a Phenomenex Aeris Peptide XB-C18 3.6 µm column (250 x 4.6 mm) over a gradient of 5-50% B (15 min) at 1.5 mL/min and visualised at 220 nm. SPR assays were carried out on a GE Biosystems Biacore S200, and data analysed using the inbuilt S200 Evaluation software.

S9.1.2 Synthesis and characterisation

All peptides were synthesised by the Fmoc solid-phase peptide synthesis protocol detailed below, with all L-amino-acids (unless otherwise specified), and then *N*-terminally acetylated before cyclisation on-resin where appropriate. Peptides were subsequently cleaved from the resin (and simultaneously globally deprotected). See Schemes 1-5 and the following general procedures.

General solid-phase peptide synthesis protocol

Rink Amide PL resin (0.2 mmol, 644 mg, 0.31 mmol/g, Agilent) was swollen in 1:1 DMF/DCM (15 mL) for 15 min. The Fmoc-protecting group was removed by treatment of the resin with a solution of 20% piperidine in DMF (8 mL) for 15 min. The solution was drained and the resin washed with DMF (3 x 8 mL). Amino-acid couplings were achieved by addition of a solution of Fmoc-protected amino-acid (5 equiv), HCTU (5 equiv) and DIPEA (10 equiv) in DMF (8 mL) to the resin and stirred intermittently for 1 h. The solution was drained and the resin washed with DMF (5 x 8 mL). Deprotection and coupling steps were repeated until the desired sequence was achieved. A TNBS test was used to verify each coupling (negative/colourless) and deprotection (positive/red) step, with steps repeated as necessary. The final *N*-terminal Fmoc-protecting group was removed by treatment of the resin with a solution of 20% piperidine in DMF (8 mL) for 10 min, and the solution was drained and the resin washed with DMF (3 x 8 mL). The *N*-terminal amine was then protected with an acetyl functionality by reaction with acetic anhydride (470 µL) and DIPEA (870 µL) in DMF (10 mL) for 15 min. The peptide-loaded resin was then split in two (2 x 0.1 mmol portions), where one portion was washed with DCM (3 x 5 mL) and dried by washing with diethyl ether (3 x 5 mL), then cleaved from the resin as described by *General procedure for cleavage and isolation*. The second portion was subjected to a cyclisation protocol as described in the individual compound sections, then peptide was cleaved from the resin as described by *General procedure for cleavage and isolation*. The peptides were then purified by semi-preparative RP-HPLC using a Phenomenex Aeris Peptide C18 Column (10 x 250 mm) over a linear gradient as specified in the individual compound section, at 4 mL/min and visualised at 220 and 254 nm.

TNBS Test:[1] A small spatula of swollen resin was taken out and 1 drop each of TNBS (100 µL 5% w/v picrylsulfonic/trinitrobenzenesulfonic acid in H_2O, added to 900 µL of DMF) and DIPEA solutions (100 µL in 900 µL of DMF) added and allowed to develop for 1 min. Clear/yellow beads indicated no free amine (negative), while red/orange beads showed free amine was present (positive).

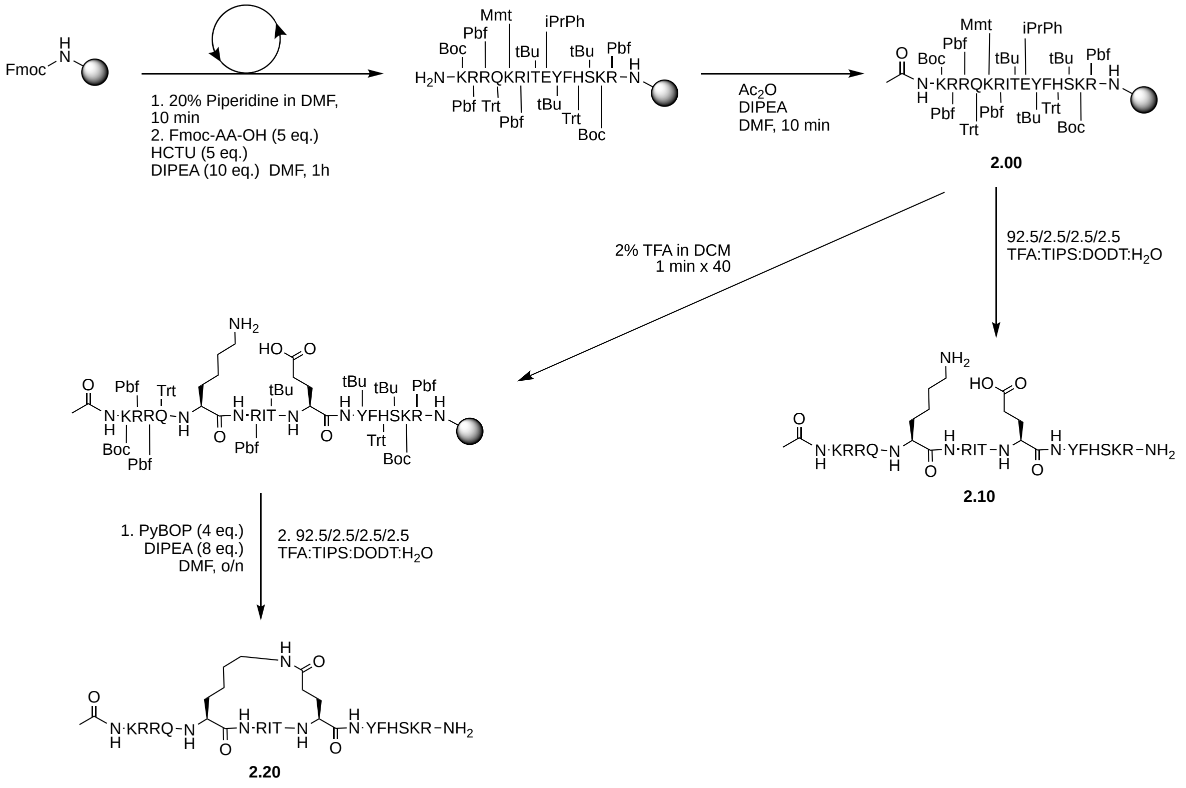
General procedure for cleavage and isolation:

The peptide was cleaved from the resin by addition of 95:2.5:2.5 TFA/TIPS/H_2O (8 mL) to the resin and the suspension rocked for 1 h. The TFA solution was pipetted away 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 -20°C. The precipitate was pelleted by centrifugation (7600 rpm, 10 min), and the supernatant was decanted. The pellet was dried under a nitrogen stream, then dissolved in 1:1 ACN/H_2O, before being syringe filtered (0.2 µm) and lyophilised.

Ac-KRRQTRITEYFHSKR-NH_2.

Peptide 1.00. The linear peptide was assembled following the *General solid-phase peptide synthesis protocol*, then cleaved from the resin following the *General procedure for cleavage and isolation.* The peptide was then purified by semi-preparative RP-HPLC using a Phenomenex Aeris Peptide C18 Column (10 x 250 mm) over a linear gradient of 25-50% B over 15 min (R_t = 5.5 min). Pure fractions were combined and lyophilised to give the final purified peptide **1.00** as a white fluffy powder. HRMS (ESI+) Expected $[M+5H]^{5+}$ for $C_{89}H_{147}N_{33}O_{23}$ (2046.1348): 410.2348, observed: $[M+5H]^{5+}$ 410.2244.

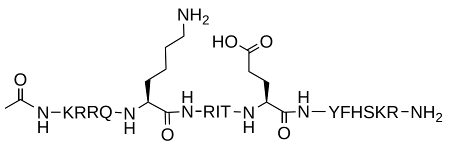
Lactam peptides


Fmoc
1. 20% Piperidine in DMF, 10 min
2. Fmoc-AA-OH (5 eq.)
HCTU (5 eq.)
DIPEA (10 eq.) DMF, 1h
H_2N-KRRQKRITEYFHSKR-N
Mmt
iPrPh
Pbf
Boc
tBu
Trt
Ac$_2$O
DIPEA
DMF, 10 min
2.00
2% TFA in DCM
1 min x 40
92.5/2.5/2.5/2.5
TFA:TIPS:DODT:H$_2$O
N·KRRQ-N
N·RIT-N
N·YFHSKR-N
N·YFHSKR-NH$_2$
HO
NH$_2$
2.10
1. PyBOP (4 eq.)
DIPEA (8 eq.)
DMF, o/n
2. 92.5/2.5/2.5/2.5
TFA:TIPS:DODT:H$_2$O
2.20

Scheme 1: General synthesis scheme for the synthesis of the lactam precursor peptides and lactam-constrained peptides

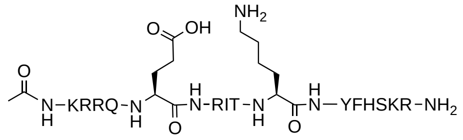
Lactam precursor peptides

The linear peptide was assembled following the *General solid-phase peptide synthesis protocol*, with the following modifications. K5 and E9 (**2.00**) or E5 and K9 (**3.00**) were introduced using Fmoc-Lys(Mmt)-OH and Fmoc-Glu(iPrPh)-OH. Following assembly of the linear sequence, the peptide was washed with DCM (3 x 5 mL) and dried by washing with diethyl ether (3 x 5 mL) then cleaved from the resin following the *General Procedure for Cleavage and Isolation.*


NH$_2$
HO
O
N-KRRQ-N
N-RIT-N
N-YFHSKR-NH$_2$

Peptide 2.10. Ac-KRRQKRITEYFHSKR-NH_2 The peptide was then purified by semi-preparative RP-HPLC using a Phenomenex Aeris Peptide C18 Column (10 x 250 mm) over a linear gradient of 10-40% B over 20 min (R_t=9.6 min).

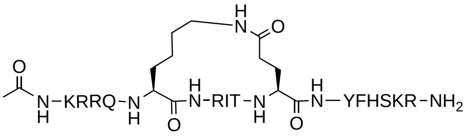
Pure fractions were combined and lyophilised to give the final purified peptide **2.10** as a white fluffy powder. HRMS (ESI+) Expected $[M+5H]^{5+}$ for $C_{91}H_{152}N_{34}O_{22}$ (2073.1820): 415.6442, observed: $[M+5H]^{5+}$ 415.6446.



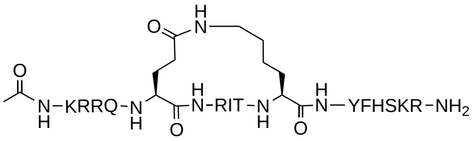
Peptide 3.10. Ac-KRRQERITKYFHSKR-NH2 The peptide was then purified by semi-preparative RP-HPLC using a Phenomenex Aeris Peptide C18 Column (10 x 250 mm) over a linear gradient of 10-40% B over 20 min (R_t=9.3 min). Pure fractions were combined and lyophilised to give the final purified peptide **3.10** as a white fluffy powder. HRMS (ESI+) Expected $[M+5H]^{5+}$ for $C_{91}H_{152}N_{34}O_{22}$ (2073.1820): 415.6442, observed: $[M+5H]^{5+}$ 415.6448.

On-resin lactamisation[2]

The linear peptide was assembled following the *General solid-phase peptide synthesis protocol*, with the following modifications. K5 and E9 (**2.00**) or E5 and K9 (**3.00**) were introduced using Fmoc-Lys(Mmt)-OH and Fmoc-Glu(iPrPh)-OH. Following assembly of the linear sequence, the Mmt and iPrPh groups were selectively removed with repetitive one minute treatments of 2% TFA in DCM (5 mL), and the resin washed with DCM (2 x 5 mL) between treatments. TFA treatments were repeated until the solution no longer turned yellow upon addition to the resin (~40x). The resin was then washed thoroughly with DMF (5 x 5 mL). The resin was transferred to a falcon tube and a solution of PyBOP (4 equiv, 208 mg) and DIPEA (4 equiv, 75 µL) in DMF (5 mL) added and the mixture rocked overnight (18.5 h). The solution was drained from the resin, and the resin washed with DMF (3 x 5 mL). A TNBS test was conducted and the cyclisation reaction repeated if a positive (red) test was returned. The peptide was then washed with DCM (3 x 5 mL) and dried by washing with diethyl ether (3 x 5 mL). The peptide was then cleaved from the resin following the *General Procedure for Cleavage and Isolation.*



Peptide 2.20 (cyclo-5,9)-Ac-KRRQK(-NHCO)RITE(-)YFHSKR-NH2 The peptide was then purified by semi-preparative RP-HPLC using a Phenomenex Aeris Peptide C18 Column (10 x 250 mm) over a linear gradient of 15-50% B over 15 min (R_t=7.2 min). Pure fractions were combined and lyophilised to give the final purified peptide **2.20** as white fluffy powder. HRMS (ESI+) Expected $[M+5H]^{5+}$ for $C_{91}H_{150}N_{34}O_{21}$ (2055.1715): 412.0421, observed: $[M+5H]^{5+}$ 412.0425.

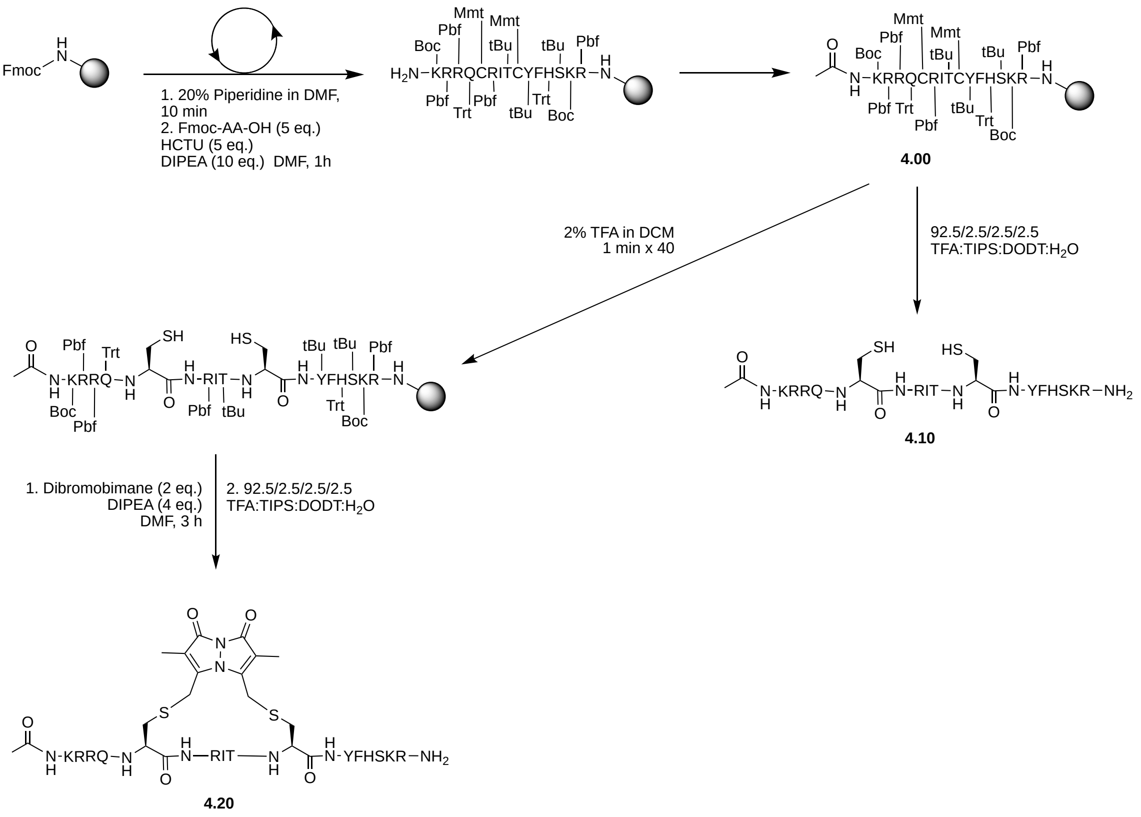


Peptide 3.2X: (cyclo-5,9)-Ac-KRRQK(-NHCO)RITE(-)YFHSKR-NH2 The peptide was then purified by semi-preparative RP-HPLC using a Phenomenex Aeris Peptide C18 Column (10 x 250 mm) over a linear gradient of 10-40% B over 20 min. Two peaks with the mass corresponding to the cyclised peptide were identified. Pure fractions were combined and lyophilised to give the final purified peptide **3.2X** as a white fluffy powder.

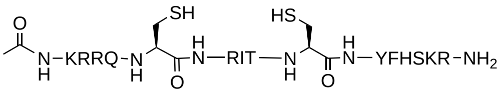
Peptide 3.20 (R_t=10.3 min) HRMS (ESI+) Expected $[M+5H]^{5+}$ for $C_{91}H_{150}N_{34}O_{21}$ (2055.1715): 412.0421, observed: $[M+5H]^{5+}$ 412.0448.

Peptide 3.21 (R_t=11.1 min) HRMS (ESI+) Expected $[M+5H]^{5+}$ for $C_{91}H_{150}N_{34}O_{21}$ (2055.1715): 412.0421, observed: $[M+5H]^{5+}$ 412.0434.

Bimane peptide

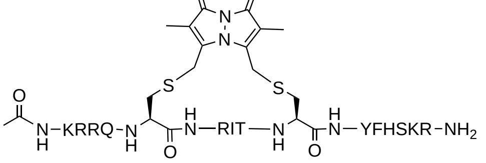

Fmoc
1. 20% Piperidine in DMF, 10 min
2. Fmoc-AA-OH (5 eq.)
HCTU (5 eq.)
DIPEA (10 eq.) DMF, 1h
4.00
2% TFA in DCM
1 min x 40
92.5/2.5/2.5/2.5
TFA:TIPS:DODT:H2O
4.10
1. Dibromobimane (2 eq.)
DIPEA (4 eq.)
DMF, 3 h
2. 92.5/2.5/2.5/2.5
TFA:TIPS:DODT:H2O
4.20

Scheme 2: General synthesis scheme for the synthesis of the cysteine precursor peptides and bimane-constrained peptides



Peptide 4.10. Ac-KRRQCRITCYFHSKR-NH_2 The linear peptide (**4.00**) was assembled following the *General solid-phase peptide synthesis protocol*, with the following modifications. C5 and C9 were introduced using Fmoc-Cys(Mmt)-OH. Following assembly of the linear sequence, the peptide was washed with DCM (3 x 5 mL) and dried by washing with diethyl ether (3 x 5 mL) then cleaved from the resin following the *General Procedure for Cleavage and Isolation.* The peptide was then purified by semi-preparative RP-HPLC using a Phenomenex Aeris Peptide C18 Column (10 x 250 mm) over a linear gradient of 10-40% B over 20 min (R_t=11.6 min). Pure fractions were combined and lyophilised to give the final purified peptide **4.10** as a white fluffy powder. HRMS (ESI+) Expected $[M+5H]^{5+}$ for $C_{86}H_{143}N_{33}O_{20}S_2$ (2022.0629): 405.4204, observed: $[M+5H]^{5+}$ 405.4191.

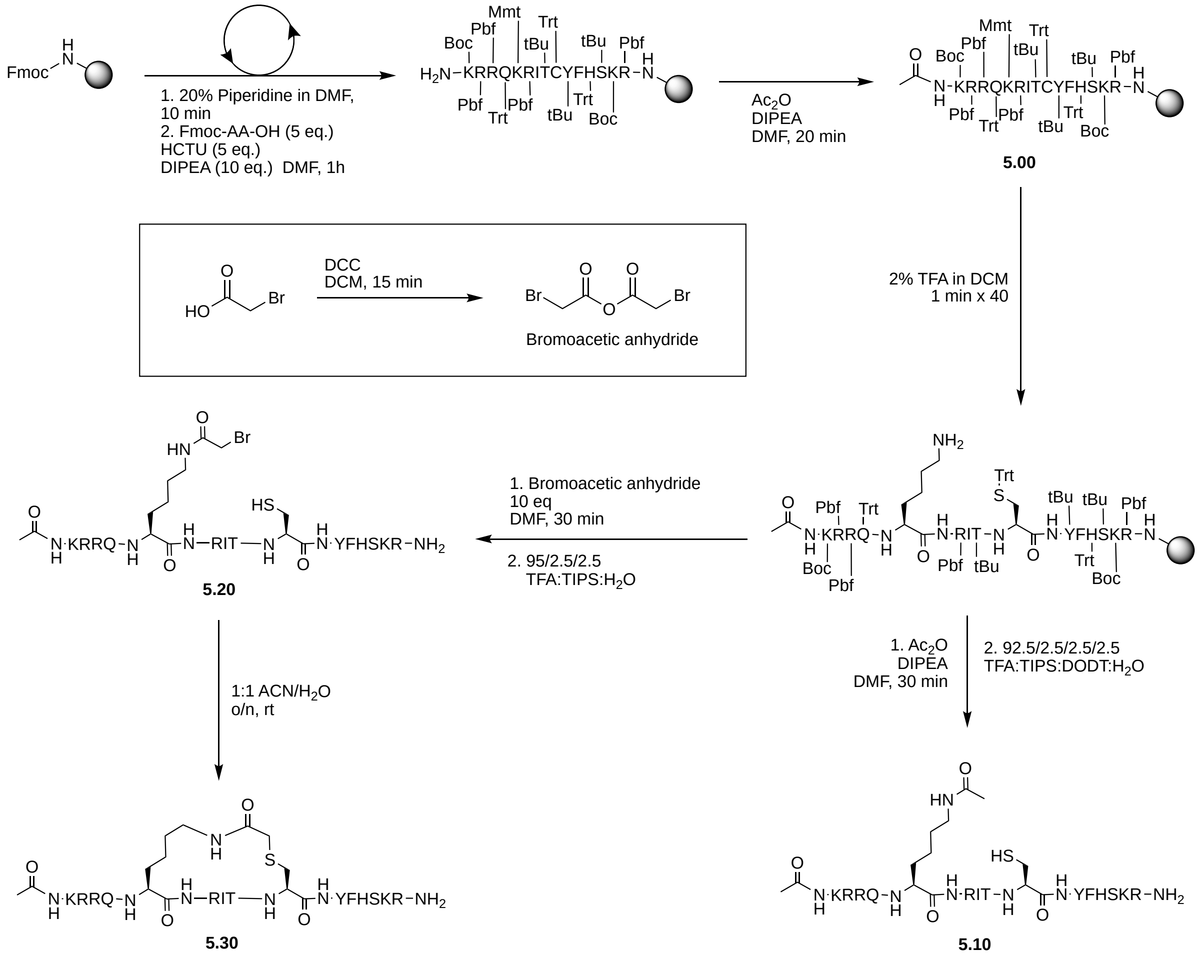
On-resin cyclisation with bimane linker[3]



Peptide 4.20 (cyclo-5,9)-Ac-KRRQC(Bim-)RITC(-)YFHSKR-NH_2 The linear peptide (**4.00**) was assembled following the *General Solid-Phase Peptide Synthesis Protocol*, with the following modifications. C5 and C9 were introduced using Fmoc-Cys(Mmt)-OH. Following assembly of the linear sequence, the Mmt and iPrPh groups were selectively removed with repetitive one minute treatments of 2% TFA in DCM (5 mL), and the resin washed with DCM (2 x 5 mL) between treatments. TFA treatments were repeated until the solution no longer turned yellow upon addition to the resin (~40x).

The resin was then washed thoroughly with DMF (5 x 5 mL). A solution of dibromobimane (2 equiv, 70 mg) and DIPEA (4 equiv, 174 µL) in DMF (6 mL) was added to the resin, and reacted for 3 h with intermittent stirring. The solution was then removed and the resin washed with DMF (5 x 5 mL) and DCM (5 x 5 mL), then dried with diethyl ether (3 x 5 mL). The peptide was then cleaved from the resin following the *General procedure for cleavage and isolation.* The peptide was then purified by semi-preparative RP-HPLC using a Phenomenex Aeris Peptide C18 Column (10 x 250 mm) over a linear gradient of 10-40% B over 20 min (R_t=16.7 min). Pure fractions were combined and lyophilised to give the final purified peptide **4.20** as a pale yellow fluffy powder. HRMS (ESI+) Expected $[M+5H]^{5+}$ for $C_{96}H_{151}N_{35}O_{22}S_2$ (2210.1214): 443.0321, observed: $[M+5H]^{5+}$ 443.0324.

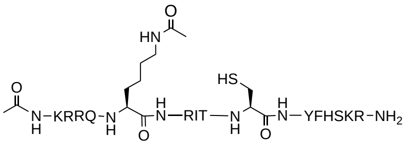
Thioether peptides


Fmoc
1. 20% Piperidine in DMF, 10 min
2. Fmoc-AA-OH (5 eq.)
HCTU (5 eq.)
DIPEA (10 eq.) DMF, 1h
H2N-KRRQKRITCYFHSKR-N
Mmt
Pbf
Boc
Trt
tBu
Ac2O
DIPEA
DMF, 20 min
5.00
2% TFA in DCM
1 min x 40
DCC
DCM, 15 min
Bromoacetic anhydride
1. Bromoacetic anhydride
10 eq
DMF, 30 min
2. 95/2.5/2.5
TFA:TIPS:H2O
5.20
1:1 ACN/H2O
o/n, rt
1. Ac2O
DIPEA
DMF, 30 min
2. 92.5/2.5/2.5/2.5
TFA:TIPS:DODT:H2O
5.30
5.10

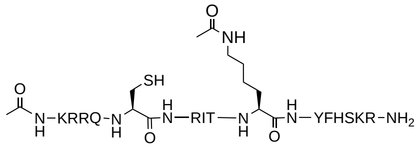
Scheme 3: General synthesis scheme for the synthesis of the acyclic *N*-acetyl peptides and thioether-constrained peptides.

Lysine acetylated peptides

The linear peptide was assembled following the *General solid-phase peptide synthesis protocol*, with the following modifications. K5 (**5.00**) or K9 (**6.00**) was introduced using Fmoc-Lys(Mmt)-OH. Following assembly of the linear sequence, the Mmt groups were selectively removed with repetitive one minute treatments of 2% TFA in DCM (5 mL), and the resin washed with DCM (2 x 5 mL) between treatments. TFA treatments were repeated until the solution no longer turned yellow upon addition to the resin (~20x). The resin was then washed thoroughly with DMF (5 x 5 mL). Quenching resin with DIPEA, and insufficient DMF washing, can result in di-functionalised (acetylated) amines. A solution of acetic anhydride (470 µL) and DIPEA (870 µL) in DMF (10 mL) was then added to the resin and left to react for 15 min. The solution was then removed and the resin washed with DMF (5 x 5 mL) and DCM (5 x 5 mL), then dried with diethyl ether (3 x 5 mL). Following assembly of the linear sequence, the peptide was washed with DCM (3 x 5 mL) and dried by washing with diethyl ether (3 x 5 mL) then cleaved from the resin following the *General procedure for cleavage and isolation.*



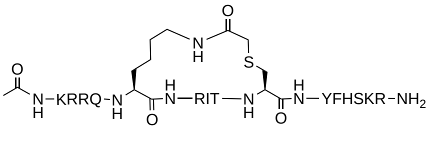
Peptide 5.10. Ac-KRRQK(Ac)RITCYFHSKR-NH$_2$ The peptide was then purified by semi-preparative RP-HPLC using a Phenomenex Aeris Peptide C18 Column (10 x 250 mm) over a linear gradient of 10-40% B over 20 min (R$_t$=10.7 min). Pure fractions were combined and lyophilised to give the final purified peptide **5.10** as a white fluffy powder. HRMS (ESI+) Expected [M+3H]$^{3+}$ for $C_{91}H_{152}N_{34}O_{21}S$ (2089.1592): 697.3942, observed: [M+3H]$^{3+}$ 697.3898.



Peptide 6.10. Ac-KRRQCRITK(Ac)YFHSKR-NH$_2$ The peptide was then purified by semi-preparative RP-HPLC using a Phenomenex Aeris Peptide C18 Column (10 x 250 mm) over a linear gradient of 10-40% B over 20 min (R$_t$=10.8 min). Pure fractions were combined and lyophilised to give the final purified peptide **6.10** as a white fluffy powder. HRMS (ESI+) Expected [M+3H]$^{3+}$ for $C_{91}H_{152}N_{34}O_{21}S$ (2089.1592): 697.3942, observed: [M+3H]$^{3+}$ 697.3845.

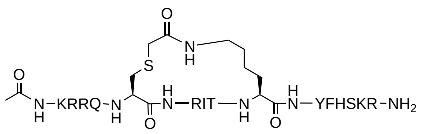
Bromoacetyl-modified peptides – precursor to thioether bridged peptides

The linear peptide was assembled following the *General solid-phase peptide synthesis protocol*, with the following modifications. K5 (**5.00**) or K9 (**6.00**) was introduced using Fmoc-Lys(Mmt)-OH. Following assembly of the linear sequence, the Mmt groups were selectively removed with repetitive one minute treatments of 2% TFA in DCM (5 mL), and the resin washed with DCM (2 x 5 mL) between treatments. TFA treatments were repeated until the solution no longer turned yellow upon addition to the resin (~20x). The resin was then washed thoroughly with DMF (5 x 5 mL). Quenching resin with DIPEA, and insufficient DMF washing, can result in di-functionalised (bromoacetylated) amines. A solution of bromoacetic acid (20 equiv, 2 mmol, 288 mg) and DCC (10 equiv, 1 mmol, 207 mg) were dissolved in DCM (5 mL) and stirred for 15 min. The resulting precipitate was removed by filtration. Solvent was removed from the filtrate *in vacuo* and the resulting residue (bromoacetic anhydride) was dissolved in DMF (6 mL), and immediately added to the resin for 30 min with intermittent stirring. The solution was then removed and the resin washed with DMF (3 x 5 mL), then DCM (5 x 5 mL), and dried with diethyl ether (3 x 5 mL).



Peptide 5.30. (cyclo-5,9)-Ac-KRRQK(COCH$_2$-)RITC(-)YFHSKR-NH$_2$ The peptide was then cleaved from the resin following the *General procedure for cleavage and isolation* to give **5.20** which was purified by semi-preparative RP-HPLC using a Phenomenex Aeris Peptide C18 Column (10 x 250 mm) over a linear gradient of 15-50% B over 15 min (R$_t$=6.7 min). Pure fractions* were combined and lyophilised, then redissolved in 50% aq. ACN and rocked overnight to allow reaction of the cysteine thiol with the bromoacetylated lysine residue to yield the final purified peptide, and then lyophilised to give **5.30** as a white fluffy powder.

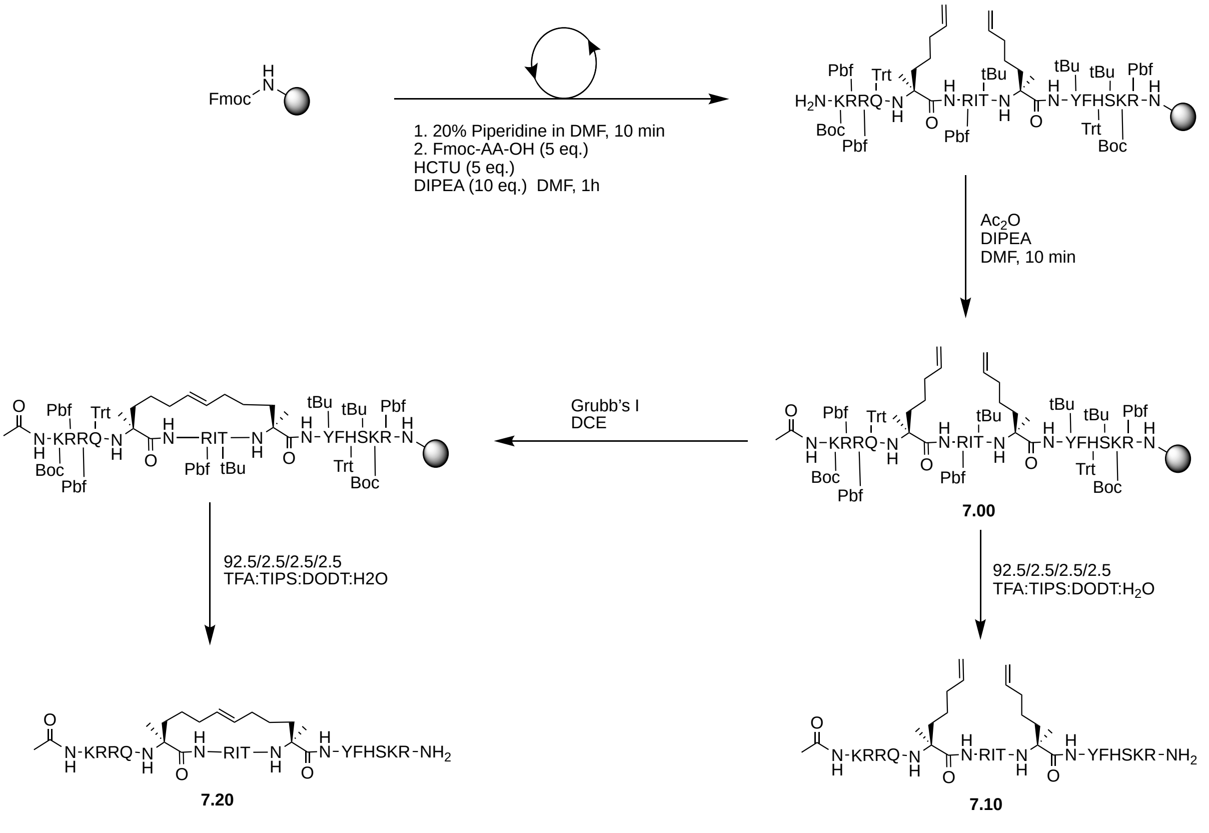
*Analysis of HPLC fractions by MS revealed some spontaneous cyclisation of modified lysine and cysteine; fractions containing pure uncyclized or cyclised peptide were combined together. HRMS (ESI+) Expected [M+3H]$^{3+}$ for $C_{91}H_{150}N_{34}O_{21}S$): 697.7223, observed: [M+3H]$^{3+}$ 697.7213.



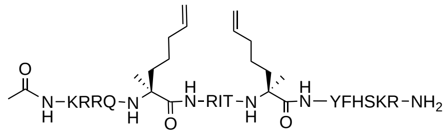
Peptide 6.30. (cyclo-5,9)-Ac-KRRQK(COCH$_2$-)RITC(-)YFHSKR-NH$_2$ The peptide was then cleaved from the resin following the *General procedure for cleavage and isolation* to give **6.20** which was purified by semi-preparative RP-HPLC using a Phenomenex Aeris Peptide C18 Column (10 x 250 mm) over a linear gradient of 10-40% B over 20 min (R_t=11.4 min). Pure fractions* were combined and lyophilised, then redissolved in 50% aq. ACN and rocked overnight to allow reaction of the cysteine thiol with the bromoacetylated lysine residue to yield the final purified peptide, and then lyophilised to give **6.30** as a white flaky powder.

*Analysis of HPLC fractions by MS revealed some spontaneous cyclisation of modified lysine and cysteine; fractions containing pure uncyclized or cyclised peptide were combined together. HRMS (ESI+) Expected $[M+3H]^{3+}$ for $C_{91}H_{150}N_{34}O_{21}S$): 697.7223, observed: $[M+3H]^{3+}$ 697.7274.

All-hydrocarbon peptides

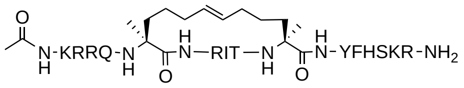


Scheme 4: General synthesis scheme for the synthesis of the all-hydrocarbon precursor peptides and metathesis all hydrocarbon-constrained peptides



Peptide 7.10. Ac-KRRQXRITXYFHSKR-NH$_2$ (where, X=S-pentenylalanine) The linear peptide (**7.00**) was assembled following the *General solid-phase peptide synthesis protocol*, with the following modifications. X5 and X9 were introduced using Fmoc-S-pentenylalanine-OH. Following assembly of the linear sequence, the peptide was washed with DCM (3 x 5 mL) and dried by washing with diethyl ether (3 x 5 mL), then cleaved from the resin following the

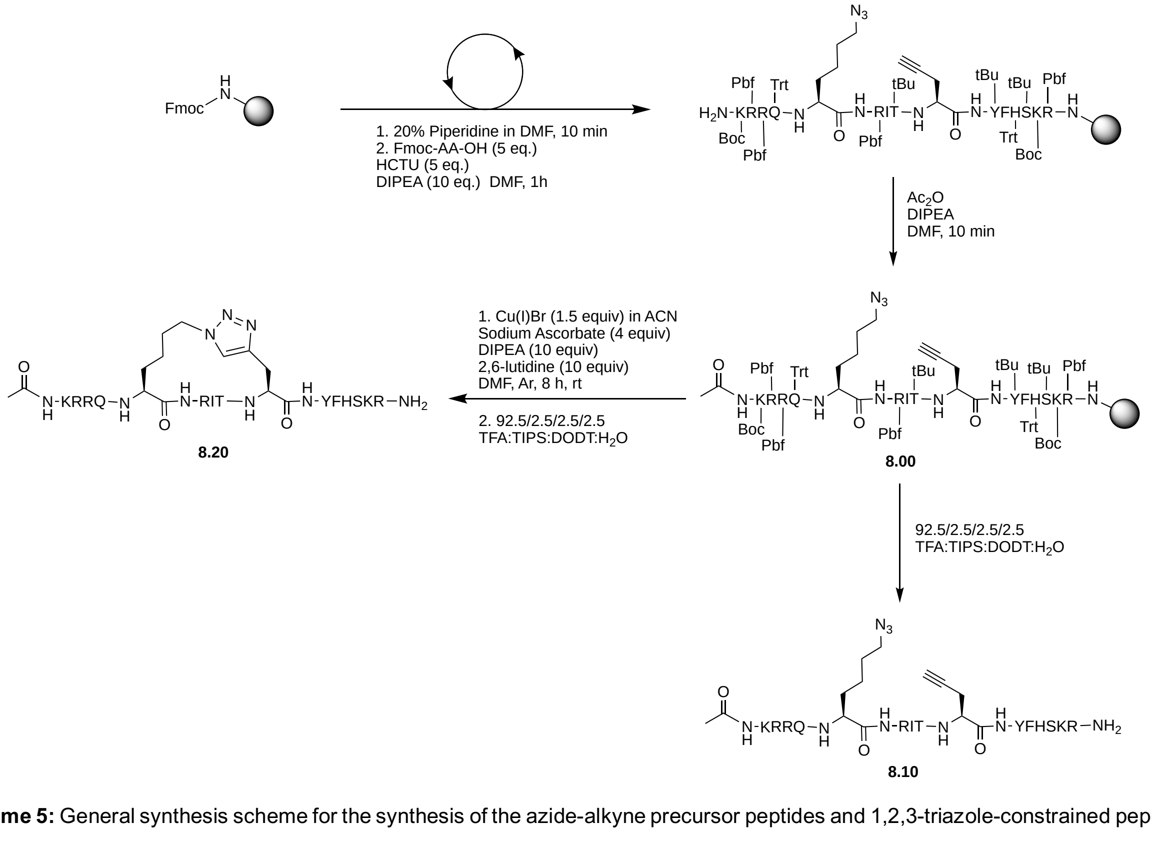
General procedure for cleavage and isolation. The peptide was then purified by semi-preparative RP-HPLC using a Phenomenex Aeris Peptide C18 Column (10 x 250 mm) over a linear gradient of 15-50% B over 25 min (R_t=9.8 min). Pure fractions were combined and lyophilised to give the final purified peptide **7.10** as a white fluffy powder. HRMS (ESI+) Expected $[M+4H]^{4+}$ for $C_{94}H_{155}N_{33}O_{20}$ (2066.2126): 517.5610, observed: $[M+4H]^{4+}$ 517.5596.


O
N-KRRQ-N
H
H
H
N—RIT—N
O
H
O
H
N-YFHSKR-NH2

On-resin metathesis[4]

Peptide 7.20. (cyclo-5,9)-Ac-KRRQX(-)RITX(-)YFHSKR-NH_2 Following assembly of the linear sequence (**7.00**), the resin was washed with DCM (3 x 5 mL) and dried by washing with diethyl ether (3 x 5 mL), then split into two 0.05 mmol portions. The remaining ether was allowed to evaporate overnight. The resin was then swelled in argon-degassed and sieve-dried DCE (3 mL) for 15 min, under an argon atmosphere, then Grubbs first generation catalyst (Grubbs Catalyst® M102, Sigma) (20 mol%, 8.26 mg) in degassed DCE (2 mL) was added and the mixture stirred at rt for 2 h. The resin was then removed and washed with DCM (4 x 5 mL), DMF (3 x 5 mL), DCM (3 x 5 mL) and dried by washing with diethyl ether (3 x 5 mL). A small cleave (one small spatula of resin) of the peptide from the resin was performed to determine if the reaction was complete (by MS and HPLC), and the cyclisation reaction repeated if necessary. The peptide was then cleaved from the resin following the *General procedure for cleavage and isolation.* The peptide was then purified by semi-preparative RP-HPLC using a Phenomenex Aeris Peptide C18 Column (10 x 250 mm) over a linear gradient of 15-50% B over 15 min (R_t=6.8 min). Pure fractions were combined and lyophilised to give the final purified E-isomer of peptide **7.20** as a white fluffy powder. HRMS (ESI+) Expected $[M+5H]^{5+}$ for $C_{93}H_{153}N_{33}O_{20}$ (2052.1970): 411.4472, observed: $[M+5H]^{5+}$ 411.4463.

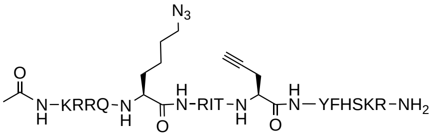
Azide + Alkyne peptides


Fmoc
H
N
1. 20% Piperidine in DMF, 10 min
2. Fmoc-AA-OH (5 eq.)
HCTU (5 eq.)
DIPEA (10 eq.) DMF, 1h
N3
Pbf Trt tBu tBu tBu Pbf
H2N-KRRQ-N
N-RIT-N
N-YFHSKR-N
Boc Pbf Pbf Trt Boc
Ac2O
DIPEA
DMF, 10 min
1. Cu(I)Br (1.5 equiv) in ACN
Sodium Ascorbate (4 equiv)
DIPEA (10 equiv)
2,6-lutidine (10 equiv)
DMF, Ar, 8 h, rt
2. 92.5/2.5/2.5/2.5
TFA:TIPS:DODT:H2O
N-KRRQ-N
N-RIT-N
N-YFHSKR-NH2
8.20
8.00
92.5/2.5/2.5/2.5
TFA:TIPS:DODT:H2O
8.10

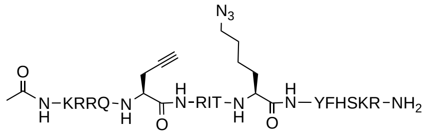
Scheme 5: General synthesis scheme for the synthesis of the azide-alkyne precursor peptides and 1,2,3-triazole-constrained peptides

Linear Azide/Alkyne Peptides

The linear peptide was assembled following the *General solid-phase peptide synthesis protocol*, with the following modifications. Lys(N_3)5 and Pra9 (**8.00**) and Pra5 and Lys(N_3)9 (**9.00**) were introduced using Fmoc-Lys-ε-azide-OH and Fmoc-L-propargylglycine-OH. Following assembly of the linear sequence, the peptide was washed with DCM (3 x 5 mL) and dried by washing with diethyl ether (3 x 5 mL) then cleaved from the resin following the *General Procedure for Cleavage and Isolation.*



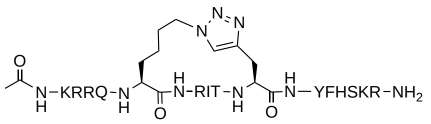
Peptide 8.10. Ac-KRRQK(N_3)RITPraYFHSKR-NH_2 The peptide was then purified by semi-preparative RP-HPLC using a Phenomenex Aeris Peptide C18 Column (10 x 250 mm) over a linear gradient of 15-50% B over 15 min (R_t=6.8 min). Pure fractions were combined and lyophilised to give the final purified peptide **8.10** as a white fluffy powder. HRMS (ESI+) Expected $[M+4H]^{4+}$ for $C_{91}H_{148}N_{36}O_{20}$ (2065.1671): 517.2996, observed: $[M+4H]^{4+}$ 517.2999.



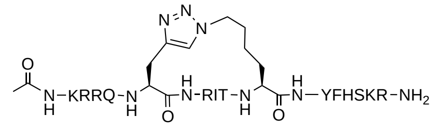
Peptide 9.10: Ac-KRRQPraRITK(N_3)YFHSKR-NH_2 The peptide was then purified by semi-preparative RP-HPLC using a Phenomenex Aeris Peptide C18 Column (10 x 250 mm) over a linear gradient of 10-40% B over 20 min (R_t=9.5 min). Pure fractions were combined and lyophilised to give the final purified peptide **9.10** as a white fluffy powder. HRMS (ESI+) Expected $[M+4H]^{4+}$ for $C_{91}H_{148}N_{36}O_{20}$ (2065.1671): 517.2960, observed: $[M+4H]^{4+}$ 517.2999.

On-resin Huisgen 1,3-dipolar cycloaddition[5]

Following assembly of the linear sequence (**8.00** or **9.00**), the resin was washed with DCM (3 x 5 mL) and dried by washing with diethyl ether (3 x 5 mL), then split into two 0.05 mmol portions. The remaining ether was allowed to evaporate overnight. The resin was then swelled in argon-degassed anhydrous DMF (3 mL) for 15 min, under an argon atmosphere. DIPEA (10 equiv, 87.1 µL) and 2,6-lutidine (10 equiv, 57.9 µL) were added to the resin, followed by a suspension of sodium ascorbate (4 equiv, 39.6 mg) in argon-degassed anhydrous DMF (2 mL), and finally a solution of Cu(I)Br (1.5 equiv, 10.8 mg) in acetonitrile (1.5 mL). The mixture was left to react under an argon-atmosphere stirred at rt for 8 h. The resin was then removed and washed with DMF (3 x 5 mL), MeOH (3 x 5 mL), H_2O (3 x 5 mL) and DCM (3 x 5 mL), then dried by washing with diethyl ether (3 x 5 mL). A small cleave (a small spatula of resin) of the peptide was performed to determine if the reaction had gone to completion (by HPLC), and the reaction repeated if necessary. The peptide was then cleaved from the resin following the *General Procedure for Cleavage and Isolation.*



Peptide 8.20. (cyclo-5,9)-Ac-KRRQK(-)RITPra(-)YFHSKR-NH_2 The peptide was then purified by semi-preparative RP-HPLC using a Phenomenex Aeris Peptide C18 Column (10 x 250 mm) over a linear gradient of 15-50% B over 15 min (R_t=6.4 min). Pure fractions were combined and lyophilised to give the final purified peptide **8.20** as a white fluffy powder. HRMS (ESI+) Expected $[M+4H]^{4+}$ for $C_{91}H_{148}N_{36}O_{20}$ (2065.1671): 517.2996, observed: $[M+4H]^{4+}$ 517.3000.



Peptide 9.20. (cyclo-5,9)-Ac-KRRQPra(-)RITK(-)YFHSKR-NH_2 The peptide was then purified by semi-preparative RP-HPLC using a Phenomenex Aeris Peptide C18 Column (10 x 250 mm) over a linear gradient of 15-50% B over 15 min (R_t=6.1 min). Pure fractions were combined and lyophilised to give the final purified peptide **9.20** as a white fluffy powder. HRMS (ESI+) Expected $[M+4H]^{4+}$ for $C_{91}H_{148}N_{36}O_{20}$ (2065.1671): 517.2996, observed: $[M+4H]^{4+}$ 517.3001.

S9.1.3 SPR Protocol

The running buffer used for ligand attachment and analyte binding experiments was 10 mM HEPES buffer with 300 mM NaCl, 3 mM EDTA and 0.05% Tween20, adjusted to pH 7.4 with 2M NaOH. A GE CM5 (series S) sensor chip was primed with running buffer and preconditioned per the manufacturer's recommendation with successive 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, respectively. The surface was then activated with an injection of 0.2 M EDC and 50 mM NHS (600 s, 10 µL/min). PCNA (5 µL, 12 mg/mL) was diluted into running buffer (245 µL). Only once the preactivation was complete was the protein further diluted to a final concentration of 25 µg/mL in 10 mM NaAc (~pH 4.6) by addition of PCNA/HEPES (50 µL) to a solution of 100 mM NaAc (50 µL) and water (400 µL). This solution was immediately injected over only one flow cell (10 µL/min) until ~1500 RU was reached at stabilisation. Both flow cells were then blocked with 1.0 M ethanolamine pH 8.5 (600 s, 10 µL/min). The chip was left to stabilise for two hours before sample injections commenced and a final chip loading of 1505 RU of protein was achieved.

Peptides (approx. 2 mg by weight) were dissolved in milliQ H_2O (50 µL) and centrifuged (7800 rpm, 10 min) to remove any particulate. The peptide stock concentration was determined by 205 nm absorbance (A_{205}), where 2 µL of the stock was further diluted in water (20-40 fold) and a measurement taken in triplicate with a Nanodrop2000 and baselined to 750 nm absorbance. The ε_{280} for each peptide was calculated using an online calculator (http://nickanthis.com/tools/a205.html),[6] or ε_{380} used for bimane containing peptides as reported in Shen 2011.[7] The peptide stock solution concentration was then calculated per $c = (A_\lambda/\varepsilon_\lambda.l).DF$ where concentration is in molar, A_λ is absorbance at λ nm calculated as an average of three readings, λ is the appropriate wavelength (here, 280 or 380 nm), l is the pathlength in cm (1 mm for Nanodrop), and ε_{205} is the molar absorptivity at 205 nm and DF is the dilution factor. The stock concentrations are tabulated in Table S2. The peptides were then diluted into running buffer before further dilution as necessary.

Steady state affinity experiments were conducted at a flow rate of 30 µL/min, with a starting contact time of 40 s and dissociation of 60 s, and extended if a steady state could not be reached. All peptides were diluted 1 in 2, 8x and injected sequentially from lowest to highest concentration, preceded by a buffer only blank injection. Following each injection the surface was regenerated with 2 M NaCl (2 x 30 s, 30 µL/min). After an optimal concentration range was found the series of injections were repeated to ensure reproducibility. The top concentration for the final concentration range for each peptide in listed in Table S2. All data was analysed using the GE Biosystems Biacore S200 Evaluation Software. All data are summarised in Table S2 and Figure 2.

S9.2 Supplementary information

S9.2.1 SPR

Table S1: SPR Data

	Name	λ	ε_λ	Top Conc	Affinity K_D (nM)	K_D SE (nM)	Chi^2	Ass/Diss time (s)	Purity (220 nm)
1.00	p21OM3	280	1490	100	2.488	0.790	1.07	40/60	87.2
2.10	p21OM3-KE4	280	1490	2000	115.8	21/0	3.54	40/60	91.1
3.10	p21OM3-EK4	280	1490	500	19.03	4.40	6.75	40/60	98.5
2.20	p21OM3-KE4L	280	1490	200	8.627	2.09	2.27	40/60	92.4
3.20	p21OM3-EK4L1	280	1490	2000	217.0	46.0	6.53	40/60	95.6
3.21	p21OM3-EK4L2	280	1490	500	38.68	7.00	3.65	40/60	82.4
4.10	p21OM3-CC4	280	1490	500	NS*			40/60	94.4
4.20	p21OM3-CC4B	380	4694	500	91.24	14.0	1.99	40/60	85.2
5.10	p21OM3-KC4A	280	1490	500	NS*			40/60	90.0
6.10	p21OM3-CK4A	280	1490	200	NS*			40/60	93.8
5.30	p21OM3-KC4T	280	1490	250	24.10	7.90	7.66	40/60	88.2
6.30	p21OM3-CK4T	280	1490	1000	71.12	14.0	3.74	40/60	90.0
7.10	p21OM3-SS4	280	1490	5000	1.031	160	2.73	40/60	81.9
7.20	p21OM3-SS4M	280	1490	500	45.70	7.10	1.61	40/60	84.2
8.10	p21OM3-PK4	280	1490	500	NS*			40/60	93.4
9.10	p21OM3-KP4	280	1490	500	31.62	6.40	3.46	40/60	80.1
8.20	p21OM3-PK4H	280	1490	500	94.26	2.20	3.65	40/60	84.0
9.20	p21OM3-KP4H	280	1490	100	15.67	1.40	0.586	40/60	88.3
*largely non-specific interaction with the sensor chip or protein									

REFERENCES

1. W. S. Hancock and J. E. Battersvy, A new micro-test for the detection of incomplete coupling reactions in solid-phase peptide synthesis using 2,4,6-trinitrobenzene-sulphonic acid, *Anal. Biochem.*, 1976, **71**, 260-264.

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

3. A. J. Horsfall, K. R. Dunning, K. L. Keeling, D. B. Scanlon, K. L. Wegener and A. D. Abell, A bimane‐based peptide staple for combined helical induction and fluorescent imaging, *ChemBioChem*, 2020, **21**, 3423-3432.

4. Y.-W. Kim, P. S. Kutchukian and G. L. Verdine, Introduction of All-Hydrocarbon i,i+3 staples into alpha-helices via ring-closing olefin metathesis, *Org. Lett.*, 2010, **12**, 3046-3049.

5. B. B. Metaferia, M. Rittler, J. S. Gheeya, A. Lee, H. Hempel, A. Plaza, W. G. Stetler-Stevenson, C. A. Bewley and J. Khan, Synthesis of novel cyclic NGR/RGD peptide analogs via on resin click chemistry, *Bioorg. Med. Chem. Lett.*, 2010, **20**, 7337-7340.

6. N. J. Anthis and G. M. Clore, Sequence-specific determination of protein and peptide concentrations by absorbance at 205 nm, *Protein Sci.*, 2013, **22**, 851-858.

7. X. Shen, C. B. Pattillo, S. Pardue, S. C. Bir, R. Wang and C. G. Kevil, Measurement of plasma hydrogen sulfide in vivo and in vitro, *Free Radical Biol. Med.*, 2011, **50**, 1021-1031.

Appendix 1.

AN INHERENTLY FLUORESCENT PEPTIDE CONSTRAINT TO DEFINE SECONDARY STRUCTURE: MOVING AWAY FROM AUXILIARY TAGS

Focus Article: The Australian Journal of Chemistry, **2021**, In press

An inherently fluorescent peptide constraint to define secondary structure: Moving away from auxiliary tags

Aimee J. Horsfall[1], and Andrew D. Abell[1*]

[1] ARC Centre of Excellence for Nanoscale BioPhotonics, Institute of Photonics and Advanced Sensing, School of Physical Sciences, The University of Adelaide, Adelaide, South Australia, 5005, Australia

*Corresponding author

© 2021 Australian Journal of Chemistry, CSIRO Publishing

First published online: 16th August 2021

STATEMENT OF AUTHORSHIP

Title of Paper	An inherently fluorescent peptide modification to influence structure: Moving away from auxiliary fluorescent tags
Publication Status	☐ Published ☐ Accepted for Publication ☐ Submitted for Publication ☒ Unpublished and Unsubmitted work written in manuscript style
Publication Details	**Invited Focus Article:** A. J. Horsfall, and A. D. Abell, *Australian Journal of Chemistry,* 2021

Principal Author

Name of Principal Author (Candidate)	Aimee J Horsfall		
Contribution to the Paper	Conceived, wrote and edited manuscript		
Overall percentage (%)	80%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	15/02/2021

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Andrew D Abell		
Contribution to the Paper	Supervised AJH, discussed and edited manuscript.		
Signature		Date	15/2/2021

A1.1 Background

Peptides and peptidomimetics have an innate ability to bind and inhibit large surface area interactions, such as those that occur between proteins. Peptides are often constrained into a well-defined geometry to yield peptidomimetics that favour target binding (see Reviews[1-4] for comprehensive examples). This can have the added effect of enhancing proteolytic stability and cell permeability. One approach to access peptidomimetics with a pre-formed geometry is to covalently link two appropriately-distanced amino-acid side-chains. An investigation of such peptidomimetics within a biological environment, either *in vitro* or *in vivo*, generally requires subsequent attachment of an auxiliary tag, such as a fluorophore to enable tracking by fluorescence microscopy. However, the introduction of a fluorophore, to an otherwise optimised peptidomimetic scaffold, can have unpredictable effects on target binding, as well as impacting secondary structure and cell permeability.[5-8] New approaches have endeavoured to incorporate fluorescent moieties in ways that have less impact on the function or structure of the molecule of interest (e.g.[9-11]) but are not truly bifunctional in nature. This highlights the need for new methodology that can stabilise peptide secondary structure and allow concomitant fluorescence-based imaging. This approach avoids the need for further derivatisation (i.e. synthetic effort), and allows cellular or imaging assays to be conducted directly on the therapeutically useful molecule. *In vitro* assays then more accurately reflect the behaviour of the peptidomimetic as it progresses through the drug development process. This Focus Article provides a brief summary of a bimane-based linker as a bifunctional peptide modification that stabilises secondary structure while also fluorescing at 480 nm (ex. 380 nm).[12, 13]

A1.2 Method and Applications

The fluorescent bimane linker can be selectively introduced into thiol-containing peptides either on-resin, immediately following solid-phase peptide synthesis, or on an unprotected peptide in solution. The dibromobimane reagent reacts with cysteine residues, separated by three amino-acids (*i*-*i*+4), to define a helical structure (Figure 1, left).[12, 13] The bimane-containing peptides can be imaged by fluorescence microscopy, without the need for further derivatisation.[12, 14]

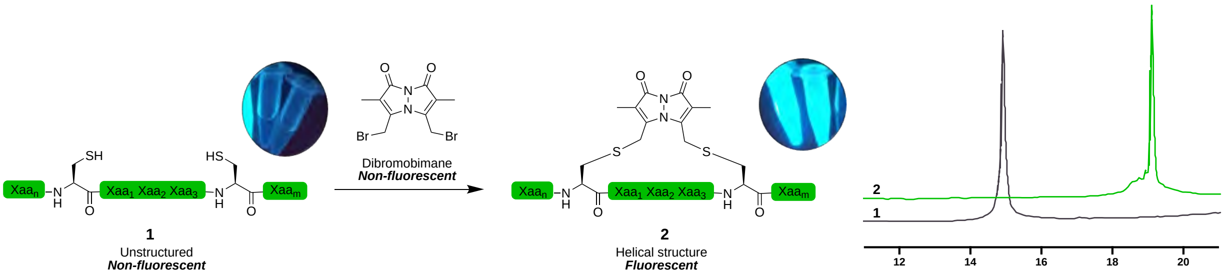
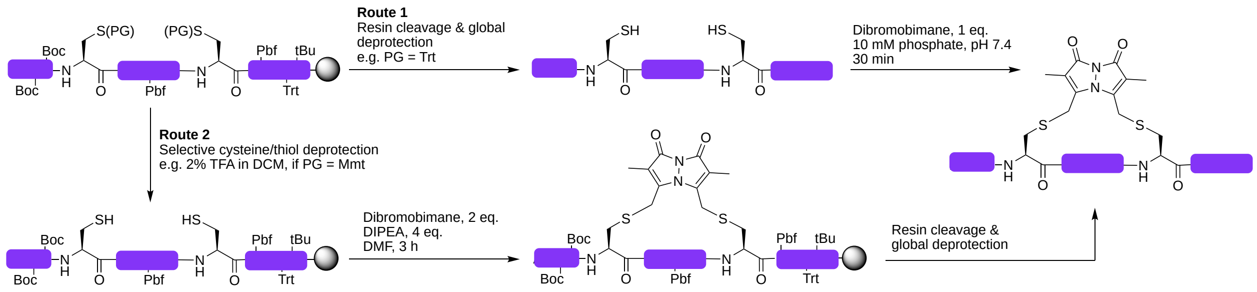


Figure 1: Macrocyclisation of *i*-*i*+4 separated cysteines by reaction with dibromobimane defines helical structure into short peptides and results in a highly fluorescent product. No significant by-products are generally observed, shown by the general RP-HPLC trace on the right, for the sequence Ac-WKCRILCRLLQE.


Route 1
Resin cleavage & global deprotection
e.g. PG = Trt
Dibromobimane, 1 eq.
10 mM phosphate, pH 7.4
30 min
Route 2
Selective cysteine/thiol deprotection
e.g. 2% TFA in DCM, if PG = Mmt
Dibromobimane, 2 eq.
DIPEA, 4 eq.
DMF, 3 h
Resin cleavage &
global deprotection
S(PG)
(PG)S
SH
HS
Boc
Pbf
tBu
Trt

Scheme 1: General synthetic methods to incorporate a bimane into a cysteine containing peptide. Route 1 demonstrates macrocyclization of an unprotected peptide in solution. Route 2 demonstrates on-resin reaction with dibromobimane where the cysteine thiols are selectively unmasked while the peptide remains attached to the resin (e.g. PG = Mmt, and the resin bears a high acid labile Rink amide or Wang linker).

On-resin attachment of bimane

The bimane can be introduced into resin-bound peptides, where the sequence contains orthogonally protected cysteines such that the protecting group/s can be liberated without unmasking other amino-acid protecting groups or cleavage of the peptide from the resin. For example, in Fmoc-solid-phase peptide synthesis on high-acid labile resin (e.g. Rink amide or Wang resin), cysteines protected with 4-methoxytrityl (Mmt), are readily unmasked on treatment with 2% TFA in DCM (Scheme 1, if PG = Mmt). Following deprotection of the cysteine side-chains, the free thiol/s are reacted with two equivalents of di- or monobromobimane, in the presence of four equivalents of DIPEA in DMF for 3 hours. The bimane-containing peptide is cleaved from the resin under standard resin acidolysis conditions, and purified by RP-HPLC.

Solution-phase bimane attachment:

The bimane is introduced into unprotected peptides by reaction of thiol-groups with one equivalent of di- or monobromobimane in 10 mM phosphate buffer at pH 7.4, which must be maintained throughout the reaction time of 30 min (Scheme 1, e.g PG = Trt). Macrocyclization is best performed at dilute peptide conditions to avoid oligomerisation, with 0.5 mg/mL proving to be optimal. The reaction is typically complete within 5 minutes, but is left for 30 minutes to assure completion. The reaction progress can be monitored by HPLC, LCMS or fluorescence (ex. 380 nm, em. 480 nm). The final solution is lyophilised prior to purification of the peptide by RP-HPLC.

Funding

The research was supported by the ARC Centre of Excellence in Nanoscale BioPhotonics (CNBP) (CE140100003). A.J.H. is supported by an Australian Government Research Training Program Stipend (RTPS).

Conflict of Interest

There are no conflicts of interest to declare.

REFERENCES

1. T. A. Hill, N. E. Shepherd, F. Diness and D. P. Fairlie, Constraining cyclic peptides to mimic protein structure motifs, *Angew. Chem. Int. Ed. (English)*, 2014, **53**, 13020-13041.
2. M. Pelay-Gimeno, A. Glas, O. Koch and T. N. Grossmann, Structure-Based Design of Inhibitors of Protein-Protein Interactions: Mimicking Peptide Binding Epitopes, *Angew Chem Int Ed Engl*, 2015, **54**, 8896-8927.
3. Y. H. Lau, P. de Andrade, Y. Wu and D. R. Spring, Peptide stapling techniques based on different macrocyclisation chemistries, *Chem. Soc. Rev.*, 2015, **44**, 91-102.
4. D. P. Fairlie and A. D. de Araujo, Stapling peptides using cysteine crosslinking, *Biopolymers*, 2016, **106**, 843-852.
5. H. H. Szeto, P. W. Schiller, K. Zhao and G. Luo, Fluorescent dyes alter intracellular targeting and function of cell-penetrating tetrapeptides, *The FASEB Journal*, 2005, **19**, 118-120.
6. M. P. Luitz, A. Barth, A. H. Crevenna, R. Bomblies, D. C. Lamb and M. Zacharias, Covalent dye attachment influences the dynamics and conformational properties of flexible peptides, *PLoS One*, 2017, **12**, e0177139.
7. H. Wang, R. S. Dawber, P. Zhang, M. Walko, A. J. Wilson and X. Wang, Peptide-based inhibitors of protein-protein interactions: biophysical, structural and cellular consequences of introducing a constraint, *Chem Sci*, 2021, **12**, 5977-5993.
8. D. Birch, M. V. Christensen, D. Staerk, H. Franzyk and H. M. Nielsen, Fluorophore labeling of a cell-penetrating peptide induces differential effects on its cellular distribution and affects cell viability, *Biochim. Biophys. Acta. Biomembr.*, 2017, **1859**, 2483-2494.
9. Y. Wu, M. T. Bertran, J. Rowley, E. D. D. Calder, D. Joshi and L. J. Walport, Fluorescent Amino Acid Initiated de novo Cyclic Peptides for the Label-Free Assessment of Cell Permeability*, *ChemMedChem*, 2021, DOI: 10.1002/cmdc.202100315.
10. L. Peraro, K. L. Deprey, M. K. Moser, Z. Zou, H. L. Ball, B. Levine and J. A. Kritzer, Cell Penetration Profiling Using the Chloroalkane Penetration Assay, *J. Am. Chem. Soc.*, 2018, **140**, 11360-11369.
11. N. Assem, D. J. Ferreira, D. W. Wolan and P. E. Dawson, Acetone-Linked Peptides: A Convergent Approach for Peptide Macrocyclization and Labeling, *Angew. Chem. Int. Ed. (English)*, 2015, **54**, 8665-8668.
12. A. J. Horsfall, K. R. Dunning, K. L. Keeling, D. B. Scanlon, K. L. Wegener and A. D. Abell, A bimane‐based peptide staple for combined helical induction and fluorescent imaging, *ChemBioChem*, 2020, **21**, 3423-3432.
13. A. J. Horsfall, D. P. McDougal, D. B. Scanlon, J. B. Bruning and A. D. Abell, Approches to introduce helical structure in cysteine-containing peptides with a bimane group, *ChemBioChem*, 2021, **In press, doi.org/10.1002/cbic.202100241**.
14. A. J. Horsfall, B. A. Vandborg, Z. Kikhtyak, D. B. Scanlon, W. D. Tilley, T. E. Hickey, J. B. Bruning and A. D. Abell, A short, cell permeable bimane-constrained PCNA-interacting peptide *Advance Article, doi.org/10.1039/D1CB00113B*, 2021.

Appendix 2.

OTHER CONTRIBUTIONS

A SILK-BASED FUNCTIONALIZATION ARCHITECTURE FOR SINGLE FIBER IMAGING AND SENSING

Patrick K. Capon[#], Jiawen Li[*#], **Aimee J. Horsfall**, Suliman Yagoub, Erik P. Schartner, Asma Khalid, Rodney W. Kirk, Malcolm S. Purdey, Kylie R. Dunning, Robert A. McLaughlin,* Andrew D. Abell*

[#]co-first authors , *corresponding author

Published online at: Advanced Functional Materials **2021**, 2010713, doi.org/10.1002/adfm.202010713.

A2.1 Background

A new fiber functionalization architecture for single-fiber imaging and sensing is presented. 5(6)Carboxy-SNARF2 (a fluorescent pH sensor) is attached to a silk-binding peptide and the complex added to aqueous silk fibroin protein. These bind with a K_d of 36 μM as determined by a fluorescence polarization assay. The fiber is dip-coated into the silk and peptide mixture, and scanning electron microscopy images reveal a uniform silk coating on the fiber tip. The coating is stable to repeated washes and does not affect the imaging light emitted from the fiber, which allows concurrent optical coherence tomography (OCT) imaging and pH sensing. Oocytes were metabolically stimulated with $CoCl_2$ to produce lactic acid, and a pH reduction of 0.04 was measured using the probe. The distance between fiber tip and oocyte was monitored by simultaneous OCT acquisitions to precisely position the probe. Lastly, OCT imaging of an ovary revealed the presence/absence of an oocyte within a follicle, an important step toward improving patient outcomes during IVF, by limiting the number of invasive follicle punctures required. These results demonstrate the utility of this new coating to enable simultaneous OCT imaging and sensing, which provides significant insight into complex biological systems.

Contribution:

Conception of using solid-binding peptide to adhere a sensor to silk-coated optical fibres. Design of fluorophore conjugated silk-binding peptide for sensing applications and aid in synthesis and troubleshooting of peptide synthesis and purification. Designed and carried out fluorescence polarisation binding assay of fluorophore-functionalised silk-binding peptide to silk fibroin. Contributed to manuscript preparation and editing.

CARBON-BINDING PEPTIDES TO FUNCTIONALISE NANODIAMOND

Patrick K. Capon, Philipp Reineck, **Aimee J. Horsfall**, Wioleta Kowalczyk, Marco D. Torelli, Olga A. Shendarova, Malcolm S. Purdey, Andrew D. Abell*

*corresponding author

Manuscript in preparation

A2.2 Background

Functionalisation of nanodiamonds (NDs) with an organic compound can provide a hybrid nanoparticle with advantageous properties for applications in drug delivery, cell imaging, and sensing, among others. Here four carbon-binding peptides (**1-DLC**, **2-CN**, **3-DF** and **4-GF**) are employed to non-covalently coat carboxylated NDs, with the presence of each peptide on the washed ND surface confirmed through characteristic peaks in infrared spectra and the mass loss observed during thermogravimetric analysis. A colorimetric assay was used to assess retention of each peptide on NDs after extensive washing in buffer. A diamond-like carbon binding peptide (**1-DLC**) found to be the most effective peptide to coat NDs, with 88% and 40% retention on detonation and high-pressure high-temperature NDs respectively. This shows that our peptide-based method is generalisable between ND sources. **1-DLC** was found to bind to NDs predominantly through electrostatic interactions, while a carbon nanotube binding peptide (**2-CN**) was found to bind largely through non-polar interactions. This study is an important contribution to the ND functionalisation field as it provides a new, highly adaptable approach to functionalise NDs using carbon-binding peptides.

Contribution

Conception of diamond-binding peptides for nanodiamond functionalisation. Peptide design and synthesis. Manuscript editing.

A2.3 Synthesis and characterisation of peptides

All starting materials were purchased from commercial sources and used without further purification, unless otherwise indicated. All peptides were synthesised by Fmoc/tBu solid-phase peptide synthesis detailed below, with all L-amino-acids (unless otherwise specified) purchased from Chem-Impex International: Fmoc-Ala-OH, Fmoc-Arg(Pbf)-OH, Fmoc-Asp(tBu)-OH, Fmoc-Asn(Trt)-OH, Fmoc-Cys(Trt)-OH, Fmoc-Gln(Trt)-OH, Fmoc-Glu(tBu)-OH, Fmoc-Gly-OH, Fmoc-His(Trt)-OH, Fmoc-Ile-OH, Fmoc-Leu-OH, Fmoc-Lys(Boc)-OH, Fmoc-Met-OH, Fmoc-Thr(tBu)-OH, Fmoc-Tyr(tBu)-OH, Fmoc-Trp(Boc)-OH, Fmoc-Ser(tBu)-OH, Fmoc-Phe-OH, Fmoc-Pro-OH, Fmoc-Val-OH and Fmoc-ε-Ahx-OH. Peptides were purified by semi-preparative RP-HPLC using a Gilson GX-Prep RP-HPLC equipped with a Phenomenex Aeris Peptide XB-C18 (250 x 10 mm, 5 µm) column at 4 mL/min and visualised at 220 and 254 nm. Purity of all compounds was confirmed by analytical RP-HPLC on an Agilent 1260 HPLC equipped with a Phenomenex Luna C18(2) column (250 x 4.6 mm, 5 µm) over a gradient of 5-50% B (15 min) at 1.5 mL/min, and visualised at 220 nm. RP-HPLC solvents were A: H_2O with 0.1% TFA and B: ACN with 0.1%

TFA. High-resolution mass spectra were collected using an Agilent 6230 ESI-TOF via direct injection in ACN with 0.1% formic acid as the running buffer. Characterisation data for all peptides in recorded in Table 1.

General procedure for peptide synthesis:

All peptides were synthesised by the Fmoc/tBu solid-phase peptide synthesis protocol on a CEM Liberty Blue Automated Microwave Peptide Synthesiser (CEM Corp., Matthews, NC, USA) using the standard manufacturer's conditions. The peptides were assembled on 0.2 mmol scale on Chem Impex International Rink Amide AM resin (0.47 mmol/g) or Mimotopes Rink Amide resin (0.456 mmol/g). The resin was initially swollen in DCM (15 mL, 15 min) and then the resin washed with DMF (2 × 8 mL). The resin-bound Fmoc-groups were deprotected with a mixture of 20% piperidine and 0.1 M OxymaPure in DMF using the standard microwave deprotection method with a maximum temperature of 90°C. Couplings were performed with Fmoc-protected amino-acids (0.2 M in DMF, 4 equiv), OxymaPure (1 M in DMF, 4 equiv) and DIC (0.5 M in DMF, 4 equiv) under the 'Standard Coupling' microwave method with a maximum temperature of 90°C, except for coupling of Fmoc-His(Trt)-OH which was coupled using a maximum 50°C 10 min coupling procedure; and Fmoc-Arg(Pbf)-OH which used the default 'Arginine Double Coupling' microwave method which included two couplings steps – the first at room temperature and the second at a maximum of 75°C. Following assembly of the desired sequence the *N*-terminal protecting group was removed. The resin was then removed from the synthesiser, washed with DCM (3 × 8 mL) and then diethyl ether (3 × 8 mL) and air dried with suction.

The resin-bound peptide (0.025 mmol) was reswollen in 1:1 DCM/DMF (5 mL) for 15 min, the solution drained and the resin washed with DMF (3 × 5 mL). A solution of FITC (2 equiv, 0.5 mmol, 22 mg) and DIPEA (4 equiv, 0.1 mmol, 17.5 µL) in DMF (1 mL) was added to the resin and reacted for 1.5 h, with intermittent stirring. The solution was drained and the resin thoroughly washed with DMF (3 × 5 mL), then DCM (3 × 5 mL), diethyl ether (3 × 5 mL) and air dried with suction. The peptide was subsequently cleaved from the resin and the side-chain protecting groups simultaneously globally deprotected by treatment of the resin with 92.5/2.5/2.5/2.5 TFA:TIPS:DODT:H_2O (5 mL) for 2 h. The cleavage solution was pipetted from the resin and concentrated under a stream of nitrogen to 0.5-1 mL. The peptide was then precipitated by addition of diethyl ether (10 mL) and the mixture cooled at -20°C. The precipitate was pelleted by centrifugation (7600 rpm, 10 min), and the supernatant decanted. The pellet was dried under a nitrogen stream, and then dissolved in 1:1 ACN/H_2O, then syringe filtered (0.2 µm) and lyophilised to yield the crude peptide as a fluffy white powder. The peptide was purified by semi-preparative RP-HPLC using a Phenomenex Aeris Peptide C18 column (10 × 250 mm), over a linear ACN/H_2O gradient optimised for each peptide sample. The final peptide purity and identity were confirmed by RP-HPLC and HRMS. Characterisation data for all peptides in recorded in Table 1.

Table 1: Characterisation data for carbon-binding peptides

Peptide	Sequence	Formula	Exact mass	Calc X+ $[M+XH]^{X+}$	Found	Purity % 220 nm	Ref.
DLC	FITC-Ahx-KNSAPQKSENRKVPFYQHQ	$C_{127}H_{179}N_{35}O_{35}S_1$	2786.3024	3+ 929.7753	929.7746	91	[1]
CN	FITC-Ahx-HWKHPWGAWDTL	$C_{102}H_{119}N_{23}O_{21}S_1$	2033.8672	3+ 678.9635	678.9629	92	[2]
DF	FITC-Ahx-GVGGLTTVNYSR	$C_{79}H_{109}N_{19}O_{23}S_1$	1723.7664	2+ 862.8910	862.8917	84	[3]
GF	FITC-Ahx-MVTESSDYSSY	$C_{80}H_{100}N_{14}O_{28}S_2$	1768.6273	2+ 885.3215	885.3225	84	[4]

References

1. B. Gabryelczyk, G. R. Szilvay, M. Salomaki, P. Laaksonen and M. B. Linder, Selection and characterization of peptides binding to diamond-like carbon, *Colloids and Surfaces B: Biointerfaces*, 2013, **110**, 66-73.
2. S. Wang, E. S. Humphreys, S. Y. Chung, D. F. Delduco, S. R. Lustig, H. Wang, K. N. Parker, N. W. Rizzo, S. Subramoney, Y. M. Chiang and A. Jagota, Peptides with selective affinity for carbon nanotubes, *Nat. Mater.*, 2003, **2**, 196-200.
3. S. Swaminathan and Y. Cui, Recognition of diamond with phage display peptides, *RSC Adv.*, 2016, **6**, 49127-49129.
4. M. J. Penna, M. Mijajlovic, C. Tamerler and M. J. Biggs, Molecular-level understanding of the adsorption mechanism of a graphite-binding peptide at the water/graphite interface, *Soft Matter*, 2015, **11**, 5192-5203.

PROTEIN DETECTION ENABLED USING FUNCTIONALISED SILK-BINDING PEPTIDES ON A SILK-COATED OPTICAL FIBRE

Patrick K. Capon, **Aimee J. Horsfall**, Jiawen Li, Erik P. Schartner, Asma Khalid, Malcolm S. Purdey, Robert A. McLaughlin, Andrew D. Abell*

*corresponding author

Published online at: RSC Advances, 2021 **11**, 22334-22342, doi.org/10.1039/D1RA03584C.

A2.4 Abstract

We report the detection of AlexaFluor-532 tagged streptavidin by its binding to D-biotin, which is functionalised onto an optical fibre *via* incorporation in a silk fibroin fibre coating. The D-biotin was covalently attached to a silk-binding peptide to provide SBP-Biotin, which adheres the D-biotin to the silk-coated fibre tip. These optical fibre probes were prepared by two methods. The first involves dip-coating the fibre tip into a mixture of silk fibroin and SBP-Biotin, which distributes the SBP-Biotin throughout the silk coating (method A). The second method uses two steps, where the fibre is first dip-coated in silk only, then SBP-Biotin added in a second dip-coating step. This isolates SBP-Biotin to the outer surface of the silk layer (method B). A series of fluorescence measurements revealed that only the surface bound SBP-Biotin detects streptavidin with a detection limit of 15 μg/mL. The fibre coatings are stable to repeated washing and long-term exposure to water. Formation of silk coatings on fibres using commercial aqueous silk fibroin was found to be inhibited by a lithium concentration of 200 ppm, as determined by atomic absorption spectroscopy. This was reduced to less than 20 ppm by dialysis against water, and was found to successfully form a coating on optical fibres.

Contribution

Conception and design of modified silk-binding peptides, and assistance in peptide synthesis, purification and troubleshooting. Assistance with experimental design and troubleshooting and protein handling. Manuscript editing and figure design.

DEFINING HIGH AFFINITY INVERSE AGONISM IN PPARγ

Rebecca L. Frkic, Blagojce Jovcevski, Elise Bruning, Mi Ra Chang, **Aimee J. Horsfall**, Wioleta Kowalczyk, Bruce Pascal, Theodore M. Kamenecka, Jeffrey Harmer, Partick R. Griffin, and John B. Bruning*

*corresponding author

Manuscript in preparation

A2.5 Abstract

Despite recent advances in non-activating antidiabetic ligands of PPARγ with lowered side effects, there still exists complications associated with these compounds, such as poor pharmacokinetic profiles, low affinity, and off target effects. Our previous studies with the inverse agonist SR10171 and antagonist SR11023 have shown their capacity to shift H12 of PPARγ into the inactive conformation which enables the recruitment of corepressors to limit transactivation of genes contributing to side-effects. Here we present a biochemical and structural characterisation of the structurally related analogue of these compounds, SR10221, which retains the capacity to shift H12 into the inactive conformation with higher affinity than its predecessors. Our crystal structures, in combination with HDX and EPR, showed that while SR10221 causes a shift in H12, the addition of a corepressor peptide causes H12 to lose a distinct conformation and sample a non-homogenous plethora of conformations. We present SR10221 as a model inverse agonist of PPARγ, which should be used as a tool for investigating non-activating ligands of PPARγ for the development of antidiabetic drugs.

Contribution

Synthesis and purification of two PPARγ coactivator derived peptides, NCOR1-3 and SMRT-2.

General method for peptide synthesis and characterisation

All peptides were synthesised by the Fmoc/tBu solid-phase peptide synthesis protocol on a CEM Liberty Blue Automated Microwave Peptide Synthesiser using the standard manufacturer's conditions. The peptides were assembled on 0.1 mmol scale on Chem Impex (Int'l Ltd.) Rink Amide resin (0.57 mmol/g). The resin was initially swollen in DCM (10 mL, 15 min) and then the resin washed with DMF (2 x 5 mL) and transferred to the microwave reaction vessel. The resin-bound Fmoc-groups were deprotected with a mixture of 20% piperidine and 0.1 M OxymaPure® in DMF using the standard microwave deprotection method with a maximum temperature of 90°C. Couplings were performed with Fmoc-protected amino-acids (0.2 M in DMF, 5 equiv), OxymaPure® (1 M in DMF, 5 equiv) and DIC (0.5 M in DMF, 5 equiv) under the 'Standard Coupling' microwave method with a maximum temperature of 90°C, except for coupling of Fmoc-Arg(Pbf)-OH which used the default 'Arginine Double Coupling' microwave method which included two couplings steps – the first at room temperature and the second at a maximum of 75°C. Between each coupling and deprotection step the resin was washed with DMF. Following complete assembly of the peptide and deprotection of the final Fmoc group the peptides were subsequently cleaved from the resin and the side-chain protecting groups simultaneously globally deprotected by treatment of the resin with 92.5/2.5/2.5/2.5 TFA:TIPS:DODT:H_2O (5 mL) for 2 h. The TFA mixture was pipetted from the resin and concentrated under a stream of nitrogen to 0.5-1 mL. The peptide was then precipitated by addition of diethyl ether (10 mL) and the mixture cooled to -20°C. The precipitate was pelleted by centrifugation (7600 rpm, 10 min), and the supernatant decanted. The pellet was dried under a nitrogen stream, and then dissolved in 1:1 ACN/H_2O, before being syringe filtered (0.2 μm) and lyophilised to yield the crude peptide as a fluffy white powder. The peptides were purified by semi-preparative RP-HPLC on a Gilson GX-Prep system using a

Phenomenex Aeris Peptide C18 column (10 x 250 mm), over an ACN/H_2O gradient as specified in the individual compound sections below. RP-HPLC solvents were (A) H_2O with 0.1% TFA and (B) ACN with 0.1% TFA. The product containing fractions were pooled and lyophilised. The identity of the final compounds was confirmed by High Resolution Mass Spectrometry on an Agilent 6230 ESI-TOF LCMS. Purity of the peptides was confirmed by analytical RP-HPLC on an Agilent 1260 HPLC equipped with a Phenomenex Luna C18(2) column (250 x 4.6 mm) over a gradient of 5-50% B (15 min) at 1.5 mL/min.

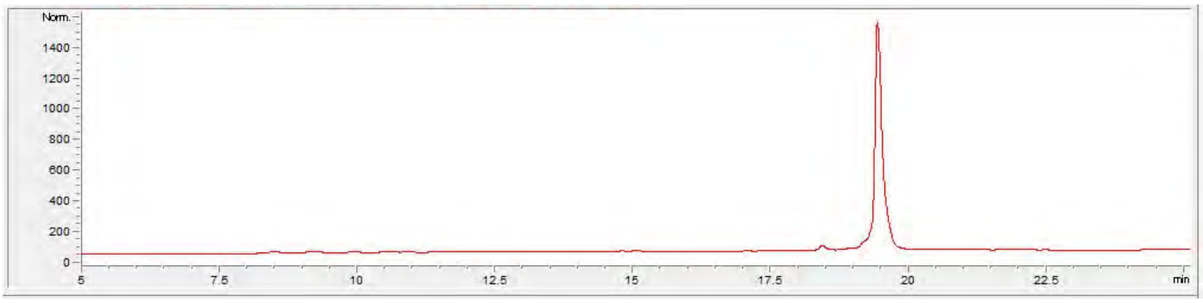
NCOR1-3 *ASNLGLEDIIRKALMGSFD*

The peptide was prepared as described in the general method and purified by RP-HPLC over a linear gradient of 30-60% B over 15 minutes to give the final peptide a fluffy white powder (15.0 mg). R_t –14.5 min (5-50% B, 15 min). Purity (215 nm RP-HPLC) >95%. HRMS (ESI+) Expected $[M+2H]^{2+}$ for $C_{89}H_{147}N_{25}O_{28}S_1$: 1024.0357, observed: $[M+2H]^{2+}$ 1024.0572.

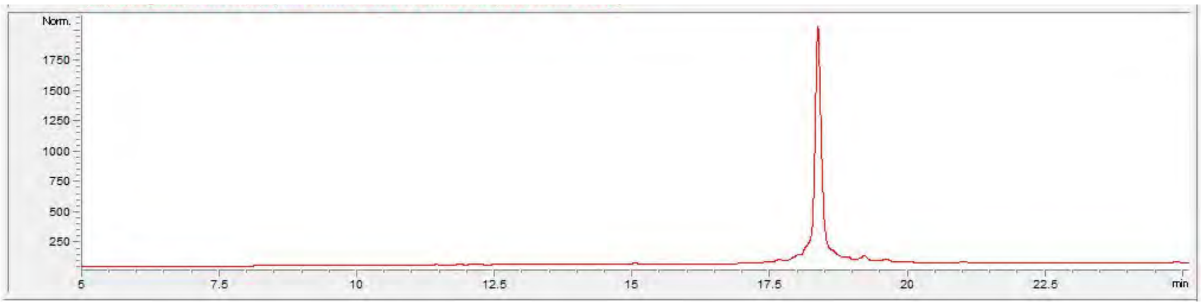
SMRT-2 *TNMGLEAIIRKALMGKYDQWEE*

The peptide was prepared as described in the general method and purified by RP-HPLC over a linear gradient of 35-70% B over 20 minutes to give the final peptide a fluffy white powder (13.9 mg). R_t - 13.4 min (5-50% B, 15 min). Purity (215 nm RP-HPLC) 86%. HRMS (ESI+) Expected $[M+4H]^{4+}$ for $C_{114}H_{181}N_{31}O_{34}S_2$: 649.0780, observed: $[M+3H]^{3+}$ 649.0813.

HPLC Traces



RP-HPLC of **NCOR1-3** peptide, run on a Phenomenex Luna C18(2) column (250 x 4.6 mm) over a linear gradient of 5-50% B (15 min), visualised at 220 nm. The linear gradient starts at 5 minutes.



RP-HPLC of **SMRT-2** peptide, run on a Phenomenex Luna C18(2) column (250 x 4.6 mm) over a linear gradient of 5-50% B (15 min), visualised at 220 nm. The linear gradient starts at 5 minutes.

IDENTIFICATION AND CHARACTERISATION OF A NOVEL L-AMINO-ACID LIGASE FROM A PATHOGENIC BACTERIUM

Jordan L. Pederick, Nan Hao, **Aimee J. Horsfall**, Andrew D. Abell, Keith E. Shearwin, John B. Bruning*

*corresponding author

Manuscript in preparation

A2.6 Background:

L-amino-acid ligases (LALs) have mostly been studied for industrial applications due to their ability to synthesise a wide range of dipeptide products, dependant on the LAL substrate specificity. To date, the isolation and characterisation of LALs has been largely restricted to non-pathogenic bacterial species, and as such, their physiological roles remain poorly understood. In this work, a hypothetical LAL from a pathogenic bacterial species (LAL^P) was cloned, purified and screened for ligase activity with the 20 proteinogenic amino-acids as substrates. This revealed that LAL^P possesses high activity specifically for the L-Met and L-Asp substrate pair, supporting formation of an Asp-Met or Met-Asp dipeptide. Isolation and subsequent characterisation of the reaction product confirmed L-Aspartyl-L-Methionine was formed. Having confirmed the activity and specificity of LAL^P, its physiological role in amino-acid metabolism is currently being investigated.

Contribution:

Isolation of dipeptide product from a novel L-amino-acid ligase (LAL) enzyme mixture by HPLC, and characterisation of dipeptide connectivity by NMR.

A2.7 Methods

Enzyme reaction

25 mM L-Met, 40 mM L-Asp.NaOH, 5 mM $MgCl_2$, 50 mM Tris pH 8.0, 20 mM ATP, and 100 nM LAL were combined and reacted at 37°C for 4 h, and then stored at -20 °C.

Dipeptide derivatisation

TNBS (200 µL, 5% w/v in H_2O) was added to a sample of the enzyme reaction mixture (2 mL) and reacted for 20 min.

Isolation

Purification of MD reaction mixture by semi-preparative RP-HPLC was carried out on a Gilson GX-Prep system using a Phenomenex Aeris Peptide C18 column (250 x 10 mm) over a linear gradient of 10-40% ACN with 0.1% TFA, in H_2O with 0.1% TFA (20 min) at 4 mL/min, and visualised at 220 nm.

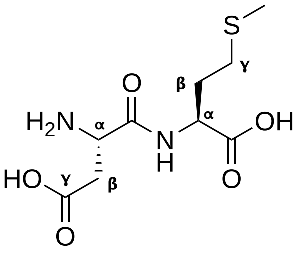
Mass spectrometry

Mass spectra were collected on a Bruker HCT Ultra connected to an Agilent 1100 LC system, equipped with a Phenomenex Kinetex C18 column (3.2 µm, 100 x 3 mm) using a solvent system of water with 0.1% formic acid and acetonitrile with 0.1% formic acid.

NMR Collection

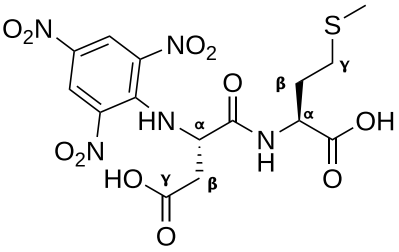
NMR spectra were collected on an Oxford Varian 600 MHz NMR spectrometer with samples dissolved in 10% aq. D_2O ~pH 4, and referenced to DSS at 0 ppm. A waterES suppression sequence was applied and zTOCSY and ROESYAD were collected with 256 increments and 8 scans. gHMBCAD and gHSQCAD were collected with 256 increments and 4 scans.

A2.8 Compound characterisation



AspMet

Product **A** was isolated from the '*Enzyme reaction'* mixture as described in the 'Isolation' protocol to give a white powder identified as **AspMet**. R_t (C18, 10-40% B 20 min) ~3.2 min. MS (ESI+) expected for $C_9H_{16}N_2O_5S$, 264.3; found 265.3 $(M+H)^+$, 529.7 $(2M+H)^+$. ^1H NMR (599 MHz, 10% aq. D_2O) δ = 8.53 (d, *J*=7.6, 2H, Met NH), 4.81 – 4.54 (m, 1H, Met αH), 4.21 – 4.10 (m, 1H, Asp αH), 3.08 – 2.95 (m, 2H, Asp βH), 2.69 – 2.60 (m, 1H, Met γH), 2.60 – 2.49 (m, 1H, Met γH), 2.24 – 2.12 (m, 1H, Met βH), 2.11 (s, 3H, Met CH_3), 2.06 – 1.96 (m, 1H, Met βH) ppm. ^{13}C NMR (151 MHz, 10% aq. D_2O) δ 178.5 (Met CO), 174.9 (Asp γCO), 174.0 (Asp CO), 54.7 (Met αC), 53.4 (Asp αC), 37.2 (Asp βC), 32.7(Met γC), 32.0 (Met βC), 16.7 (Met CH_3) ppm. 2D spectra are provided in Figures A1-3.



TNB-AspMet

Product **B** was isolated from the '*Dipeptide derivatisation'* mixture as described in the 'Isolation' protocol to give a yellow powder identified as **TNB-AspMet**. R_t (C18, 10-40% B 20 min) ~19 min. MS (ESI+) expected for $C_{15}H_{17}N_5O_{11}S$, 475.4; found 476.5 $(M+H)^+$. ^1H NMR (599 MHz, 10% aq. D_2O) δ = 9.40 (d, *J*=8.9, 1H Asp NH), 9.17 (s, 1H), 8.51 (d, *J*=7.6, 1H, Met NH), 4.43 – 4.34 (m, 1H, Met αH), 4.25 – 4.19 (m, 1H, Asp αH), 2.99 – 2.84 (m, 2H, Asp βH), 2.58 – 2.50 (m, 1H, Met γH), 2.48 – 2.39 (m, 1H, Met γH), 2.15 – 2.06 (m, 1H, Met βH), 2.06 (s, 3H, Met CH_3), 1.96 – 1.87 (m, 1H, Met βH) ppm. 2D spectra are provided in Figure A4.

A2.9 Results

Mass spec (MS) analysis of product **A** isolated from the '*Enzyme reaction'* mixture at 3.2 min (RP-HPLC, 10-40%, 20 min) revealed a mass of 265 which corresponds to the M+H of a dipeptide comprised of aspartic acid and methionine. Resonances in the ^1H NMR were also consistent with presence of both methionine and aspartic acid, where two signals at 4.51 and 4.41 ppm correspond to the α-protons and a diagnostic three proton singlet at 2.10 ppm corresponding to the methionine methyl group. A one proton doublet at 8.53 ppm was the only signal apparent above 5 ppm. Cross-peaks in the TOCSY spectrum were observed between the 8.53 ppm doublet αH at 4.51 ppm and five signals upfield of 3 ppm, including the three-proton singlet, indicating these resonances are part of the methionine spin system. Therefore the 4.41 ppm signal was assigned at the aspartic acid αH, and a TOCSY crosspeak with the remaining two proton signal at 2.90 ppm is consistent with this assignment. A cross-peak in the ROESY spectrum between the aspartic acid αH signal (4.41 ppm) and methionine amide signal (8.52 ppm) suggests the dipeptide connectivity is AspMet and not MetAsp. Crosspeaks are also observed in the HMBC spectrum between both Asp βH (3.02 ppm) and αH (4.41 ppm), and a signal at 174.0 ppm, which is therefore assigned as the Asp CO resonance. This signal (174.0, Asp CO) displays an additional cross peak with the amide signal at 8.52 ppm, in further support of the dipeptide AspMet.

A sample of the LAL '*Enzyme reaction*' mixture was also reacted with trinitrobenzenesulfonic acid (TNBS) that is known to react readily with amine groups, in order to further derivatise the dipeptide and confirm its identity. A peak with strong 254 nm absorbance was isolated by RP-HPLC to yield a yellow powder of product **B**, where MS analysis revealed a mass of 476 which corresponds to a M+H of a methionine and aspartic acid dipeptide with a trinitrobenzene group (**TNB-AspMet**). Analysis of the ^{1}H 1D NMR spectrum of this sample overall was similar however all the signals were shifted downfield, consistent with introduction of the strongly electron withdrawing trinitrobenzene group. Two additional signals above 5 ppm were apparent – a one proton doublet at 9.40 ppm and two proton singlet at 9.16 ppm. Crosspeaks in the TOCSY spectrum were observed between the Asp βH (2.91 ppm) and αH (4.22 ppm) resonances, and the one proton doublet at 9.40 ppm indicating this signal corresponds to the aspartic acid NH. This is consistent with a dipeptide connectivity of AspMet, where the aspartic acid amine reacted with TNBS to produce an aryl amine with a large downfield shift relative to the methionine amide. This further confirms assignment of the enzyme X product as AspMet.

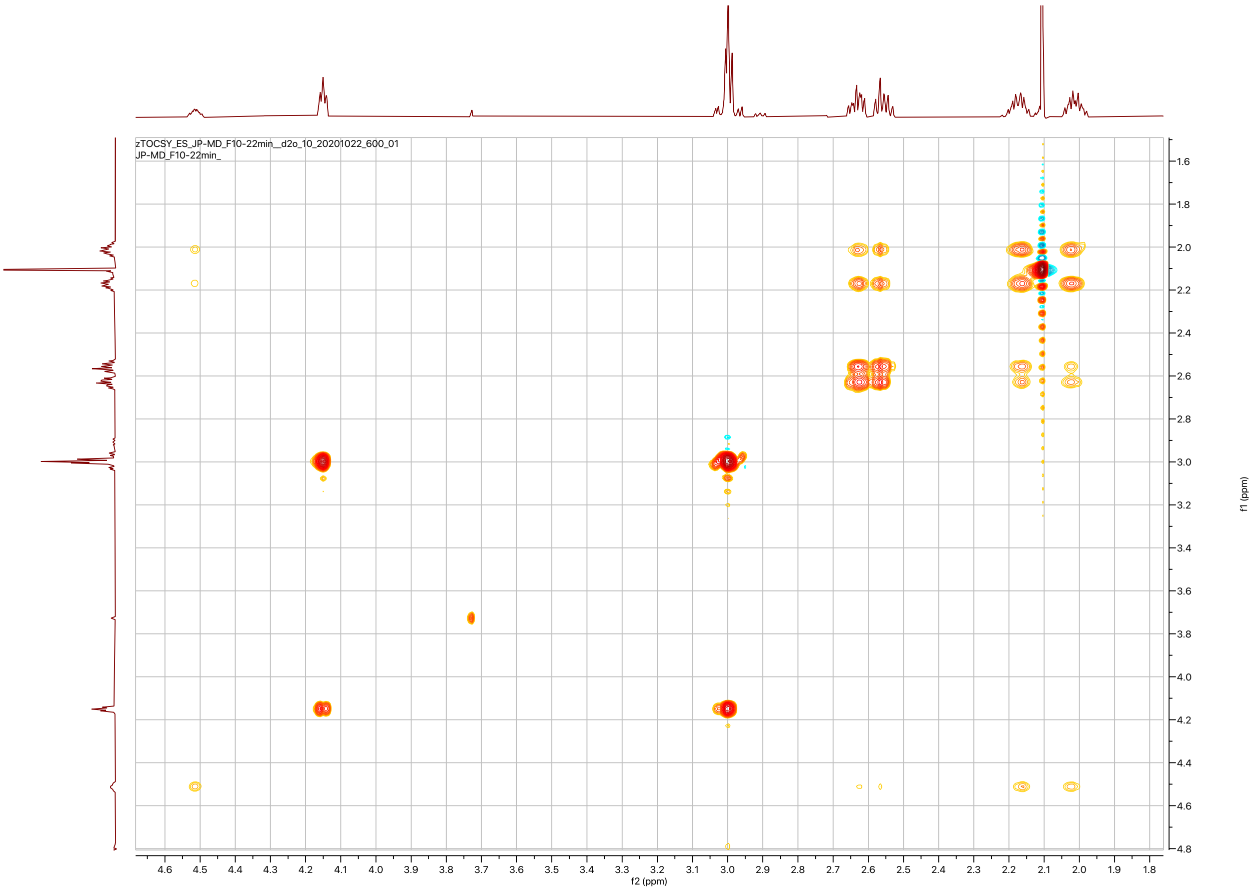


Figure A1: Alkyl region of zTOCSY spectrum of product **A** assigned as AspMet.

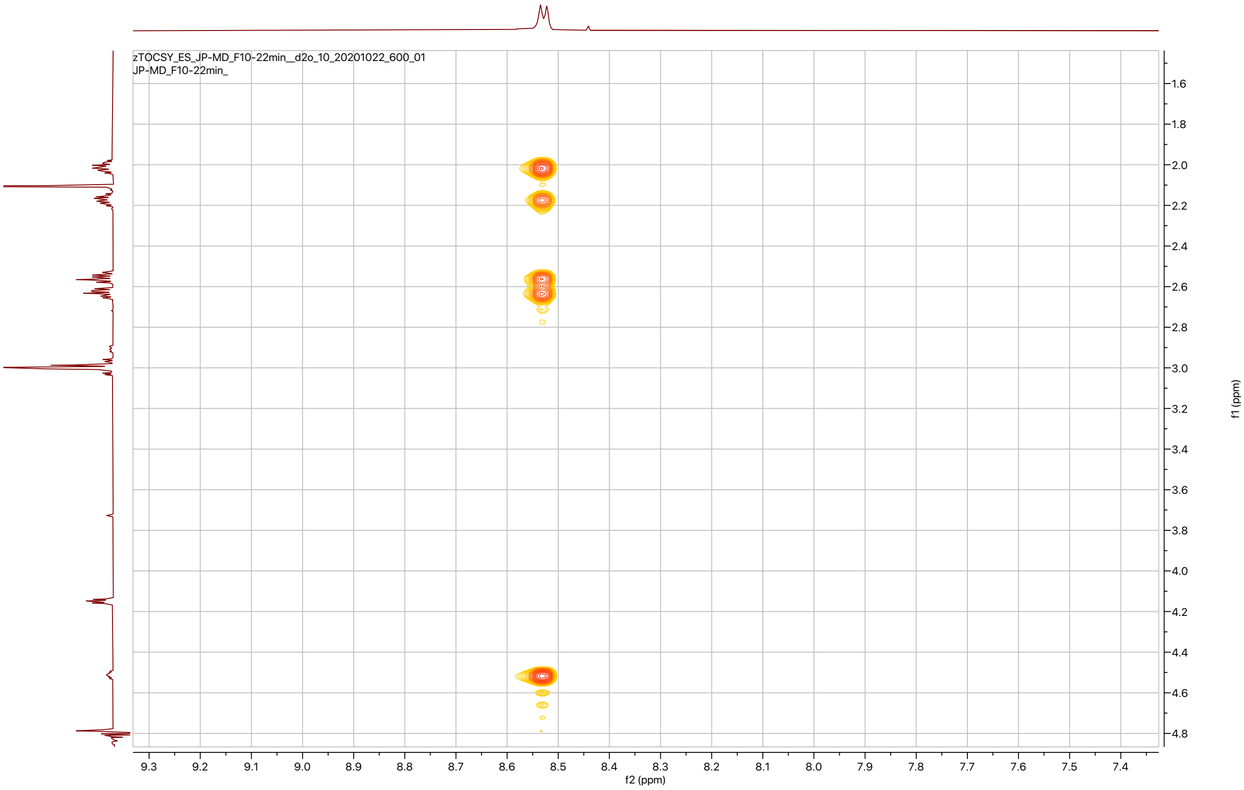


Figure A2: A zTOCSY spectrum of product **A** assigned as AspMet, showing the alkyl cross-peaks of the lone amide signal.

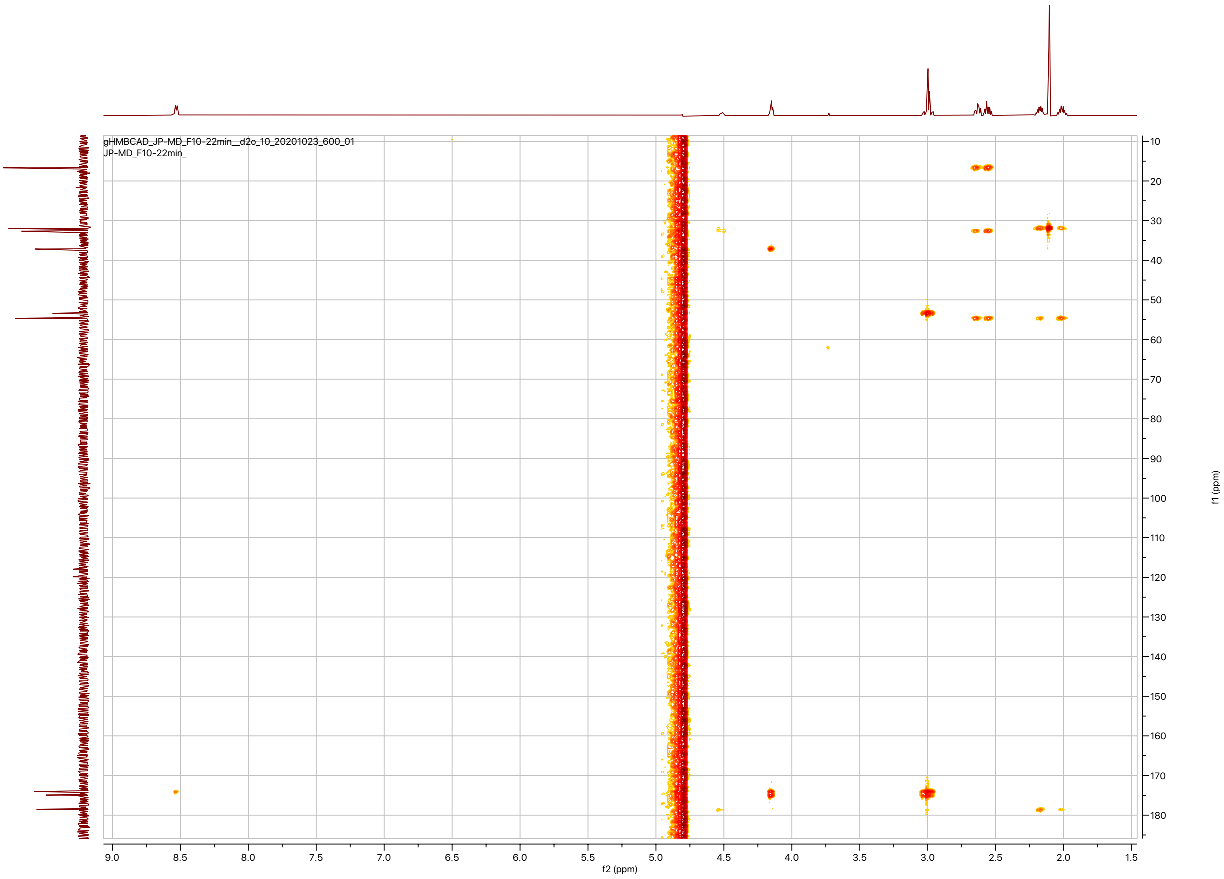


Figure A3: A gHMBCAD spectrum of product **A** assigned as **AspMet**.

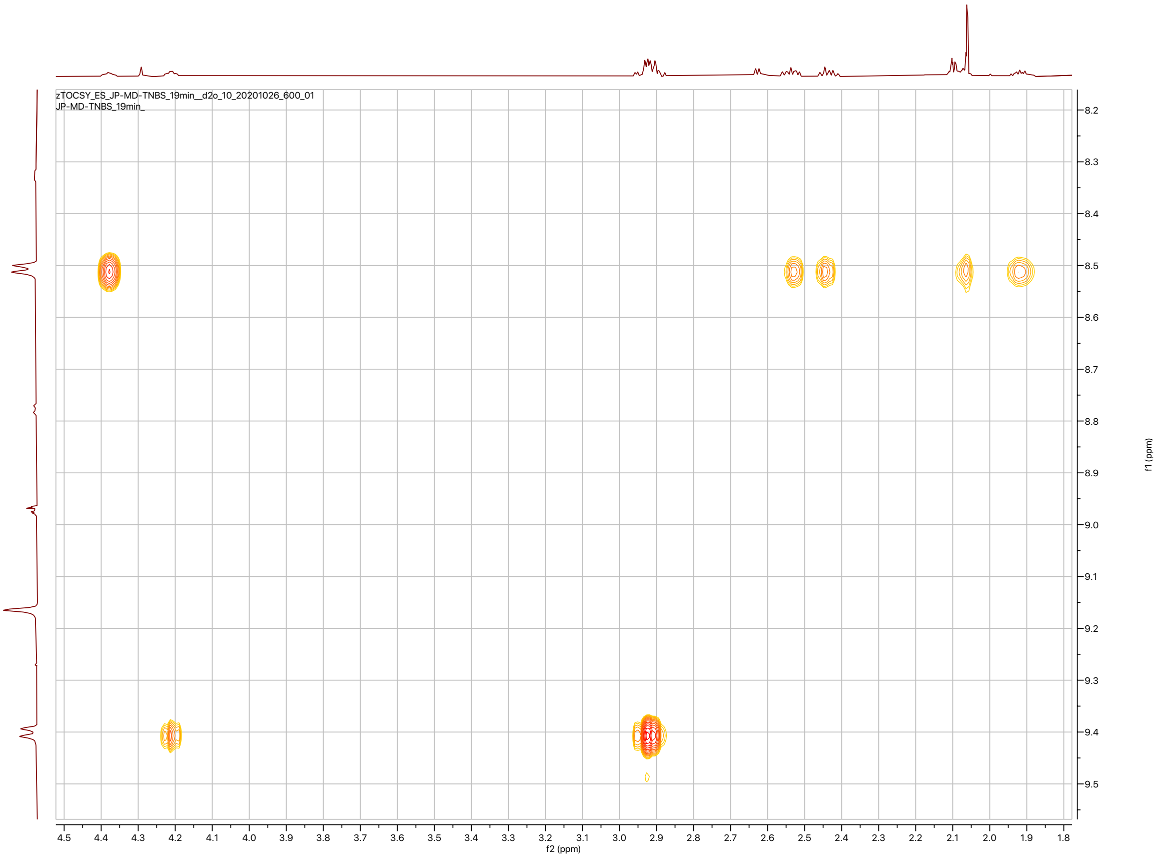


Figure A4: A zTOCSY spectrum of product **B** assigned as **TNB-AspMet**, showing the crosspeaks between the aromatic/amide signals and the alkyl region.

SYNTHESIS OF MODIFIABLE PHOTOCLEAVABLE MALDI MASS TAGS USING COPPER CLICK CHEMISTRY

Katherine G. Stevens, **Aimee J. Horsfall**, River J. Pachulicz, Andrew D. Abell and Tara L. Pukala*

Published as: Katherine G. Stevens, Lewis O. McFarlane, Kirsten Platts, Neil O'Brien-Simpson, Wenyi Li, Anton Blencowe, Paul J. Trim, and Tara L. Pukala*, *Retro Diels-Alder fragmentation of fulvene-maleimide bioconjugates for mass spectrometric detection of biomolecules*, Analytical Chemistry **2021**, In press, doi.org/10.1021/acs.analchem.1c00193.

*corresponding author

A2.10 Background

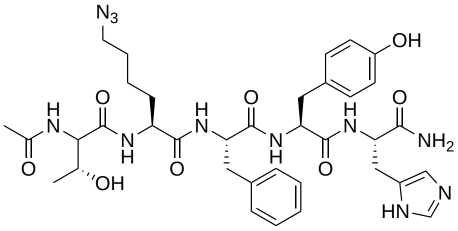
Growing recognition of the independent limitations of both conventional immunoassays and mass spectrometry for biomarker detection and quantification has led to the development of hyphenated techniques, which combine antibody-based biomarker targeting with MS detection. One branch of this broad area of research is focused on the development of chemical labels, or 'mass tags,' which improve the detection of large biomolecules, such as antibodies, to better enable their ionization and quantification using standard mass spectrometry instrumentation. A key consideration for this strategy is the ease with which these mass tags can be modified to enable variation in their structure and subsequent mass-to-charge ratio during analysis. The ability to easily synthesize a wide array of different tags would enable the simultaneous detection of multiple biomarkers. In the present study, we employ copper-catalyze azide-alkyne cycloaddition chemistry to conjugate an azide-labelled peptide or fluorophore mass tag to proteins via a photocleavable linker, which releases the tag during matrix-assisted laser desorption/ionization mass spectrometric analysis. This linker is successfully used for the fluorescence and MALDI detection of an antibody, demonstrating its potential utility for identifying large intact biomolecules in mass spectrometry-based immunoassays.

Contribution

Synthesis and characterization of a short azide-labelled peptide.

A2.11 Synthesis and characterisation

Unless otherwise indicated, all starting materials were purchased from commercial sources and used without further purification. All peptides were synthesised by the Fmoc solid-phase peptide synthesis protocol detailed below, with all L-amino-acids (unless otherwise specified) purchased from Chem Impex Int'l. RP-HPLC solvents were buffer A: H_2O with 0.1% TFA, and buffer B: ACN with 0.1% TFA; 0.2 µm filtered. Purity of all compounds was determined by analytical RP-HPLC on an Agilent 1260 Infinity HPLC equipped with a Phenomenex Luna C18(2) column (250 x 4.6 mm) over a linear gradient of 0-100% B (15 min). Mass spectra for characterisation were collected using a Bruker HCT Ultra via direct injection, and ACN with 0.1% formic acid as the running buffer.



Chemical Formula: $C_{36}H_{47}N_{11}O_8$
Molecular Weight: 761.8410

Solid-phase peptide synthesis of peptide Ac-TK(N_3)FYH-NH_2

The following Fmoc-protected amino-acids were used for all peptides unless otherwise specified. All amino-acids were purchased from Chem-Impex Int'l.: Fmoc-L-Thr(tBu)-OH, Fmoc-L-Lys(ε-N_3)-OH, Fmoc-L-Phe-OH, Fmoc-L-Tyr(tBu) -OH, Fmoc-L-His(Trt)-OH. Rink Amide PL resin (0.1 mmol, 322 mg, 0.31 mmol/g, Agilent) was swollen in 1:1 DMF/DCM (10 mL) for 15 min. The mixture was transferred to a fritted syringe, the solution drained, and the resin washed with DMF (3 x 5 mL). The Fmoc-protecting group was removed by treatment of the resin with a solution of 20% piperidine in DMF (5 mL) for 15 min. The solution was drained and the resin washed with DMF (3 x 5 mL). Amino-acid couplings were achieved by addition of a solution of Fmoc-protected amino-acid (5 equiv), HATU (5 equiv) and DIPEA (10 equiv, 174 µL) in DMF (5 mL), to the resin and stirred intermittently for 1 h. The solution was drained and the resin washed with DMF (5 x 5 mL). The *N*-terminal Fmoc-protecting group was removed by treatment of the resin with a solution of 20% piperidine and in DMF (5 mL) for 10 min, the solution was drained and the resin washed with DMF (5 x 5 mL). A TNBS test* was used to verify each coupling (negative/colourless) and deprotection (positive/red) step, with steps repeated as necessary. Successive couplings and Fmoc-deprotections were repeated to achieve the desired sequence. After the final Fmoc-deprotection, the *N*-terminal amine was protected with an acetyl functionality by reaction with acetic anhydride (470 µL) and DIPEA (870 µL) in DMF (10 mL) for 15 min. The resin was washed with DMF (3 x 5 mL), and DCM (5 x 5 mL) then dried with Et_2O (3 x 5 mL). The peptide was cleaved from the resin by addition of 95:2.5:2.5 TFA/TIPS/H_2O (6 mL) to the resin and rocked for 1 h. The TFA solution was pipetted from the resin and concentrated to 0.5-1 mL under a nitrogen stream, then the peptide was precipitated with diethyl ether (10 mL) and the mixture cooled to -20°C. The precipitate was pelleted by centrifugation (7600 rpm, 10 min), the supernatant decanted. The pellet was dried under a nitrogen stream, and then dissolved in 1:1 ACN/H_2O, before syringe filtering (0.2 µm) and lyophilised to give the crude peptide as a white fluffy powder (~40% crude purity). MS (ESI+) Expected $[M+H]^+$ for $C_{36}H_{47}N_{11}O_8$: 762.8, observed: $[M+H]^+$ 762.6; R_t (C18, 0-100%, 20 min) 12.0 min.

TNBS Test: A small spatula of swollen resin taken out and 1 drop each of TNBS (100 µL 5% w/v picrylsulfonic/trinitrobenzenesulfonic acid in H_2O added to 900 µL of DMF) and DIPEA solutions (100 µL in 900 µL of DMF) added and allowed to develop for 1 min. Clear/yellow beads indicated no free amine (negative), while red/orange beads showed free amine was present (positive).

DESIGN AND SYNTHESIS OF BONE-TARGETING PRODRUGS FOR THE SELECTIVE TREATMENT OF BREAST CANCER BONE METASTASIS

Rouven Becker, Stephen Fitter, **Aimee J. Horsfall**, Andrew D. Abell and Andrew C. W. Zannettino.

*corresponding author

Manuscript in preparation

A2.12 Background

Breast cancer is the most prevalent malignant tumour, where an estimated 85% of all patients diagnosed with breast cancer are suffering from tumour metastasis to the bone, known as osteolytic bone metastasis. To reduce tumour progression and bone resorption, it is a viable option to conjugate anticancer drugs to a bone-targeting system, to create a pro-drug with a high affinity to calcium-rich bone surfaces. Acidic oligopeptides, (poly-Asp or poly-Glu) are one such methodology to target a drug to the bone, with lower toxicity than comparable bone-targeting compounds such as *bis*-phosphonates. Here, a poly-D-Asp peptide is conjugated to a first-in-class inhibitor of the kinase activity of mammalian target of rapamycin complex 1 and 2 (mTORC1 and mTORC2), AZD8055. Inhibition of the mTOR pathway which is associated with cell survival, proliferation, angiogenesis and growth, has shown to be a promising strategy to treat breast cancer. However, clinical studies towards the safety and tolerability of AZD8055 in advanced solid tumours have identified dose limiting adverse effects upon administration of AZD8055, which led to discontinuation of the drugs development.

The inhibitors excellent *in vivo* and *in vitro* activity combined with dose-limiting adverse-effects in patients, led to the conjugation of AZD8055 to a bone-targeting acidic oligopeptide via a cleavable ester-linker to form an inactive pro-drug (**Figure**), which has an increased affinity to the bone. After delivery to the desired tissue, local esterases have the ability to liberate the active compound to facilitate localised treatment at the bone, thereby mitigating side-effects due to off-target AZD8055 activity.

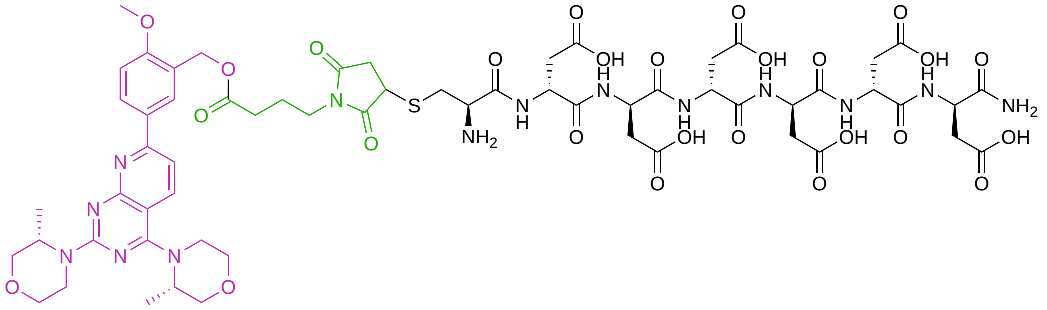


Figure 1. Bone-targeting conjugate: Active Drug, AZD8055 (purple); Linker, 3-Maleimidopropionic acid (green); bone-targeting acidic oligopeptide (black).

Contribution

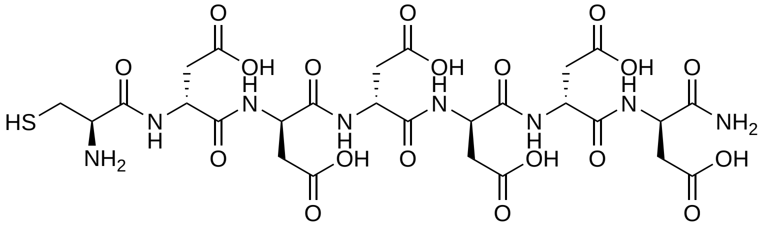
Design and synthesis of poly-D-aspartic acid peptides for the synthesis of bone-targeting prodrugs for the selective treatment of breast cancer bone metastasis.

A2.13 Synthesis and characterisation

Analytical methods

Peptides were purified by semi-preparative RP-HPLC on a Gilson GX-Prep system using a Phenomenex Aeris Peptide XB-C18 column (10 x 250 mm, 5 µm) at 4 mL/min and visualised at 220 and 254 nm. RP-HPLC solvents were (A) H_2O with 0.1% TFA and (B) ACN with 0.1% TFA. The purity of the final peptide was confirmed by analytical RP-HPLC on an Agilent 1260 HPLC equipped with a Phenomenex Aeris Peptide XB-C18 column (250 x 4.6 mm, 5 µm) over a gradient of 5-50% B (15 min) at 1.5 mL/min. Identity of the final compounds was confirmed by High Resolution Mass Spectrometry on an Agilent 6230 ESI-TOF LCMS.

*Solid-phase peptide synthesis of peptide **Cd6**: Cdddddd*



The following Fmoc-protected amino-acids were used for all peptides unless otherwise specified. All amino-acids were purchased from Chem-Impex Int'l.: Fmoc-D-Cys(Trt)-OH, and Fmoc-D-Asp(tBu)-OH. Rink Amide AM resin (0.2 mmol, 408 mg, 0.49 mmol/g, Chem Impex Int'l) was swollen in 1:1 DMF/DCM (10 mL) for 15 min. The mixture was transferred to a fritted syringe, the solution drained, and the resin washed with DMF (3 x 8 mL). The Fmoc-protecting group was removed by treatment of the resin with a solution of 20% piperidine in DMF (5 mL) for 10 min. The solution was drained and the resin washed with DMF (3 x 8 mL). Amino-acid couplings were achieved by addition of a solution of Fmoc-protected amino-acid (5 equiv), HCTU (5 equiv) and DIPEA (10 equiv, 348 µL) in DMF (8 mL), to the resin and stirred intermittently for 20 min. The solution was drained and the resin washed with DMF (5 x 8 mL). The *N*-terminal Fmoc-protecting group was removed by treatment of the resin with a solution of 20% piperidine and in DMF (5 mL) for 10 min, the solution was drained and the resin washed with DMF (5 x 8 mL). Successive couplings and Fmoc-deprotections were repeated to achieve the desired sequence. After the final Fmoc-deprotection, the resin was washed with DMF (3 x 5 mL), and DCM (5 x 5 mL) then dried with Et_2O (3 x 5 mL). The peptide was cleaved from the resin by addition of 95:2.5:2.5 TFA/TIPS/H_2O (6 mL) to the resin and rocked for 1 h. The TFA solution was pipetted from the resin and concentrated to 0.5-1 mL under a nitrogen stream, then the peptide was precipitated with diethyl ether (10 mL) and the mixture cooled to -20°C. The precipitate was pelleted by centrifugation (7600 rpm, 10 min), and the supernatant decanted. The pellet was dried under a nitrogen stream, then dissolved in 1:1 ACN/H_2O, syringe filtered (0.2 µm) and lyophilised to yield the crude peptide as a fluffy white powder. The peptide was purified by semi-preparative RP-HPLC using a Phenomenex Aeris Peptide XB-C18 column over an isocratic gradient of 2% B (20 min) at 4 mL/min, and the product containing fractions were pooled and lyophilised to give **Cd6** as a white fluffy powder. R_t (C18, 0-20%, 20 min) 12.0 min. HRMS (ESI+) Expected $[M+H]^+$ for $C_{27}H_{38}N_8O_{19}S$ (810.1974): 811.2052, observed: $[M+H]^+$ 811.2047.

That’s all folks….!