

Insights into Peptidomimetic Approaches for Inhibiting the DNA Sliding Clamps

A thesis submitted for the degree of Doctor of Philosophy

Bethiney Vandborg

**Department of Molecular and Cellular Biology,
School of Biological Sciences, the University of Adelaide**

November 2023



THE UNIVERSITY
of **ADELAIDE**

Table of Contents

Citation List.....	4
Thesis Summary.....	5
Thesis Statement of Declaration.....	6
Acknowledgments	7
Chapter 1: Introduction.....	8
1.1 Cancer.....	9
1.2 Current cancer treatments.....	9
1.3 DNA replication as a drug target.....	10
1.4 Processive DNA replication and repair is only possible through PCNA.....	11
1.5 Current PCNA inhibitors in research.....	12
1.6 DNA replication and repair regulation through protein p21.....	14
1.7 Peptidomimetic inhibitors: a new route of PCNA inhibition.....	16
1.8 p21 peptidomimetics create a new avenue of biomedical research into fungal infection treatments: enter <i>Aspergillus fumigatus</i>	17
1.9 Basic principles of Protein X-ray Crystallography.....	18
1.10 Basis of Fluorescence Polarisation.....	19
1.11 A structural approach to advance understanding of p21 peptidomimetic PCNA inhibitors.....	20
Chapter 2: Publication.....	21
Statement of authorship.....	22
“Production of recombinant human proliferating cellular nuclear antigen (PCNA) for structural and biophysical characterization.”	24
Abstract and Introduction	24
Materials and Methods.....	25
Results.....	27
Discussion.....	28
References.....	29
Supplementary Information	31
Chapter 3: Publication	35
Statement of authorship.....	36
“Unlocking the PIP-box: A peptide library reveals interactions that drive high-affinity binding to human PCNA.....	38
Abstract and Introduction	38
Results.....	40
Discussion.....	43
Conclusions.....	49
References.....	51
Supplementary Information	52
Chapter 4: Publication	78

Statement of authorship.....	79
<i>“A cell permeable bimane-constrained PCNA-interacting peptide.”</i>	81
Abstract and Introduction	81
Results & Discussion.....	82
Conclusions.....	89
References.....	89
Supplementary Information	91
Chapter 5: Publication	112
Statement of authorship.....	114
Graphical Abstract.....	115
<i>“Towards a High Affinity Peptidomimetic Targeting PCNA from Aspergillus fumigatus.”</i>	
Abstract.....	116
Introduction.....	116
Materials and Methods.....	118
Results.....	121
Discussion.....	126
References.....	129
Supplementary Information	132
8GJ5 Validation report.....	141
8GJF Validation report.....	156
Chapter 6: Manuscript.....	173
Statement of authorship.....	174
<i>“AI generated cyclized peptide for sliding clamp peptidomimetics for cancer therapeutic”</i>	176
Introduction.....	176
Materials and Methods.....	177
Results	179
Discussion and conclusion.....	184
Chapter 7: Discussion	188
Chapter 8: Bibliography of chapters 1, 6 and 7.....	192
Chapter 9: Appendix.....	197
<i>“Designing fluorescent nuclear permeable peptidomimetics to target proliferating cell nuclear antigen.”</i>	198
Supplementary Information	208

Citation List

Vandborg, B., Holroyd, DL, Pukala, T & Bruning, JB. Production of recombinant human proliferating cellular nuclear antigen (PCNA) for structural and biophysical characterization. . *Protein Expr Purif.* (2023) 212, 10.1016/j.pep.2023.106353.

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

Horsfall, A. J., **Vandborg, B. A.**, Kikhtyak, Z., Scanlon, D. B., Tilley, W. D., Hickey, T. E., Bruning, J. B., and Abell, A. D. A cell permeable bimeane-constrained PCNA-interacting peptide. *RSC Chemical Biology* (2021) 2, 1499-1508, 10.1039/D1CB00113B.

Vandborg, B. C., Horsfall, A. J., Pederick, J. L., Abell, A. D., and Bruning, J. B. Towards a High-Affinity Peptidomimetic Targeting Proliferating Cell Nuclear Antigen from *Aspergillus fumigatus*. *Journal of Fungi* (2023) 9, 1098,

Appendix

Horsfall, A. J., Chav, T., Pederick, J. L., Kikhtyak, Z., **Vandborg, B. C.**, Kowalczyk, W., Scanlon, D. B., Tilley, W. D., Hickey, T. E., Abell, A. D., and Bruning, J. B. Designing Fluorescent Nuclear Permeable Peptidomimetics to Target Proliferating Cell Nuclear Antigen. *Journal of Medicinal Chemistry* (2023) 66, 10354-10363, 10.1021/acs.jmedchem.3c00471.

Thesis Summary

In 2020, there were 48,266 deaths from cancer in Australia, a bewildering assortment of diseases that's caused from uncontrollable proliferation. One of the main problems of conventional cancer therapy is the low specificity of chemotherapeutic drugs for carcinogenic cells over healthy cells. This highlights the necessity for effective cancer treatments and therapeutics.

Human proliferating cell nuclear antigen (hPCNA) is a DNA replication processivity factor, which acts as a docking platform, allowing proteins to have access to the replication fork, and increasing the affinity of these DNA interacting proteins. PCNA is a promising target for the inhibition of cancer cell growth as it is critical for cell survival. The trimer structure of PCNA forms a ring-shaped oligomer allowing DNA to pass through the middle and interacting proteins to dock on the outside of the ring. Without this structural formation, there is a loss of DNA replication and repair in the cell. Due to the location of subunit-subunit termini, the addition of a purification tag can hamper crystallography and biophysical experiments, as the trimer complex folding can be impeded. To avoid these complications, a tag-less, step-wise purification was implemented, allowing the further structural research of this protein for cancer treatments.

The known regulator protein p21 has been used to investigate new approaches for targeting and inhibiting PCNA for the development of cancer therapeutics. The p21 protein has been shown in the human system to block binding to the docking platform and halt cell replication. A peptide of the binding domain of p21 has been shown to bind and maintain the affinity of the full protein.

Structural modifications of peptides guided by rational design and molecular modelling have been established to develop novel synthetic approaches, this is peptidomimetic research. This can be done using the PIP-box of p21, the PCNA interacting protein box, a sequence which allows tight interaction with PCNA. A peptide using the PIP-box of p21 has been investigated in previous literature and research for a PCNA inhibitor. A mutagenesis approach has been used to identify what characteristics increase or decrease peptide affinity for PCNA. This is so the differences in PIP-boxes can be shown at a residue level, to identify structural points which form a tightly bound confirmation. This has been done to reveal a new peptide sequence with a stronger binding affinity than the native protein.

This isn't the full extent of the use of this peptidomimetic research, it can be used on other species, as PCNA has been shown to be present in almost all forms of life. *Aspergillus fumigatus* is abundant in the environment and the most common cause of invasive fungal infection. Central nervous system aspergillosis has mortality rates of ~90%. As PCNA has important roles in replication and cell survival, it is hypothesised that it is a potential target for antifungal treatments, illustrating the importance of investigating the crystal structure of *A. fumigatus* PCNA, to understand how interacting peptides bind to aid rational drug design. This was done to reveal a newly designed mimetic with a new secondary structure with high affinity binding which can be utilised in further fungal mimetic research.

This thesis presents a number of peer-reviewed publications which shed light on peptidomimetic use on the protein PCNA for a number of therapeutic pathways. The result is critical developments in the beginning of the drug developmental pipeline. In addition, this thesis includes publications highlighting the merit of repurposing the peptidomimetics synthesised for hPCNA in the development for a fungal treatment for *A. fumigatus*.

Thesis Statement of 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. The author acknowledges that copyright of published works contained within the 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.

Bethiney Vandborg

Acknowledgements

Thank you to Dr. John Bruning for this opportunity and all your guidance, I am lucky to have been part of these projects. I've wanted to study biomedical science and do research since year 10, I am grateful to be given the opportunity to study something I find interesting and feel will make a difference since my honours year and through my PhD. You have given me every opportunity to succeed and have always looked out for my best interests, I'm so grateful to have had you as a mentor for the past years, and I hope for many more to come.

Thank you to the rest of Bruning Lab for their support, especially Jordan Pederick and Aimee Horsfall. Their help and guidance throughout my candidature has been taught me so much and I could not have done anything without you.

Thank you to Ann Patchett, the lords gift.

Thank you to my Jakeb, you're the best thing to ever happen to me and I would have achieved nothing without you.

Chapter 1:

Introduction

Introduction

1.1 Cancer; a bewildering assortment of diseases that kills 7.5 million people each year.

In 2020, there were 48,266 deaths from cancer in Australia [1]. Although the mortality rate of cancer has decreased over the past decade in both males and females, it is still the third most common cause of mortality in Australia, following coronary heart disease and dementia [2]. Various forms of cancer caused thousands of deaths in Australia in 2022, with lung cancer outranking all others at over 8,000 deaths (Figure 1). Generally, the cause of cancer development is a genetic defect in one or more proteins that regulate the cell and DNA replication cycle [3]. These genes, which in certain circumstances can transform a cell into a tumour cell, are described as oncogenes. Oncogenes can be classified into five groups, growth factors, growth factor receptors, signal transducers, transcription factors, and those including programmed cell death regulators[4]. This leads us to the question, can investigation of these areas of cell regulation lead us to a new treatment target?

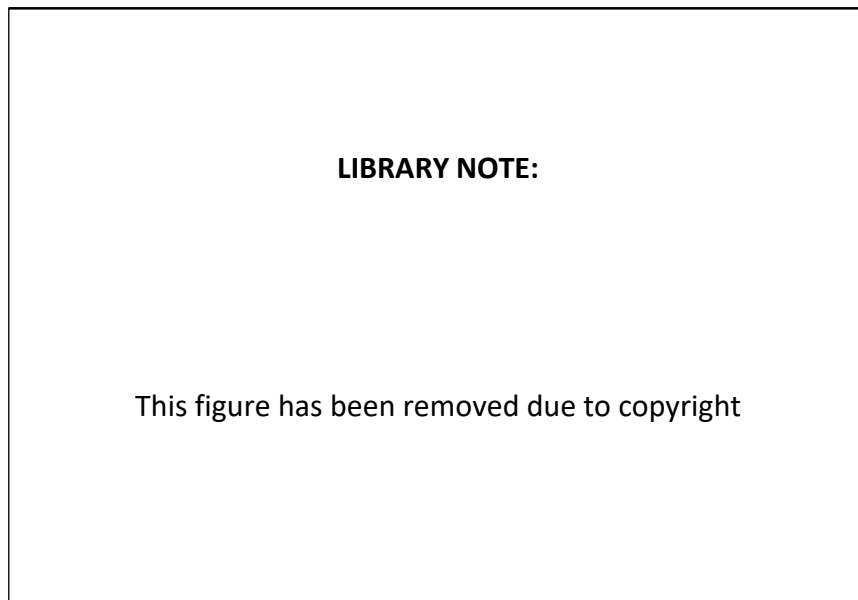


Figure 1. Estimated cancer mortality in Australia in 2022. Figure adapted from *Cancer Australia*. [1].

1.2 Current cancer treatments: lack of specificity currently thwarts a cure for every patient.

One of the main problems of conventional cancer therapy is the low specificity for carcinogenic cells. Most cancer therapeutics act both on healthy and diseased tissues, generating severe side effects for the patients [5]. During cancer progression, tumours become highly heterogeneous, creating a mixed population of cells characterised by different molecular features with various outcomes for different therapies. This difference in response to treatments informs us of how resistant phenotypes come about, which are promoted by the selective pressure from the treatment administered [6]. This occurrence of resistance from application of trusted cancer therapeutics creates a necessity for new treatments and targets to conquer cancer.

The most common form of cancer treatment is chemotherapy, the use of various agents to target the macromolecular function of neoplastic cells by interfering or inhibiting DNA or protein production and function. When this is successful this usually triggers cell death[7]. This is critical to avoid the eventual invasion and metastasis of the cancer cells if left untreated. The results of this administration can lead to toxic effects on the healthy cells of the patient, such as Vinca alkaloids and Taxanes, which block mitotic spindle formation[8]. Common toxicity produced from these agents include nausea, vomiting, diarrhea, alopecia, fatigue, sterility, and infertility[9]. Another major side effects include myelosuppression, also known as bone marrow suppression, which results in reduced production of blood cells[10]. The various blood disorders that arise from this lead to increased risk of infection for the patients.

Radiation as a cancer therapeutic is a more local treatment. At high doses, radiation therapy kills cancer cells or slows their growth by damaging their DNA, also leading to apoptosis of the cells of interest [11]. The main source of this treatment is External and Internal Beam Radiation Therapy[12]. External therapy is where the radiation is aimed at the localised cancer, whereas internal is where the radiation is given in the form of a capsule or liquid. [13]. This comes with a similar array of severe side effects depending on the location of the tumour and therefore the radiation, most common of those being fatigue, hair loss, nausea and vomiting, and potentially affecting the localised tissue and organ, because although more localised, is still not carcinogenetic specific [14]. This presents again a need for more specific treatments.

1.3 DNA replication as a drug target.

As mentioned, the key to cell survival is DNA replication. At the replication fork, two DNA strands are synthesised through two different mechanisms. The leading strand is replicated by 5'-3' polymerases. The other strand, termed the lagging strand, is replicated through the Okazaki fragments in a 3'-5' fashion[15]. Primase creates an RNA primer as a template for the DNA synthesis by polymerase α . Replication factor C (RFC) binds and catalyses the loading of the ring protein, proliferating cell nuclear antigen (PCNA), which allows the association of Pol δ and Pol ϵ to orchestrate DNA lagging strand and leading strand synthesis, respectively [16] (Figure 2). PCNA enhances the processivity of these enzymes, which carry out the bulk of DNA synthesis. These polymerases contain the 3'-5' exonucleases, reducing the misincorporation of nucleotides. In lagging strand synthesis, when the polymerases reach an end from a previous Okazaki fragment, it partially displaces this fragment by ongoing DNA synthesis, and a flap structure is generated, this recruits flap structure-specific endonuclease-1 (FEN-1), and the resulting nick from the endonuclease activity is sealed by DNA ligase I in the repair process [17].

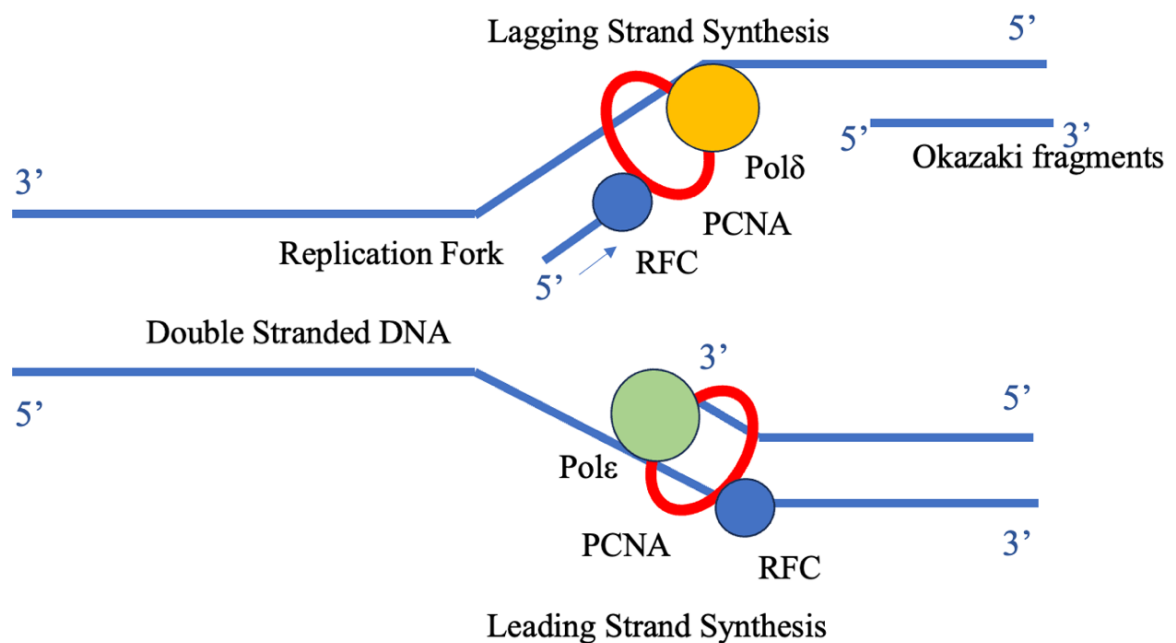


Figure 2. PCNA during DNA replication and synthesis, allowing DNA polymerases to associate at the leading and lagging strands. DNA polymerase δ and ϵ are shown in orange and green respectively. PCNA is shown as a red ring. RFC (Replication Factor C) shown in blue. DNA shown as blue lines, including Okazaki fragments.

1.4 Processive DNA replication and repair is only possible through PCNA.

Proliferating cell nuclear antigen (PCNA) is a DNA processivity factor. PCNA is attached to DNA in an ATP dependent manner by clamp loader proteins [18]. It is required for DNA replication by allowing access of polymerases and a host of other proteins required for DNA replication and repair. It does this by acting as a docking platform, and increasing the affinity of DNA polymerase to DNA by over 60,000- fold [19].

PCNA is a ring shaped homotrimer and is therefore composed of three subunits. Each subunit has two domains of two alpha helices, and nine beta sheets. The two domains are held together by the interdomain connector loop (IDCL), which spans amino acid residues M121-Y133 [19]. The ring structure is critical to form a central cavity of 35.4 Å, through which DNA slides (Figure 3) [19].

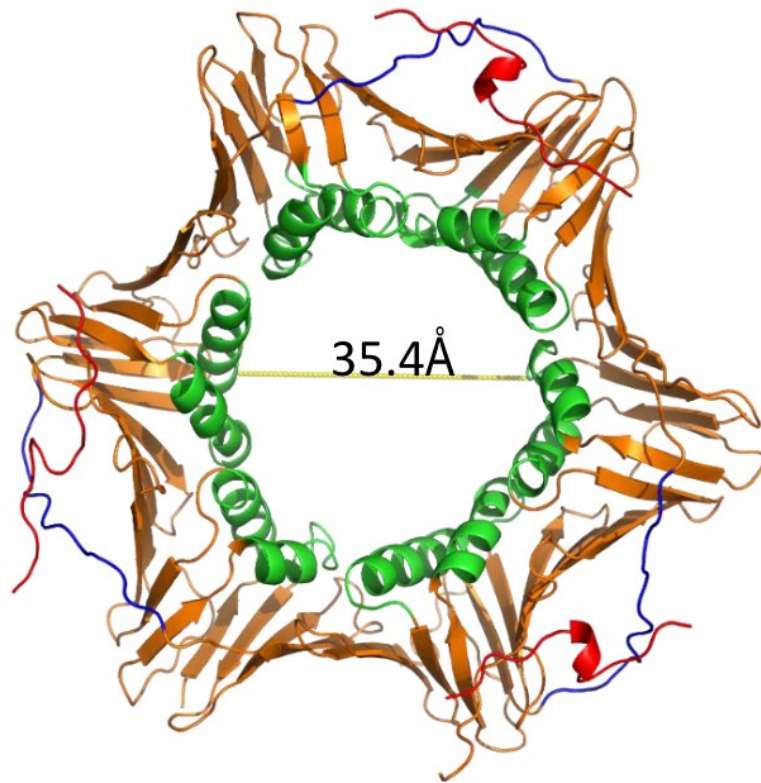


Figure 3. Structure of human PCNA bound to a peptide. A peptide shown in red indicates binding domain of each homotrimer at the interdomain connecting loop (IDCL) shown in blue. The inside of the protein lined with positively charged arginine and lysine residues are shown in the green alpha helices, this 35.4 Å cavity through which DNA slides. The monomers of the homotrimer consist of beta-sheets shown in orange (PDB:7KQ1).

How is it related to human health? It was discovered by gene knockout of PCNA in mice, that PCNA is critical for survival, as the result is embryonic lethality[20]. It has been discovered that PCNA acts as a marker for malignant cells and tumour, as PCNA is overexpressed in nearly all forms of cancer, at a six-fold increase [20]. It has also been reported that the malignancy of the tumour correlates with the level of PCNA protein[21].

Due to its critical role, inhibitors of major proteins of the replication process are commonly used in anti-cancer and anti-viral therapies[22]. This leads us to the research being done on the use of PCNA as a target for cancer therapeutics.

1.5 Current PCNA inhibitors in research.

hPCNA targeting inhibitors have been shown to have more profound inhibitory effects on tumour cells versus normal cells, providing a therapeutic window (Figure 4a and b) [21]. This has been shown with the use of antisense oligonucleotides targeting messenger RNA encoding hPCNA that hybridise to inhibit gene expression. This provides a significant advantage in contrast to many chemotherapeutic drugs used. This differential sensitivity to cancerous cells is likely due to higher demands of hPCNA in tumour cells, and tumour cells are also genetically more unstable due to their accumulation of gene mutations and require more hPCNA for DNA repair to survive [21]. These cells

are therefore more vulnerable to a hPCNA-targeting agent. This information leads to the hypothesis that PCNA could be a potential drug target to treat cancerous cells.

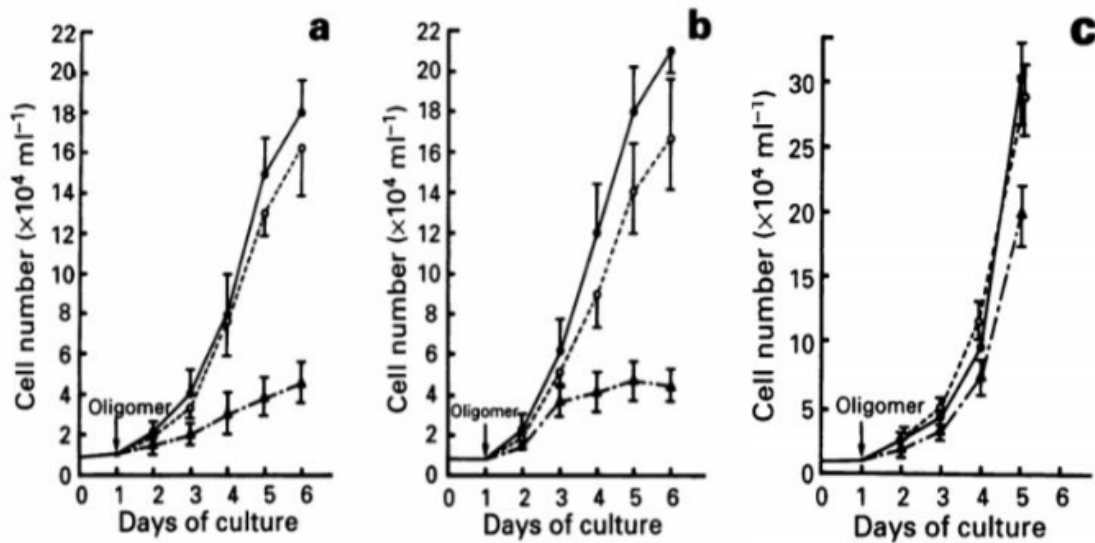


Figure 4. Growth of gastric cell lines and control cell line post dose dependent application of antisense oligonucleotide (triangle) targeting mRNA encoding PCNA. Sense oligonucleotides with random sequence was also tested and showed no inhibition (white circle). Phosphate saline (black circle) was used as a negative control. A) Gastric cell line. B) Gastric cell line. C) Human fibroblast cell line. Figure adapted from Sakakura et al 1994.

Previous work on PCNA inhibitory compounds used the PCNA IDCL as a basis for a binding inhibitor. This compound being based specifically on the IDCL site of cancer specific PCNA. In a broad range of cancer cells and carcinogenic tissues, an isoform of PCNA is expressed (termed caPCNA) which has a specific antigenic site in the IDCL between residues Leu126 and Tyr133 [23]. As previously mentioned, this is also the major binding site of DNA interacting proteins. This compound termed R9-caPep (caPCNA Peptide), is an eight amino acid residue sequence corresponding to the sequence of the IDCL mentioned attached to the C terminus of a nine D-arginine sequence with a spacer of two cysteine residues to achieve a cell permeable peptide [23].

In cellular experimentation, it was found that R9-caPep was selectively toxic to breast, lung and pancreatic cancer cell lines and neuroblastoma cell growth when compared to that of non-cancerous cell lines tested [23]. As previously mentioned, the mechanism for this selectivity of PCNA specific therapeutic designs is not clear.

Experimentation using sequence and confocal microscopy to assess the localisation of specific DNA interacting proteins and PCNA in the nucleus was used after application of the inhibiting peptide and a control of the same residues but in a scrambled sequence. This revealed R9-caPep blocked the localization of Flap endonuclease-1 (fen-1) and DNA ligase (LIG1). However, it did not block the interaction of p21 or the recruitment of DNA polymerase (POLD3) to PCNA (Figure 6) [23].

The differences in the ability of R9-caPep to block the recruitment of different PCNA interacting proteins may reflect the different affinities of these proteins for PCNA, suggesting that the compound may have the ability to stall replication and repair at the replication fork [23].

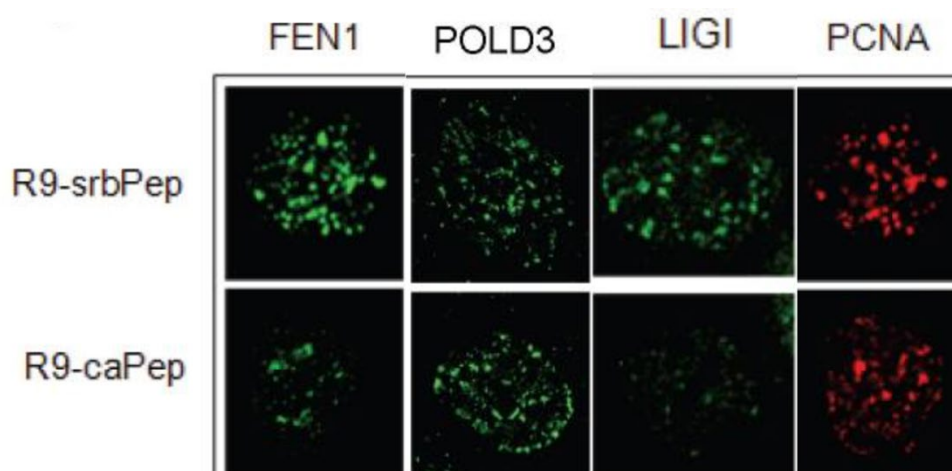


Figure 5. Fluorescence confocal microscopy of nuclear localisation following DAPI counterstaining using mouse anti-FEN1, mouse anti-POLD3, mouse anti-LIG1, and goat anti-PCNA antibodies. Figure adapted from Gu *et al* 2014 [23].

Another method of investigation into PCNA inhibition comes from blocking PCNA binding to chromatin. In silico screening found docking partner structures to synthesise into compounds that could potentially selectively bind to PCNA trimers to reduce the level of chromatin-associated PCNA in cancer cells[24]. These compounds termed PCNA-Is, consisted of the small molecule PCNA-I1, which was found to bind PCNA and reduce the trimer stabilisation. By reducing the association to chromatin it induces cancer cell arrest at the S and G2/M phases. This caused the growth of tumour cells of various tissue types to decrease. However, this did also affect the growth of normal nontransformed cells, so this faces the same problem of specificity as other cancer treatments, but this only occurred at significantly higher of 1.6 μ M[24].

1.6 DNA replication and repair regulation through protein p21.

p21 is a tumour suppressor protein which acts as a competitive inhibitor for the binding site on PCNA, as it's affinity for PCNA is much higher than other associating proteins at 83 nM [19]. It regulates the cell cycle by inhibiting cyclin dependent kinases, which control the cell cycle progression. It is upregulated when DNA damage increases[25]. When p21 binds to PCNA, it blocks its function and the access of DNA replication proteins, and therefore halts processes at the replication fork.

Cancer cells may exhibit increased concentrations of PCNA to high levels as a mechanism that may saturate the pool of p21 available to bind, allowing a significant portion of uninhibited PCNA to continue DNA replication. This raises the question, how can PCNA be inhibited for therapeutic purposes, so that the excessive replication of cancerous cells can be stopped? Many PCNA inhibitors are small molecules that have only sub-micromolar potency, and there is no information to whether further off target effects or further toxicity will occur [26]. Small molecules are often

degraded by proteases, and do not survive in the bloodstream. Perhaps as new avenue for a PCNA inactivation could be through our native inhibitor, protein p21?

The PIP-box is a conserved sequence found in PCNA interacting proteins, which allows it to interact and bind the PCNA surface, depending on the conservation of the sequence: QXX(M/L/I)XF(Y/F) (Table 1) [27]. Of the data on PIP-boxes sequences and affinities for interacting with PCNA, the binding data varies depending on the sequence and conservation of residues. These differences are hypothesised to be essential for the interacting roles these proteins play together during DNA replication and repair, and for moving proteins in and out of the replication fork, allowing a correlation to be observed between the proteins function and the tightly binding of affinity for PCNA [19].

Table 1. Various PIP-box sequences among major PCNA interacting proteins, including FEN1 (flap endonuclease 1) and DNA mismatch repair protein, MutS Homolog 3 (MSH3). Conserved positions 1, 4, 7, and 8 that match the consensus sequence are underlined.

Protein	PIP-Box
p21	<u>Q</u> T <u>S</u> M <u>T</u> D <u>F</u> <u>Y</u>
FEN1	<u>Q</u> G <u>R</u> L <u>D</u> D <u>F</u> <u>F</u>
p66 (DNA pol δ)	<u>Q</u> V <u>S</u> I <u>T</u> G <u>F</u> <u>F</u>
DNA pol η	M <u>Q</u> <u>T</u> <u>L</u> E <u>S</u> <u>F</u> <u>F</u>
DNA pol κ	K <u>H</u> <u>T</u> <u>L</u> D <u>I</u> <u>F</u> <u>F</u>
DNA pol ι	K <u>K</u> <u>G</u> <u>L</u> I <u>D</u> <u>Y</u> <u>Y</u>
MSH3	<u>Q</u> A <u>V</u> L <u>S</u> R <u>F</u> <u>F</u>

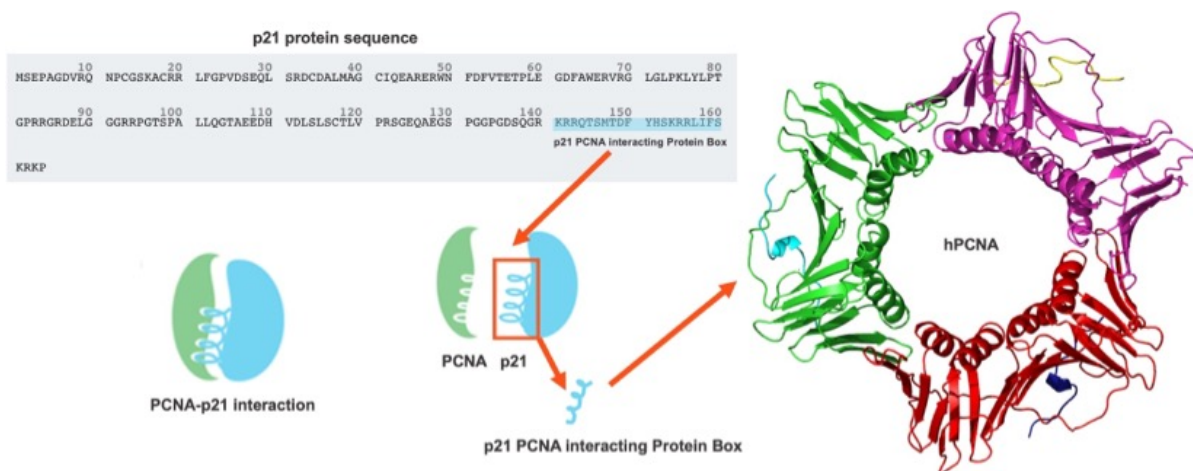


Figure 6. Schematic demonstrating the p21 protein PIP box location in sequence and the isolation of the PCNA interacting Protein Box in a short sequence demonstrated in Gulbis et al 1996, which binds to the Interdomain Connecting Loop of PCNA. PCNA monomer shown in green, then in trimer with other monomers in magenta and red. P21 isolated peptide shown in cyan, and then in trimer with other peptide in yellow and dark blue (PDB:7KQ1). P21 protein sequence adapted from UniProt (UniProt:P38936).

The significance of the PIP-box of p21 was demonstrated through the research done in Gulbis et al 1996. Through the isolation of the p21 PIP-box with limited flanking residues, a 22 amino acid peptide was synthesised which maintained high affinity binding for PCNA at a binding constant (K_D) of approximately 2.5 nM, as measured by biosensor analysis and efficiently inhibits DNA synthesis *in vitro* [28]. The C-terminus of the protein which contains the PIP-box forms a 3_{10} helix and inserts into the hydrophobic pocket on PCNA next to the IDCL, like several other PIP-box peptides (Figure 6) [19]. By forming this helix when bound to PCNA, it exposes the side chains of the peptide residues, which makes interactions with the hydrophobic patch and each other. The C-terminus of the PIP-box forms an antiparallel β -sheet structure with the IDCL [19]. This finding that a relatively short peptide can perhaps capture the binding and inhibitory capacity of the full p21 protein leads to the hypothesis that these characteristics are important for interaction of PCNA with a therapeutic. Perhaps this peptide could be further improved for a cancer treatment? [27].

1.7 Peptidomimetic inhibitors: a new route of PCNA inhibition.

Small protein-like chains designed to mimic peptides can be synthesized and could be used as an hPCNA inhibitor and cancer therapeutic [29]. Compounds whose elements mimic a natural protein and which retain the ability to interact with a biological target and produce a biological effect are peptidomimetics[30]. Peptidomimetics can be modified to overcome problems of previous inhibitors by improving the drug likeness and permeability [29].

Cell penetrating peptide (a type of nuclear localisation sequence)-fused elastin like peptides have been used previously to deliver a peptide mimetic of p21 to PCNA in several cancer models *in vitro* [31]. The polypeptide was taken up by HeLa cervical cells and localized to the cytoplasm. This study demonstrates HeLa cells treated with CPP-fused ELP with p21 fused to the C-terminus had an antiproliferative effect, but not control polypeptides.

By understanding the interaction between hPCNA and binding proteins, the affinity of an inhibitory peptide can be optimised. Other PIP-box variations may lead to development of a small peptide inhibitor mimetic which contains a series of advantageous residues and chemical improvements, forming interactions that increase affinity.

One of these improvements may be the permanent holding of secondary structure, as p21 in aqueous solution does not have a defined secondary structure[32]. It forms a 3_{10} helix upon binding to PCNA allowing it to bind tightly[27]. By maintaining this structure in solution it may be able to bind to PCNA with a lower entropic cost and therefore more readily[26]. In order to give the p21 peptide this permanent secondary structure, it requires peptidomimetic chemistry to covalently constrain it into the critical 3_{10} helix using a $i, i+4$ lactam bridge. Lactam bridges link amino acid residues that are multiple residues apart in the sequence, giving a flexible method for introducing conformational constraints into a peptide structure [33]. The lactam bridge will link side chains to constrain the helix confirmation. Residues Thr145 and Asp149 were found [26] to be ideal candidate peptides that the lactam bridge be designed to constrain. These amino acids are ideal for modification as the side chains

are not involved in interactions with PCNA, therefore no alteration to the binding will be made, and are within the right distance. Two peptidomimetics have previously been formed [26]. Firstly, ACR1, which has residue 145 substituted to L-lysine, and 149 to L-glutamate, giving it an eight-atom length. Secondly, ACR2, which has residue 145 substituted with DAB (diaminobutyric acid, a shorter analogue of lysine), and 149 with glutamate, giving this a six-atom length. These substitutions were made to avoid any extensive irreversible aspartimide formation problems.

These two analogues were synthesized, purified, and then tested using Nuclear Magnetic Resonance (NMR) to confirm the intended helical structure. The crystal structure of PCNA in complex with ACR2 was solved, and research showed that the peptidomimetics constrained the 3_{10} helical secondary structure, and that it binds PCNA similarly to that of the wildtype p21 protein. Therefore, this provides a template for further mimetic optimisation.

In attempts to optimise the mimetic further, four N-terminal residues (139-GRKR-142) were removed. However, this resulted in a significant decrease in affinity, to a K_D in the low mM range. This is evidence of the importance of the electrostatic interactions of these residues in the N-terminus [27]. This all suggests that for a peptidomimetic to be successful, the mimetic must include the four N-terminal residues, the constrained 3_{10} helix structure, and the constrained β -sheet structure. These are the characteristics required for a peptidomimetic to mimic p21.

The sequence of the PIP-box of p21 has also been investigated regarding improving its characteristics for its affinity for PCNA [19]. It has been found that there is a tyrosine amino acid in place of a phenylalanine residue in the eighth conserved position [19] which is present in many PIP-boxes. The significance of the hydroxyl group at the specific position has been investigated. This was done through mutagenesis. This residue was changed to phenylalanine, and co-crystallised in complex with PCNA, to compare the structure's affinity data, measured using isothermal titration calorimetry (ITC), to the same complex containing the wildtype p21 peptide [19].

The angle of the phenylalanine side chain in the p21 PIP-box differs from the wildtype residue by a 38° angle. This change of angle discourages the tight packaging into the pocket on PCNA. The wildtype residue of Tyr151 of the p21 PIP-box is shown to sit tightly in the pocket [19]. It was discovered there is a 3-fold difference between these affinities [19]. The decrease in affinity of the mutated peptide is due to the loss of the hydrogen bonding provided from the tyrosine side chain oxygen atom. The residue Tyr151 acts to tether the peptide to PCNA, increasing tight packaging and recognition of the peptide. This increases the affinity of p21 to PCNA.

This development sheds light on the mutagenic approach that could be used to develop a small peptide inhibitor mimetic. This could contain a series of advantageous residues, forming affinity increasing interactions. A mutagenesis approach can be used to identify what characteristics increase or decrease peptide affinity for PCNA. This is so the differences in PIP-boxes can be shown at a residue level, to identify structural points which form a tightly bound confirmation. [19].

1.8 p21 peptidomimetics create a new avenue of biomedical research into fungal infection treatments: enter *Aspergillus fumigatus*.

Invasive fungal infections are a prevalent cause of death in immunocompromised patients [34]. A major fungal pathogen causing such infections is *Aspergillus fumigatus*, a filamentous fungus usually present in decaying organic matter [35]. The conidia of *A. fumigatus* inhaled from the environment can be cleared from the lungs by a healthy immune system, however when this fails, the conidia can reach the lower airways and evade host immune cells [36]. This can result in infection of the bronchi and sinuses, and dissemination to the brain and other vital organs through the circulatory system [37]. This is known as invasive aspergillosis. When invasive aspergillosis invades the nervous system, it has mortality rates of up to 90% [38].

As the infection rates of *A. fumigatus* increase there have been more negative clinical outcomes, largely due to antifungal drug resistance[39]. Although many antifungals are available, the mortality rates remain high due to the development of drug resistance in *A. fumigatus* [40]. Current treatments of invasive aspergillosis target components of the cell membrane, for example the antibiotic amphotericin B [41]. Unfortunately, in addition to increasing resistance rates, amphotericin B is associated with severe side effects, the most notable being kidney and liver toxicity[42]. As such, there is an urgent need for improved diagnostic protocols and a broader range of antifungal options.

As discussed previously, PCNA is a protein with important roles in replication and has been shown to be crucial to cell survival [43], demonstrating its value as a target for cancer treatments, and now also as a target for antifungal treatments, illustrating the importance of investigating the crystal structure of *A. fumigatus* PCNA, to understand how interacting peptides bind to aid rational drug design. Previous studies have solved the X-ray crystal structure of PCNA from *A. fumigatus* (afumPCNA) to 2.6Å resolution [44].

Comparison of afumPCNA to hPCNA reveals a trimeric tertiary structure identical to that of hPCNA, this revealed a root-square standard deviation (RMSD) of 0.939 Å[44]. Despite this, the proteins only have a sequence similarity of 53% [44]. This is seen in the changes in residues involved at the IDCL and therefore interaction with binding partners.

Fluorescence polarisation experiments of afumPCNA and the PIP-box of human p21 protein in Marshall et al 2017 has proven their interaction, suggesting that afumPCNA interacts with DNA binding proteins using a conserved PIP-box mechanism like hPCNA [44]. Further investigation therefore into the human PIP-box sequence may indicate the characteristics for high affinity binding to the afumPCNA PIP-box binding domain. Investigation also into *A. fumigatus* DNA binding proteins sequences to search for a PIP-box like sequence may shed light into the mechanism of binding to the afumPCNA surface. Proteins known to be important in DNA replication and repair cycle such as flap structure-specific endonuclease 1 (FEN1) and replication factor C (RFC) and DNA ligases and polymerases would be essential to include a PIP-box binding mechanism to afumPCNA, and where this investigation will begin [45].

Given that *A. fumigatus* does not have a known p21 equivalent, and this p21 derived peptide shows a high affinity interaction, further investigation into the human PIP-box sequence interactions with afumPCNA may allow for the design of a high affinity mimetic.

1.9 Basic Principles of Protein X-ray Crystallography

X-ray crystallography allows crystalline atoms to diffract X-rays, the intensity and angle of the diffraction generates three-dimensional electron density data captured as images [46]. Using this, the positions of atoms in a crystal and chemical bond lengths are determined through statistical analysis and computational methods[47]. The basic steps of this experimentation to produce protein structures include: Protein purification, protein crystallisation, X-ray diffraction, data collection and processing, scaling of data, molecular replacement, structure building and refinement through electron density map refinement and validation[48].

During processing and scaling of data, certain parameters are used to achieve a desirable resolution, and decide the cut off for data inclusion before refinement. The most important for this is considered the CC1/2 [49]. This is Pearson's correlation coefficient, where the multiple collections of data for each reflection is split in half; then a correlation is calculated between the average of both data sets. The higher this parameter, the more correlation there is between the two data sets and is therefore considered more desirable data, this value should not be lower than 0.3. Another important indicator of data quality is the completeness, which is the percentage of possible reflections measured that were genuinely measured[50]. For every protein, each space group has a theoretical number of

reflections that should be observed. The completeness percentage of this data for the overall and highest shell should be 95% or higher. The highest shell should never aim less than 80% and the overall percentage should never aim less than 93%.

The I/sigma parameter, although not as essential as achieving a high completeness and CC1/2 value, represents the ratio of data signal to noise. It is desirable to achieve an I/sigma value of at least 2 in the highest shell. The higher the signal to the level of background data in the higher shell, ordinarily indicates higher quality of data[50].

Following data scaling, the next major step is molecular replacement. This is a method of trial and error through model refinement to find a suitable model and solve the crystal structure. A key indicator of the success of molecular replacement is if the new refined map shows new structural information that was not present in the starting model, i.e. new side chains, secondary structures such as loops that differ to the starting model, and perhaps a bound ligand. During refinement in the later stages of structure solving, model bias is a large problem to overcome. Using a high resolution structure, model bias can be removed using density modification methods. When using data with a low resolution, which is considered worse than 3 Å, refinement can be unsuccessful in removing model bias[51]. The solution may be correct, but the poor model may not be able to interpret new areas of the map for refinement of results [52].

1.10 Basis of Fluorescence Polarisation

The primary methodology of chapter 6 is the use of fluorescence polarisation experimentation to investigate different sequences of cyclical peptidomimetics interactions with PCNA. Surface plasmon resonance (SPR) has been the primary experimental method of determining indicating values and interpreting peptide sequences for binding to PCNA surface throughout this thesis. However, when it came to these newly synthesised peptides, which do not use the same secondary structure as the p21 peptide from which we have used as the template, the data came back irregular as if the peptides did not bind at all to the PCNA surface.

PCNA can be sensitive to acidic solutions, and due to the trimeric nature of PCNA, sensitive to DMSO addition which can easily disrupt the tertiary structure and therefore affect the binding site in experimentation [53]. Upon use of different techniques, such as this fluorescence polarisation technique utilised in Wegener et al 2018 [26], which does not require buffer reagents that could impact the PCNA structure, results showed the expected curves representing protein interaction. The probe, the peptide of interest, is two-fold serially diluted into solutions containing a constant concentration of a fluorescently labelled control peptide, which is the competition for the peptide of interest to outcompete, and purified hPCNA. This interaction is tested using fluorescence polarisation competition binding. The fluorescence polarization is measured by a plate reader.

Competing ligand concentrations from serial dilutions and polarisation binding values obtained are then plotted on the on X and Y axis respectfully using GraphPad prism analysis tools. The resulting polarisation values were used to calculate the K_i , which represents the equilibrium binding affinity for a ligand that reduces the activity of its binding partner, in this case hPCNA.

K_i value determination was done by the following equation determined by Nikolovska-Coleska 2004 [54]:a

$$K_i = I_{50} / L_{50} / K_d + P_0 / K_d + 1$$

The concentration of the free inhibitor at 50% inhibition (I_{50}) (also known as IC50). The concentration of the free labelled ligand at 50% inhibition (L_{50}). the concentration of the free protein at 0% inhibition (P_0). The dissociation constant of the protein-ligand complex (K_d), defined in previous research [26]. This equation was derived from the basic principles of a competitive binding assay and

is not restricted by the concentrations of the protein and the ligand[54]. This is explained in more detail and with more information in chapter 6.

1.11 A structural approach to advance understanding of p21 peptidomimetic PCNA inhibitors: Aims and Scope of the Project.

This thesis focuses on the mechanism of p21 peptidomimetics interaction with PCNA for more favourable clinical potential over current cancer and fungal treatments. A structural approach is used, along with biochemical validation techniques, to understand certain modification in peptide design.

A new purification method which aimed to provide a replicable system to obtain tag-less protein, allowing the further structural research of PCNA for cancer treatments, was utilised in proceeding research. Using this new method, new PCNA targeting mimetic structures have been found. Investigation of the p21 native sequence has the aim to synthesise rationally designed peptides, which lead to the highest affinity found of any p21 peptidomimetic.

A foundational step in peptidomimetics aims to allow the peptide to have cellular permeability, which was investigated in macrocyclic peptides which bound to hPCNA with high affinity, and initial research suggests this linker affords cellular permeability. Another aim of interest was to determine new AI generated sequences for binding PCNA, the resultant candidates from Pepticom and of this preliminary research have been identified through fluorescence polarisation experimentation. And finally, investigating the crystal structure of *A. fumigatus* PCNA bound with fungal inspired mimetics, aims to understand how interacting peptides bind to aid rational drug design for fungal infections. This research in particular fills the gap of knowledge in fungal PIP-box and PCNA binding mechanisms.

A number of peer review published articles are presented, which show an improvement in peptide mechanism and design, as well as some new secondary structure designs for future research. This research reveals a new avenue for p21 peptidomimetics as drug candidates for PCNA.

Chapter 2: Publication

“Production of recombinant human proliferating cellular nuclear antigen (PCNA) for structural and biophysical characterization.”

Statement of Authorship

Title of Paper	Production of recombinant human proliferating cellular nuclear antigen (PCNA) for structural and biophysical characterization.
Publication Status	<input checked="" type="checkbox"/> Published <input type="checkbox"/> Accepted for Publication <input type="checkbox"/> Submitted for Publication <input type="checkbox"/> Unpublished and Unsubmitted work written in manuscript style
Publication Details	Vandborg, B., Holroyd, DL, Pukala, T & Bruning, JB. Production of recombinant human proliferating cellular nuclear antigen (PCNA) for structural and biophysical characterization. . Protein Expr Purif. (2023) 212, 10.1016/j.pep.2023.106353.

Principal Author

Name of Principal Author (Candidate)	Bethiney Vandborg
Contribution to the Paper	Protein synthesis and purification, validation experiments, discussed results and wrote and edited manuscript.
Overall percentage (%)	70%
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by regulations or contractual agreements with a Re: I am the primary author of this paper. thir
Signature	ie 29/08/2023

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 to 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	Dayna Holroyd
Contribution to the Paper	Validation experiments, discussed results and wrote and edited manuscript.
Signature	Date 8/09/2023

Name of Co-Author	Tara Pukala
Contribution to the Paper	Validation experiments, discussed results and wrote and edited manuscript.
Signature	Date 8/09/2023

Please cut and paste additional co-author panels here as required.

Name of Co-Author	John Bruning		
Contribution to the Paper	Supervised BM and DH, discussed results and edited manuscript.		
Signature		Date	8/09/2023



Production of recombinant human proliferating cellular nuclear antigen (PCNA) for structural and biophysical characterization

B. Vandborg^{a,b}, D.L. Holroyd^b, T. Pukala^c, J.B. Bruning^{a,b,*}

^a Institute of Photonics and Advanced Sensing (IPAS), School of Biological Sciences, The University of Adelaide, Adelaide, South Australia, 5005, Australia

^b School of Biological Sciences, The University of Adelaide, Adelaide, South Australia, 5005, Australia

^c Department of Chemistry, School of Physical Sciences, The University of Adelaide, Adelaide, South Australia, 5005, Australia

ARTICLE INFO

Keywords:

hPCNA
DNA replication proteins
Purification
Non-tag purification
Biophysical characterization
Dynamic light scattering
X-ray crystallography

ABSTRACT

Human proliferating cell nuclear antigen (hPCNA) is a DNA replication processivity factor, which acts as a docking platform, allowing proteins to have access to the replication fork and increasing the affinity of DNA interacting proteins, making it critical for cell survival. The trimer forms a ring-shaped oligomer allowing DNA to pass through the middle and interacting proteins to dock on the outside of the ring. Without this structural formation, there is a loss of DNA replication and repair in the cell. Due to the location of subunit-subunit termini, the addition of a purification tag can hamper crystallography and biophysical experiments, as the trimer complex folding can be impeded. To avoid these complications, a tag-less, step-wise purification was implemented, which resulted in 17.6 mg from 2 L culture of pure hPCNA with a 260 nm/280 nm value of 0.43. The produced crystal structure reveals a correctly formed oligomer. The clear depletion of the tracer binding and probe protein interaction in a fluorescence polarisation competition-based assay demonstrates the purification method produces a protein structure with a functional binding site. This purification method presents a reliable and simple method for producing hPCNA for biophysical characterisation.

1. Introduction

Human proliferating cell nuclear antigen (hPCNA) is a DNA processivity factor required for DNA replication and repair processes [1,2]. In an ATP-dependent manner, clamp loader proteins are required to load hPCNA onto DNA, allowing DNA polymerases and a host of other proteins to interact with the replication fork [3,4]. By providing this docking platform, hPCNA allows DNA polymerase binding affinity for DNA to increase by over 60,000-fold [5–7]. A heterozygous gene knockout of hPCNA in mice resulted in embryonic lethality, suggesting that this function of hPCNA is critical for cell survival [8–10].

hPCNA is a toroidal ring-shaped protein, with three identical subunits [11]. Each subunit has two domains, and each domain has two alpha helices, and nine beta strands. The two domains are held together by the interdomain connector loop (IDCL), which spans amino acid residues M121-Y133 [12]. The three subunits form a ring with a central cavity of approximately 30 Å; the inside of the ring is arginine- and lysine-rich, giving a positive charge and allowing double-stranded DNA

to interact [13]. The outside of the ring is made of negatively charged curved beta sheets, allowing it to interact with positively charged DNA interacting proteins [14].

hPCNA is of critical importance in cellular functions because it is the centre of many afflictions [15]. PCNA-associated DNA repair disorder (PARD) is caused by homozygous mutations in the genetic material, particularly characterised by an associated single mutation of the common protein interacting site S228I, and characterized by neurodegeneration, prelingual sensorineural hearing loss, and premature aging [16,17]. It has been discovered that hPCNA acts as a marker for malignant cells and tumours, it is overexpressed up to 6-fold in nearly all forms, and the malignancy of the tumour correlates with the level of hPCNA [10]. Down regulation of hPCNA messenger RNA and protein levels displays inhibition of cancerous cell growth [18]. This leads to the hypothesis that hPCNA is a potential drug target to treat cancerous cells [19].

Complete protein purity is key for successful *in vitro* structural studies such as X-ray crystallography, which enable the understanding

Abbreviations: (PCNA), Proliferating cell nuclear antigen.

* Corresponding author. Institute of Photonics and Advanced Sensing (IPAS), School of Biological Sciences, The University of Adelaide, Adelaide, South Australia, 5005, Australia.

E-mail address: john.bruning@adelaide.edu.au (J.B. Bruning).

<https://doi.org/10.1016/j.pep.2023.106353>

Received 21 June 2023; Received in revised form 14 August 2023; Accepted 14 August 2023

Available online 18 August 2023

1046-5928/© 2023 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

of protein function in [signalling pathways](#) and identification of novel drug targets. These investigations require mg quantities of pure protein for experimentation.

Purification tags are introduced through cloning vectors to the N- or C-terminus of a protein of interest and act as a purification tool for recombinant proteins. It facilitates selective binding of the protein to an affinity column. It is an advantage when the cell lysate contains unwanted proteins.

Purification tags may alter solubility or increase aggregation of the purified protein [20]. The tag may hinder crystallization and may influence the native structure, particularly at the terminus where the tag is attached [20]. This is believed to be the case in human PCNA crystallography. The tag is likely flexible and may cover an exposed area of the protein surface which could be an interaction or assembly site [21]. Although PCNA of other species has been crystallised with a purification tag attached, no crystal structure of human PCNA has been achieved with a purification tag attached. The trimer subunit-subunit dynamics of interactions could be affected by addition of a tag addition to the N- or C-terminus [22]. Tag introduction is additionally problematic in the introduced steps of tag cleavage for crystallization [20], which may not be completely successful and leave some uncleaved or additional amino acids behind. Without the use of a tag, additional steps of chromatography become numerous, time consuming, and costly; there is a need for an effective purification method that overcomes these challenges. We present here an optimised method for the purification and biophysical characterisation of hPCNA.

2. Materials and methods

2.1. Materials

2.1.1. General materials

Purification columns HiTrap DEAE (Diethylaminoethyl cellulose) FF 5 ml Column, HiTrap Phenyl FF (high sub) 5 ml Column, HiTrap Q FF 5 ml Column were obtained from GE Healthcare (Little Chalfont, UK). 96-well Intelli-Plate were obtained from Hampton Research (Aliso Viejo, CA, USA) and INDEX screen was obtained from Hampton Research (Aliso Viejo, CA, USA). 4–12% Bis-Tris polyacrylamide precast gels were obtained from Invitrogen (Carlsbad, CA, USA).

2.1.2. Chemical reagents

Isopropyl- β -D-thiogalactoside (IPTG) was obtained from Sigma-Aldrich (St Louis, MO, USA). Ammonium sulphate was obtained from Ajax Finechem (Taren Point, Australia).

2.1.3. Bacterial strains and growth media

For overexpression of recombinant hPCNA, *E. coli* BL21 (λ DE3) was used and grown in LB Lennox Broth composed of NaCl 5 g/L, Bacto yeast extract 5 g/L, Bacto Tryptone 10 g/L.

2.1.4. Solutions

Coomassie Blue Stain was made using 0.2% (w/v) Coomassie brilliant blue, 10% (v/v) methanol, 10% (v/v) acetic acid. Coomassie destain was made using 10% (v/v) methanol, 10% (v/v) acetic acid. SDS-PAGE Loading Buffer was made using 0.1 M Tris pH 7.0, 4% (w/v) SDS, 0.2% (w/v) bromophenol blue, 20% (w/v) glycerol.

2.1.5. Peptides

Purity >95%, purified by HPLC.

5FAM-p21-22mer [(5FAM)-GRKRRQTSMTDFYHSKRLLIFS-OH] and p21 μ [p21 μ -15mer] [KRRQTSMTDFYHSKR-OH] were synthesised by Genscript Biotech, Singapore.

2.2. Expression of recombinant hPCNA

A glycerol stock of *E. coli* BL21 (λ DE3) cells carrying an *E. coli*

optimised hPCNA-pMCSG19 plasmid were grown in a 100 mL overnight culture. Two 1 L baffled flasks containing LB with 100 μ g/mL of ampicillin were inoculated with 50 mL of the overnight culture. Cultures were incubated at 37 °C for 2 h until OD₆₀₀ = 0.7 and expression induced with a final concentration of 0.5 mM IPTG. Cultures were grown overnight at 16 °C with shaking at 200 rpm. Cultures were pelleted at 5000 \times g for 20 min. After removing the supernatant, pellets were resuspended in 20 mL 20 mM Tris pH 7.5, 20 mM NaCl, 2 mM DTT, then lysed by sonication at 70% amplification for 20 s with 40 s waiting period for 25 min. Lysate was clarified through pelleting at 45,000 \times g for 45 min and the supernatant was collected for purification.

2.3. Purification of recombinant hPCNA by chromatography series

Buffer solutions were filtered before being used. Clarified lysate containing hPCNA was first purified at 4 °C by fast protein liquid chromatography (FPLC), via anion exchange chromatography using two DEAE columns in series (HiTrap DEAE FF 5 mL column), equilibrated in Buffer A (20 mM Tris pH 7.5, 20 mM NaCl, 2 mM DTT), and hPCNA eluted using a linear gradient (0.02 M–0.7 M NaCl). Fractions containing hPCNA were pooled and ammonium sulphate added dropwise to a final concentration of 1.5 M from a stock solution of 3 M ammonium sulphate. The sample was allowed to stir gently for 1 h at 4 °C to allow DNA-protein dissociation, and then applied to hydrophobic interaction chromatography (HiTrap Phenyl FF (high sub) 5 mL column), equilibrated in Buffer B (20 mM Tris pH 7.5, 20 mM NaCl, 2 mM DTT, 0.5 mM EDTA, 1.5 M ammonium sulphate) and eluted in Buffer C (20 mM Tris pH 7.5, 2 mM DTT, 0.5 mM EDTA) with a reverse linear gradient (1.5 M–0 M ammonium sulphate). Fractions containing hPCNA were pooled and dialyzed overnight in Buffer A. hPCNA was then applied to a second anion exchange step. The Q Sepharose column (5 ml Q Sepharose FF column (GE)) was equilibrated in Buffer A and protein was eluted using a linear gradient (0.02 M–0.7 M NaCl). Fractions containing hPCNA were pooled and dialyzed overnight in 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.

2.4. Mass spectrometry

Protein mass measurements were carried out under denaturing conditions using an Agilent 6230 time-of-flight instrument coupled to an Agilent 1260 Infinity II LC System. Protein sample (5 μ L) was injected and electrosprayed using 50% aqueous acetonitrile/0.01% formic acid at a flow rate of 0.2 mL/min, without chromatographic separation. ESI-MS conditions were: positive-ion mode; capillary voltage, 3500 V; nozzle voltage 2000 V; fragmentor, 175 V; gas 13 L/min; gas temperature, 325 °C; sheath gas 11 L/min; and sheath gas temperature, 350 °C, *m/z* range, 500–3200.

2.5. Dynamic light scattering (DLS)

The DLS method calculates the [diffusion coefficient](#), [hydrodynamic radius](#), relative [molar mass](#), and [polydispersity](#) of the samples investigated. This was done using a Zetasizer nano-ZS instrument (Malvern Instruments, Malvern, UK) using 100 μ L of 1 mg/mL concentration protein samples in a low-volume (100 μ L) cuvette.

2.6. Biorad Protein (Bradford) assay

The dye reagent was prepared by diluting 1-part Dye Reagent Concentrate with 4 parts DDI water and filtered through a 0.25 μ m filter to remove particulates. A Bovine Serum Albumin (BSA) protein standard was prepared in concentrations 0.1–0.5 mg/ml and assayed in duplicate. 10 μ L of each standard and the sample solution was pipetted into separate wells of a microtiter plate and treated with 200 μ L of the

Table 1

Purification data of protein samples following chromatography steps measuring the concentration of protein by Bio-rad Protein Assay and the 260 nm/280 nm ratio indicating purity of protein from nucleic acids, standard curves can be found in supplementary material.

Step	Sample Volume	Sample Concentration	Sample Protein Yield	260/280 Value
Initial Sample - Post lyse and clarification	100 ml	8.79 mg/ml	879 mg	260/280: 1.69.
1. Post anion exchange column (DEAE)	25 ml	1.62 mg/ml	40.5 mg	260/280: 1.60.
2. Post hydrophobic interaction column (Phenyl)	20 ml	1.38 mg/ml	27.6 mg	260/280: 0.67.
3. Post anion exchange (Q Sepharose)	15 ml	0.905 mg/ml	13.5 mg	260/280: 0.59.
Final Sample	12 ml	1.47 mg/ml	17.64 mg	260/280: 0.43.

prepared diluted dye reagent. These were mixed by pipetting up and down using the multichannel pipette. The plate was incubated at room temperature for at least 5 min before measuring the absorbance at 595 nm.

2.7. Protein crystallization

Crystals were grown by sitting drop vapor diffusion method in 96-well Intelli-Plate (Hampton Research) containing 80 μ L well solution, by mixing 1 μ L hPCNA at 10 mg/ml concentration with equal volume of well solution [23–25]. Diffracting crystals of hPCNA were formed in 0.2 M Magnesium chloride hexahydrate, 0.1 M HEPES pH 7.5, 25% w/v polyethylene glycol 3,350 at 16 °C after 7 days. Crystals were mounted on cryo-loops, cryoprotected using paratone-N, and flash cooled in liquid nitrogen [23–25]. Data was collected at 100 K using the MX1 beamline at the Australian Synchrotron (Clayton, Vic).

Diffraction data were indexed and integrated using X-ray Detector Software [26]. Pointless (CCP4i) [27] was used to create an mtz file for scaling. Data was scaled using Aimless (CCP4i) [28,29] to a resolution of 2.85 Å. Phasing was solved by molecular replacement using Phaser MR [30] using a search model of human PCNA (PDB: 1AXC) [31]. Solutions were refined in phenix.refine [32] in iterative rounds with manual rebuilding in Coot [33,34].

The final structure of hPCNA is deposited on the RCSB database under accession number PDB: 8GCJ.

2.8. Fluorescence polarisation competition binding assay

The fluorescence polarisation competition binding assay was performed as described in Wegener et al., 2018. Briefly, the probe p21 μ was two-fold serially diluted starting from 108 μ M into solutions containing a constant concentration of 5FAM-p21-22mer, the tracer (50 nM), and purified hPCNA (175 nM), and interaction tested using fluorescence polarisation competition binding. This was performed using a buffer of

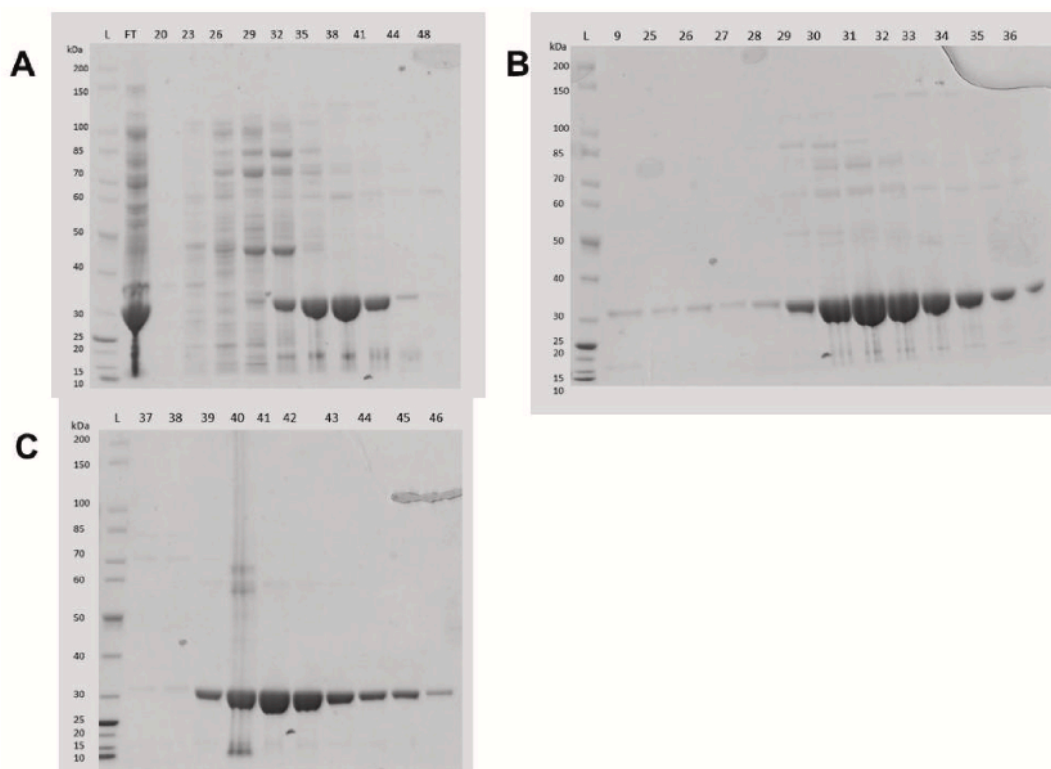


Fig. 1. SDS-PAGE of hPCNA following chromatography step-wise purification. Lane 1 labelled L contains a ladder indicator. Fractions containing hPCNA have a large band peak around 30 kDa. **A)** Anion exchange chromatography purification. Lane 2 labelled FT contains Flow Through. Fractions 35 to 48 were collected for further analysis. **B)** Hydrophobic interaction chromatography purification. Lane 2 labelled 9 contains Flow Through from the column. Fractions 30 to 36 were collected for further analysis. **C)** Q Sepharose anion exchange chromatography purification. Fractions 42–46 were collected for further analysis.

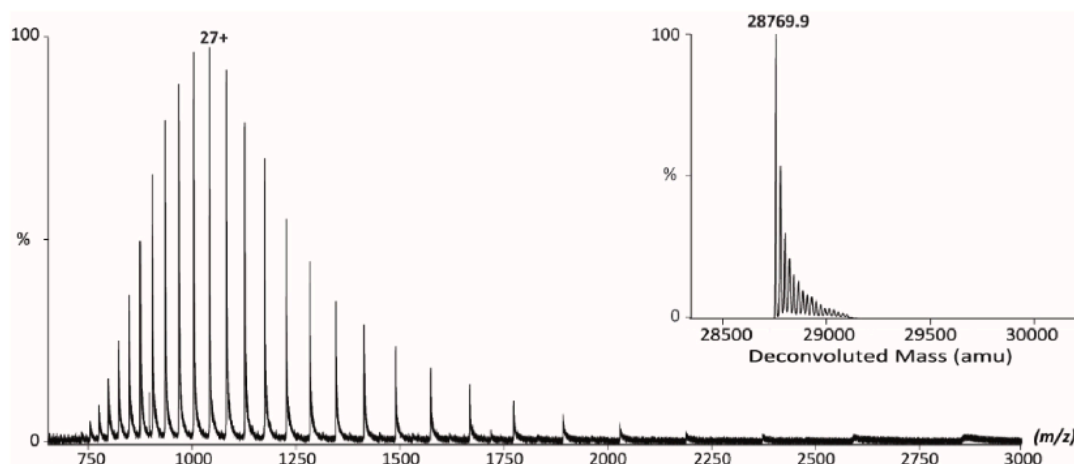


Fig. 2. Biophysical characterization of purified hPCNA protein. Mass spectrometry after purification of hPCNA protein reveals a single species with a molecular weight of 28.769 kDa representing the intact, pure monomeric protein (theoretical molar mass 28.769 kDa).

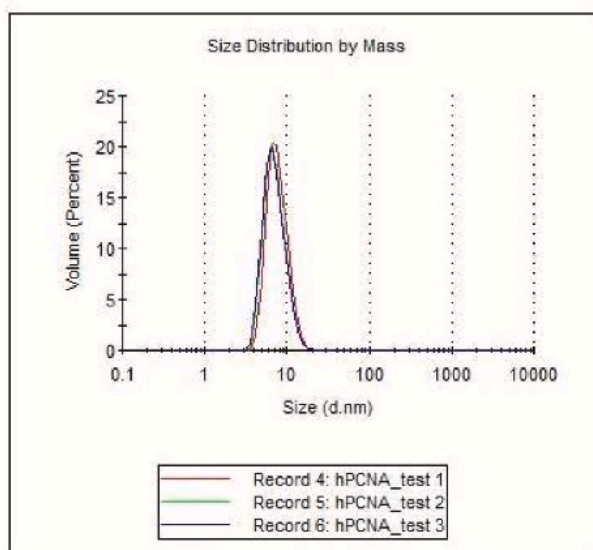


Fig. 3. Dynamic light scattering of purified recombinant hPCNA size distribution by mass. The hPCNA protein exists as higher order oligomers in the purified solution used for crystallography and functionality competition assays.

20 mM Tris pH 7.5, 100 mM NaCl, 0.01% Triton x-100, 12 mM DTT, 10% glycerol. A blank was used to remove background absorbance.

3. Results

3.1. Expression and purification of recombinant hPCNA

Crude lysate was applied to anion exchange purification through a diethylaminoethyl cellulose (DEAE) column. The resulting purification sample was measured using UV absorbance to give a 260 nm/280 nm ratio indicating the DNA to protein purity; a value of 0.6 is considered pure protein [35]. This revealed a value of 1.60 (Table 1). Ammonium sulphate was used to disrupt binding DNA to positively charged surfaces of PCNA before being applied to a hydrophobic interaction (Phenyl) column to release the hydrophobic molecules in a gradient. This removes a significant amount of DNA, indicated by the reduced 260 nm/280 nm value to 0.67 (Table 1). The SDS-PAGE gel (Fig. 1B) shows contamination among the PCNA containing fractions, leading to further purification using a stronger anion exchange column, Q Sepharose, to allow removal of final contaminants, as shown by fractions (Fig. 1C) and final 260 nm/280 nm ratio of 0.43 (Table 1).

A Biorad Protein (Bradford) Assay was used to quantify the concentration of protein at each purification step. The final purified sample of yield was calculated to 17.64 mg (Table 1).

3.2. Biophysical characterization of final pure hPCNA

Mass spectrometry and Dynamic light scattering (DLS) was utilized after purification. The mass spectrum reveals a single charge state distribution corresponding to a molar mass of 28.76 kDa representing the monomeric protein and showing exceptional protein purity (Fig. 2). DLS

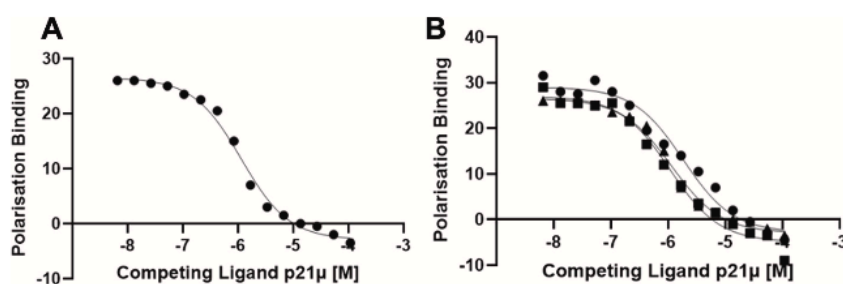


Fig. 4. Interactions of 5FAM-p21-22mer and p21 μ with hPCNA, as determined by fluorescence polarisation. A) Experiment A, $R^2 = 0.9927$. B) Experiment A shown as circles, IC50 value of 1870 nM. Experiment B shown as squares, IC50 value of 1066 nM. Experiment C shown as triangles, IC50 value of 1141 nM. Competition binding show the displacement of the 5FAMp21-22mer with competing ligand p21 μ . The data was plotted as absorbance of the tracer binding to hPCNA as a function of logarithmic ligand concentration. Experiments were conducted in triplicates.

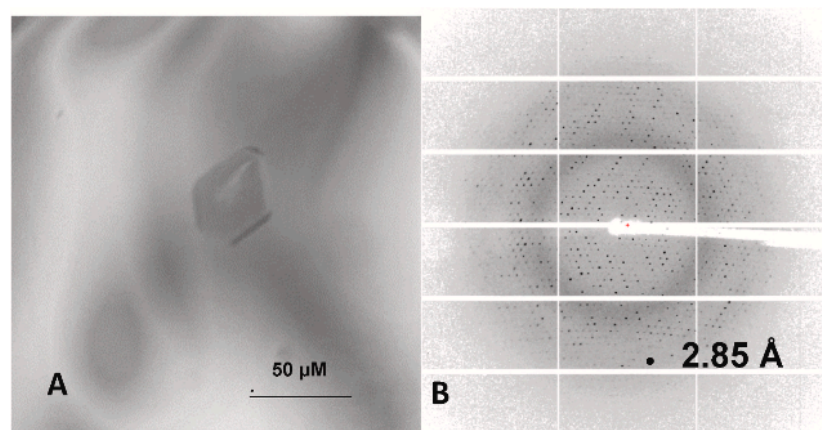


Fig. 5. A) Crystal grown after 7 days by sitting with vapor diffusion in condition 0.2 M Magnesium chloride hexahydrate, 0.1 M HEPES pH 7.5, 25% w/v polyethylene glycol 3,350. B) Diffraction image of crystal at resolution 2.85 Å.

Table 2

Data collection and refinement statistics of apo hPCNA structure (PDB: 8GCJ). Statistics for the highest-resolution shell are shown in parentheses.

Parameter	Value
Wavelength	0.9537
Resolution range	44.55 - 2.85 (2.952 - 2.85)
Space group	C 1 2 1
Unit cell	147.104 85.186 149.806 90 116.865 90
Total reflections	75727 (7582)
Unique reflections	38186 (3798)
Multiplicity	2.0 (2.0)
Completeness (%)	97.8 (98.3)
Mean 1/sigma(I)	17.5 (5.4)
Wilson B-factor	60.2
^a R-merge	0.02 (0.1)
R-meas	0.03 (0.1)
^b R-pim	0.02 (0.1)
CCL/2	0.9 (0.9)
CC*	1 (0.9)
Reflections used in refinement	37981 (3789)
Reflections used for R-free	1989 (200)
^c R-work	0.2 (0.3)
^d R-free	0.2 (0.3)
CC(work)	0.9 (0.9)
CC(free)	0.9 (0.7)
Number of non-hydrogen atoms	11135
macromolecules	11121
solvent	14
Protein residues	1514
RMS(bonds)	0.009
RMS(angles)	1.3
Ramachandran favoured (%)	91.4
Ramachandran allowed (%)	7.6
Ramachandran outliers (%)	0.8
Rotamer outliers (%)	0.3
Molprobrity Clashscore	29.9
Average B-factor	74.7
macromolecules	74.8
solvent	64.4

$$^a R_{merge} = \sum |I - \langle I \rangle| / \sum I$$

$$^b R_{pim} = \sum h [I / (I_h - 1)]^{1/2} \sum |<I_h> - I_{h,i}| / \sum h \sum_i I_{h,i} (2)$$

$$^c R_{work} = \sum |F_o - F_c| / \sum |F_o| \text{ for all data excluding data used to calculate Rfree.}$$

$$^d R_{free} = \sum |F_o - F_c| / \sum |F_o| \text{ for all data.}$$

assesses the molecular size distribution (molar mass, M_r) and oligomeric state by measuring the time-dependent fluctuations of the scattered light intensity by the hPCNA solution. The hPCNA protein exists as higher order oligomers in the purified solution used for crystallography and functionality competition assays and contained no aggregated protein or

salt aggregates (Fig. 3).

The protein functionality was tested using a fluorescence polarisation competition binding assay. The abilities of the protein-peptide interaction using untagged p21μ probe peptide to inhibit the hPCNA and tracer 5FAM-p21-22mer interaction were demonstrated (Fig. 4). This displayed a clear depletion of the tracer binding as probe protein interaction increases, determining the protein structure allows function of the binding site during the competition assay.

This suggested the purification method is effective at removing impurities that may affect subsequent experimental analysis.

3.3. Structural study of the purified protein

The protein was crystallised using the sitting-drop method using 0.2 M magnesium chloride hexahydrate, 0.1 M HEPES pH 7.5, 25% w/v polyethylene glycol 3,350 (Fig. 5). The data diffracted to a resolution of 2.85 Å (Table 2), the resolution cut off considered medium atomic resolution [36]. The phase problem was overcome by means of molecular replacement. The protein X-ray crystal structure demonstrates purified PCNA using the protocol presented here adopts the previously identified tertiary and quaternary structure (Fig. 6).

4. Discussion

The goal of this study was to generate large amounts of functionally intact hPCNA at high concentrations through a tag-less purification method suitable for structural and biophysical characterization, and to study a biophysical analysis of the product.

Current purification methods for hPCNA detail using FPLC and a range of different chromatography techniques to achieve separation without the use of a tag [5]. A method of hPCNA purification beginning with hydrophobic interaction chromatography by a 5 mL HiTrap Phenyl FF column, anion exchange purification by a HiTrap Q FF column (GE), followed by repeating the initial hydrophobic interaction chromatography, and applied to a HiPrep 26/60 Sephacryl S-300 HR column (GE) for size exclusion chromatography has been successful previously [5]. Analysis of the initial purification showed a large amount of hPCNA unbound from the column rather than separated in the elution (Fig. 1A). The column may be able to bind the protein more effectively if more resin is available. To address this, two 5 ml columns were stacked for the initial anion exchange step for the lysate to be purified to collect more protein.

The addition of salt by 1.5 M ammonium sulphate and hydrophobic interaction chromatography (HIC) is a critical step in purification of the

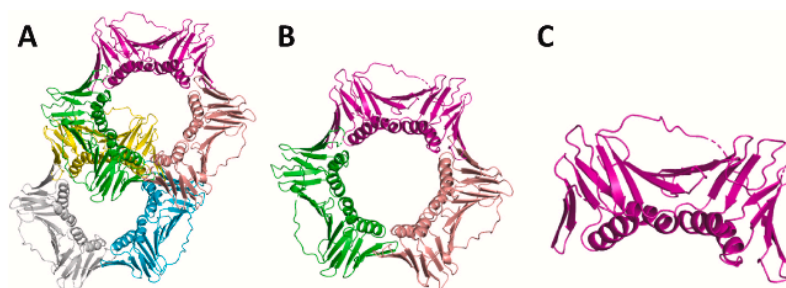


Fig. 6. Apo Human hPCNA structure (PDB: 8GCJ). **A)** Double Trimer. **B)** Single Trimer. **C)** Monomer. Made using Pymol [37].

DNA binding protein. During expression, DNA and DNA binding proteins interact and bind to the surface of hPCNA, interfering with biophysical and structural studies if not completely removed [8]. The salt of ammonium sulphate was used to disrupt the binding of DNA to positively charged surfaces, and proteins to the negatively charged surfaces and hydrophobic pockets of hPCNA. Once removed from each other, the hydrophobic interaction column can then elute them gradient-wise to separate in solution. The purity of the solution before and after this step is shown by the UV absorbance measurement 260 nm/280 nm ratio indicating the ratio of DNA to protein, which has dropped from 1.60 to 0.67 [36](Table 1). The SDS-PAGE gel post hydrophobic interaction indicates a little contamination left in hPCNA concentrated fractions (Fig. 1B), requiring a final purification step.

Although the hydrophobic interaction achieves substantial removal of contaminants, the second hydrophobic interaction chromatography step of the previous method has been replaced with instead another anion exchange step in the interest of time optimisation, as hydrophobic chromatography requires excessive steps of dropwise addition of salt and post-purification dialysis to remove excess salt. This optimised method has been used previously for crystallography [38,39].

This anion exchange step was performed using the Q Sepharose anion exchanger, which has quaternary ammonium moieties with more alkyl groups attached to the nitrogen atom than the tertiary amines of DEAE. These have lower formal charge on the nitrogen atom than the quaternary ammonium salts which have a formal charge of +1 [40]. These have narrower pH ranges to be fully ionized rather than Q Sepharose which is always cationic independently from applying pH via mobile phase flow.

Post the Q Sepharose step, the size exclusion column was eliminated due to no longer being required since final protein sample 260 nm/280 nm ratio reached 0.43, indicating pure protein [36]. The advantage of size exclusion purification is the removal of protein aggregates that can affect crystallography and biochemical experiments. This is not directly shown in the 260 nm/280 nm measurement or SDS-PAGE, however it can be indicated from dynamic light scattering (DLS) [41]. The sample was examined using DLS, the results indicated a lack of aggregation of protein material (Fig. 3), justifying the removal of this purification step.

Successful protein purification is key for successful *in vitro* structural studies such as X-ray crystallography, which enable the understanding of protein function and identification of novel drug targets. The crystallization and diffraction data shows the protein was purified to crystallography standard. Structure solution further demonstrates that the protein is in trimeric form, indicating the correct folding of the protein (Fig. 6). This is supported by the results of the competition binding assay. Untagged p21 μ probe peptide to inhibit the hPCNA and tracer 5FAM-p21-22mer interaction were demonstrated (Fig. 4), displaying a clear depletion of the tracer binding as probe protein interaction increases, determining the protein structure allows function of the binding site.

The optimised purification scheme provided has improved the purity and yield of the DNA binding protein in the absence of a tag. These

results demonstrate a more effective purification method of these sliding clamp proteins, to produce protein samples with little to no contamination from bound proteins and DNA.

Author statement

Bethiney Vandborg: Conceptualization; Methodology; Software; Validation; Formal analysis; Validation; Writing - Original Draft Writing - Review & Editing; Visualisation. Dayna L. Holroyd: Investigation; Methodology; Validation; Writing - Review & Editing. Tara Pukala: Investigation; Methodology; Validation; Writing - Review & Editing. John B. Bruning: Conceptualization; Methodology; Validation; Formal analysis; Investigation; Writing - Original Draft; Writing - Review & Editing; Visualisation; Project administration; Funding acquisition.

Funding sources

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of competing interest

The authors declare that there are no conflicts of interest.

Data availability

The crystallization data has been submitted to the PDB

Acknowledgments

B. Vandborg is a recipient of an Australian Government Research Training Program stipend scholarship.

This research was undertaken in part using the MX1 and MX2 beamlines at the Australian Synchrotron, part of ANSTO, and made use of the Australian Cancer Research Foundation detector.

Special thanks to William Walters, Paul Jr Walters and Polly Walters for their crystallography consultation.

Special thanks to Jakeb Vandborg for their scientific consultation.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.pep.2023.106353>.

References

- [1] K.L. Wegener, Amy E. McGrath, Nicholas E. Dixon, Aaron J. Oakley, Denis B. Scanlon, Andrew D. Abell, John B. Bruning, Rational design of a 3 10 -helical PIP-box mimetic targeting hPCNA, the human sliding clamp, *Eur. J. Chem.* 24 (2018) 11325–11331.
- [2] G.D. Bowman, M. O'Donnell, J. Kuriyan, Structural analysis of a eukaryotic sliding DNA clamp-clamp loader complex, *Nature* 429 (2004) 724–730.

- [3] I. Bruck, M. O'Donnell, The ring-type polymerase sliding clamp family, *Genome Biol.* 2 (2001), 3001.3001–3001.3003.
- [4] E. Warbrick, PCNA binding through a conserved motif, *Bioessays* 20 (1998) 195–199.
- [5] A.J. Kroker, John B. Bruning, p21 exploits residue Tyr151 as a tether for high-affinity PCNA binding, *Biochemistry* 54 (2015) 3483–3493.
- [6] A. Horsfall, A. Abell, J. Bruning, Targeting PCNA with peptide mimetics for therapeutic purposes, *Chembiochem* 21 (2019) 442–450.
- [7] G. Prelich, C.K. Tan, M. Kostura, M.B. Mathews, A.G. So, K.M. Downey, B. Stillman, Functional identity of proliferating cell nuclear antigen and a DNA polymerase-delta auxiliary protein, *Nature* 326 (1987) 517–520.
- [8] T. Tsurimoto, PCNA binding proteins, *Front. Biosci.* 4 (1999) 849–858.
- [9] S. Roa, E. Avdievich, J.U. Peled, T. MacCarthy, U. Werling, F.L. Kuang, C. Zhao, A. Berman, P.E. Cohen, W. Edelman, M.D. Schnarf, Ubiquitylated PCNA plays a role in somatic hypermutation and class-switch recombination and is required for meiotic progression, *Proc. Natl. Acad. Sci. U.S.A.* 105 (2008) 16248–16253.
- [10] C. Sakakura, A. H. H. Tsujimoto, K. Ozaki, T. Sakakibara, T. Oyama, M. Ogaki, T. Takahashi, The anti-proliferative effect of proliferating cell nuclear antigen-specific antisense oligonucleotides on human gastric cancer cell lines, *Surg. Today* 25 (1995) 184–186.
- [11] Z. Tan, M. Wortman, K.L. Dillehay, W.L. Seibel, C.R. Evelyn, S.J. Smith, L. H. Malkas, Y. Zheng, S. Lu, Z. Dong, Small-molecule targeting of proliferating cell nuclear antigen chromatin association inhibits tumor cell growth, *Mol. Pharmacol.* 81 (2012) 811–819.
- [12] M. De March, N. Merino, S. Barrera-Vilarmau, R. Crehuet, S. Onesti, F.J. Blanco, A. De Blasio, Structural basis of human PCNA sliding on DNA, *Nat. Commun.* 8 (2017).
- [13] J. Majka, P.M.J. Burgers, The PCNA–RFC families of DNA clamps and clamp loaders, *Prog. Nucleic Acid Res. Mol. Biol.* 78 (2004) 227–260.
- [14] A. Gonzalez-Magana, F.J. Blanco, Human PCNA structure, function and interactions, *Biomolecules* 10 (2020) 570.
- [15] C. Duffy, B.J. Hilbert, B.A. Kelch, A disease-causing variant in PCNA disrupts a promiscuous protein binding site, *J. Mol. Biol.* 428 (2016) 1023–1040.
- [16] S. Naryzhny, H. Lee, Characterization of proliferating cell nuclear antigen (PCNA) isoforms in normal and cancer cells: there is no cancer-associated form of PCNA, *FEBS (Fed. Eur. Biochem. Soc.) Lett.* 528 (2007) 4917–4920.
- [17] R.H.C. Wilson, Antonio J. Biasutto, Lihao Wang, Roman Fischer, Emma L. Baple, Andrew H. Crosby, Erika J. Mancini, Catherine M. Green, PCNA dependent cellular activities tolerate dramatic perturbations in PCNA client interactions, *DNA Repair* 50 (2017) 22–35.
- [18] A. Prestel, N. Wichmann, J.M. Martins, R. Marabini, N. Kassem, S.S. Broendum, M. Otterlei, O. Nielsen, M. Willemoes, M. Ploug, W. Boomsma, B.B. Kragelund, The PCNA interaction motifs revisited: thinking outside the PIP-box, *Cell. Mol. Life Sci.* 76 (2019) 4923–4943.
- [19] J. Taylor, *The Synthesis and Study of Side-chain Lactam-bridged Peptides*, vol. 66, American Peptide Society, 2004, pp. 49–75.
- [20] M. Carson, D.H. Johnson, H. McDonald, C. Brouillette, L.J. Delucas, His-tag impact on structure, *Acta Crystallogr. D Biol. Cryst.* 63 (2007) 295–301.
- [21] I. Stoimenov, T. Helleday, PCNA on the crossroad of cancer, *Biochem. Soc. Trans.* 37 (2009) 605–613.
- [22] A. McPherson, J. Gavira, Introduction to protein crystallisation, *Acta Crystallogr., Sect. F: Struct. Biol. Commun.* 70 (2013) 2–20.
- [23] J. Bruning, Y. Shamoo, Structural and thermodynamic analysis of human PCNA with peptides derived from DNA polymerase-delta p66 subunit and flap endonuclease-1, *Structure* 12 (2004) 2209–2219.
- [24] A.C. Marshall, A.J. Kroker, L.A. Murray, K. Gronthos, H. Rajapaksha, K.L. Wegener, J.B. Bruning, Structure of the sliding clamp from the fungal pathogen *Aspergillus fumigatus* (AfumPCNA) and interactions with human p21, *FEBS J.* 284 (2017) 985–1002.
- [25] R.L. Frkic, B.S. Chua, Y. Shin, B.D. Pascal, S.J. Novick, T.M. Kamenecka, P. R. Griffin, J.B. Bruning, Structural and dynamic elucidation of a non-acid PPARγ partial agonist, *Nucl. Recept. Res.* 5 (2018).
- [26] W. Kabsch, XDS (X-ray detector software), *Acta Crystallogr. D Biol. Crystallogr.* 66 (2010) 125–132.
- [27] P. Evans, Scaling and assessment of data quality, *Acta Crystallogr. D Biol. Crystallogr.* 62 (2006) 72–82.
- [28] 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, Overview of the CCP4 suite and current developments, *Acta Crystallogr. D Biol. Crystallogr.* 67 (2011) 235–242.
- [29] E. Potterton, P.J. Briggs, M.G.W. Turkenburg, E. Dodson, A graphical user interface to the CCP4 program suite, *Acta Crystallogr. D Biol. Crystallogr.* 59 (2003) 1131–1137.
- [30] A.J. McCoy, R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni, R. J. Read, Phaser crystallographic software, *J. Appl. Crystallogr.* 40 (2007) 658–674.
- [31] J.M. Gulbis, Z. Kelman, J. Hurwitz, M. O'Donnell, J. Kuriyan, Structure of the C terminal region of p21(WAF1/CIP1) complexed with human PCNA, *Cell* 87 (1996) 297–306.
- [32] 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, Phenix: macromolecular structure determination using X-rays, neutrons and electrons: recent developments in phenix, *Acta Crystallogr. D Biol. Crystallogr.* 75 (2019) 861–877.
- [33] 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, P.D. Adams, Towards automated crystallographic structure refinement with phenix.Refine, *Acta Crystallogr. D Biol. Crystallogr.* 68 (2012) 352–367.
- [34] P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics, *Acta Crystallogr. D Biol. Crystallogr.* 60 (2004) 2126–2132.
- [35] V. Dubach, A. Guskov, The resolution in X-ray crystallography and single-particle cryogenic electron microscopy, *Crystals* 10 (2020) 580.
- [36] J. Glasel, Validity of nucleic acid purities monitored by 260nm/280nm absorbance ratios, *Biotechniques* 18 (1995) 62–63.
- [37] L. Schrödinger, The PyMOL Molecular Graphics System, 2015. Version 1.2r3pre.
- [38] A.J. Horsfall, B.A. Vandborg, W. Kowalczyk, T. Chav, D.B. Scanlon, A.D. Abell, J. B. Bruning, Unlocking the PIP-box: a peptide library reveals interactions that drive high-affinity binding to human PCNA, *J. Biol. Chem.* 296 (2021), 100773.
- [39] A.J. Horsfall, B.A. Vandborg, Z. Kikhtyak, D.B. Scanlon, W.D. Tilley, T.E. Hickey, J. B. Bruning, A.D. Abell, A cell permeable bimane-constrained PCNA-interacting peptide, *RSC Chem. Biol.* 2 (2021) 1499–1508.
- [40] M. Andersson, I. Drevin, B. Johansson, Characterization of the chemical and functional stability of DEAE Sepharose, *Process Biochem.* 28 (1993) 223–230.
- [41] B. Lorber, F. Fischer, M. Bailly, H. Roy, D. Kern, Protein analysis by dynamic light scattering: methods and techniques for students, *Biochem. Mol. Biol. Educ.* 40 (2012) 372–382.

Supplementary material

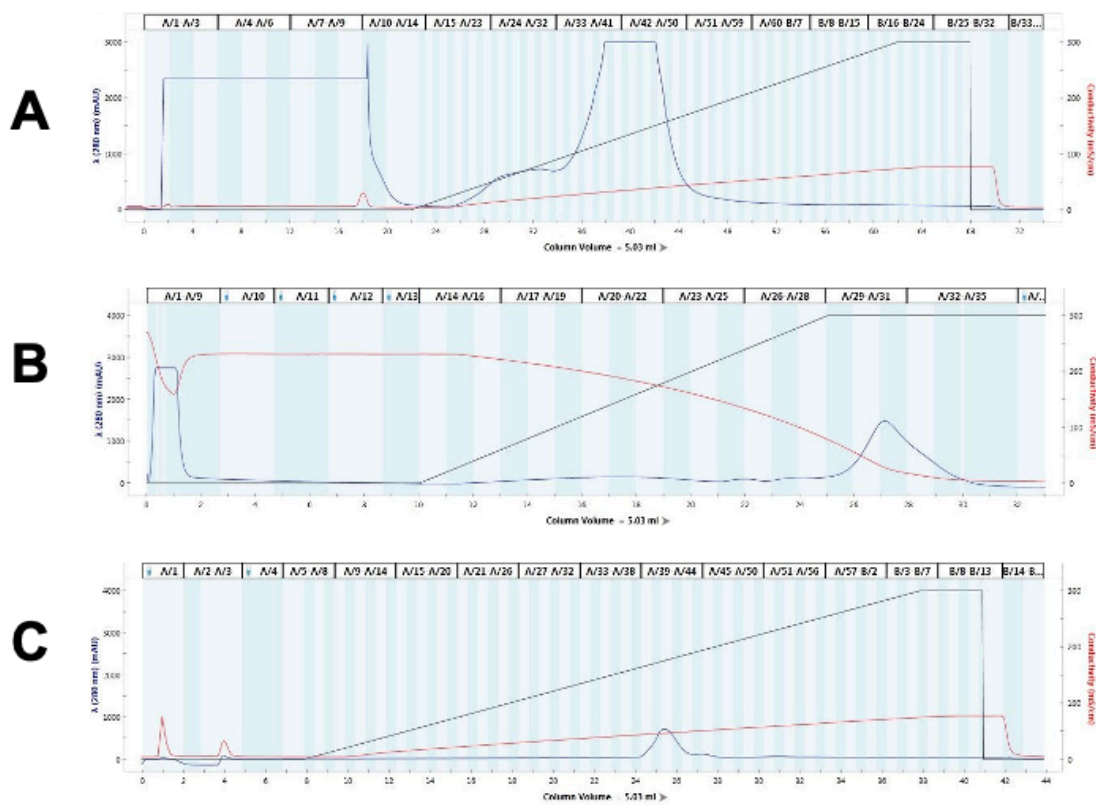


Figure S1. FPLC chromatogram of hPCNA purification steps. **A**) Anion exchange (DEAE column). **B**) Hydrophobic exchange (Phenyl column). **C**) Anion exchange (Q Sepharose column). Blue trace indicates absorbance at 280nm. Red trace indicates conductivity. Black trace indicates percentage of elution buffer.

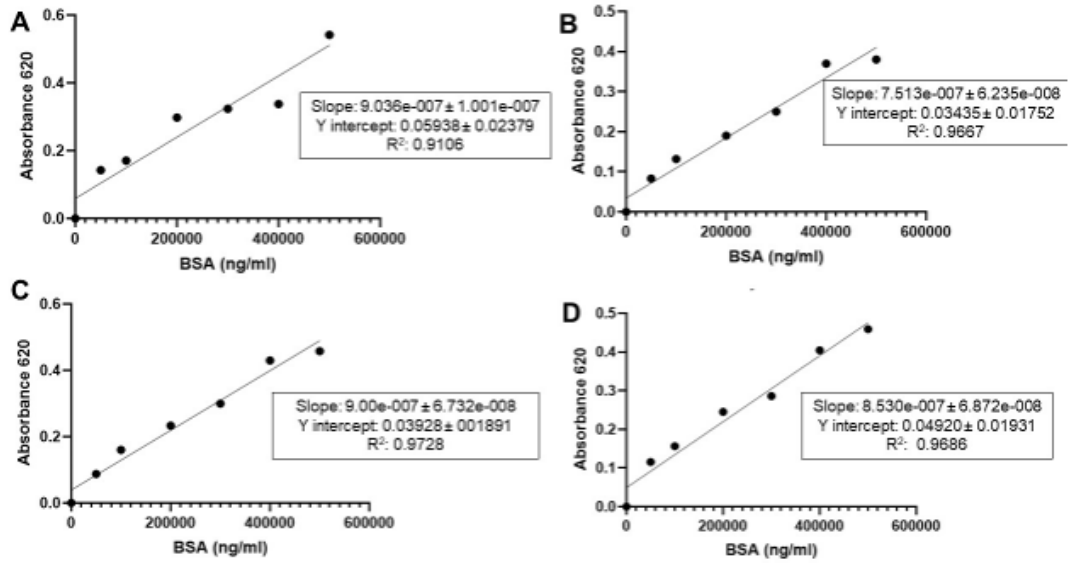


Figure S2. Biorad protein assay standards for protein concentration quantification. **A)** Post Lyse and Clarification Standard. **B)** Post DEAE and Phenyl Column Standard. **C)** Post Q Sepharose Column Standard. **D)** Post Dialysis and Final Standard.

Table S1. Protein parameter biochemical data of human PCNA provided by Expsy server Protein Identification and Analysis Tools[1].

Amino Acid Sequence:

MFEARLVQGSILKKVLEALKDLNEACWDISSSGVNLQSMDSHVSLVQLTLRSEGFDTYRCDRNLAM
 GVNLTSMKILKCAGNEDIITLRAEDNADTLALVFEAPNQEKVSDYEMKMLMDLDVEQLGIPEQEYSCV
 VKMPSGEFARICRDLSHIGDAVVISCAKDGVKFSASGELGNGNIKLSQTSNVDKKEEAVTIEMNEPVQ
 LTFALRYLNFFTKATPLSSTVTLSMSADVPLVVEYKIADMGHLYLAPKIEDEEGS

Number of amino acids: 261

Molecular weight: 28768.78

Theoretical pI: 4.57

Amino acid composition

Ala (A)	19	7.3%
Arg (R)	8	3.1%
Asn (N)	12	4.6%
Asp (D)	18	6.9%
Cys (C)	6	2.3%
Gln (Q)	8	3.1%
Glu (E)	23	8.8%
Gly (G)	14	5.4%

His (H)	3	1.1%
Ile (I)	14	5.4%
Leu (L)	29	11.1%
Lys (K)	16	6.1%
Met (M)	10	3.8%
Phe (F)	8	3.1%
Pro (P)	7	2.7%
Ser (S)	25	9.6%
Thr (T)	12	4.6%
Trp (W)	1	0.4%
Tyr (Y)	7	2.7%
Val (V)	21	8.0%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%
(B)	0	0.0%
(Z)	0	0.0%
(X)	0	0.0%

Total number of negatively charged residues (Asp + Glu): 41

Total number of positively charged residues (Arg + Lys): 24

Atomic composition:

Carbon	C	1257
Hydrogen	H	2020
Nitrogen	N	328
Oxygen	O	408
Sulfur	S	16

Formula: C₁₂₅₇H₂₀₂₀N₃₂₈O₄₀₈S₁₆

Total number of atoms: 4029

Extinction coefficients:

Extinction coefficients are in units of M⁻¹ cm⁻¹, at 280 nm measured in water.

Ext. coefficient 16305

Abs 0.1% (=1 g/l) 0.567, assuming all pairs of Cys residues form cystines

Ext. coefficient 15930

Abs 0.1% (=1 g/l) 0.554, assuming all Cys residues are reduced

Estimated half-life:

The N-terminal of the sequence considered is M (Met).

The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).

>20 hours (yeast, in vivo).

>10 hours (Escherichia coli, in vivo).

Instability index:

The instability index (II) is computed to be 45.15

This classifies the protein as unstable.

Aliphatic index: 94.87

Grand average of hydropathicity (GRAVY): -0.095

References

- [1] Gasteiger E., H. C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A. Protein Identification and Analysis Tools on the ExPASy Server, *(In) John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press (2005). 571-607.*

Chapter 3: Publication

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

Statement of Authorship

Title of Paper	Unlocking the PIP-box: Understanding factors for high affinity binding to human PCNA.
Publication Status	<input checked="" type="checkbox"/> Published <input type="checkbox"/> Accepted for Publication <input type="checkbox"/> Submitted for Publication <input type="checkbox"/> 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)	Bethiney Vandborg
Contribution to the Paper	Protein synthesis and purification, protein crystallography, structure analysis, discussed results and edited manuscript.
Overall percentage (%)	40%
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by subject to any obligations or contractual agreements with a inclusion in this thesis. I am the primary author of this paper.
Signature	Date 29/08/2023

Co-Author Conti

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

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

Name of Co-Author	Aimee J Horsfall
Contribution to the Paper	Designed and synthesised all peptides, SPR assays, analysed structure results, discussed results, wrote and edited manuscript.
Signature	Date 29/08/2023

Name of Co-Author	Wioleta Kowalczyk
Contribution to the Paper	Peptide synthesis, discussed results and edited manuscript
Signature	Date 18/09/2023

Please cut and paste additional co-author panels here as required.

Name of Co-Author	Theresa Chav		
Contribution to the Paper	SPR assays, discussed results and edited manuscript		
Signature		Date	18/9/23

Name of Co-Author	Denis B Scanlon		
Contribution to the Paper	Peptide synthesis, discussed results and edited manuscript		
Signature	Supervisor JBB	Date	18/9/23

Name of Co-Author	Andrew D. Abell		
Contribution to the Paper	Supervised AJH, discussed results and edited manuscript.		
Signature		Date	18/9/23

Name of Co-Author	John B Bruning		
Contribution to the Paper	discussed results and edited manuscript.		
Signature		Date	18/9/23



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

Received for publication, February 10, 2021, and in revised form, May 2, 2021. Published, Papers in Press, May 11, 2021, <https://doi.org/10.1016/j.jbc.2021.100773>

Aimee J. Horsfall¹, Beth A. Vandborg², Wioleta Kowalczyk³, Theresa Chav¹, Denis B. Scanlon¹, Andrew D. Abell^{1,*}, and John B. Bruning^{2,*}

From the ¹ARC Centre of Excellence for Nanoscale BioPhotonics, Institute of Photonics and Advanced Sensing, School of Physical Sciences, ²Institute of Photonics and Advanced Sensing, School of Biological Sciences, The University of Adelaide, Adelaide, South Australia, Australia; ³CSIRO Manufacturing, Clayton South, Victoria, Australia

Edited by Patrick Sung

The human sliding clamp, Proliferating Cell Nuclear Antigen (hPCNA), interacts with over 200 proteins through a conserved binding motif, the PIP-box, to orchestrate DNA replication and repair. It is not clear how changes to the features of a PIP-box modulate protein binding and thus how they fine-tune downstream processes. Here, we present a systematic study of each position within the PIP-box to reveal how hPCNA-interacting peptides bind with drastically varied affinities. We synthesized a series of 27 peptides derived from the native protein p21 with small PIP-box modifications and another series of 19 peptides containing PIP-box binding motifs from other proteins. The hPCNA-binding affinity of all peptides, characterized as K_D values determined by surface plasmon resonance, spanned a 4000-fold range, from 1.83 nM to 7.59 μ M. The hPCNA-bound peptide structures determined by X-ray crystallography and modeled computationally revealed intermolecular and intramolecular interaction networks that correlate with high hPCNA affinity. These data informed rational design of three new PIP-box sequences, testing of which revealed the highest affinity hPCNA-binding partner to date, with a K_D value of 1.12 nM, from a peptide with PIP-box QTRITEYF. This work showcases the sequence-specific nuances within the PIP-box that are responsible for high-affinity hPCNA binding, which underpins our understanding of how nature tunes hPCNA affinity to regulate DNA replication and repair processes. In addition, these insights will be useful to future design of hPCNA inhibitors.

Human Proliferating Cell Nuclear Antigen (hPCNA) 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 to cope with the increased demand for DNA replication. A hPCNA KO is lethal (6, 7), which reflects its importance in cell cycle progression. hPCNA is a toroidal-shaped homotrimer with 6-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 (Fig. 1, A 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 hPCNA 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 hPCNA, which span four orders of magnitude (3, 10–12). Such proteins compete to bind hPCNA to gain access to the replication fork, where successful binding is ultimately dictated by hPCNA affinity. The cell-cycle regulator protein, p21^{CIP1/WAF1} (referred to herein as p21), is the highest affinity hPCNA-interacting protein (PIP) known (K_D 2.5–90 nM, (13–15)), and upon binding hPCNA, shuts down replication. This consequently stalls cell-cycle progression to provide a necessary checkpoint for healthy proliferation. p21 requires high affinity for hPCNA to fulfil this role and outcompete other proteins from interacting with hPCNA and hence the replication fork. In contrast, the Y-family translesion polymerases pol λ , pol κ , and pol ι bind hPCNA with micromolar affinity (16), and a peptide derived from the major processive polymerase (pol $\delta_{p66\ 452-466}$) has an affinity for hPCNA of 15.6 μ M (14). A fundamental understanding of the molecular-level nuances that dictate hPCNA binding affinity is required to understand the regulation of DNA-replication and DNA-repair processes. Here, we begin to unpack subtle sequence changes that influence hPCNA-binding affinity with an investigation of the structure–activity relationship of peptides that bind hPCNA.

The majority of hPCNA partners bind on the surface of a hPCNA subunit located between the two subunit domains, nestled under the interdomain connecting loop (Fig. 1C). PIPs (or hPCNA-interacting peptides) interact with hPCNA through a consensus sequence aptly named the PIP-box motif. The PIP-

* For correspondence: John B. Bruning, john.bruning@adelaide.edu.au; Andrew D. Abell, andrew.abell@adelaide.edu.au.



Unlocking the PIP-box

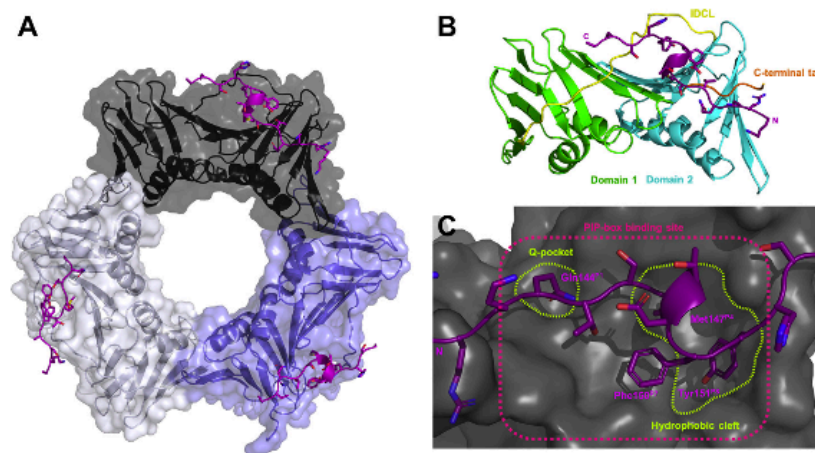


Figure 1. Cocrystal structure of p21_μ (7KQ1) bound to hPCNA solved by X-ray crystallography. p21_μ peptide (purple) shown in cartoon form with side chains as sticks bound to each hPCNA subunit. A, hPCNA shown in cartoon form with transparent surface representation with each subunit indicated in white, gray, and pale purple. B, a hPCNA 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 hPCNA in orange. C, a zoom-in of a p21_μ peptide (purple) bound to hPCNA (gray, 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 labeled in purple. IDCL, interdomain connecting loop; p21, p21^{CIP1/WAF1}; p21_μ, p21 sequence 141 to 155.

box motif is defined as Qxxφxxψ, 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 positions 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 hPCNA-binding motifs as many noncanonical PIP-box peptides, which are defined by the absence of a glutamine at position 1, bind hPCNA with low micromolar or nanomolar affinity (11, 17). The noncanonical definition will be used here for simplicity of discussion. The highly canonical PIP-box motif of p21, ¹⁴⁴QTSMTDEY, was initially defined following an alanine mutation scan from Lys141 to Ser160 and revealed Gln144^{P1}, Met147^{P4}, Phe150^{P7}, and Tyr151^{P8} as the conserved binding residues (positions 1, 4, 7, and 8 (18, 19)). A 22-mer peptide, derived from p21 (residues 139–160), that contains this PIP-box sequence binds to hPCNA as a single well-defined 3₁₀-helical turn that is anchored to the hPCNA surface by a hydrophobic triad (Met, Phe, and 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 hPCNA (22). The nonconserved PIP-box residues (positions 2, 3, 5, and 6) of p21_{139–160} only participate in three 3₁₀-helical stabilizing 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 hPCNA. 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, to begin to understand how affinity for hPCNA plays a role in the regulation of DNA-replication and DNA-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 hPCNA-binding peptides and determine the secondary interactions that fine-tune affinity for hPCNA. 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 hPCNA-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 than point modifications and allows direct comparison of the effect of the PIP-box sequence on hPCNA affinity, as the peptides contain the same sequence flanking the PIP-box. The hPCNA-binding affinity of each peptide was determined by surface plasmon resonance (SPR). The changes in the structure of the peptides bound to hPCNA were studied by X-ray crystallography and computational modeling studies, and correlated to the hPCNA-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 hPCNA affinity modified p21 peptides. This comprehensive and systematic structure–activity investigation advances the ability to predict hPCNA-binding affinity from PIP-box sequence. In addition, it provides insight into how nature has fine-tuned affinity of PIP-box sequences for hPCNA, which can be leveraged to further investigate how hPCNA regulates the DNA-replication and

DNA-repair process or develop inhibitors of these interactions.

Results

Fifty-one peptides were synthesized by solid-phase peptide synthesis (see [Experimental procedures](#)) to allow systematic study of the structure–activity relationship of peptides that bind hPCNA ([Table S1](#)). These peptides comprise two control p21 peptides known to bind hPCNA; a series of 27 point-modified p21 peptides; a second series of 19 peptides, where the PIP-box sequence from a number of hPCNA-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 to 155 (referred to herein as p21_μ) was chosen as it has been reported the p21_{139–160} sequence can be shortened without drastically impacting affinity (13). Interestingly, the preparation of this shorter sequence gave less aspartimide formation compared with the longer 22mer, with a M-18 peak in the mass spectrum not apparent for p21_μ, in contrast to p21_{139–160} (unpublished work), thereby improving synthetic yields. All subsequent peptides were based on this shorter p21_μ peptide (residues 141–155).

Five p21_μ peptides with a 144^{P1} point mutation were prepared: p21_μ-Q144K, p21_μ-Q144M, p21_μ-Q144D, p21_μ-Q144S, and p21_μ-Q144N. Four of these were inspired by the following noncanonical 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_μ-Q144N). Five p21_μ peptides, with a single amino-acid modification at the conserved position 4 of the PIP-box, were prepared and include valine (p21_μ-M147V), isoleucine (p21_μ-M147I), or leucine (p21_μ-M147L), which are commonly observed in native PIP-box sequences, as well as the smaller but still hydrophobic alanine (p21_μ-M147A) and the nonpolar aromatic tryptophan (p21_μ-M147W). Aromatic residues, such as phenylalanine and tyrosine, are commonly observed at conserved positions 7 and 8 of the PIP-box. p21_μ contains the Phe^{P7}-Tyr^{P8} combination and consequently three peptides, each with a different permutation of phenylalanine and tyrosine, were prepared (p21_μ-Y151F, p21_μ-F150Y, and p21_μ-FY150/151YF). In addition, a p21_μ-F150H peptide was prepared, inspired by the PARG PIP-box. Positively charged residues are commonly observed in native PIP-box sequences at nonconserved positions 2 and 3 (e.g., XPG, FEN1, WRN, PARG, p15, Cdt1). Consequently, eight p21_μ peptides with arginine or lysine residues at PIP-box position 2 and/or 3 were prepared (p21_μ-T145K, p21_μ-T145R, p21_μ-S146K, p21_μ-S146R, p21_μ-TS145/146KK, p21_μ-TS145/146KR, p21_μ-TS145/146RK, and p21_μ-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_μ-T148D, p21_μ-T148E, p21_μ-TD148/149DE, p21_μ-TD148/149EE, and p21_μ-D149E). Finally, to investigate a variety of PIP-box sequence combinations, a variety of native PIP-box sequences replaced the p21 PIP-box in a p21_μ peptide. Ten canonical PIP-box sequences from hPCNA-binding proteins Cdt1, Pogo, XPG, DNALig1, MCMT, pol δ_{p66}, FEN1, p15, WRN, and RecQ5 were chosen, along with eight peptides with noncanonical 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 hPCNA affinity and is not influenced by the native flanking sequence.

The binding affinity (K_D) of each peptide for hPCNA was determined by SPR, and the results are summarized in [Figure 2](#) and [Table S2](#). A representative sample of SPR sensorgrams are included in [Fig. S1](#). p21_{139–160}, prepared as a positive control, bound to hPCNA with a K_D value of 4.32 nM ([Fig. 2A](#), [Table S2](#)), which agrees with previous literature (13–15). p21_μ (residues 141–155) bound to hPCNA with 12.3 nM affinity, which indicates the truncated sequence was well tolerated, although the peptide is seven amino acids shorter than p21_{139–160} ([Fig. 2A](#), [Table S2](#)). All modified p21 peptides are compared with p21_μ throughout the discussion, unless otherwise indicated. The K_D value of p21_μ 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_μ peptides spanned four orders of magnitude, with the best binding peptide of these two series having a K_D value of 1.83 nM (p21_μ-TS145/146RR) and the lowest affinity was 8.14 μM (p21_μ-pol λ) ([Fig. 2, B and C](#), [Table S2](#)). The binding affinity of p21_μ-Cdt2 and p21_μ-RecQ5 could not be determined because of nonspecific binding to the sensor chip. Interestingly, the analysis revealed seven peptides with higher affinity than p21_μ: p21_μ-TS145/146RR (1.83 nM), p21_μ-S146R (4.30 nM), p21_μ-Cdt1 (8.76 nM), p21_μ-Pogo (8.82 nM), p21_μ-Y151F (10.6 nM), and p21_μ-M147I (11.1 nM) ([Fig. 2, B and C](#)). The affinity of canonical PIP-box p21_μ peptides (with Gln144^{P1}) here range from 1.83 nM to 3.57 μM ([Fig. 2C](#), ‘canonical’). In general, the noncanonical peptides as a group, unsurprisingly, bound hPCNA with lower affinity than the canonical p21_μ peptides (see [Fig. 2C](#)).

The conformation of the peptides bound to hPCNA was investigated to uncover potential reasons for differences in affinity. Cocrystal structures of p21_μ (PDB ID: 7KQ1, 3.30 Å) and p21_μ-F150Y bound to hPCNA (PDB ID: 7KQ0, 2.40 Å) were solved. Both served as important controls for our modeling studies: p21_μ showed that the shorter peptide binds hPCNA in a similar manner (position on hPCNA surface and secondary structure) as the 22mer p21_{139–160} peptide (PDB ID: 1AXC), and p21_μ-F150Y demonstrated a proof of concept that our docking studies represented the crystal structures accurately. The resulting structures are shown in [Figure 3](#), [Figs. S1 and S2](#), and the data collection and refinement statistics are summarized in [Table S3](#). The cocrystal structure of p21_μ

Unlocking the PIP-box

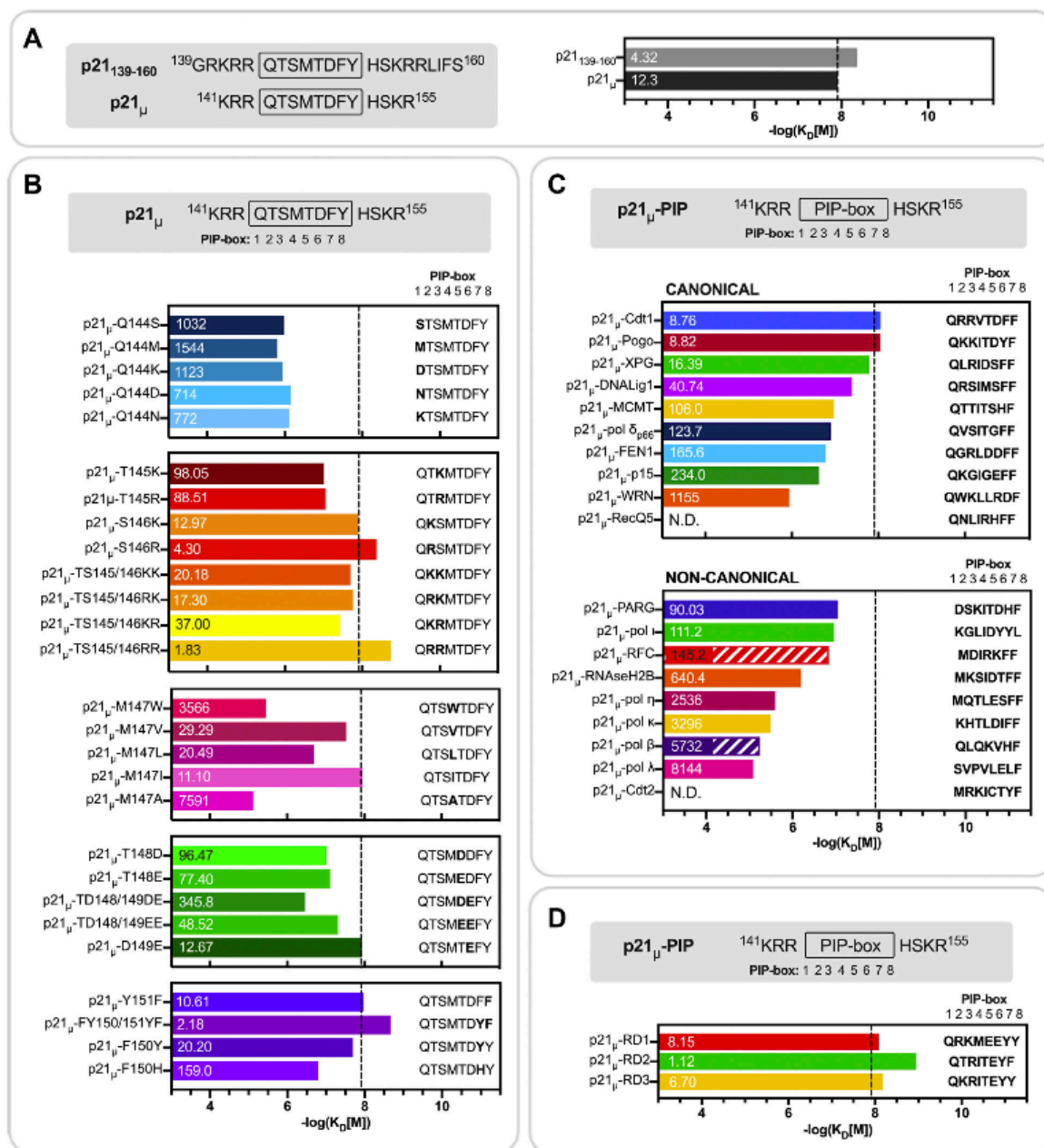


Figure 2. Binding affinity (K_D) of p21 peptides to PCNA determined by SPR. K_D values were calculated using the in-built Biacore Evaluation S200 software and are shown in nanomolar over the respective bar. The bars graphically represent the K_D values (in molar, M) as $-\log(K_D[M])$. All experiments were repeated to ensure reproducibility, and additional SPR information including fitting errors can be found in Table S2. The dashed line on all panels represents the affinity of p21_μ (12.3 nM, $-\log(K_D[M]) = 7.91$) to which all modified p21_μ 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 hPCNA-binding partner (top) or native noncanonical PIP-box from an alternate hPCNA-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. N.D., could not be determined; p21, p21^{CIP1/WAF1}; SPR, surface plasmon resonance.

bound to hPCNA (Fig. 1) revealed six intermolecular and three intramolecular hydrogen bonds (Fig. 3C, Fig. S2 and Table S4). The PIP-box residues 144 to 151 all retain similar conformations to p21₁₃₉₋₁₆₀ (1AXC, Fig. 3, A 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 of the PIP-box residues of p21_μ compared with p21₁₃₉₋₁₆₀ in 1AXC (Table S19). Gln144^{P1} of p21_μ makes a hydrogen bond with Ala252 and Pro253 within the Q-pocket, but not Ala208, in contrast to Gln144^{P1} of p21₁₃₉₋₁₆₀ that formed three hydrogen

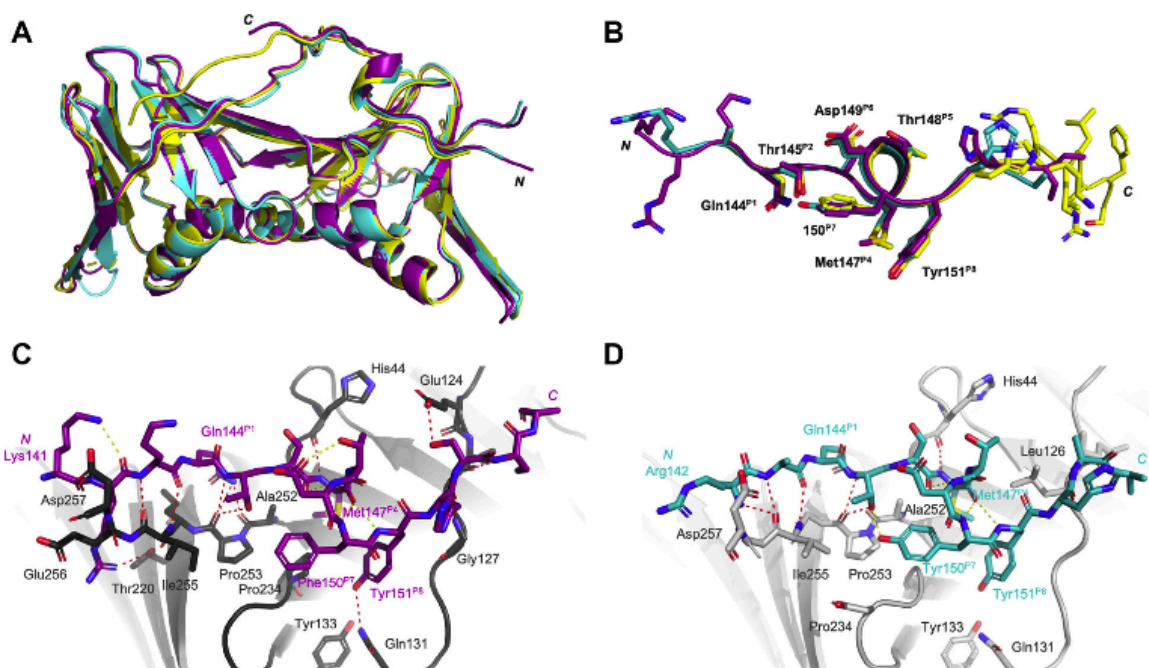


Figure 3. Cocystal structures of p21_μ (7KQ1, purple) and p21_μ-F150Y (7KQ0, blue) bound to hPCNA solved by X-ray crystallography, compared with p21_{139–160} (1AXC, yellow). *A*, overall binding mode and structure of p21 peptides bound to hPCNA is the same. Peptide and hPCNA shown are cartoons. *B*, an overlay of peptide structures shows the PIP-box residues (labeled) adopt similar orientations. *C* and *D*, intramolecular (yellow dashed lines) and intermolecular (red dashed lines) polar interactions of p21_μ (purple, *C*) and p21_μ-F150Y (blue, *D*) shown as sticks bound to hPCNA in gray. Conserved PIP-box residues, and the N-terminal residue, are labeled in the corresponding peptide color. p21, p21^{CIP1/WAF1}; p21_μ, p21 sequence 141 to 155.

bonds. The side chain of Met147^{P4} was packed slightly differently to the analogous Met147^{P4} in p21_{139–160} (1AXC). The C-terminal residues of hPCNA (Lys254–Asp257) and N-terminal residues of p21_μ (141–143) gave well-defined electron density, in contrast to the p21_{139–160} structure (1AXC) that revealed an electrostatic interaction between Lys141 and Arg142 of p21_μ, and Glu256 and Asp257 of hPCNA (Fig. 3C).

After this, a larger collection of the modified p21_μ peptides were selected for further investigation by computational modeling of the hPCNA-bound peptides, to allow high-throughput analysis of more structures. The crystal structure of p21_μ (PDB ID: 7KQ1) bound to hPCNA was used as a starting point for modeling as this peptide is the most similar to the modeled peptides, and the structure was overall similar to the p21_{139–160} structure bound to hPCNA (PDB ID: 1AXC). The peptides chosen for computational analysis were p21_μ-S146R, p21_μ-M147I, p21_μ-D149E, and p21_μ-FY150/151YF as these modifications resulted in the largest increase in affinity for the respective PIP-box position. In addition, a representative group containing native PIP-box sequences was also modeled bound to hPCNA: two canonical PIP-box peptides, p21_μ-Pogo and p21_μ-pol δ_{p66} ; two noncanonical PIP-box peptides p21_μ-pol ι and p21_μ-PARG, as well as seven amino-acid PIP-box peptide p21_μ-RFC (Figs. S2–S15). The resulting computationally modeled peptide structures bound to hPCNA were compared with the cocystal structure of p21_μ bound to hPCNA, which revealed the complexed structures were similar overall (Fig. 4A), represented by an

average RMSD value of 0.223 Å (Table S18). This in turn suggests that the difference in affinity for hPCNA is due to subtle structural changes. The conserved residues all adopted similar conformations between all peptide structures (Fig. 4B); however, the nonconserved residues (positions 2, 3, 5, and 6) and residues flanking the PIP-box were much more varied (Fig. 4, C and D). Intramolecular (peptide-protein) and intermolecular (peptide-peptide) interactions for each hPCNA-bound peptide structure are summarized in Tables S4, S5–S17.

In addition, the hPCNA-bound peptides p21_μ, p21_μ-pol δ_{p66} , p21_μ-pol ι , p21_μ-PARG, and p21_μ-Pogo were compared with the co-crystal structure reported for the analogous native peptide bound to hPCNA, to investigate how the sequence flanking the PIP-box motif influences binding conformation. These structures are p21_{139–160} (1AXC, (15)), pol δ_{p66} 452–466 (1U76, (14)), pol ι 415–437 (2ZVM, (16)), PARG_{402–420} (5MAV, (23)), and the Pogo PIP-box included in the mutant Pogo-Ligase peptide (1VYJ, (24)). The hPCNA-bound peptide cocystal structures were overlaid (Fig. S16), and the overall peptide conformation (RMSD values in Table S20) and secondary interactions were 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 hPCNA with RMSD values of 0.51, 0.69, 0.49, and 0.69 Å, respectively (Table S20, Fig. S16). However, the pol ι PIP-box-containing peptides bound to hPCNA adopt very different

Unlocking the PIP-box

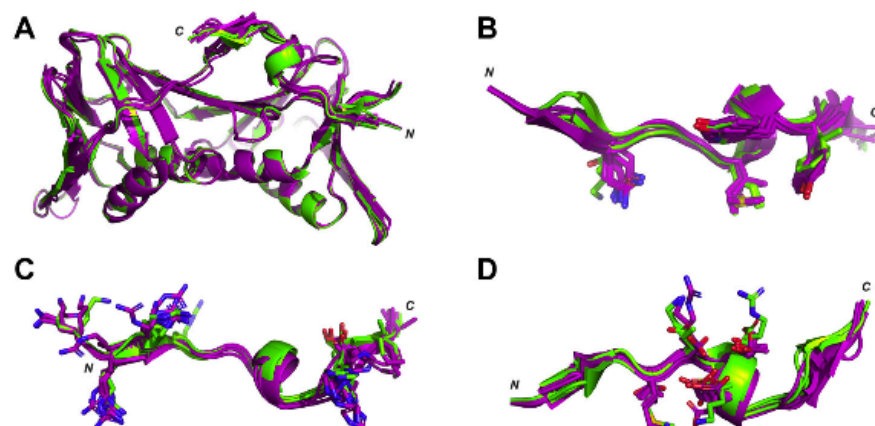


Figure 4. Superimposition of cocrystallized or computationally modeled peptide:hPCNA structures. Canonical PIP-box peptides are shown in purple, and noncanonical PIP-box peptide structures are shown in green. Canonical: p21_μ, p21_μ-F150Y, p21_μ-S146R, p21_μ-M147I, p21_μ-D149E, p21_μ-FY150/151YF, p21_μ-Pogo, and p21_μ-pol δ₆₆₆. Noncanonical: p21_μ-pol ι, p21_μ-PARG, and p21_μ-RFC. *A*, all peptides adopt a single 3₁₀-helical turn upon binding hPCNA. *B*, conserved PIP-box residues shown with sticks show a high degree of similarity. *C*, flanking residues shown with sticks are oriented in similar directions but are still flexible. *D*, nonconserved PIP-box residues shown with sticks and adopt a large variety of orientations. p21, p21^{CIP1/WAF1}; p21_μ, p21 sequence 141 to 155.

conformations, represented by an RMSD of 1.13 Å (Fig. S16 and Table S20).

Finally, the affinity and structural information for these two series of peptides was harnessed to design three new PIP-box sequences in an attempt to test our understanding of the requirements for high-affinity hPCNA binding. These new PIP-box sequences were designed to mimic favorable secondary interactions observed for native PIP-box sequences or point-modified peptides (Fig. 5). These sequences were designed with particular primary sequence combinations in mind that correlated with high hPCNA affinity, or particular secondary interactions that appeared important to stabilize the binding conformation which are highlighted throughout the discussion. These sequences were then synthesized within the p21_μ scaffold, as before, to give p21_μ-RD1–3. The binding affinity to hPCNA was determined by SPR (Fig. 2D) to reveal that all three peptides bind with higher affinity than p21_μ and helped inform and refine the interaction map shown in Figure 6. The peptides were then computationally modeled bound to the hPCNA surface (Fig. 7), which indicates that these peptides all bind hPCNA in a similar conformation to p21_μ (RMSD 0.181–0.270 Å, Table S18), where a main-chain hydrogen bond between residues 149^{P6} and 146^{P3} defines the 3₁₀-helical binding conformation (Fig. 7A). The best-performing peptide p21_μ-RD2 binds hPCNA with 1.12 nM affinity, which is remarkably the highest affinity hPCNA-binding peptide or protein reported (Fig. 2, Table S2).

Discussion

The binding affinities of the p21_μ peptides for hPCNA, and where appropriate, the structure-based information were compared and contrasted to highlight interactions that correlated to changes in hPCNA-binding affinity. Individual modifications at each PIP-box position were first examined for simplicity, then the cooperative interactions that arise were summarized, and finally how these observations pertain

to the design of the peptides p21_μ-RD1–3 are discussed below.

Conserved glutamine position 1: Glutamine enhances hPCNA affinity

Gln^{P1} is known to contribute significantly to hPCNA-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 because of the prevalence of noncanonical PIP-box sequences (11, 17). Here, the contribution of Gln^{P1} to hPCNA-binding affinity is quantified in comparison with other common position 1 residues.

Gln144^{P1} binds in the Q-pocket, and the orientation was unchanged between the canonical computationally modeled or cocrystallized hPCNA-bound peptide structures (Fig. 4B), with hydrogen bonds to hPCNA residues Ala252 and Pro253 within the Q-pocket ranging from 3.0 to 3.6 Å. The Q-pocket of hPCNA went unfilled in the hPCNA-bound noncanonical peptide structures (p21_μ-PARG, p21_μ-pol ι, p21_μ-RFC), where the position 1 residue extended over the Q-pocket to make diverse contacts with the hPCNA surface (see Figs. S10–S12). The inclusion of longer and bulkier side chains at PIP-box position 1, such as lysine (p21_μ-Q144K) and methionine (p21_μ-Q144M), resulted in peptides with the lowest affinity for hPCNA (K_D values of 1.12 and 1.54 μM, respectively). p21_μ-pol η, also with Met144^{P1}, bound with similar affinity to p21_μ-Q144M at 2.54 μM. p21_μ-RNaseH2B, containing a Met144^{P1}, shows 2.4-fold improved affinity compared with p21_μ-Q144M (640 nM *cf.* 1.54 μM). The Lys144^{P1} containing peptide, p21_μ-pol κ, bound hPCNA with 3.30 μM affinity, whereas p21_μ-pol ι that also contains a Lys144^{P1}, in contrast, gave a K_D value for hPCNA of 111 nM.

Peptides with the point modifications of aspartic acid (p21_μ-Q144D), asparagine (p21_μ-Q144N), and serine (p21_μ-Q144S) at position 1 of the PIP-box displayed significantly lower affinity (50- to 80-fold) than the p21_μ peptide, with K_D

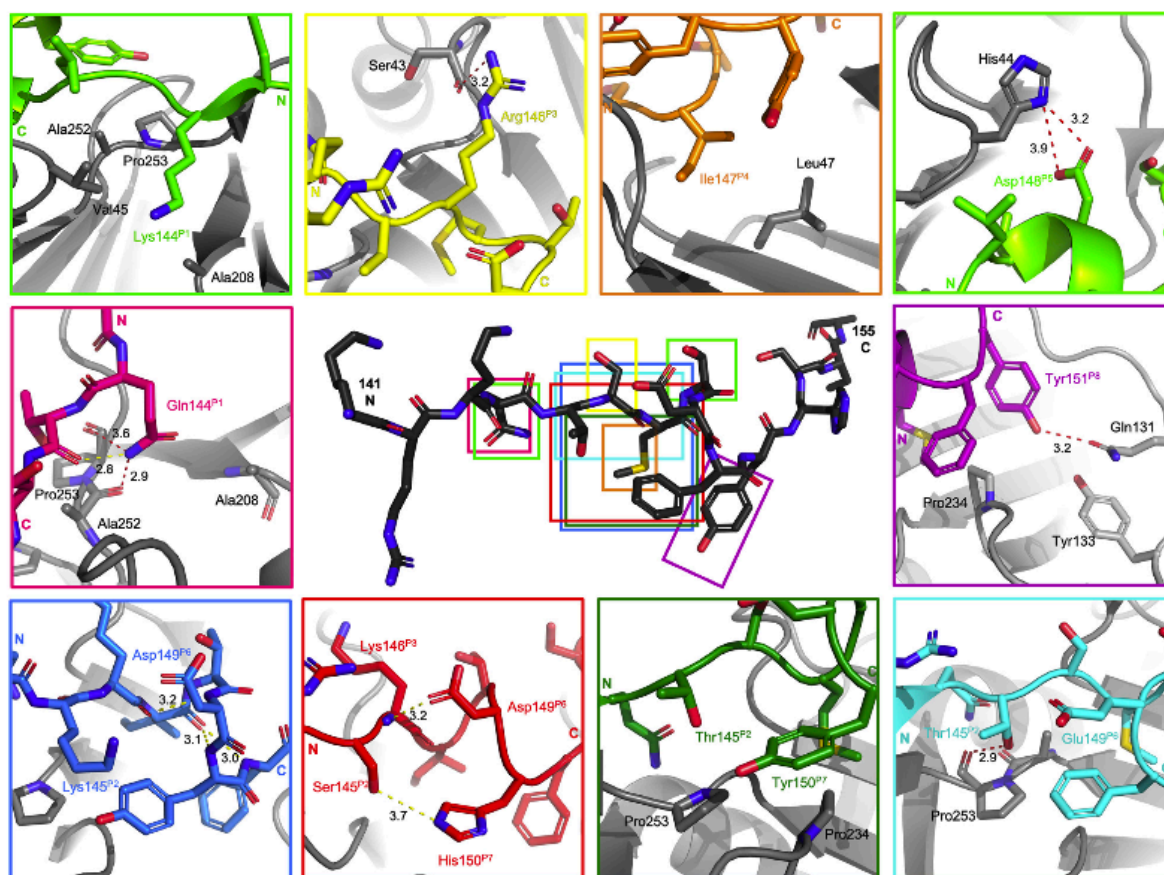


Figure 5. Representative examples of key intermolecular and intramolecular interactions. Structures shown in cartoon format with side chains as sticks. p21_μ structure shown in the center in black and the area of key interactions shown in colored rectangles. hPCNA is shown in gray. Computationally modeled peptides bound to hPCNA: p21_μ-pol ι, light green; p21_μ-S146R, yellow; p21_μ-M147I, orange; p21_μ-pol δ₉₆₆, purple; p21_μ-PARG, red; p21_μ-D149E, light blue; p21_μ-Pogo, dark blue; and p21_μ-FY150/151YF, dark green. Cocrystal structures of peptides bound to hPCNA: p21_μ-F150Y, pink and p21_μ, purple. Intramolecular hydrogen bonds are indicated as a yellow dashed line, and intermolecular hydrogen bonds are indicated as a red dashed line. Distances are indicated in angstroms. Elemental coloring: nitrogen, blue; oxygen, red; sulfur, yellow. p21, p21^{CIP1/WAF1}; p21_μ, p21 sequence 141 to 155.

values of 714 and 772 nM and 1.03 μM, respectively (see Fig. 2B, Table S2). p21_μ-pol λ, containing Ser144^{P1}, gave a K_D value of 8.14 μM, compared to p21_μ-PARG with Asp144^{P1} that binds hPCNA with significantly higher affinity at 90.0 nM. Asp144^{P1} of p21_μ-PARG does not, however, make any clear secondary interactions to explain the higher hPCNA affinity.

p21_μ-RFC, with a seven amino-acid PIP-box, was one of the highest affinity noncanonical PIP-box peptides, with a similar K_D value to p21_μ-PARG and p21_μ-pol ι, at 145 nM (Fig. 2C, 'non-canonical'). It was assumed here that Ile, Phe, and Phe of p21_μ-RFC formed a hydrophobic triad (and hence PIP-box positions 4, 7, and 8) to insert into the hydrophobic cleft of hPCNA, which results in an arginine positioned near the Q-pocket (Fig. S12). p21_μ-RFC was computationally modeled on the hPCNA 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_μ-pol ι. The other seven amino acid PIP-box in p21_μ-pol β, in contrast, was one of the worst performing peptides at 5.73 μM (Fig. 2C, Table S2), where all potential binding modes likely

force a polar/charged residue into one of the hydrophobic pockets to create unfavorable interactions and lower binding affinity of p21_μ-pol β for hPCNA.

Conserved hydrophobic position 4: Not too big and not too small

The central hydrophobic PIP-box residue plays a key role to stabilize the hPCNA-bound binding structure and anchors the peptide to the hPCNA 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 (Tables S2, S4–S13). The position 4 PIP-box residue, in all peptides here, either computationally modeled or cocrystallized with hPCNA, is entirely buried within the hydrophobic pocket indicated by a buried surface area of 100% (Table S19).

The hPCNA-bound peptide structures analyzed here all contain a methionine or isoleucine at position 4 of the PIP-box. There was no significant difference between the overall

Unlocking the PIP-box

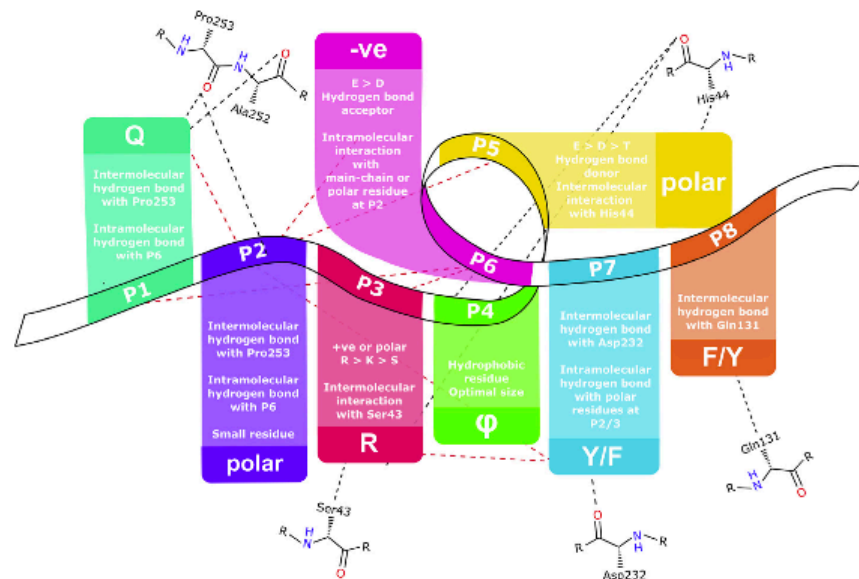


Figure 6. Summary of the guideline to design a high-affinity PIP-box and the interactions each residue participates in when bound to hPCNA. hPCNA residues shown as *line* structures. Intermolecular interactions to side chain or main chain shown as *black dashed lines*, intramolecular interactions from side chain or main chain shown as *red dashed lines*.

structures of p21_μ and p21_μ-M147I, confirmed by the RMSD value of 0.191 Å (Table S18). p21_μ-M147I has a marginally higher affinity for hPCNA than p21_μ, at 11.1 nM (Fig. 2B). The orientation of Met147^{P4} differed only by slight rotations of the

side chain because of its flexibility and large pocket size, and made side-chain interactions with hydrophobic residues Val45, Leu47, and Val48 of hPCNA. A main-chain hydrogen bond interaction between peptide residue Met147^{P4} and hPCNA

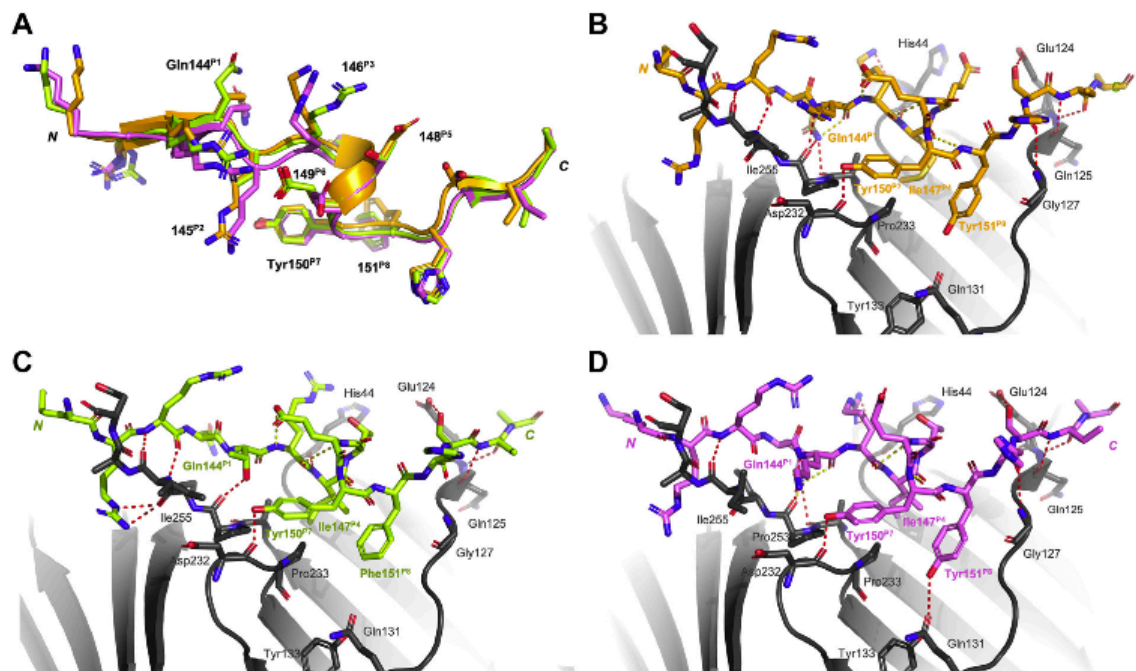


Figure 7. Computational modeling of rationally designed PIP-box peptides p21_μ-RD1, p21_μ-RD2, and p21_μ-RD3 on hPCNA. hPCNA shown in *gray* as a *cartoon*, with interacting amino acids shown as *sticks* and labeled in *gray*. Intermolecular polar interactions are shown as *red dashed lines* and intramolecular polar interactions shown as *yellow dashed lines*. *A*, overlay of the structure of the rationally designed PIP-box peptides. *B*, p21_μ-RD1 shown in *orange* and conserved residues labeled. *C*, p21_μ-RD2 shown in *green* and conserved residues labeled. *D*, p21_μ-RD3 shown in *pink* and conserved residues labeled. p21_μ, p21 sequence 141 to 155; p21, p21^{CIP1/WAF1}.

residue His44 was evident in all eight of the Met147^{P4}-containing modeled peptides bound to hPCNA. The Ile147^{P4} peptides p21_μ-pol ι and p21_μ-PARG similarly make this contact.

Hydrophobic amino-acid substitutions in p21_μ-M147V and p21_μ-M147L were well tolerated and bound hPCNA with 29.3 nM and 20.5 nM affinity, respectively. p21_μ-M147W, with the bulkier Trp147^{P4}, revealed lower affinity for hPCNA at 3.57 μM. However, the peptide containing alanine (p21_μ-M147A), with the smallest side chain of this subset, resulted in even weaker affinity for hPCNA, at 7.59 μM. These results reiterate the necessity for a hydrophobic residue at position 4, with an optimal size to fill the hydrophobic pocket and anchor the peptide to the hPCNA surface.

Conserved aromatic positions 7/8: Hydrogen bonds create tighter interactions between peptide and hPCNA

These aromatic residues complete the PIP-box hydrophobic triad that anchors a PIP-box sequence onto the hPCNA surface and stabilizes the single 3₁₀-helical turn characteristic of the hPCNA-partner binding conformation (Fig. 4A). The p21_μ PIP-box position 8 residue inserts into a hydrophobic pocket on the hPCNA surface formed by Gln131, Ile128, Pro234, Tyr133, and Tyr250 (Fig. 3C). This is exemplified by the Phe150^{P7}Tyr151^{P8} combination observed in p21_μ, where the tyrosine phenol forms a 3.4 Å hydrogen bond to hPCNA Gln131. This interaction is also seen for Tyr151^{P8} in p21_μ-S146R and p21_μ-M147I that both hydrogen bond to Gln131 at 2.8 Å with hPCNA affinities of 4.20 and 11.1 nM, respectively.

The p21_μ-Y151F peptide has comparable affinity to p21_μ (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 with the analogous native peptide (residues 139–160) (22). This may be due to a difference in binding-affinity assay (SPR *versus* isothermal calorimetry), difference in the assay conditions (e.g., buffer), or a series of small conformational changes that allow p21_μ to bind with relatively high affinity, despite being notably shorter than p21_{139–160}. 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_μ-Cdt1 to 3.30 μM for p21_μ-pol κ (Fig. 2C).

p21_μ-F150Y displayed a slight decrease in affinity for hPCNA, compared with p21_μ, with a *K_D* value of 20.2 nM (Fig. 2B, Table S2). The cocrystal structure of p21_μ-F150Y bound to hPCNA indicates the Tyr150^{P7} side chain of p21_μ-F150Y forms a 3.7 Å hydrogen bond to Pro253, that is located at the edge of the hydrophobic cleft (Figs. 3D and 5, pink). The phenol group introduced from the Phe150Tyr^{P7} modification participates in a 3.6 Å hydrogen bond with Thr145^{P2} of the peptide (Fig. 3D), which stabilizes the 3₁₀-helical binding conformation of the peptide and may enhance hPCNA binding affinity. Tyr^{P8} often makes a hydrogen bond with Gln131, which is not observed for p21_μ-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, although 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_μ-FY150/151YF (Fig. 2B). The computationally modeled structure of p21_μ-FY150/151YF bound to hPCNA 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 (cf. 3.4 Å in p21_μ), and there is also no hydrogen bond evident between Tyr150^{P7} and Thr145^{P2}. However, a strong intermolecular hydrogen bond from Tyr150^{P7} to the Ala232 carbonyl may improve hPCNA affinity. The high-affinity p21_μ-Pogo (8.82 nM) also displays the Tyr^{P7}Phe^{P8} combination.

A sequence modification to include the aromatic histidine, p21_μ-F150H, displayed lower affinity for hPCNA at 159 nM, compared with p21_μ, with a *K_D* value of 4.32 nM (Fig. 2B). The computationally modeled structure of p21_μ-PARG bound to hPCNA shows His150^{P7} positioned similarly to the analogous Phe150^{P7} in p21_μ. His150^{P7} is angled toward the N terminus of the peptide and makes a secondary interaction with Ser145^{P2} (Fig. 5, red), which may improve the affinity of the noncanonical PIP-box sequence. Other nonaromatic 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_μ-WRN was the worst performing native canonical PIP-box peptide with an affinity for hPCNA of 1.15 μM (Fig. 2C), whereas p21_μ-pol ι, which contains a Leu151^{P8}, bound hPCNA with 111 nM affinity. The computationally modeled structure of p21_μ-pol ι bound to hPCNA suggests Leu151^{P8} is lifted out of the hydrophobic pocket, compared with Tyr151^{P8} in p21_μ, indicated by an increased distance to Gln131 (3.2 Å in p21_μ, to 6.4 Å in p21_μ-pol ι). This may explain the 9-fold lower affinity of p21_μ-pol ι (111 nM) than p21_μ (12.3 nM).

Nonconserved 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 hPCNA affinity (18). Ser146^{P3} in the cocrystal structure of p21_μ bound to hPCNA is orientated toward the hPCNA surface and makes a 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 (Fig. 3C) in all three repeats (PDB ID: 7KQ1). Inclusion of lysine or arginine at 145^{P2} (p21_μ-T145K and p21_μ-T145R) resulted in 8- and 7-fold decreased affinity for hPCNA, respectively, relative to p21_μ (Fig. 2B), whereas p21_μ-S146K, with Lys146^{P3}, revealed affinity for hPCNA comparable with p21_μ. However, an Arg146^{P3} instead resulted in a 2.8-fold improvement in the *K_D* value, to 4.30 nM for p21_μ-S146R (Fig. 2B). The computationally modeled structure of p21_μ-

Unlocking the PIP-box

S146R bound to hPCNA indicates Arg146^{P3} main-chain amide makes a 3.2 Å intramolecular hydrogen bond to the carbonyl of the Asp149^{P6} side chain (Fig. S6B), and a 2.8 Å intermolecular hydrogen bond to Ser43 on the hPCNA surface (Fig. 5, yellow). These interactions may together enhance the hPCNA-binding affinity of p21_μ-S146R. The high-affinity peptides p21_μ-XPG and p21_μ-FEN1 also contain an Arg146^{P3} and bind hPCNA with 16.4 and 166 nM affinity, respectively (Fig. 2C). p21_μ-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 hPCNA binding affinity when included in the p21_μ peptide, p21_μ-TS145/146RR, which has a K_D value of 1.83 nM.

p21_μ-Pogo binds hPCNA with slightly higher affinity (8.82 nM) than p21_μ-TS145/146KK, which contains the same Lys^{P2}Lys^{P3} motif (20.2 nM, Fig. 2C). The computationally modeled structure of p21_μ-Pogo bound to hPCNA shows Lys146^{P3} makes an analogous 3.2 Å intramolecular hydrogen bond to Arg146^{P3} in p21_μ-S146R, with the 149^{P6} residue, but no intermolecular interactions (Fig. 5, dark blue). The two other permutations of arginine and lysine as dual modifications at positions 145 and 146 (p21_μ-TS145/146RK, p21_μ-TS145/146KR) both resulted in lower K_D values for hPCNA, than p21_μ, at 17.3 and 37.0 nM (Fig. 2B).

Negatively charged residues are not commonly observed at positions 2 and 3; however, Asp146^{P3} is present in the RFC PIP-box sequence. The computationally modeled p21_μ-RFC peptide bound to hPCNA indicates Asp146^{P3} makes a 3.2 Å intramolecular interaction with the positively charged Lys149^{P6} (Fig. S12). This complementary interaction is equivalent to the interaction seen for peptides such as p21_μ-S146R where Arg146^{P3} and Asp149^{P6} interact. Such intramolecular interactions may contribute to the unexpectedly high affinity of p21_μ-RFC (145 nM) by stabilizing the 3₁₀-helical binding conformation.

The PIP-box sequences of pol ι and pol δ_{p66} contain neutral residues at PIP-box positions 2 and 3, ⁴²¹GlyLeu and ⁴⁵⁷ValSer, respectively. Both p21_μ-pol ι and p21_μ-pol δ_{p66} bind hPCNA with lower affinity than p21_μ at 111 and 124 nM, respectively, and lack the ability to make a PIP-box position 2 side-chain hydrogen bond with hPCNA, as in p21_μ. This observation reinforces the conclusion that positively charged or polar residues at these nonconserved positions lead to enhanced hPCNA-binding affinity.

Nonconserved 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 hPCNA affinity observed because of 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 cocrystal structure of p21₁₃₉₋₁₆₀ (PDB ID: 1AXC) highlights two intramolecular interactions of Asp149^{P6} and one of Thr145^{P5} (10, 15). The cocrystal structure of p21_μ bound to hPCNA (PDB ID: 7KQ1)

here shows the Asp149^{P6} side chain (hydrogen bond acceptor) makes a hydrogen bond with the Thr145^{P5} side chain in 1 of the 3 monomer repeats. Residues Asp149^{P6} and Ser146^{P3} in p21_μ participate in an intramolecular main-chain hydrogen bond to define a 3₁₀-helix (Fig. 3C). These two hydrogen bonds are the only interactions of the PIP-box position 5 or 6 residues in p21_μ. The analogous interaction is made in the computationally modeled structure of p21_μ-PARG and p21_μ-Pogo bound to hPCNA. The Asp149^{P6} side chain in p21_μ-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_μ-T148E and p21_μ-T148D) resulted in ~6- to 8-fold lower affinity for hPCNA, than p21_μ. The K_D value of p21_μ-TD148/149DE was notably worse than p21_μ, at 28-fold lower affinity. However, this was recovered to only 4-fold lower than p21_μ for a Glu^{P5}Glu^{P6} modification (for p21_μ-TD148/149EE) (Fig. 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 the His44 side chain of hPCNA where an Asp^{P5} or Glu^{P5} acts as a hydrogen bond donor (e.g., Fig. 5, light green) and is more common in noncanonical 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 3₁₀-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, Fig. 2C). In p21_μ-pol ι , the unusual Tyr149^{P6} points away from the surface in the computationally modeled structure bound to hPCNA, and the 3₁₀-helix stabilizing 3.3 Å intramolecular main-chain hydrogen bond is instead made between position 6 and Ile147^{P4} (Fig. S11).

Individual PIP-box modifications must be considered in the context of the entire PIP-box sequence

All p21_μ peptides with greater than 20 nM affinity for hPCNA display charge or hydrogen-bond complementary residues on either side of the hydrophobic residue at position 4, supporting the idea that this configuration aids stabilization of the 3₁₀-helical binding conformation and in turn enhances affinity. This is seen in p21_μ-Pogo and p21_μ-XPG that bind with higher affinity than the p21_μ peptide (Fig. 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 noncanonical PIP-box peptides. The RFC seven amino-acid PIP-box contains a charge complementary pair on 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, noncanonical

PARG PIP-box in the p21_μ-PARG peptide displays relatively high hPCNA affinity at 90.0 nM despite the lack of a canonical Gln^{P1} and contains His150^{P7} rather than phenylalanine or tyrosine. Interestingly, this peptide binds with higher affinity than p21_μ-F150H, suggesting the expected affinity loss due to the absence of the Gln144^{P1} is overcome by the remaining nonconserved residues, such as the charge complementary Lys146^{P3} and Asp149^{P6}, and hydrogen bond between His150^{P7} and Ser145^{P2} (Fig. 5, red). The PARG PIP-box also contains the Thr^{P5}Asp^{P6} motif observed in the high-affinity hPCNA partners p21 and Pogo-Ligase. The highly divergent PIP-box of the p21_μ-pol ι peptide also shows relatively high hPCNA affinity (111 nM), in alignment with the higher-than-expected affinity displayed by the native pol ι₄₁₅₋₄₃₇ 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 stabilize the binding conformation and improve affinity in the absence of these traditional elements, for example, a hydrogen bond between Asp425^{P5} and His44 of hPCNA (cf. Fig. 5, light green) and an intramolecular Tyr427^{P8} and Lys421 hydrogen bond (Fig. S16F, yellow).

Some peptides such as p21_μ-pol δ_{p66} and p21_μ-FEN1 bind with unexpectedly lower affinity than p21_μ despite strict canonical sequences. p21_μ-FEN1 is particularly interesting as it contains not only 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 hPCNA binding affinity when introduced as a point modification in p21_μ (p21_μ-S146R, 4.30 nM). The other sequence differences in p21_μ-FEN1 (compared with p21_μ) are the Phe^{P7}Phe^{P8} motif (seen in p21_μ-Y151F, 10.6 nM) and Leu147^{P4} (seen in p21_μ-M147L, 20.5 nM) that did not drastically impact affinity in the point-modified peptides. Perhaps, the lower affinity of p21_μ-FEN1 is owed in part to the Asp^{P5}Asp^{P6} combination, as the p21_μ-T148D peptide showed a notably lower hPCNA affinity relative to p21_μ (96.5 nM, cf. 4.32 nM for p21_μ). 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 hPCNA-binding affinity have been summarized 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 hPCNA affinity, extends to the sequence that flanks the PIP-box. The p21_μ 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_μ peptides containing a native PIP-box was compared with the reported affinity of the analogous native peptide sequence to reveal an interesting trend: All five of the p21_μ native canonical PIP-box hybrid peptides display higher affinity than the native peptide (Table 1). The most significant difference is greater than 820-fold for the Cdt1 PIP-box peptides, where p21_μ-Cdt1 binds with 8.76 nM and Cdt1₂₋₁₅ 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 with the p21_μ with four arginine residues. In contrast, four of the six noncanonical PIP-box:p21_μ hybrids showed lower affinity relative to the native sequence (Table 1).

Rationally designed PIP-box sequences: Designing the highest affinity hPCNA 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_μ peptides in conjunction with the structural observations from cocrystal structures, computationally modelled peptides, and native PIP-box structures, which are summarized graphically in Figures 5 and 6. Three PIP-box sequences were synthesized to trial whether amino-acid combinations not previously observed together could give rise to high hPCNA 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 hPCNA affinity of p21_μ-RD1.

Three peptides were synthesized, the affinity for hPCNA determined by SPR, and the hPCNA-bound structures computationally modelled to reveal, remarkably, that all three

Table 1

The effect of changing the sequence flanking the PIP-box motif, on the hPCNA-binding affinity

Name	Sequence	Affinity K_D		Native peptides	Affinity K_D	Ref.
p21 _μ -Pogo	KRRQKKITDYFHSKR	8.82 nM	<	SAVLQKKITDYFHPKK ¹	100 nM	(21, 24)
p21 _μ -Cdt1	KRRQRRVTDFHFSKR	8.76 nM	<	² MEQRRVTDFARRR ¹⁵	7.20 μM	(26)
p21 _μ -pol δ _{p66}	KRRQVSITGFFHFSKR	124 nM	<	⁴⁵² KANRQVSITGFFQRK ⁴⁶⁶	15.6 μM	(14)
p21 _μ -FEN1	KRRQGRLLDDFFHFSKR	166 nM	<	³³¹ SRQGSTQGRLLDDFFKVTGSL ³⁵⁰	59.9 μM	(14)
p21 _μ -p15	KRRQKGGIEFFHFSKR	234 nM	<	⁴¹ APVCVRPTPKWQKGGIEFFAA ⁷²	5.56 μM	(46)
p21 _μ -PARG	KRRDSKITDHFHFSKR	90.0 nM	<	⁴⁰² QHGGKDSKITDHFMRPKA ⁴²⁰	3.3 μM	(23)
p21 _μ -pol ι	KRRKGLIDYLYLHFSKR	111 nM	<	⁴¹⁵ ALNTAKKGLIDYLYLMPSLSTTSR ⁴³⁷	0.39 μM	(16)
p21 _μ -pol η	KRRMQTLESFFHFSKR	2.54 μM	>	⁶⁹³ CKRPRPEGMQTLESFFKPLTH ⁷¹³	0.4 μM	(16)
p21 _μ -pol κ	KRRKHTLDIFFHFSKR	3.30 μM	>	⁸⁶¹ PKHTLDIFFK ⁸⁷⁴ PLTH ⁹	4.9 μM	(16)
p21 _μ -Cdt2	KRRMRKICTYFHSKR	ND ^a	>>	⁷⁰⁴ SSMRKICTYFHRKS ⁷¹⁷	57 nM	(26)
p21 _μ -RNaseH2B	KRRMKSIDTFFHFSKR	640 nM	<	²⁹⁰ DKSGMKSIDTFFGVNKKKIGKV ³¹²	35 μM	(47)

PL, Pogo-Ligase.

A comparison of p21_μ:PIP-box hybrid peptides and the analogous native peptides from which the PIP-box derives.

^a PL peptide is a mutant hybrid peptide and not an entirely native sequence.

^b PLTH is not part of the native sequence and was added to the sequence to improve binding affinity.

Unlocking the PIP-box

peptides bound with higher affinity than the native p21_μ (Fig. 2D, Table S2). The highest affinity peptide p21_μ-RD2 bound hPCNA with 1.12 nM affinity, which is 11-fold higher affinity than p21_μ and 3-fold higher than p21₁₃₉₋₁₆₀ despite being seven amino acids shorter. This makes p21_μ-RD2, a PIP-box peptide rationally designed amino acid by amino acid, the highest affinity hPCNA-binding peptide or protein to date, displaying a significant improvement in affinity for hPCNA. The computationally modeled structures of these peptides bound to hPCNA were analyzed to confirm whether the hypothesized 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_μ-RD1 and p21_μ-RD3, the bond between Tyr150^{P7} and Pro253 is stronger than that of p21_μ, at distances of 3.5 Å and 3.4 Å, respectively, which may contribute to the improved affinity for hPCNA. The highest affinity peptide p21_μ-RD2 (*K_D* 1.12 nM) has four defined intramolecular hydrogen bond interactions (Fig. 7C), which all involve PIP-box residues and may stabilize the 3₁₀-helical binding conformation and lead to the enhanced hPCNA affinity. In particular, the Glu149^{P6} forms a main-chain hydrogen bond to Arg146^{P3}. p21_μ-RD2 contains an Arg146^{P3} that makes a hydrogen bond to Ser43 in the computationally modeled structure bound to hPCNA, which is also observed for p21_μ-S146R that bound hPCNA with high affinity (4.30 nM, Fig. 5, yellow). In addition, Thr145^{P3} side chain of p21_μ-RD2 hydrogen bonds to the Pro253 carbonyl, and the Tyr150^{P7} phenol hydrogen bonds with Asp232 main-chain carbonyl (Fig. 7C). The 146Arg^{P3} modification resulted in increased affinity in p21_μ-S146R and was included in p21_μ-RD2 and p21_μ-RD3; however, the affinity of p21_μ-RD3 was not significantly different to p21_μ and suggests the other modifications may be working to oppose an increase in affinity. Interestingly, p21_μ-RD1 with the Tyr^{P7}-Tyr^{P8} motif only makes an intermolecular hydrogen bond between Tyr150^{P7} and Asp232, as seen in p21_μ-F150Y; however, p21_μ-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.

Conclusions and outlook

This work provides a comprehensive study using a set of hPCNA-binding peptides and highlights a series of interactions that synergistically contribute to high hPCNA affinity, as summarized in Figure 6. The glutamine residue and hydrophobic triad are essential to enhance hPCNA-binding affinity and adhere a peptide to the hPCNA surface. The PIP-box position 1 glutamine residue and its hydrogen bond network contribute notably to hPCNA binding affinity, although glutamine is not essential to confer hPCNA binding. The PIP-box position 4 hydrophobic residue, which fills the hydrophobic cleft, is essential to a high-affinity hPCNA 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, to stabilize the

characteristic 3₁₀-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 hPCNA within the hydrophobic cleft. The cocrystal structure of p21_μ-F150Y bound to hPCNA is the first Tyr^{P7}Tyr^{P8} peptide reported to interact with hPCNA. 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 on either side of the hydrophobic residue at position 4 stabilize the 3₁₀-helix binding conformation and appear particularly important to increase hPCNA affinity of noncanonical PIP-box sequences. Introduction of a positively charged residue at position 3 is more favorable than at position 2 to increase hPCNA-binding affinity. The nonconserved PIP-box combinations that lead to increased hPCNA-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 hPCNA surface. Negatively charged residues at position 5 more commonly hydrogen-bond intermolecularly with hPCNA, in particular with His44, whereas position 6 residues often form intramolecular hydrogen bonds. Finally, 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 nonconserved PIP-box combinations on hPCNA-binding affinity and has identified key modifications that increase affinity such as the PIP-box position 3 arginine.

p21_μ-RD2, with PIP-box QTRITEYF, is the highest affinity hPCNA-binding partner reported to date (1.12 nM) and is seven amino acids shorter than the previous title holder in p21₁₃₉₋₁₆₀ (4.32 nM). This peptide conforms explicitly to the guidelines for secondary interactions that lead to high-affinity hPCNA binding set out in Figure 6. It is important to note that p21_μ-RD2 provides just one solution to the puzzle of obtaining high affinity to hPCNA and many other combinations may also lead to high hPCNA binding affinity. This study has highlighted the complexity of the secondary interaction network that gives rise to high-affinity binding of short peptides to hPCNA and has made significant progress to define the rules of high-affinity hPCNA interaction. The importance of these interactions, in particular the intermolecular interactions with hPCNA, may be further probed by mutation of key hPCNA residues such as His44. Furthermore, it emphasizes that to predict hPCNA-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 hPCNA-binding peptides or tune the affinity of hPCNA peptides and proteins to probe interactions important in DNA replication

and DNA repair. Consequently, this study provides an insight into how nature has fine-tuned affinity of the native hPCNA-binding proteins to allow tight regulation of DNA-replication and DNA-repair processes. This knowledge may be further leveraged to design inhibitors of human hPCNA interactions for therapeutic applications.

Experimental procedures

Peptide synthesis

All peptides were synthesized using Fmoc/t-Bu Solid-Phase Peptide Synthesis (27). Six peptides were synthesized using Rink Amide resin on a Liberty Blue peptide synthesizer (CEM Corp); 31 peptides were synthesized on a Prelude peptide synthesizer (PTI); five peptides were synthesized manually as detailed in the [Supplementary information](#). The remaining ten peptides were purchased from Shanghai Royobiotech at >95% purity. Synthesized 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 high resolution mass spectrometry using an Agilent 6230 ESI-TOF MS. Peptide purity was characterized 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 to 50% acetonitrile with 0.1% TFA and in water with 0.1% TFA, over 15 min at 1.5 ml/min and visualized at 220 nm. Detailed methods are included in the [Electronic Supplementary Information](#) and characterization data are listed in [Table S1](#).

Protein expression and purification

A glycerol stock of *Escherichia coli* BL21 (ΔDE3) cells carrying a hPCNA-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 an absorbance at 600 nm 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 5000g 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) and then lysed by five rounds of cell disruption by a microfluidics cell disrupter. Lysate was pelleted at 45,000g for 60 min, and the supernatant was collected for purification.

hPCNA was purified at 4 °C by fast protein liquid chromatography, using an anion exchange diethylaminoethyl column (HiTrap diethylaminoethyl 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 analyzed 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 analyzed by SDS-PAGE, and those of interest were dialyzed 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 a 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 analyzed 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 analyzed 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.

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% Tween-20, adjusted to pH 7.4 with 2 M NaOH. A GE CM5 (series S) sensor chip was primed with the 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% H₃PO₄, and glycine, pH 9.5, respectively. The surface was then activated with an injection of 0.2 M 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide and 50 mM *N*-hydroxysuccinimide (600 s, 10 μl/min). hPCNA (5 μl, 12 mg/ml) was diluted into the 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 hPCNA/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 stabilization. Both flow cells were then blocked with 1.0 M ethanolamine, pH 8.5 (600 s, 10 μl/min). The chip was left to stabilize for 2 h before sample injections commenced.

Peptides (approx. 2 mg) were dissolved in MilliQ water 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>, (28)); however, an additional glycine residue was added to each peptide sequence to account for the terminal amide of the synthesized peptides. The peptide stock solution concentration was then calculated per $c = (A_{205}/\epsilon_{205} \times l) \times DF$, where the concentration is in molar, A_{205} is absorbance at 205 nm calculated as an average of three readings, l

Unlocking the PIP-box

is the pathlength in centimeter (0.1 cm for Nanodrop), 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 the running buffer before further dilution as necessary.

Steady-state affinity experiments were conducted at a flow rate of 30 μ l/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, eight times, was performed for each peptide, and the resulting solutions were injected sequentially from the lowest to highest concentration, preceded by a buffer-only blank injection. After each injection, the surface was regenerated with 2 M NaCl (2 \times 30 s, 25 μ l/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 were analyzed using the GE Biosystems Biacore S200 Evaluation Software. All data are summarized in Table S2 and Figure 2.

Protein-peptide cocrystallization experiments

hPCNA was mixed with peptide of interest at 1:1.2 M ratio, and after incubation on ice for 30 min, the sample was pelleted at 16,000g 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 linbro plates containing 500 μ l well solution, by mixing 1 μ l protein and peptide with an equal volume of the well solution (29–31). Initial cocrystallization screens with all p21 $_{\mu}$ -modified peptides and hPCNA were attempted. Diffracting crystals of hPCNA bound to p21 $_{\mu}$ (residues 141–150) were formed in 8% Tacsimate and 20% PEG at 16 $^{\circ}\text{C}$ after 4 weeks. Diffracting crystals of hPCNA 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 (29–31). Data were collected at 100 K using the MX1 beamline at the Australian Synchrotron (32). Diffraction data were indexed and integrated using X-ray Detector Software (33). Pointless (CCP4i) (34) was used to create an mtzfile for scaling. Data were scaled using Aimless (CCP4i) (35, 36) 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) (37) using a search model for p21 $_{\mu}$ of (PDB ID: 1AXC, human (15)) and for p21 $_{\mu}$ -F150Y (PDB ID: 4RJF, human (22)). Solutions were refined in phenix.refine (38) in iterative rounds with manual rebuilding in Coot (39, 40). Data collection and refinement statistics for hPCNA in complex with p21 $_{\mu}$ or p21 $_{\mu}$ -F150Y are summarized in Table S3. The final structures are deposited on the RCSB database under accession numbers 7KQ1 and 7KQ0, respectively.

Computational modelling

Models of hPCNA and p21 analogue peptide structures were constructed using the solved structure of hPCNA bound with p21 $_{\mu}$ peptide as a starting template (PDB ID: 7KQ1). The peptides analyzed 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, and p21 $_{\mu}$ -RD3. The residue(s) being investigated were mutated to the amino acid of interest, and unresolved side chains of residues were modeled into the computational structure. Energy minimization/annealing ($n = 30$) for refinement was carried out in ICM-Pro Molsoft (41, 42). Refined models were analyzed using PyMOL to validate the model by comparing against (p21 $_{\mu}$ structure) and assess side-chain interactions (43). The resulting structures were visualized in PyMOL and are depicted in Figs. S4–S15. Additional analysis was carried out using the RING server (44) and PoseView (45).

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.

Supporting information—This article contains supporting information (14–16, 23, 24, 28, 44, 48).

Acknowledgments—This research was undertaken in part using the MX1 beamline at the Australian Synchrotron, part of Australian Nuclear Science and Technology Organisation. The facilities of the OptoFab node of the Australian National Fabrication Facility and associated Commonwealth and South Australian State Government funding are also gratefully acknowledged.

The research was supported by the Australian Research Council Centre of Excellence for Nanoscale BioPhotonics (CNBP) (CE140100003).

Author contributions—A. J. H. conceptualization, formal analysis, investigation (peptide synthesis, SPR assays), validation, writing—original draft, and writing—review and editing; B. A. V. formal analysis, investigation, (protein synthesis, protein crystallography, computational modeling), validation, writing—original draft, and writing—review and editing; W. K. resources, supervision, and investigation, (peptide synthesis), validation, and writing—review and editing; T. C. investigation, validation, and writing—review and editing; D. B. S.—supervision, investigation, (peptide synthesis), and writing—review and editing; A. D. A. resources, supervision, and writing—review and editing; J. B. B. resources, supervision, validation, and writing—review and editing.

Funding and additional information—A. J. H. and B. A. V. are supported by Australian Government Research Training Program Stipends (RTPS).

Conflict of interest—The authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations—The abbreviations used are: p21, p21^{CIP1/WAF1}; p21 $_{\mu}$, p21 sequence 141 to 155; PIP, hPCNA-interacting protein; SPR, surface plasmon resonance.

References

1. De Biasio, A., and Blanco, F. J. (2013) Proliferating cell nuclear antigen structure and interactions: Too many partners for one dancer? *Adv. Protein Chem. Struct. Biol.* 91, 1–36

2. Tsurimoto, T. (1999) PCNA binding proteins. *Front. Biosci.* **4**, D849–D858
3. Maga, G., and Hubscher, U. (2003) Proliferating cell nuclear antigen (PCNA): A dancer with many partners. *J. Cell Sci.* **116**, 3051–3060
4. Moldovan, G. L., Pfander, B., and Jentsch, S. (2007) PCNA, the maestro of the replication fork. *Cell* **129**, 665–679
5. Boehm, E. M., Gildenberg, M. S., and Washington, M. T. (2016) The many roles of PCNA in eukaryotic DNA replication. *Enzymes* **39**, 231–254
6. Stoimenov, I., and Helleday, T. (2009) PCNA on the crossroad of cancer. *Biochem. Soc. Trans.* **37**, 605–613
7. Zhongyun, D., Wortman, M., Tan, Z., and Dillehay, K. (2012) *WO 2012/033938 A2, Identification of PCNA Targeting Compounds for Cancer Therapy and PCNA Function Regulation*, World Intellectual Property Organisation, International Bureau
8. De March, M., Merino, N., Barrera-Vilarmau, S., Crehuet, R., Onesti, S., Blanco, F. J., and De Biasio, A. (2017) Structural basis of human PCNA sliding on DNA. *Nat. Commun.* **8**, 13935
9. Majka, J., and Burgers, P. M. J. (2004) The PCNA–RFC families of DNA clamps and clamp loaders. *Prog. Nucleic Acid Res. Mol. Biol.* **78**, 227–260
10. Horsfall, A. J., Abell, A. D., and Bruning, J. B. (2019) Targeting PCNA with peptide mimetics for therapeutic purposes. *ChemBioChem.* **21**, 442–450
11. Prestel, A., Wichmann, N., Martins, J. M., Marabini, R., Kassem, N., Broendum, S. S., Otterlei, M., Nielsen, O., Willemoes, M., Ploug, M., Boomsma, W., and Kragelund, B. B. (2019) The PCNA interaction motifs revisited: Thinking outside the PIP-box. *Cell. Mol. Life Sci.* **76**, 4923–4943
12. Gonzalez-Magana, A., and Blanco, F. J. (2020) Human PCNA structure, function and interactions. *Biomolecules* **10**, 570
13. Wegener, K. L., McGrath, A. E., Dixon, N. E., Oakley, A. J., Scanlon, D. B., Abell, A. D., and Bruning, J. (2018) Rational design of a 310-helical PIP-box mimetic targeting PCNA - the human sliding clamp. *Chem. Eur. J.* **24**, 11325–11331
14. Bruning, J. B., and Shamooy, Y. (2004) Structural and thermodynamic analysis of human PCNA with peptides derived from DNA polymerase-delta p66 subunit and flap endonuclease-1. *Structure* **12**, 2209–2219
15. Gulbis, J. M., Kelman, Z., Hurwitz, J., O'Donnell, M., and Kuriyan, J. (1996) Structure of the C-terminal region of p21 WAF1/CIP1 complexed with human PCNA. *Cell* **87**, 297–306
16. Hishiki, A., Hashimoto, H., Hanafusa, T., Kamei, K., Ohashi, E., Shimizu, T., Ohmori, H., and Sato, M. (2009) Structural basis for novel interactions between human translesion synthesis polymerases and proliferating cell nuclear antigen. *J. Biol. Chem.* **284**, 10552–10560
17. Boehm, E. M., and Washington, M. T. (2016) R.I.P. to the PIP: PCNA-binding motif no longer considered specific. *Bioessays* **38**, 1117–1122
18. Warbrick, E., Lane, D. P., Glover, D. M., and Cox, L. S. (1995) 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.* **5**, 275–282
19. Warbrick, E. (1998) PCNA binding through a conserved motif. *Bioessays* **20**, 195–199
20. Zheleva, D. I., Zhelev, N. Z., Fischer, P. M., Duff, S. V., Warbrick, E., Blake, D. G., and Lane, D. P. (2000) A quantitative study of the *in vitro* binding of the C-terminal domain of p21 to PCNA: Affinity, stoichiometry, and thermodynamics. *Biochemistry* **39**, 7388–7397
21. Warbrick, E. (2006) A functional analysis of PCNA-binding peptides derived from protein sequence, interaction screening and rational design. *Oncogene* **25**, 2850–2859
22. Kroker, A. J., and Bruning, J. B. (2015) p21 exploits residue Tyr151 as a tether for high-affinity PCNA binding. *Biochemistry* **54**, 3483–3493
23. Kaufmann, T., Grishkovskaya, I., Polyansky, A. A., Kostroh, S., Kukulj, E., Olek, K. M., Herbert, S., Beltzung, E., Mechtler, K., Peterbauer, T., Gotzmann, J., Zhang, L., Hartl, M., Zagrovic, B., Elsayad, K., et al. (2017) A novel non-canonical PIP-box mediates PARG interaction with PCNA. *Nucleic Acids Res.* **45**, 9741–9759
24. Kontopidis, G., Wu, S.-Y., Zheleva, D. I., Taylor, P., McInnes, C., Lane, D. P., Fischer, P. M., and Walkinshaw, M. D. (2005) 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.* **102**, 1871–1876
25. Choe, K. N., and Moldovan, G.-L. (2016) Forging ahead through darkness: PCNA, still the principal conductor at the replication fork. *Mol. Cell* **65**, 380–392
26. Hayashi, A., Giakoumakis, N. N., Heidebrecht, T., Ishii, T., Panagopoulos, A., Caillat, C., Takahara, M., Hibbert, R. G., Suenaga, N., Stadnik-Spiewak, M., Takahashi, T., Shiomi, Y., Taraviras, S., von Castelmur, E., Lygerou, Z., et al. (2018) Direct binding of Cdt2 to PCNA is important for targeting the CRL4Cdt2 E3 ligase activity to Cdt1. *Life Sci. Alliance* **1**, e201800238
27. Fields, G. B., and Noble, R. L. (1990) Solid phase peptide synthesis utilizing 9-fluorenylmethoxycarbonyl amino acids. *Int. J. Pept. Protein Res.* **35**, 161–214
28. Anthis, N. J., and Clore, G. M. (2013) Sequence-specific determination of protein and peptide concentrations by absorbance at 205 nm. *Protein Sci.* **22**, 851–858
29. Marshall, A. C., Kroker, A. J., Murray, L. A., Gronthos, K., Rajapaksha, H., Wegener, K. L., and Bruning, J. B. (2017) Structure of the sliding clamp from the fungal pathogen *Aspergillus fumigatus* (AfumPCNA) and interactions with human p21. *FEBS J.* **284**, 985–1002
30. Frkic, R. L., Chua, B. S., Shin, Y., Pascal, B. D., Novick, S. J., Kamenecka, T. M., Griffin, P. R., and Bruning, J. B. (2018) Structural and dynamic elucidation of a non-acid PPAR γ partial agonist: SR1988. *Nucl. Receptors Res.* **5**, 101350
31. Pederick, J. L., Thompson, A. P., Bell, S. G., and Bruning, J. B. (2020) d-Alanine-d-alanine ligase as a model for the activation of ATP-grasp enzymes by monovalent cations. *J. Biol. Chem.* **295**, 7894–7904
32. Cowieson, N. P., Aragao, D., Clift, M., Ericsson, D. J., Gee, C., Harrop, S. J., Mudie, N., Panjikar, S., Price, J. R., Riboldi-Tunnicliffe, A., Williamson, R., and Caradoc-Davies, T. (2015) MX1: A bending-magnet crystallography beamline serving both chemical and macromolecular crystallography communities at the Australian Synchrotron. *J. Synchrotron Radiat.* **22**, 187–190
33. Kabsch, W. (2010) XDS (X-ray detector software). *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132
34. Evans, P. (2006) Scaling and assessment of data quality. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 72–82
35. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., et al. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242
36. Potterton, E., Briggs, P. J., Turkenburg, M. G. W., and Dodson, E. (2003) A graphical user interface to the CCP4 program suite. *Acta Crystallogr. D Biol. Crystallogr.* **59**, 1131–1137
37. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674
38. Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L. W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., et al. (2019) Phenix: Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in phenix. *Acta Crystallogr. D Biol. Crystallogr.* **75**, 861–877
39. Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H., and Adams, P. D. (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367
40. Emsley, P., and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132
41. Abagyan, R., Totrov, M., and Kuznetsov, D. (1994) ICM - new method for protein modeling and design: Applications to docking and structure prediction from the distorted native conformation. *J. Comput. Chem.* **15**, 488–506
42. Abagyan, R., and Totrov, M. (1994) Biased probability Monte Carlo conformational searches and electrostatic calculations for peptides and proteins. *J. Mol. Biol.* **235**, 983–1002
43. Schrodinger, LLC (2015) *The PyMOL Molecular Graphics System, Version 1.8*

Unlocking the PIP-box

44. Piovesan, D., Minervini, G., and Tosatto, S. C. (2016) The RING 2.0 web server for high quality residue interaction networks. *Nucleic Acids Res.* **44**, W367–W374
45. Stierand, K., Maass, P. C., and Rarey, M. (2006) Molecular complexes at a glance: Automated generation of two-dimensional complex diagrams. *Bioinformatics* **22**, 1710–1716
46. De March, M., Barrera-Wilarmau, S., Crespan, E., Mentegari, E., Merino, N., Gonzalez-Magana, A., Romano-Moreno, M., Maga, G., Crehuet, R., Onesti, S., Blanco, F., J., and De Biasio, A. (2018) p15PAF binding to PCNA modulates the DNA sliding surface. *Nucleic Acids Res.* **46**, 9816–9828
47. Duffy, C. M., Hilbert, B. J., and Kelch, B. A. (2016) A disease-causing variant in PCNA disrupts a promiscuous protein binding site. *J. Mol. Biol.* **428**, 1023–1040
48. Wlodawer, A., Minor, W., Dauter, Z., and Jaskolski, M. (2013) Protein crystallography for aspiring crystallographers or how to avoid pitfalls and traps in macromolecular structure determination. *FEBS J.* **280**, 5705–5736

Electronic Supplementary Information

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

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

ABBREVIATIONS

ACN, acetonitrile; **Ala (A)**, alanine; **Arg (R)**, arginine; **Asn (N)**, asparagine; **Asp (D)**, aspartic acid; **Boc**, tert-butoxycarbonyl; **Cys (C)**, cysteine; **DCM**, dichloromethane; **DEAE**, diethylethanolamine; **DIC**, *N,N'*-diisopropylcarbodiimide; **DIPEA**, *N,N'*-diisopropylethylamine; **DMF**, *N,N'*-dimethylformamide; **DODT**, 2,2'-(ethylenedioxy)diethanethiol; **DTT**, dithiothreitol; **E. Coli**, *Escherichia Coli*; **EDC**, 1-ethyl-3-(3-diethylaminopropyl)carbodiimide; **EDTA**, ethylenediaminetetraacetic acid; **ESI-TOF**, Electrospray Ionisation Time of Flight; **Fmoc**, 9-fluorenylmethoxycarbonyl; **Gln (Q)**, glutamine; **Glu (E)**, glutamic acid; **Gly (G)**, glycine; **HATU**, (1-[bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxide hexafluorophosphate; **HEPES**, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; **His (H)**, histidine; **HOBT**, 1-hydroxybenzotriazole; **HPLC**: High Performance Liquid Chromatography; **HRMS**: High Resolution Mass Spectrometry; **Ile (I)**, isoleucine; **IPTG**, isopropyl β -D-1-thiogalactopyranoside; **LB**: Lennox broth; **LCMS**: Liquid Chromatography Mass Spectrometry; **Leu (L)**, leucine; **Lys (K)**, lysine; **Met (M)**, methionine; **NHS**, *N*-hydroxysuccinimide; **Pbf**: 2,2,4,6,7-pentamethylidihydrobenzofuran-5-sulfonyl; **PCNA**, Proliferating Cell Nuclear Antigen; **Phe (F)**, phenylalanine; **Pro (P)**, proline; **RP-HPLC**: Reverse-Phase High Performance Liquid Chromatography; **SDS**, Sodium Dodecyl Sulfate; **Ser (S)**, serine; **SPPS**: Solid-Phase Peptide Synthesis; **tBu**: tert-butyl; **TFA**: trifluoroacetic acid; **Thr (T)**, threonine; **TIPS**: trisopropylsilane; **TNBS**, 2,4,6-trinitrobenzenesulfonic acid (picrylsulfonic acid); **Tris**, trisaminomethane; **Trp (W)**, tryptophan; **Trt**, trityl; **Tyr (Y)**, tyrosine; **Val (V)**, valine;

TABLE OF CONTENTS

SYNTHESIS & CHARACTERISATION OF PEPTIDES	2
Table S1: Peptide characterisation data	4
SPR:	5
Table S2: Peptide SPR data against hPCNA.	5
COCRYSTAL EXPERIMENTS	6
Table S3: Data collection and refinement statistics of hPCNA bound with p21 _μ , and hPCNA bound with p21 _μ -F150Y.	6
Figure S1: Co-crystal structure of p21 _μ bound to PCNA monomer.	7
Table S4: Secondary Interaction Summary for co-crystal structure of p21 _μ with hPCNA.	7
Figure S2: Co-crystal structure of p21 _μ -F150Y bound to PCNA monomer.	8
Table S5: Secondary Interaction Summary for co-crystal structure of p21 _μ -F150Y with hPCNA.	8
COMPUTATIONAL MODELLING OF PCNA MONOMERS BOUND TO P21_μ PEPTIDES	9
Figure S3: Computationally modelled structure of p21 _μ -M147I on the PIP-box binding site of a hPCNA monomer.	9
Table S6: Secondary Interaction Summary for computationally modelled structure of p21 _μ -M147I with hPCNA.	9
Figure S4: Computationally modelled structure of p21 _μ -FY150/151YF on the PIP-box binding site of a hPCNA monomer.	10
Table S7: Secondary Interaction Summary for computationally modelled structure of p21 _μ -FY150/151YF with hPCNA.	10
Figure S5: Computationally modelled structure of p21 _μ -S146R on the PIP-box binding site of a hPCNA monomer.	11
Table S8: Secondary Interaction Summary for computationally modelled structure of p21 _μ -S146R with hPCNA.	11
Figure S6: Computationally modelled structure of p21 _μ -D149E on the PIP-box binding site of a hPCNA monomer.	12
Table S9: Secondary Interaction Summary for computationally modelled structure of p21 _μ -D149E.	12
Figure S7: Computationally modelled structure of p21 _μ -pol δ_{p66} on the PIP-box binding site of a hPCNA monomer.	13
Table S10: Secondary Interaction Summary for computationally modelled structure of p21 _μ -pol δ_{p66} with hPCNA.	13
Figure S8: Computationally modelled structure of p21 _μ -Pogo on the PIP-box binding site of a hPCNA monomer.	14
Table S11: Secondary Interaction Summary for computationally modelled structure of p21 _μ -Pogo with hPCNA.	14
Figure S9: Computationally modelled structure of p21 _μ -PARG on the PIP-box binding site of a hPCNA monomer.	15
Table S12: Secondary Interaction Summary for computationally modelled structure of p21 _μ -PARG with hPCNA.	15
Figure S10: Computationally modelled structure of p21 _μ -pol ι on the PIP-box binding site of a hPCNA monomer.	16
Table S13: Secondary Interaction Summary for computationally modelled structure of p21 _μ -pol ι with hPCNA.	16
Figure S11: Computationally modelled structure of p21 _μ -RFC on the PIP-box binding site of a hPCNA monomer.	17
Table S14: Secondary Interaction Summary for computationally modelled structure of p21 _μ -RFC with hPCNA.	17
Figure S12: Computationally modelled structure of p21 _μ -RD1 on the PIP-box binding site of a hPCNA monomer.	18
Table S15: Secondary Interaction Summary for computationally modelled structure of p21 _μ -RD1 with hPCNA.	18
Figure S13: Computationally modelled structure of p21 _μ -RD2 on the PIP-box binding site of a hPCNA monomer.	19
Table S16: Secondary Interaction Summary for computationally modelled structure of p21 _μ -RD2 with hPCNA.	19
Figure S14: Computationally modelled structure of p21 _μ -RD3 on the PIP-box binding site of a hPCNA monomer.	20
Table S17: Secondary Interaction Summary for computationally modelled structure of p21 _μ -RD3 with hPCNA.	20
ANALYSIS OF STRUCTURES	21
Table S18: Root-mean-squared deviation (RMSD) values of modified p21 _μ .	21
Table S19: Buried Surface Area (BSA) of modified p21 _μ .	21
COMPARISON OF STRUCTURES TO NATIVE SEQUENCES	22
Figure S15: Overlaid structures of p21 _μ :PIP-box hybrid (blue) and native PIP-box peptide (yellow)	22
Table S20: Root-mean-squared deviation (RMSD) values of alternative PIP-box modified peptides	23
Table S21: Number of interactions for p21 _μ :hybrid peptides compared to native analogues	23
REFERENCES	24

SYNTHESIS & 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 & 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 is recorded in Table S1.

Peptide Synthesis:

All peptides were synthesised using one of the following methods, excepting ten peptides that were purchased from Shanghai Royobiotec at >95% purity, with purity & identity confirmed on receipt by HPLC and HRMS. *Purchased peptides:* p21_μ-M147L, p21_μ-M147A, p21_μ-M147W, p21_μ-M147V, p21_μ-Q144D, p21_μ-Q144M, p21_μ-Q144S, p21_μ-FY150/151YF, p21_μ-p15, and p21_μ-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 & purification*' detailed below.

Peptides synthesised by this method: p21₁₃₉₋₁₆₀[‡], p21_μ, p21_μ-F150Y, p21_μ-Y151F, p21_μ-F150H, p21_μ-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 & purification' detailed below.

Peptides synthesised by this method: p21_μ-Q144K, p21_μ-Q144N, p21_μ-T145K, p21_μ-T145R, p21_μ-S146K, p21_μ-S146R, p21_μ-T148D, p21_μ-T148E, p21_μ-D149E, p21_μ-TS145/146KK, p21_μ-TS145/146RK, p21_μ-TS145/145KR, p21_μ-TS145/146RR, p21_μ-TD148/149EE, p21_μ-TS148/149DE, p21_μ-RNaseH2B, p21_μ-MCMT, p21_μ-pol λ, p21_μ-pol β, p21_μ-pol δ_{p66}, p21_μ-FEN1, p21_μ-RFC, p21_μ-RecQ5, p21_μ-DNALig1, p21_μ-WRN, p21_μ-XPG, p21_μ-Cdt2 p21_μ-Pogo, p21_μ-pol η, p21_μ-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 & 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 H₂O 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_μ-pol ι, p21_μ-PARG, p21_μ-RD1, p21_μ-RD2, p21_μ-RD3.

Cleavage from solid-support, isolation & 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:H₂O (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/H₂O, 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/H₂O gradient optimised for each peptide sample. RP-HPLC solvents were (A) H₂O 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.

Name	Sequence	Mw	MF	M+ Calc	[M+4] ⁺ Calc	ESI+ [M+4] ⁺ Found	Purity % (220 nm)
p21 _μ (141-155)	¹⁴¹ KRRQTSMTDFYHSKR	1940.20	C ₈₂ H ₁₃₄ N ₂₀ 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 _μ -Q144D	KRRDTSMTDFYHSKR	1927.16	C ₈₁ H ₁₃₁ N ₂₀ O ₂₄ S ₁	1925.9642	482.4991	482.5010	98.8
p21 _μ -Q144M	KRRMTSMTDFYHSKR	1943.27	C ₈₂ H ₁₃₅ N ₂₀ O ₂₂ S ₂	1941.9778	486.5024	486.5020	98.9
p21 _μ -Q144S	KRRSTSMTDFYHSKR	1899.15	C ₈₀ H ₁₃₁ N ₂₀ O ₂₃ S ₁	1897.9693	475.5003	475.4999	97.3
p21 _μ -Q144K	KRRKTSMTDFYHSKR	1940.24	C ₈₃ H ₁₃₈ N ₂₀ O ₂₂ S ₁	1939.0323	485.7661	485.7658	93.4
p21 _μ -Q144N	KRRNTSMTDFYHSKR	1926.18	C ₈₁ H ₁₃₂ N ₂₀ O ₂₃ S ₁	1924.9802	482.2531	482.2527	90.7
p21 _μ -T145K	KRRQKSMTDFYHSKR	1967.27	C ₈₄ H ₁₃₉ N ₂₁ O ₂₂ S ₁	1966.0432	492.5188	492.5206	90.2
p21 _μ -T145R	KRRQRSMTDFYHSKR	1995.28	C ₈₄ H ₁₃₉ N ₂₁ O ₂₂ S ₁	1994.0493	499.5203	499.5215	91.0
p21 _μ -S146K	KRRQTKMTDFYHSKR	1981.30	C ₈₅ H ₁₄₁ N ₂₁ O ₂₂ S ₁	1980.0588	496.0227	496.0235	87.0
p21 _μ -S146R	KRRQTRMTDFYHSKR	2009.31	C ₈₅ H ₁₄₁ N ₂₁ O ₂₂ S ₁	2008.0650	503.0243	503.0250	94.0
p21 _μ -TS145/146KK	KRRQKKSMTDFYHSKR	2008.37	C ₈₇ H ₁₄₆ N ₂₂ O ₂₁ S ₁	2007.1061	3+ : 670.0492	670.3799	91.3
p21 _μ -TS145/146RK	KRRQRKMTDFYHSKR	2036.38	C ₈₇ H ₁₄₆ N ₂₄ O ₂₁ S ₁	2035.1122	509.7861	509.7864	93.0
p21 _μ -TS145/146KR	KRRQKRMTDFYHSKR	2036.38	C ₈₇ H ₁₄₆ N ₂₄ O ₂₁ S ₁	2035.1122	509.7861	509.7859	91.9
p21 _μ -TS145/146RR	KRRQRMTDFYHSKR	2064.39	C ₈₇ H ₁₄₆ N ₂₆ O ₂₁ S ₁	2063.1184	516.7876	516.7877	90.9
p21 _μ -M147I	KRRQTSITDFYHSK	1765.98	C ₇₇ H ₁₂₄ N ₂₀ O ₂₂	1764.9383	442.2426	442.2419	99.1
p21 _μ -M147L	KRRQTSITDFYHSKR	1922.16	C ₈₃ H ₁₃₈ N ₂₀ O ₂₃	1921.0395	481.2679	481.2672	99.8
p21 _μ -M147A	KRRQTSATDFYHSKR	1880.08	C ₈₀ H ₁₃₀ N ₂₀ O ₂₃	1878.9925	470.7561	470.7556	99.2
p21 _μ -M147W	KRRQTSWTFYHSKR	1995.22	C ₈₈ H ₁₃₅ N ₂₁ O ₂₃	1994.0347	499.5167	499.5161	99.5
p21 _μ -M147V	KRRQTSVTFYHSKR	1908.14	C ₈₂ H ₁₃₄ N ₂₀ O ₂₃	1907.0238	477.7640	477.7632	98.5
p21 _μ -T148D	KRRQTSMDDFYHSKR	1954.19	C ₈₂ H ₁₃₂ N ₂₀ O ₂₄ S ₁	1952.9751	489.2518	489.2514	91.5
p21 _μ -T148E	KRRQTSMEDFYHSKR	1968.21	C ₈₃ H ₁₃₄ N ₂₀ O ₂₄ S ₁	1966.9908	492.7557	492.7562	90.4
p21 _μ -D149E	KRRQTSMEFYHSKR	1954.23	C ₈₃ H ₁₃₆ N ₂₀ O ₂₃ S ₁	1953.0115	489.2609	489.2620	88.5
p21 _μ -TD148/149EE	KRRQTSMEEFYHSKR	1982.24	C ₈₄ H ₁₃₈ N ₂₀ O ₂₄ S ₁	1981.0065	496.2596	496.2595	89.4
p21 _μ -TS148/149DE	KRRQTSMEFYHSKR	1968.21	C ₈₃ H ₁₃₄ N ₂₀ O ₂₄ S ₁	1966.9908	492.7557	492.7555	90.1
p21 _μ -F150H	KRRQTSMTDHYHSKR	1930.16	C ₇₉ H ₁₃₂ N ₂₂ O ₂₃ S ₁	1928.9864	483.2546	483.2542	94.4
p21 _μ -F150Y	KRRQTSMTDYYHSKR	1956.20	C ₈₂ H ₁₃₄ N ₂₀ O ₂₄ S ₁	1954.9908	489.7557	489.7552	93.8
p21 _μ -Y151F	KRRQTSMTDFYHSKR	1924.20	C ₈₂ H ₁₃₄ N ₂₀ O ₂₂ S ₁	1923.0010	481.7582	481.7576	94.2
p21 _μ -FY150/151YF	KRRQTSMTDFYHSKR	1940.20	C ₈₂ H ₁₃₄ N ₂₀ O ₂₃ S ₁	1938.9959	485.7570	485.7589	98.1
p21 _μ -DNALig1	KRRQRSIMSFFHSKR	1963.32	C ₈₅ H ₁₄₃ N ₂₁ O ₁₉ S ₁	1962.0959	491.5320	491.5311	98.2
p21 _μ -FEN1	KRRQGRLLDDFFHSKR	1945.20	C ₈₄ H ₁₃₇ N ₂₁ O ₂₁	1944.0667	487.0247	487.0241	99.2
p21 _μ -Pogo	KRRQKKTIDYFHSKR	1990.32	C ₈₈ H ₁₄₈ N ₂₂ O ₂₁	1989.1497	498.2954	498.2965	98.0
p21 _μ -XPG	KRRQLRISDFHSKR	1973.29	C ₈₇ H ₁₄₅ N ₂₃ O ₂₀	1972.1344	494.0416	494.0409	99.1
p21 _μ -MCMT	KRRQTTITSHFFHSKR	1882.14	C ₈₀ H ₁₃₆ N ₂₂ O ₂₁	1881.0558	471.2719	471.2714	93.9
p21 _μ -p15	KRRQKIGIEFFHSKR	1873.18	C ₈₃ H ₁₃₇ N ₂₁ O ₁₉	1872.0707	469.0257	469.0265	96.5
p21 _μ -Cdt1	KRRQRRVTDFFHSKR	2016.33	C ₈₇ H ₁₄₆ N ₂₀ O ₂₀	2015.1514	504.7959	504.7954	94.8
p21 _μ -pol δ ₉₆₆	KRRQVSITGFFHSKR	1846.15	C ₈₂ H ₁₃₆ N ₂₀ O ₁₉	1845.0598	462.2729	462.2727	99.2
p21 _μ -RecQ5	KRRQNLIRHFFHSKR	2022.37	C ₉₀ H ₁₄₈ N ₂₆ O ₁₈	2021.1772	506.3024	506.3019	99.2
p21 _μ -WRN	KRRQWKLLRDFHSKR	2053.42	C ₉₂ H ₁₅₃ N ₂₈ O ₁₉	2052.2082	514.0600	514.0593	99.3
p21 _μ -PARG	KRRDSKITDHFHSKR	1910.16	C ₈₁ H ₁₃₆ N ₂₂ O ₂₂	1909.0507	478.2707	478.2705	98.4
p21 _μ -pol β	KRRQLQKVHFFHSKR	1847.18	C ₈₁ H ₁₃₉ N ₂₃ O ₁₇	1846.1027	462.5337	462.5330	96.5
p21 _μ -pol λ	KRRSVPVLELFFHSKR	1851.21	C ₈₃ H ₁₄₃ N ₂₃ O ₁₉	1850.1115	463.5359	463.5353	92.4
p21 _μ -pol ι	KRRKGLIDYLLHFSKR	1932.29	C ₈₇ H ₁₄₆ N ₂₀ O ₂₀	1931.1330	483.7913	483.7919	87.6
p21 _μ -pol κ	KRRKHTLIDIFFHSKR	1968.32	C ₈₉ H ₁₄₈ N ₂₂ O ₁₉	1967.1442	492.7941	492.7939	93.5
p21 _μ -RFC _{p14}	KRRMDIRKFFHSKR	1904.30	C ₈₄ H ₁₄₂ N ₂₂ O ₁₇ S ₁	1903.0951	476.7818	476.7813	93.4
p21 _μ -RNaseH2B	KRRMKSIDTFFHSKR	1936.29	C ₈₅ H ₁₄₂ N ₂₀ O ₂₀ S ₁	1935.0737	484.7764	484.7761	97.7
p21 _μ -Cdt2	KRRMRKICTYFHSKR	2009.46	C ₈₇ H ₁₄₉ N ₂₁ O ₁₈ S ₂	2008.1198	503.0380	503.0390	93.8
p21 _μ -pol n	KRRMQTLESFFHSKR	1950.28	C ₈₅ H ₁₄₀ N ₂₀ O ₂₁ S ₁	1949.0530	488.2713	488.2719	96.6
p21 _μ -RD1	KRRQRKMEYYHSKR	2080.39	C ₈₈ H ₁₄₆ N ₂₄ O ₂₃ S ₁	2079.1021	520.7835	520.7836	91.0
p21 _μ -RD2	KRRQTRITEYFHSKR	2005.30	C ₈₇ H ₁₄₅ N ₂₃ O ₂₂	2004.1242	502.0389	502.0396	97.1
p21 _μ -RD3	KRRQKRITEYYHSKR	2048.39	C ₈₉ H ₁₅₀ N ₂₄ O ₂₂	2047.1664	512.7994	512.8001	92.3

SPR:

Table S2: Peptide SPR data against hPCNA. 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 fitting error; On/Off, indicate times for contact and dissociation phases of each run. All peptides are C-terminally amidated.

Name	Sequence	E_{295}^*	Top Conc (nM)	Affinity K_D (nM)	K_D SE (nM)	χ^2	On/Off (s)
p21 _μ (141-155)	¹⁴¹ KRRQTSMTDFYHSKR	67860	500	12.3	0.598	0.196	30/60
p21 (139-160)	¹³⁹ GRKRRQTSMTDFYHSKRRLIFS	98620	50	4.319	1.31	10.9	60/90
p21 _μ -Q144D	KRRDTSMTDFYHSKR	67460	2000	714	30.4	0.126	40/60
p21 _μ -Q144M	KRRMTSMTDFYHSKR	69290	5000	1544	159	1.57	40/60
p21 _μ -Q144S	KRRSTSMTDFYHSKR	67460	5000	1032	116	2.07	40/60
p21 _μ -Q144K	KRRKTSMTDFYHSKR	67460	5000	1123	148	4.19	40/60
p21 _μ -Q144N	KRRNTSMTDFYHSKR	67860	5000	772	117	3.07	40/60
p21 _μ -T145K	KRRQKSMTDFYHSKR	67860	2000	98.05	10.8	2.58	40/60
p21 _μ -T145R	KRRQRSMTDFYHSKR	69210	1000	88.51	13.2	2.97	40/60
p21 _μ -S146K	KRRQTKMTDFYHSKR	67860	300	12.97	1.55	0.520	60/90
p21 _μ -S146R	KRRQTRMTDFYHSKR	69210	100	4.297	1.35	6.26	60/90
p21 _μ -TS145/146KK	KRRQKKMTDFYHSKR	67860	300	20.18	4.80	6.70	60/90
p21 _μ -TS145/146RK	KRRQRKMTDFYHSKR	69210	300	17.30	4.66	7.94	60/90
p21 _μ -TS145/146KR	KRRQKRMTDFYHSKR	69210	300	37.00	11.0	9.63	60/90
p21 _μ -TS145/146RR	KRRQRMTDFYHSKR	70560	30	1.833	0.453	5.80	60/90
p21 _μ -M147I	KRRQTSITDFYHSK	66030	200	11.1	0.258	0.0251	40/60
p21 _μ -M147L	KRRQTSITDFYHSKR	66030	200	20.49	1.75	0.784	40/60
p21 _μ -M147A	KRRQTSATDFYHSKR	66030	20000	7591	626	1.43	40/60
p21 _μ -M147W	KRRQTSWTDYHSKR	86430	5000	3566	165	0.274	40/60
p21 _μ -M147V	KRRQTSVTDYHSKR	66030	500	29.29	1.83	0.326	40/60
p21 _μ -T148D	KRRQTSMDDFYHSKR	67860	2000	96.47	7.32	0.477	40/60
p21 _μ -T148E	KRRQTSMEDFYHSKR	67860	200	77.40	2.62	0.0533	60/90
p21 _μ -D149E	KRRQTSMTDFYHSKR	67860	500	12.67	1.4	0.387	60/90
p21 _μ -TD148/149EE	KRRQTSMEEFYHSKR	67860	1000	48.52	1.16	0.0475	60/90
p21 _μ -TD148/149DE	KRRQTSMDEFYHSKR	67860	2000	345.8	2.67	0.273	60/90
p21 _μ -F150H	KRRQTSMTDHYHSKR	64460	1000	159.0	62.8	0.119	40/60
p21 _μ -F150Y	KRRQTSMTDYYHSKR	65340	300	20.2	0.424	0.0286	40/60
p21 _μ -Y151F	KRRQTSMTDFYHSKR	70380	300	10.61	1.48	0.721	60/90
p21 _μ -FY150/151YF	KRRQTSMTDYFHSKR	67860	30	2.175	0.531	3.01	60/90
p21 _μ -DNALig1	KRRQRSIMSFYHSKR	71730	250	40.74	9.20	3.42	60/90
p21 _μ -FEN1	KRRQGRLLDFYHSKR	69900	2000	165.6	2.07	1.67	40/60
p21 _μ -Pogo	KRRQKKITDYFHSKR	66030	100	8.816	3.30	6.73	60/90
p21 _μ -XPG	KRRQLRIDSFYHSKR	69900	200	16.39	4.81	11.2	60/90
p21 _μ -MCMT	KRRQTTITSHFYHSKR	65150	2000	106.0	24.4	5.32	40/60
p21 _μ -p15	KRRQKGIGEFYHSKR	68550	2000	234.0	42.4	5.21	40/60
p21 _μ -Cdt1	KRRQRRTDFYHSKR	71250	150	8.760	2.22	10.1	40/60
p21 _μ -pol δ ₂₈₆	KRRQVSITGFYHSKR	68550	1000	123.7	4.24	3.21	40/60
p21 _μ -RecQ5	KRRQNLIRHFYHSKR	75500	20000	>40000**			
p21 _μ -WRN	KRRQWKLLRDFYHSKR	81700	2000	1155	67.3	0.952	40/60
p21 _μ -PARG	KRRDSKITDFYHSKR	64750	2000	90.03	18.4	2.38	40/60
p21 _μ -pol β	KRRQLQKVHFYHSKR	62770	20000	5732	822	18.6	40/60
p21 _μ -pol λ	KRRSVPVLELFYHSKR	59550	20000	8144	766	4.81	40/60
p21 _μ -pol ι	KRRKGLIDYYLHYSKR	63110	1500	111.2	5.41	0.0710	60/90
p21 _μ -pol κ	KRRKHTLIDFFYHSKR	73350	10000	3296	353	6.62	40/60
p21 _μ -RFC _{p14}	KRRMDIRKFFYHSKR	68550	1000	145.2	34.5	6.42	40/60
p21 _μ -Cdt2	KRRMRKICTFYHSKR	69500		NS***			
p21 _μ -pol η	KRRMQTLESFFYHSKR	70380	10000	2537	745	10.2	40/60
p21 _μ -RNAseH2B	KRRMKSIDTFYHSKR	69980	2000	640.4	51.1	0.851	40/60
p21 _μ -RD1	KRRQRKMEDYYHYSKR	66690	100	8.152	1.27	1.51	40/60
p21 _μ -RD2	KRRQTRITEFYHYSKR	67380	50	1.118	0.330	2.40	60/90
p21 _μ -RD3	KRRQKRITEFYHYSKR	64860	200	6.70	1.58	3.96	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 determined.
 *** Peptide bound significantly to the reference cell of sensor chip such that a K_D could not be determined

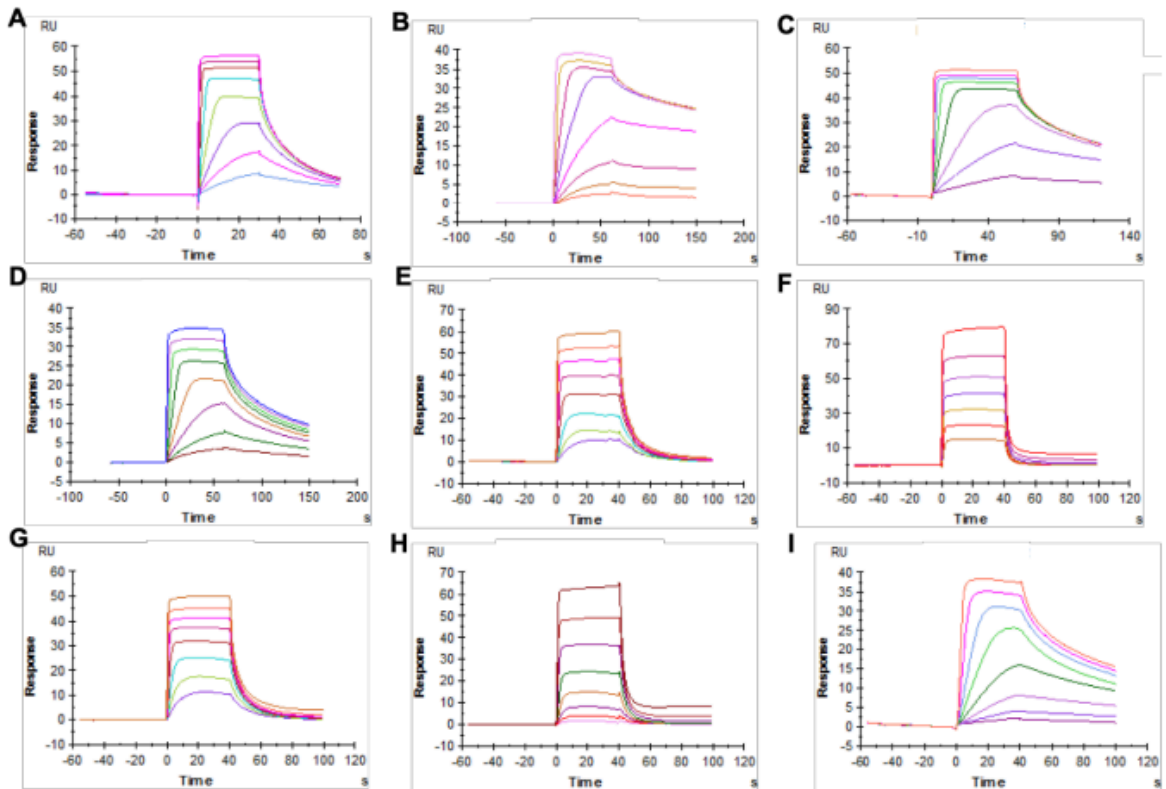


Figure S1: Representative sample of SPR sensorgrams. **A** p21 μ **B** p21 μ -S146R **C** p21 μ -FY150/151YF **D** p21 μ -Y151F **E** p21 μ -FEN1 **F** p21 μ -pol η **G** p21 μ -PARG **H** p21 μ -RNaseH2B **I** p21 μ -RD1

COCRYSTAL EXPERIMENTS

Table S3: Data collection and refinement statistics of hPCNA bound with p21 μ (RCSB PDB ID: 7KQ1), and hPCNA bound with p21 μ -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

$$^a R_{\text{merge}} = \sum |I - \langle I \rangle| / \sum I$$

$$^b R_{\text{pim}} = \sum_h [1 / (I^h - 1)]^{1/2} \sum_l \langle I^h \rangle - I_{h,l} / \sum_h \sum_l I_{h,l} \quad (2)$$

$$^c R_{\text{work}} = \sum |F_o - F_c| / \sum |F_o| \text{ for all data excluding data used to calculate } R_{\text{free}}$$

$$^d R_{\text{free}} = \sum |F_o - F_c| / \sum |F_o| \text{ for all data.}$$

Table S4: Secondary Interaction Summary for co-crystal structure of p21_u with hPCNA (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](#)

	Residue	Intermolecular		Intramolecular		Total
		VDW	H-Bond	VDW	H-Bond	
FI	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	
FI	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 (FI)		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_u-F150Y with hPCNA (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](#)

	Residue	Intermolecular		Intramolecular		Total
		VDW	H-Bond	VDW	H-Bond	
FI	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	
FI	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 (FI)		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:

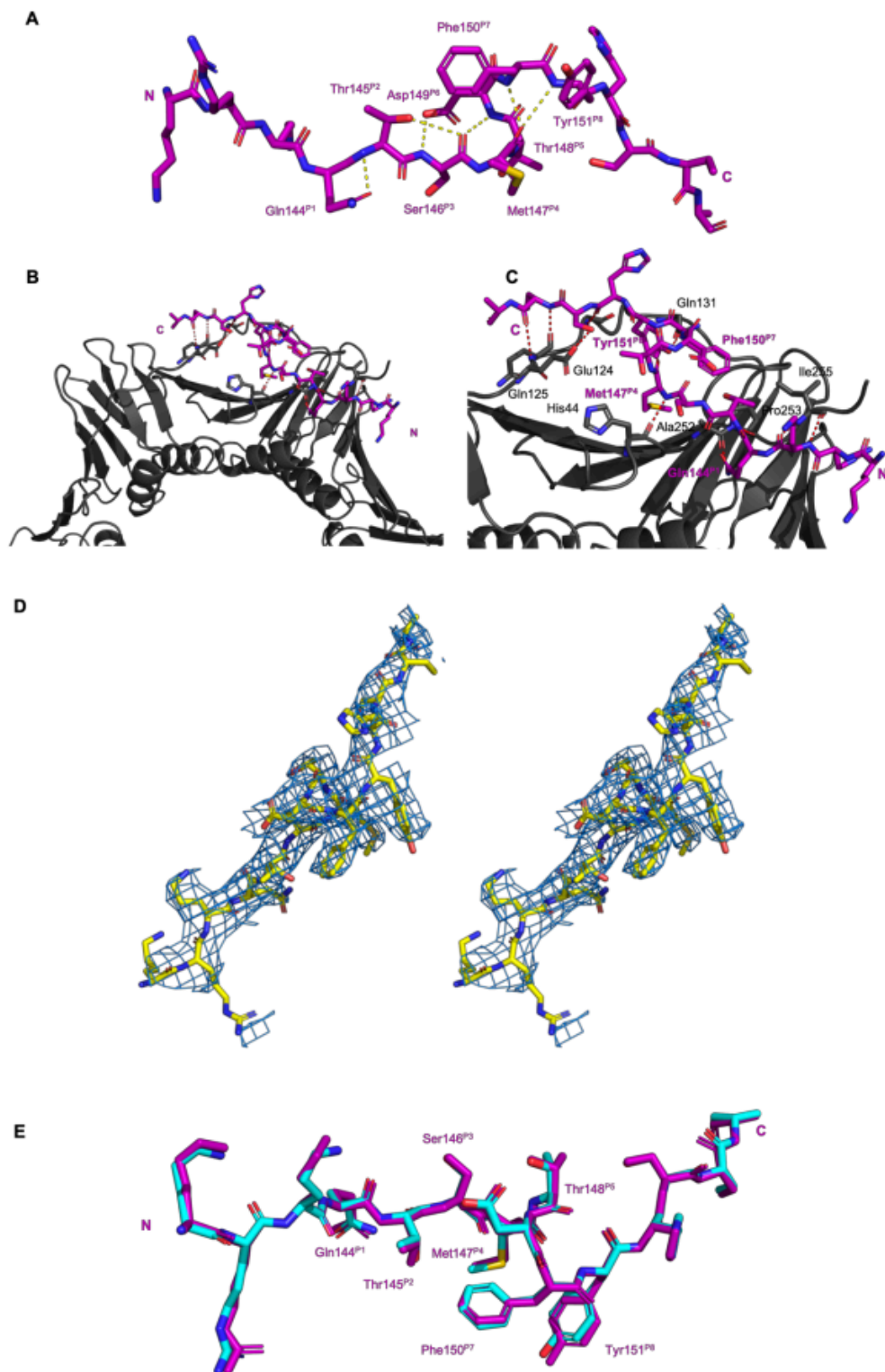


Figure S2: Co-crystal structure of p21_μ (purple, sticks) bound to PCNA monomer (grey, cartoon). Heteroatoms indicated as blue, nitrogen; red, oxygen; yellow, sulfur. **A** Intramolecular interactions shown as yellow dashes, and PIP-box residues labelled in purple. **B & C** Intermolecular interactions shown as red dashes, PCNA residues labelled in grey/white and conserved PIP-box residues labelled in purple. **D** Representative electron density of p21_μ (yellow, sticks) shown as a wall-eye stereo image 2Fo-Fc composite omit map, view contoured at 1.5σ. **E** Overlay of p21_μ (purple, sticks) bound to PCNA (not shown) co-crystal structure, with p21_μ (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.

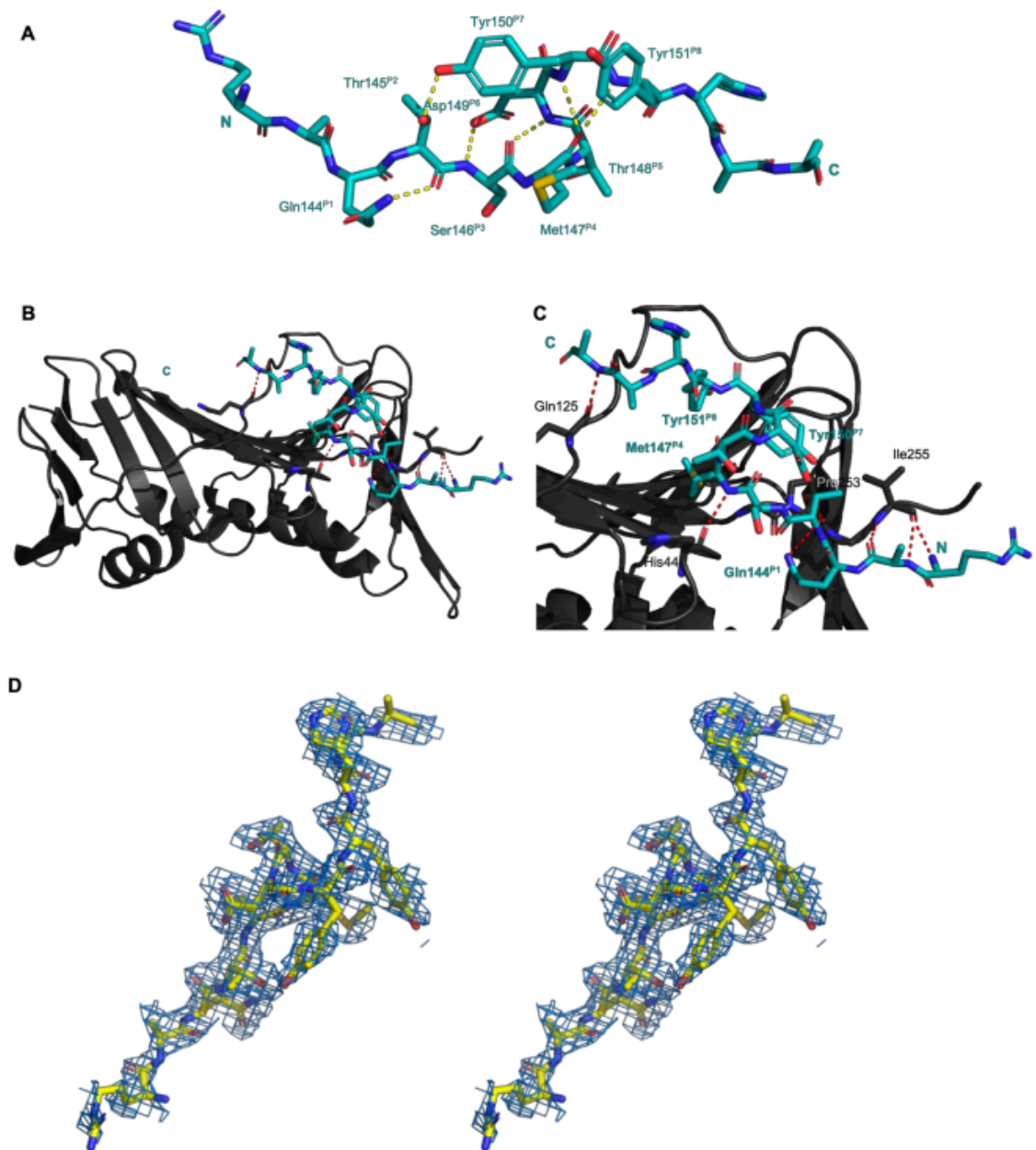


Figure S3: Co-crystal structure of p21_{F150Y} (blue, sticks) bound to PCNA monomer (grey, cartoon). Heteroatoms indicated as blue, nitrogen; red, oxygen; yellow, sulfur **A** Intramolecular interactions shown as yellow dashes, and PIP-box residues labelled in blue. **B & C** Intermolecular interactions shown as red dashes, PCNA residues labelled in grey/white and conserved PIP-box residues labelled in blue. **D** Representative electron density of p21_{F150Y} (yellow, sticks) shown as a wall-eye stereo image 2Fo-Fc composite omit map, view contoured at 1.5 σ .

COMPUTATIONAL MODELLING OF PCNA MONOMERS BOUND TO P21 μ PEPTIDES

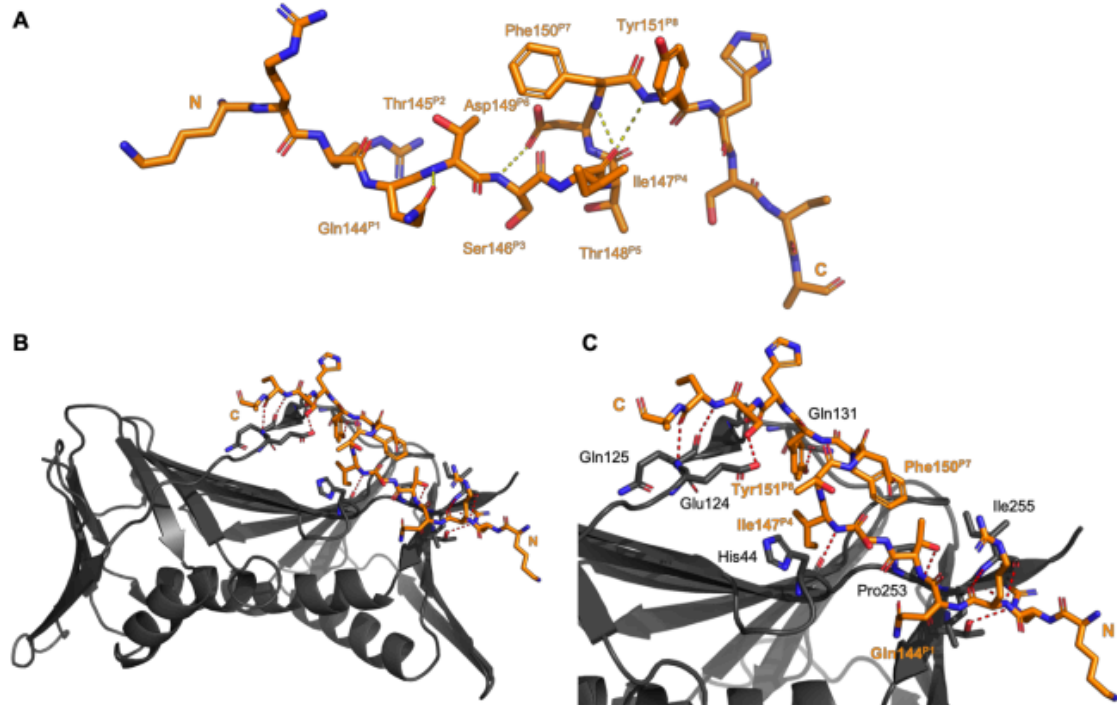


Figure S4: Computationally modelled structure of p21 μ -M1471 (orange, sticks) on the PIP-box binding site of a hPCNA monomer (grey, cartoon). Heteroatoms indicated as blue, nitrogen; red, oxygen; yellow, sulfur. **A** Intramolecular interactions shown as yellow dashes, and PIP-box residues labelled 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 structure of p21 μ -M1471 with hPCNA calculated using the RING server. Chain B interactions only. RING Session ID: [5f3b21280e9f94078ea22cfe](https://ring.scripps.edu/session/5f3b21280e9f94078ea22cfe)

	Residue	Intermolecular		Intramolecular		Total
		VDW	H-Bond	VDW	H-Bond	
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	
	152	0	1	0	0	
FI	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 B-Tyr151 and A-Tyr133

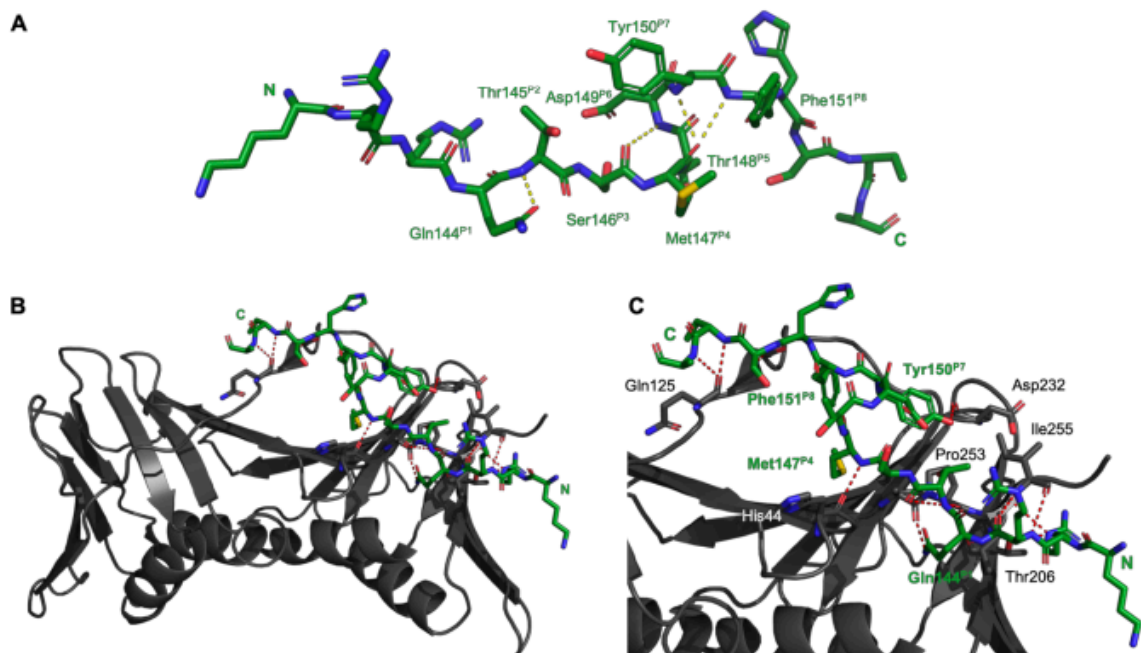


Figure S5: Computationally modelled structure of p21_μ-FY150/151YF (dark green, sticks) on the PIP-box binding site of a hPCNA monomer (grey, cartoon). Heteroatoms indicated as blue, nitrogen; red, oxygen; yellow, sulfur. **A** Intramolecular interactions shown as yellow dashes, and PIP-box residues 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 structure of p21_μ-FY150/151YF with hPCNA calculated using the RING server. Chain B interactions only. RING session ID: [5f3b22bc0e9f94078ea22d03](https://ring.rdg.ac.uk/session/5f3b22bc0e9f94078ea22d03)

Residue	Intermolecular		Intramolecular		Total
	VDW	H-Bond	VDW	H-Bond	
FI	141	0	0	0	
	142	4	2	0	
	143	1	1	0	
PIP-box	* 144	2	0	0	
	145	1	1	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
FI	152	0	0	0	
	153	1	0	0	
	154	2	1	0	
	155	0	1	0	
Total	24	8	1	3	36
PIP-box	16	3	1	3	23
Flanking (FI)	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

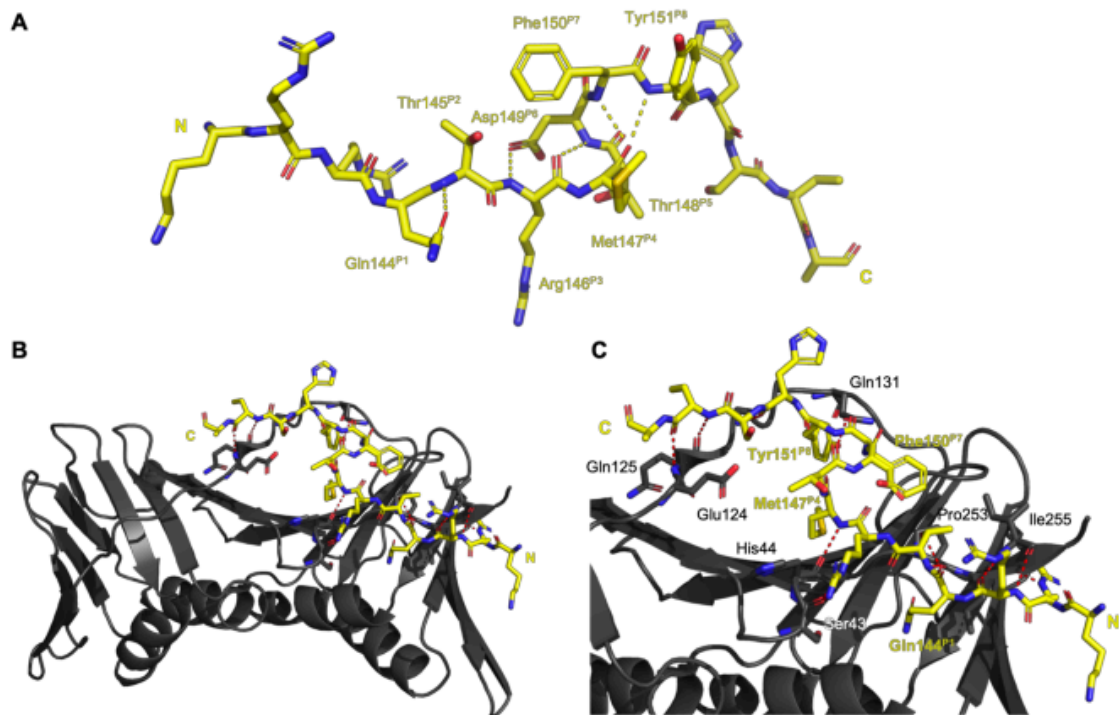


Figure S6: Computationally modelled structure of p21_μ-S146R (yellow, sticks) on the PIP-box binding site of a hPCNA monomer (grey, cartoon). Heteroatoms indicated as blue, nitrogen; red, oxygen; yellow, sulfur. **A** Intramolecular interactions shown as yellow dashes, and PIP-box residues labelled in 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 structure of p21_μ-S146R with hPCNA calculated using the RING server. Chain B interactions only. RING Session ID: [5f3b1f6b0e9f94078ea22cf9](https://ring.scripps.edu/5f3b1f6b0e9f94078ea22cf9)

	Residue	Intermolecular		Intramolecular		Total
		VDW	H-Bond	VDW	H-Bond	
FI	141	0	0	0	0	
	142	3	1	0	0	
	143	2	2	0	0	
	* 144	1	0	0	0	
PIP-box	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

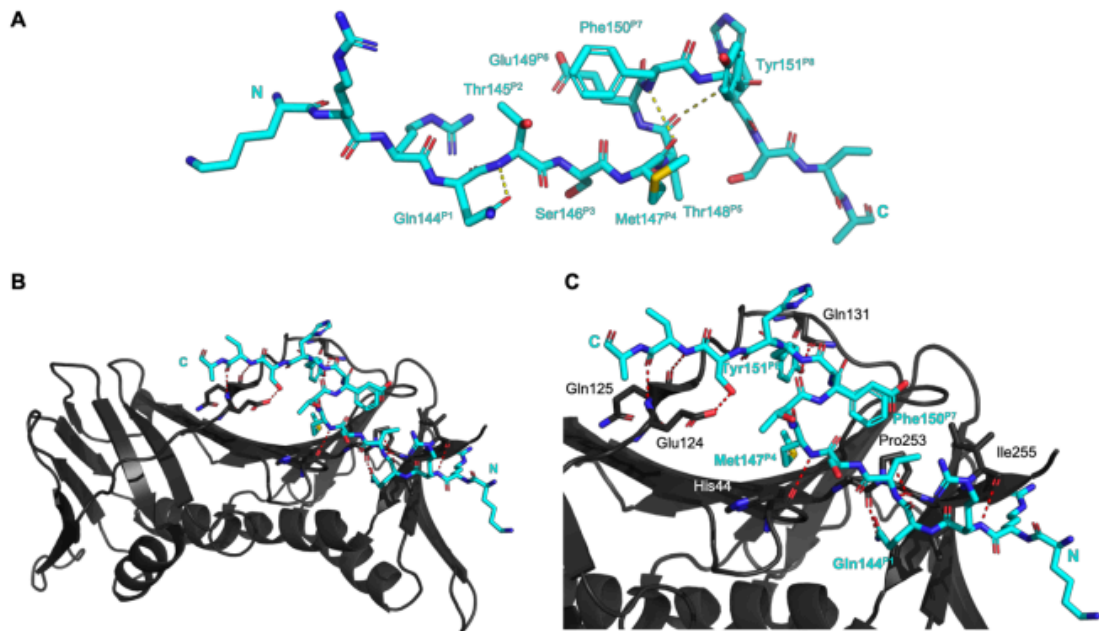


Figure S7: Computationally modelled structure of p21_L-D149E (light blue, sticks) on the PIP-box binding site of a hPCNA monomer (grey, cartoon). Heteroatoms indicated as blue, nitrogen; red, oxygen; yellow, sulfur. **A** Intramolecular interactions shown as yellow dashes, and PIP-box residues labelled 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 structure of p21_L-D149E with hPCNA calculated using the RING server. Chain B interactions only. RING Session ID: [5f3b21fe0e9f94078ea22d00](https://ring.rutgers.edu/session/5f3b21fe0e9f94078ea22d00)

	Residue	Intermolecular		Intramolecular		Total
		VDW	H-Bond	VDW	H-Bond	
FI	141	0	0	0	0	
	142	2	2	0	0	
	143	0	2	0	0	
	* 144	2	0	0	0	
PIP-box	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	
	152	1	0	0	0	
FI	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 (FI)	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

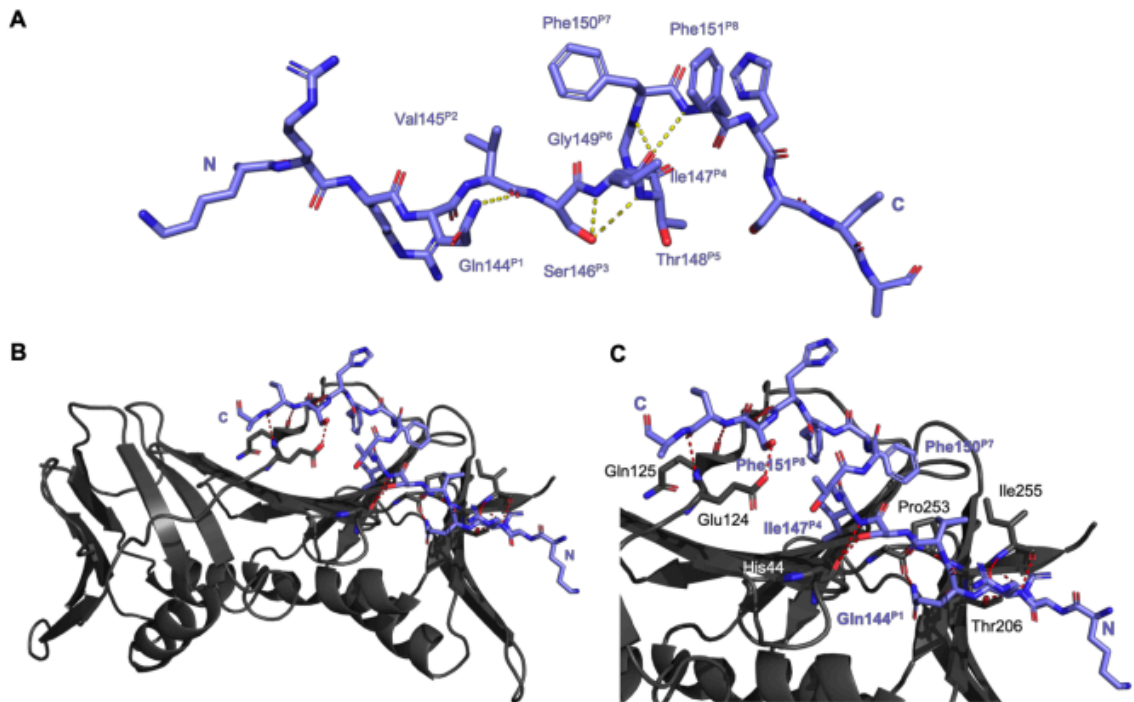


Figure S8: Computationally modelled structure of p21_μ-pol δ₆₆ (light purple, sticks) on the PIP-box binding site of a hPCNA monomer (grey, cartoon). Heteroatoms indicated as blue, nitrogen; red, oxygen; yellow, sulfur. **A** Intramolecular interactions shown as yellow dashes, and PIP-box residues labelled 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 structure of p21_μ-pol δ₆₆ with hPCNA calculated using the RING server. Chain B interactions only. RING Session ID: [5f3b202a0e9f94078ea22cfb](#)

Residue	Intermolecular		Intramolecular		Total
	VDW	H-Bond	VDW	H-Bond	
FI	141	0	0	0	
	142	3	2	0	
	143	1	2	0	
PIP-box	* 144	2	0	0	
	145	1	1	0	
	146	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	
	153	2	0	0	
	154	0	1	0	
	155	1	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

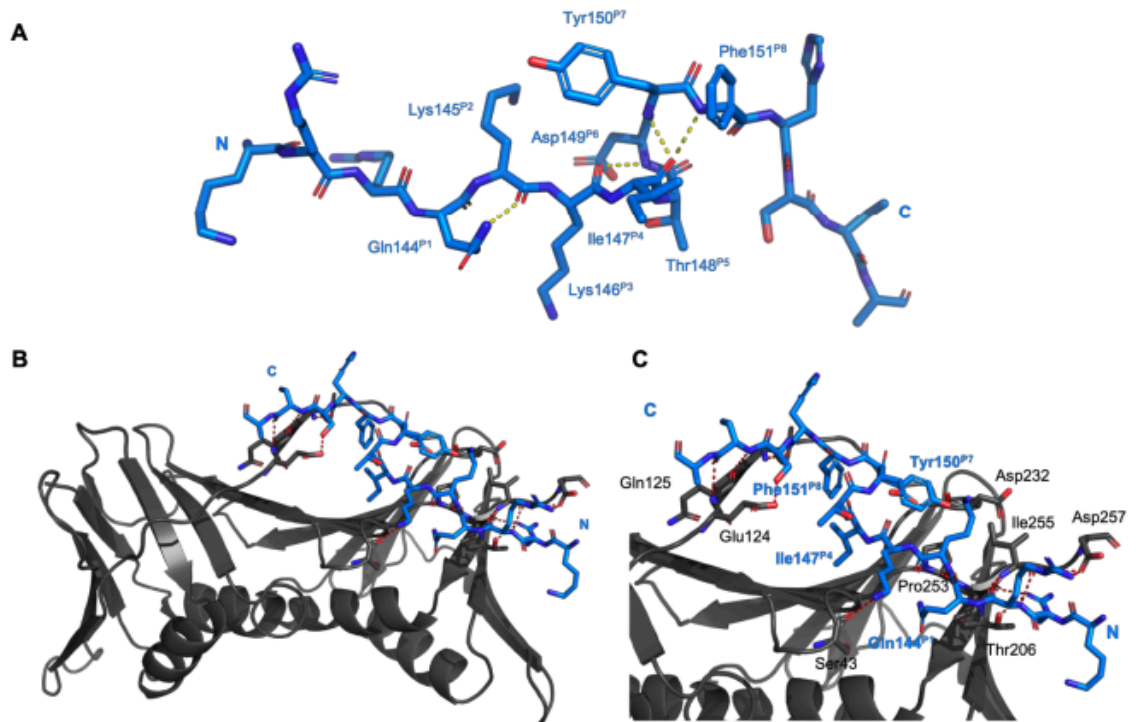


Figure S9: Computationally modelled structure of p21_Y-Pogo (dark blue, sticks) on the PIP-box binding site of a hPCNA monomer (grey, cartoon). Heteroatoms indicated as blue, nitrogen; red, oxygen; yellow, sulfur. **A** Intramolecular interactions shown as yellow dashes, and PIP-box residues labelled 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 structure of p21_Y-Pogo with hPCNA calculated using the RING server. Chain B interactions only. RING Session ID: [5f3b224d0e9f94078ea22d01](https://ring.scripps.edu/5f3b224d0e9f94078ea22d01)

	Residue	Intermolecular		Intramolecular		Total
		VDW	H-Bond	VDW	H-Bond	
FI	141	0	0	0	0	
	142	3	2	0	0	
	143	1	2	0	0	
	* 144	2	0	0	0	
PIP-box	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	
	152	0	1	0	0	
FI	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

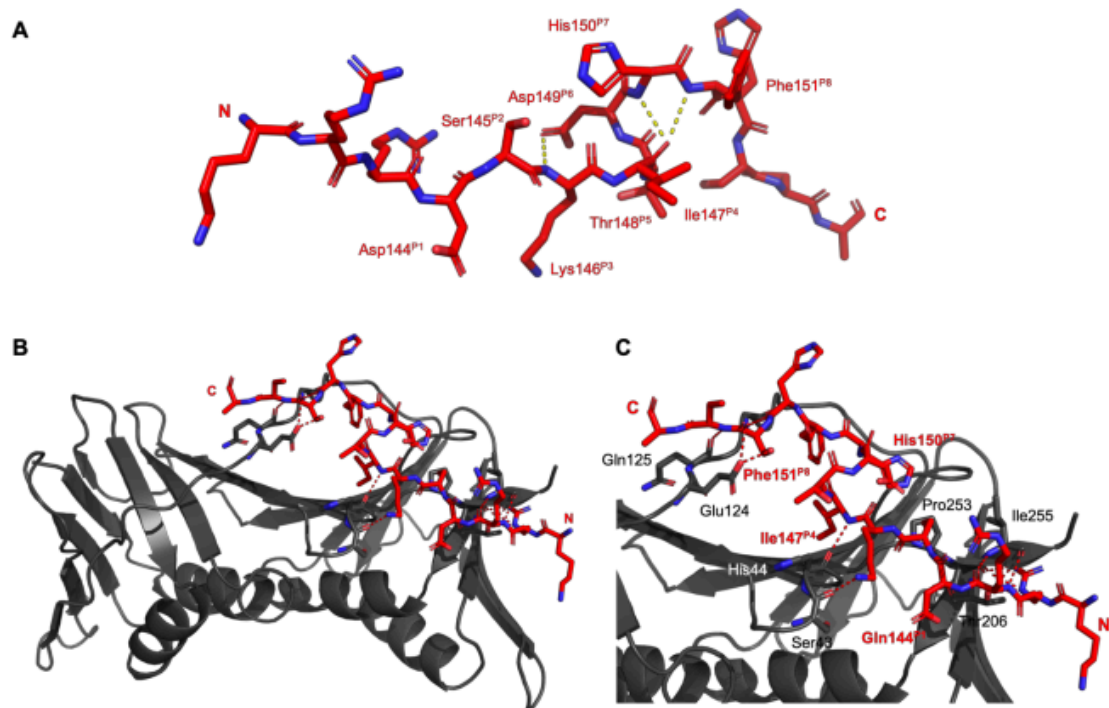


Figure S10: Computationally modelled structure of p21^Y-PARG (red, sticks) on the PIP-box binding site of a hPCNA monomer (grey, cartoon). Heteroatoms indicated as blue, nitrogen; red, oxygen; yellow, sulfur. **A** Intramolecular interactions shown as yellow dashes, and PIP-box residues labelled in red. **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^Y-PARG with hPCNA calculated using the RING server. Chain B interactions only. RING Session ID: [5f3b1de50e9f94078ea22cf7](https://ring-server.org/session/5f3b1de50e9f94078ea22cf7)

Residue	Intermolecular		Intramolecular		Total
	VDW	H-Bond	VDW	H-Bond	
FI	141	0	0	0	
	142	2	2	0	
	143	0	2	0	
PIP-box	* 144	1	0	0	
	145	1	0	0	
	146	1	0	0	
	* 147	2	1	0	
	148	0	0	0	
	149	0	0	0	
	* 150	3	0	0	
	* 151	3	0	0	
FI	152	0	1	0	
	153	1	0	0	
	154	0	1	0	
	155	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

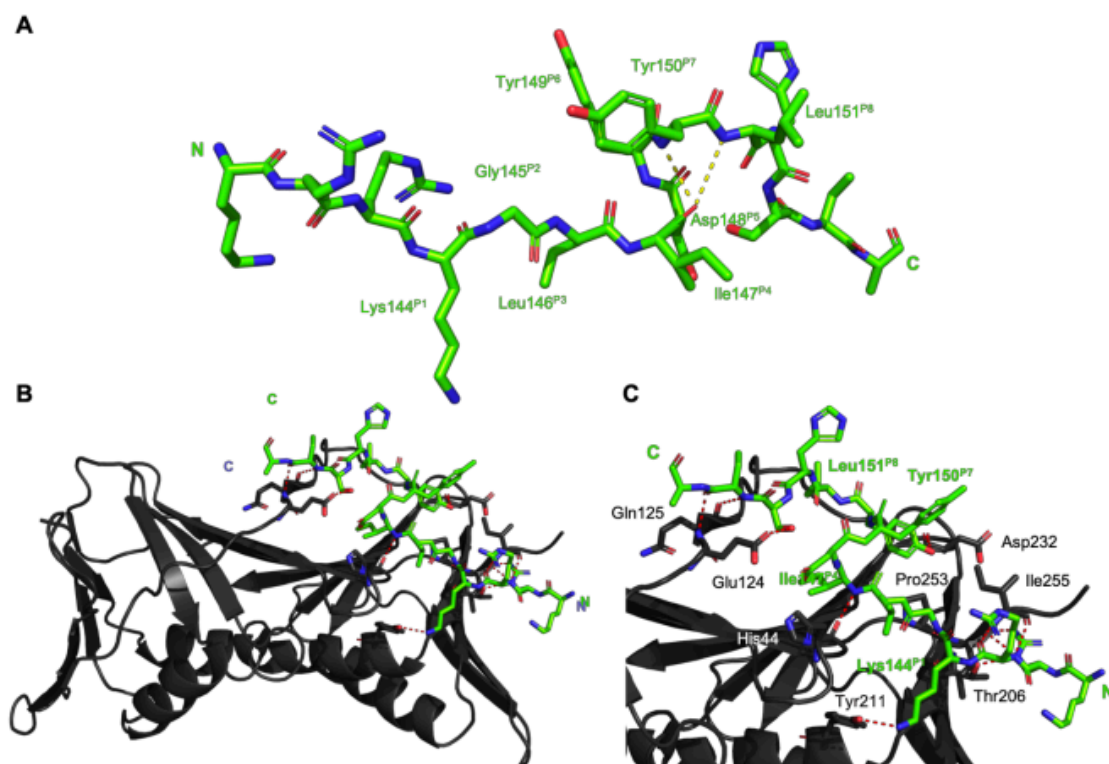


Figure S11: Computationally modelled structure of p21_{ras}-pol I (green, sticks) on the PIP-box binding site of a hPCNA monomer (grey, cartoon). Heteroatoms indicated as blue, nitrogen; red, oxygen; yellow, sulfur. **A** Intramolecular interactions shown as yellow dashes, and PIP-box residues labelled 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 structure of p21_{ras}-pol I with hPCNA calculated using the RING server. Chain B interactions only. RING Session ID: [5f3b1e580e9f94078ea22cf8](https://ring.rutgers.edu/5f3b1e580e9f94078ea22cf8)

Residue	Intermolecular		Intramolecular		Total
	VDW	H-Bond	VDW	H-Bond	
FI	141	0	0	0	
	142	2	2	0	
	143	0	2	0	
PIP-box	* 144	1	0	0	
	145	1	1	0	
	146	2	0	1	
	* 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
	152	0	1	0	0
FI	153	2	0	0	
	154	0	1	0	
	155	0	0	0	
	Total	15	9	1	1
PIP-box	11	3	1	1	16
Flanking (FI)	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:					

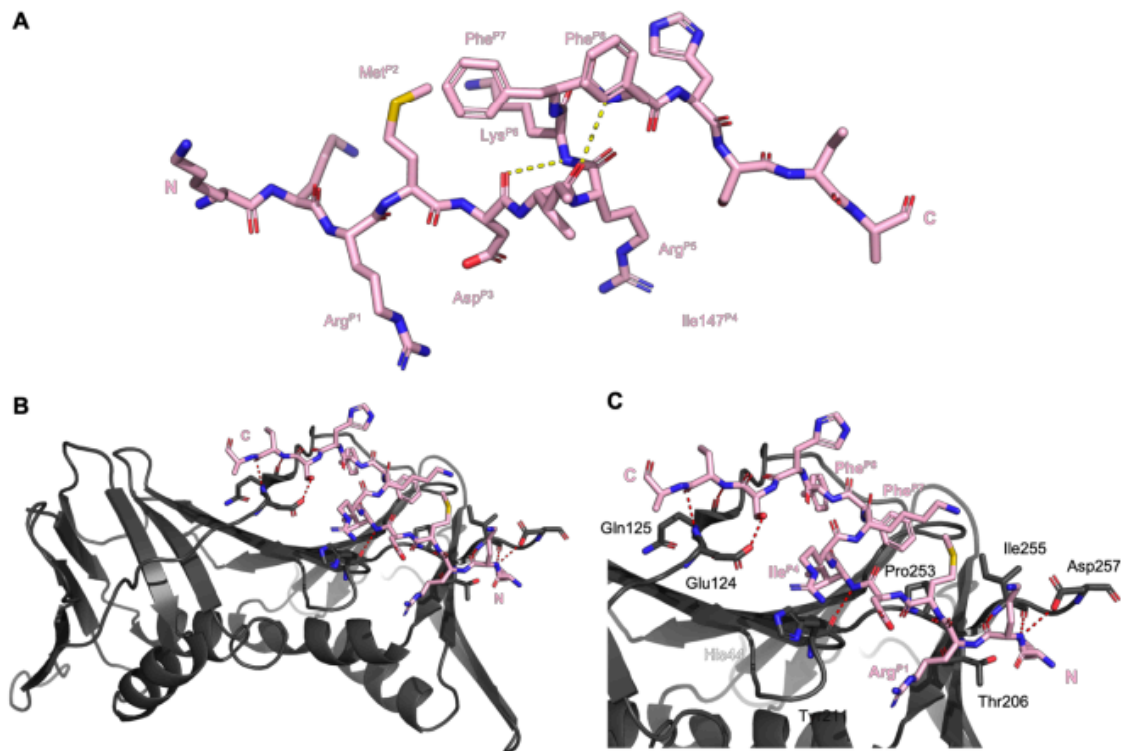


Figure S12: Computationally modelled structure of p21_μ-RFC (light pink, sticks) on the PIP-box binding site of a hPCNA monomer (grey, cartoon). Heteroatoms indicated as blue, nitrogen; red, oxygen; yellow, sulfur. **A** Intramolecular interactions shown as yellow dashes, and PIP-box residues labelled 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 structure of p21_μ-RFC with hPCNA calculated using the RING server. Chain B interactions only. RING Session ID: [5f3b20890e9f94078ea22cf](https://ring.scripps.edu/session/5f3b20890e9f94078ea22cf)

Residue	Intermolecular			Intramolecular		Total
	VDW	H-Bond		VDW	H-Bond	
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	0	1	4	27
PIP-box	11	2	0	1	4	18
Flanking (FI)	6	3	0	0	0	9
Conserved (*) PIP-box residues	6	1	0	0	2	9
Non-conserved PIP-box residues	5	1	0	1	2	9

Other: Intermolecular ionic interaction between B-Arg143 and A-Asp257; Intermolecular pi-stack between B-Tyr151 and A-Tyr250

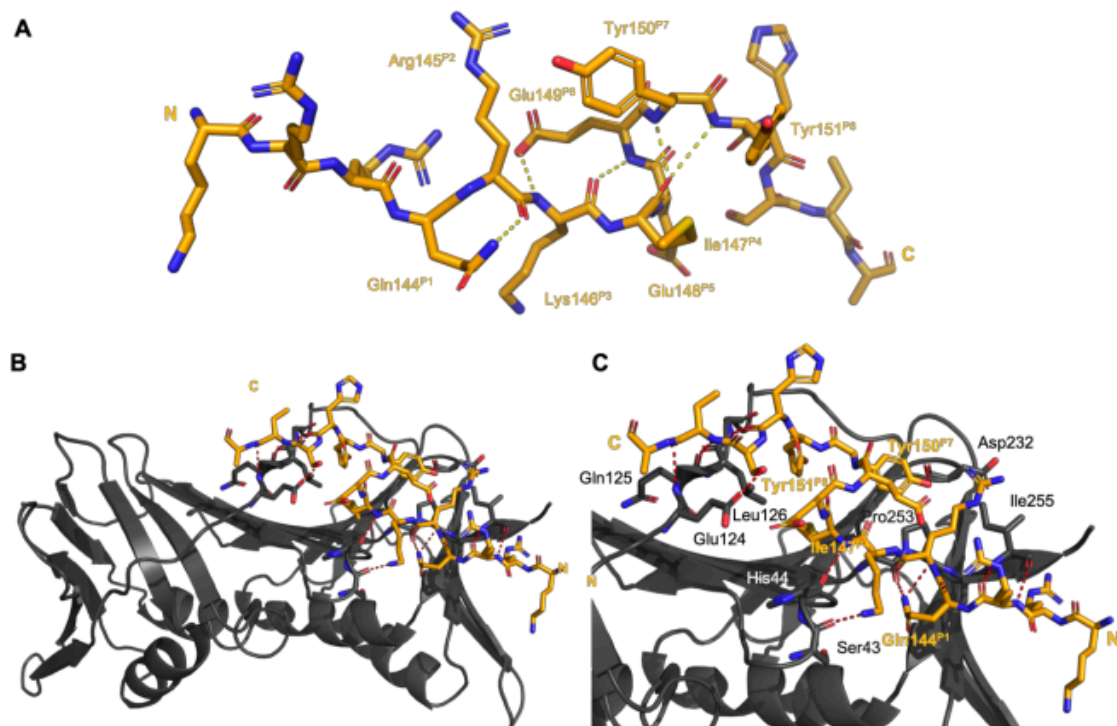


Figure S13: Computationally modelled structure of p21_μ-RD1 (light orange, sticks) on the PIP-box binding site of a hPCNA monomer (grey, cartoon). Heteroatoms indicated as blue, nitrogen; red, oxygen; yellow, sulfur. **A** Intramolecular interactions shown as yellow dashes, and PIP-box residues labelled 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 structure of p21_μ-RD1 with hPCNA calculated using the RING server. Chain B interactions only. RING Session ID: [5f3b21c10e9f94078ea22cf](#)

Residue	Intermolecular		Intramolecular		Total
	VDW	H-Bond	VDW	H-Bond	
FI	141	0	0	0	
	142	3	2	0	
	143	1	1	0	
	* 144	2	0	0	
PIP-box	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
	152	0	1	0	0
FI	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 (FI)	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

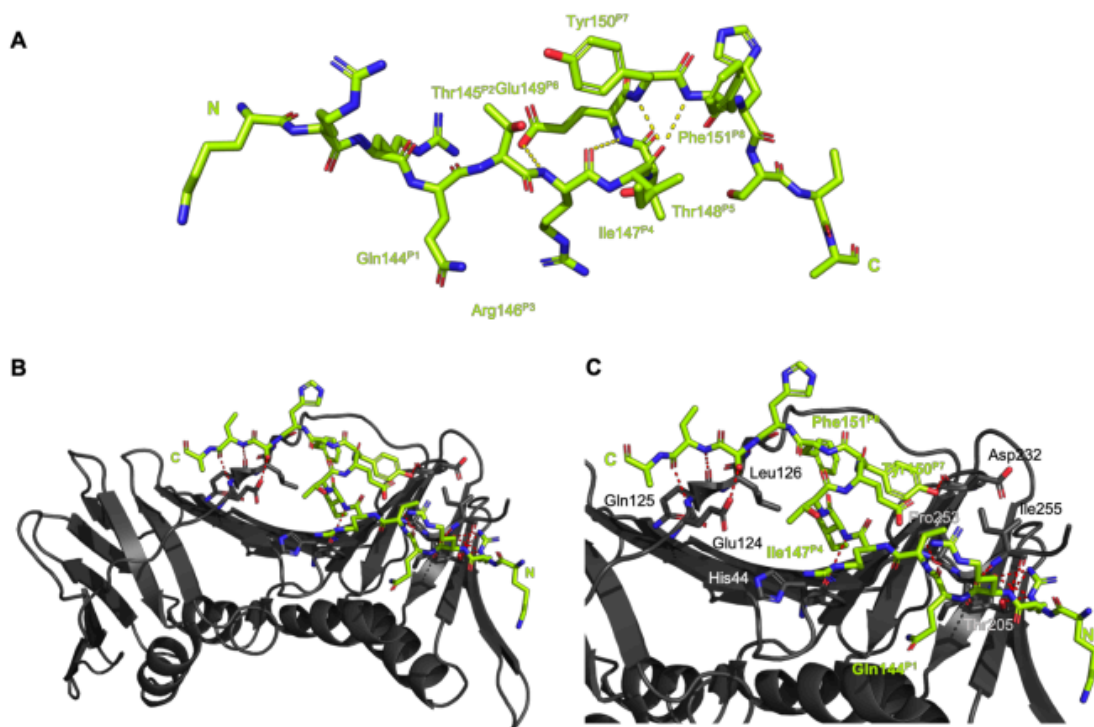


Figure S14: Computationally modelled structure of p21_μ-RD2 (green, sticks) on the PIP-box binding site of a hPCNA monomer (grey, cartoon). Heteroatoms indicated as blue, nitrogen; red, oxygen; yellow, sulfur. **A** Intramolecular interactions shown as yellow dashes, and PIP-box residues labelled in 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 structure of p21_μ-RD2 with hPCNA calculated using the RING server. Chain B interactions only. RING Session ID: [5f3b1ff30e9f94078ea22cfa](https://ring.rutgers.edu/session/5f3b1ff30e9f94078ea22cfa)

	Residue	Intermolecular		Intramolecular		Total
		VDW	H-Bond	VDW	H-Bond	
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:					

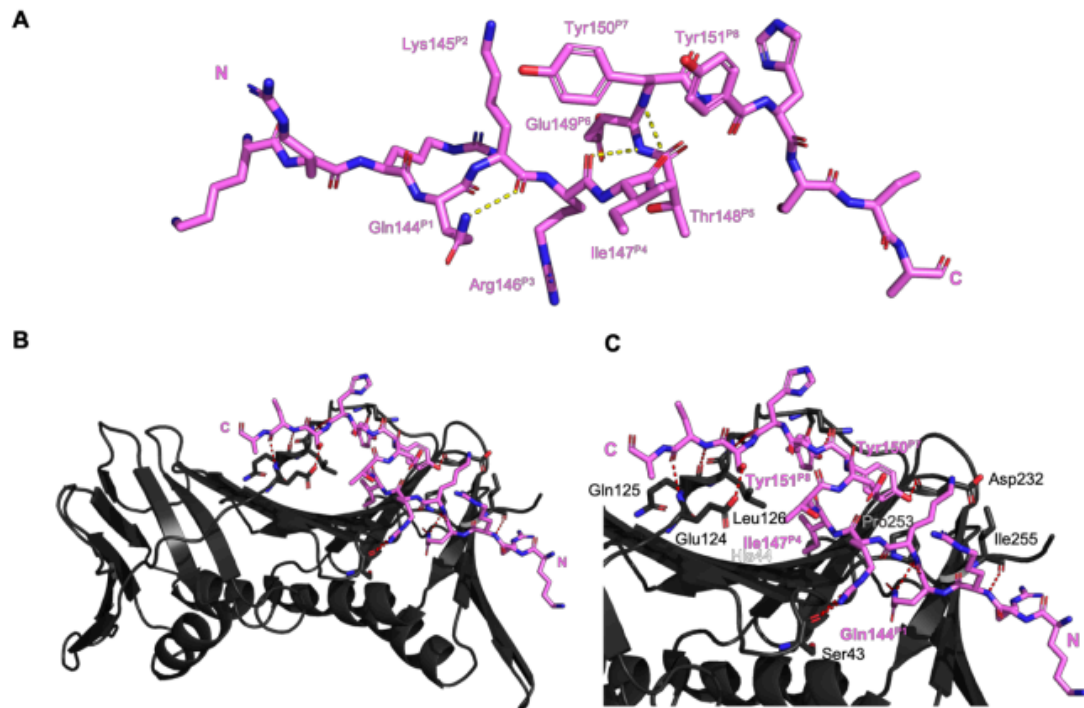


Figure S15: Computationally modelled structure of p21_μ-RD3 (pink, sticks) on the PIP-box binding site of a hPCNA monomer (grey, cartoon). Heteroatoms indicated as blue, nitrogen; red, oxygen; yellow, sulfur. **A** Intramolecular interactions shown as yellow dashes, and PIP-box residues labelled in pink. **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 structure of p21_μ-RD3 with hPCNA calculated using the RING server. Chain B interactions only. RING Session ID: [5f3b22e70e9f94078ea22d04](#)

Residue	Intermolecular		Intramolecular		Total
	VDW	H-Bond	VDW	H-Bond	
FI					
141	0	0	0	0	
142	2	1	0	0	
143	2	2	0	0	
* 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	
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 (FI)	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:					

ANALYSIS OF STRUCTURES

Table S18: Root-mean-squared deviation (RMSD) values of mutant peptides docked to the monomer of hPCNA compared to structures of p21_μ bound to hPCNA and p21 bound to hPCNA.

Name	Structure Type	RMSD value against Wild Type p21 ₁₃₉₋₁₆₉ (monomer of 1AXC)	RMSD value against monomer of p21 _μ bound to hPCNA (PDB ID: 7KQ1)
p21 _μ	Co-crystal – 7KQ1	0.511	-
p21 _μ	Computational model	0.316	0.233
p21 _μ -F150Y	Co-crystal – 7KQ0	0.342	0.451
p21 _μ -S146R	Computational model	0.570	0.176
p21 _μ -M147I	Computational model	0.571	0.191
p21 _μ -D149E	Computational model	0.556	0.197
p21 _μ -FY150/151YF	Computational model	0.559	0.183
p21 _μ -PARG	Computational model	0.571	0.192
p21 _μ -Pogo	Computational model	0.554	0.271
p21 _μ -pol δ ₉₆₈	Computational model	0.559	0.186
p21 _μ -pol ι	Computational model	0.563	0.199
p21 _μ -RFC	Computational model	0.567	0.209
p21 _μ -RD1	Computational model	0.600	0.270
p21 _μ -RD2	Computational model	0.558	0.181
p21 _μ -RD3	Computational model	0.552	0.193

Table S19: Buried Surface Area (BSA) for PIP-box residues from the cocrystal 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, Gulbis 1996)	107.86 / 70	48.31 / 70	30.39 / 60	135.10 / 100	47.60 / 60	0 / 0	59.31 / 50	132.20 / 90
p21 _μ (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 _μ -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 _μ -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 _μ -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 _μ -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 _μ -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 _μ -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 _μ -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 _μ -pol δ ₉₆₈	98.14 / 60	50.67 / 50	26.13 / 40	123.79 / 100	57.64 / 60	0 / 0	65.75 / 50	123.84 / 90
p21 _μ -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 _μ -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 _μ -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 _μ -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 _μ -RD3	88.02 / 80	22.38 / 20	63.49 / 50	137.11 / 100	60.31 / 80	0 / 0	75.71 / 50	123.64 / 90

COMPARISON OF STRUCTURES TO NATIVE SEQUENCES

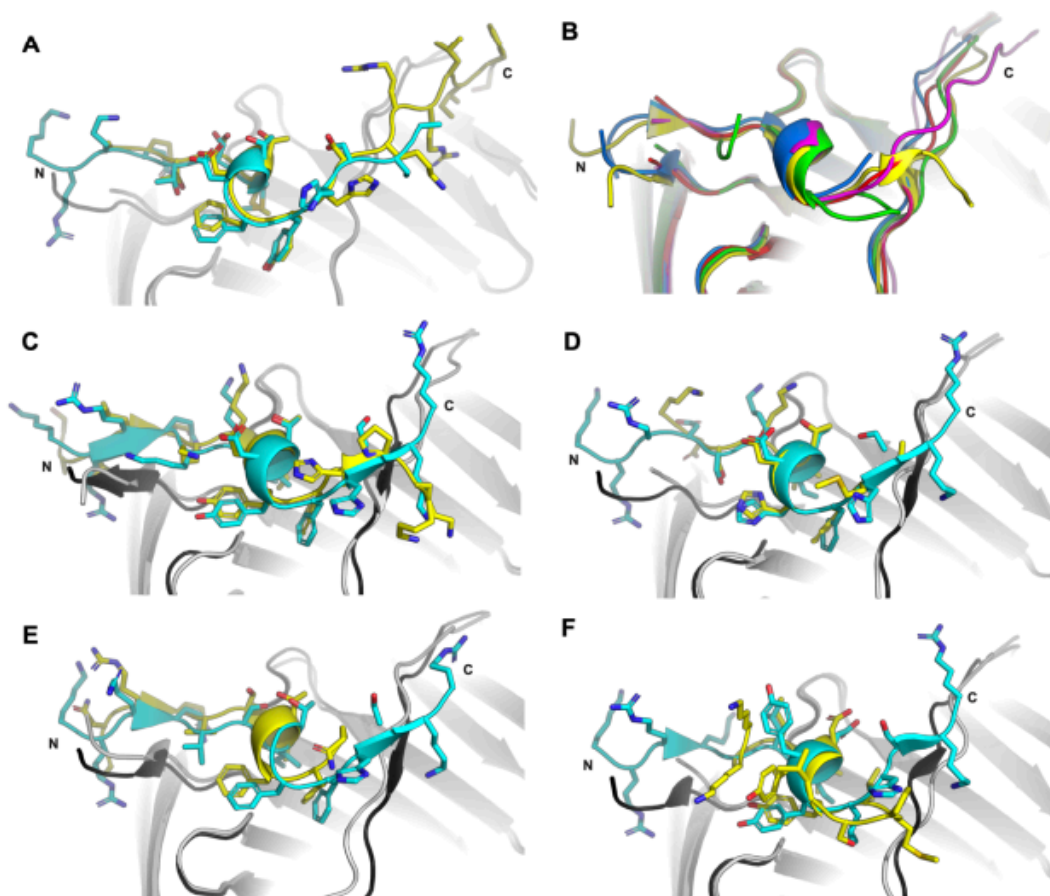


Figure S16: Overlaid structures of p21 μ :PIP-box hybrid (PCNA, white; peptide, blue) and native PIP-box peptide (PCNA, black; peptide, yellow) shown in cartoon representation and side-chains of peptide as sticks. **A** p21 μ (7KQ1) and p21¹³⁹⁻¹⁶⁰ (1AXC) **B** Overlay of Native PIP-box peptides (p21¹³⁹⁻¹⁶⁰ 1AXC, pink; PL 1VYJ, yellow; pol $\delta_{p66}^{452-466}$ 1U76, blue; PARG⁴⁰²⁻⁴²⁰ 5MAV, red; pol I ⁴¹⁵⁻⁴³⁷ 2ZVM, green **C** p21 μ -Pogo and PL (1VYJ) **D** p21 μ -PARG and PARG⁴⁰²⁻⁴²⁰ (5MAV) **E** p21 μ -pol δ and pol $\delta_{p66}^{452-466}$ (1U76) **F** p21 μ -pol I and pol I ⁴¹⁵⁻⁴³⁷ (2ZVM).

Table S20: Root-mean-squared deviation (RMSD) values of alternative PIP-box modified peptides docked to monomer of hPCNA compared to the structures of their respective native peptides bound to hPCNA

Alternative PIP box mutant peptide docked to hPCNA	Structure of native peptide bound to hPCNA	RMSD value
p21 μ -PARG	5MAV	0.496
p21 μ -Pogo	1VYJ	0.646
p21 μ -pol δ_{p66}	1U76	0.696
p21 μ -pol I	2ZVM	1.136

Table S21: Number of interactions for p21 μ :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.

Peptide	PDB ID:	Ref.	Affinity	# res	# Interactions					RING (3) session ID
					Total	PIP-box	Conserved	Non-conserved	Flanking	
p21 ¹³⁹⁻¹⁶⁰	1AXC	(4)	5.96 nM	22	60.67	25	9.33	15.67	35.67	5ef515200e9f94078ea226cb
p21 μ	7KQ1		26.1 nM	15	30.67	23.67	16.67	7	7	5ef50f2b0e9f94078ea226bd
PL ¹⁻¹⁸ [mutant]	1VYJ	(5)	100 nM	16	22.33	17.67	13.33	4.33	4.67	5ef515730e9f94078ea226cc
p21 μ -Pogo			9.12 nM	15	34	24	16	5	10	5f3b224d0e9f94078ea22d01
pol $\delta_{p66}^{452-466}$	1U76	(6)	15.6 μ M	15	30.33	24.33	17	7.33	6	5ef516b50e9f94078ea226cf
p21 μ -pol δ_{p66}			268 nM	15	29	16	11	3	13	5f3b202a0e9f94078ea22cfb
PARG ⁴⁰²⁻⁴²⁰	5MAV	(7)	3.3 μ M	19	26	19	4.83	14.17	7	5ef5164c0e9f94078ea226ce
p21 μ -PARG			401 nM	15	24	15	8	4	9	5f3b1de50e9f94078ea22cf7
pol I ⁴¹⁵⁻⁴³⁷	2ZVM	(8)	0.39 μ M	23	22.33	17.67	17.67	4.67	4.67	5ef515e00e9f94078ea226cd
p21 μ -pol I			1.42 μ M	15	26	16	9	7	10	5f3b1e580e9f94078ea22cf8

REFERENCES

1. Anthis, N. J., and Clore, G. M. (2013) Sequence-specific determination of protein and peptide concentrations by absorbance at 205 nm. *Protein Sci.* **22**, 851-858
2. Wlodawer, A., Minor, W., Dauter, Z., and Jaskolski, M. (2013) Protein crystallography for aspiring crystallographers or how to avoid pitfalls and traps in macromolecular structure determination. *FEBS J.* **280**, 5705-5736
3. Piovesan, D., Minervini, G., and Tosatto, S. C. (2016) The RING 2.0 web server for high quality residue interaction networks. *Nucleic Acids Res.* **44**, W367-374
4. Gulbis, J. M., Kelman, Z., Hurwitz, J., O'Donnell, M., and Kuriyan, J. (1996) Structure of the C-Terminal Region of p21 WAF1/CIP1 Complexed with Human PCNA. *Cell* **87**, 297-306
5. Kontopidis, G., Wu, S.-Y., Zheleva, D. I., Taylor, P., McInnes, C., Lane, D. P., Fischer, P. M., and Walkinshaw, M. D. (2005) 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.* **102**, 1871-1876
6. Bruning, J. B., and Shamooy, Y. (2004) Structural and thermodynamic analysis of human PCNA with peptides derived from DNA polymerase-delta p66 subunit and flap endonuclease-1. *Structure* **12**, 2209-2219
7. Kaufmann, T., Grishkovskaya, I., Polyansky, A. A., Kostrhon, S., Kukolj, E., Olek, K. M., Herbert, S., Beltzung, E., Mechtler, K., Peterbauer, T., Gotzmann, J., Zhang, L., Hartl, M., Zagrovic, B., Elsayad, K., Djinovic-Carugo, K., and Slade, D. (2017) A novel non-canonical PIP-box mediates PARG interaction with PCNA. *Nucleic Acids Res.* **45**, 9741-9759
8. Hishiki, A., Hashimoto, H., Hanafusa, T., Kamei, K., Ohashi, E., Shimizu, T., Ohmori, H., and Sato, M. (2009) Structural basis for novel interactions between human translesion synthesis polymerases and proliferating cell nuclear antigen. *J. Biol. Chem.* **284**, 10552-10560

Chapter 4: Publication

“A cell permeable bimanane-constrained PCNA-interacting peptide.”

Statement of Authorship

Title of Paper	A cell permeable fluorescent peptidomimetic targeting human PCNA.
Publication Status	<input checked="" type="checkbox"/> Published <input type="checkbox"/> Accepted for Publication <input type="checkbox"/> Submitted for Publication <input type="checkbox"/> 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)	Bethiney Vandborg		
Contribution to the Paper	Protein Synthesis and purification, protein crystallography, computational modelling studies, discussed results, edited manuscript.		
Overall percentage (%)			
Certification:	I certify that the above information is true and correct and was obtained during the period of my Higher Degree by me and is not the result of obligations or contractual agreements with a third party. I am the primary author of this paper.		
Signature	<table border="1"> <tr> <td>Date</td> <td>29/08/2023</td> </tr> </table>	Date	29/08/2023
Date	29/08/2023		

Co-Author Contributions

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

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

Name of Co-Author	Aimee J Horsfall.		
Contribution to the Paper	Designed and synthesised peptides, NMR assignment and analysis, analysed and discussed results, wrote and edited manuscript.		
Signature	<table border="1"> <tr> <td>Date</td> <td>29/08/2023</td> </tr> </table>	Date	29/08/2023
Date	29/08/2023		

Name of Co-Author	Zoya Kikhtyak		
Contribution to the Paper	Cell uptake assays, discussed results, and edited manuscript		
Signature	<table border="1"> <tr> <td>Date</td> <td>29/08/23</td> </tr> </table>	Date	29/08/23
Date	29/08/23		

Please cut and paste additional co-author p

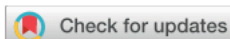
Name of Co-Author	Denis B Scanlon	
Contribution to the Paper	Peptide synthesis, discussed results and edited manuscript.	
Signature	Supervisor JBB	Date 17/9/23

Name of Co-Author	Theresa E H	
Contribution to the Paper	Supervised cell uptake assays, discussed results, edited manuscript.	
Signature		Date 19/09/2023

Name of Co-Author	Wayne D Tilley	
Contribution to the Paper	Supervised cell uptake assays, discussed results, edited manuscript.	
Signature		Date 19/09/2023

Name of Co-Author	John B Bruning	
Contribution to the Paper	AV, discussed results and edited manuscript.	
Signature		Date 17/9/23

Name of Co-Author	Andrew D. Abell	
Contribution to the Paper	Supervised AJH and discussed results and edited manuscript.	
Signature		Date 26/09/2023

Cite this: *RSC Chem. Biol.*, 2021, 2, 1499

A cell permeable bimeane-constrained PCNA-interacting peptide†

Aimee J. Horsfall,^{ib abc} Beth A. Vandborg,^{ad} Zoya Kikhtyak,^{ib e} Denis B. Scanlon,^{ab} Wayne D. Tilley,^e Theresa E. Hickey,^e John B. Bruning^{*ad} and Andrew D. Abell^{ib *abc}

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 145 and 149 to access peptidomimetics that target PCNA. The resulting macrocycles bind PCNA with K_D values ranging from 570 nM to 3.86 μ M, with the bimeane-constrained peptide **7** proving the most potent. Subsequent X-ray crystallography and computational modelling studies of the macrocyclic 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 bimeane linker. The inherent fluorescence of the bimeane 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 bimeane 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.

Received 19th May 2021,
Accepted 19th July 2021

DOI: 10.1039/d1cb00113b

rsc.li/rsc-chembio

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 (Fig. 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

^a Institute of Photonics and Advanced Sensing (IPAS), The University of Adelaide, Adelaide, South Australia, 5005, Australia. E-mail: john.bruning@adelaide.edu.au, andrew.abell@adelaide.edu.au

^b School of Physical Sciences, The University of Adelaide, Adelaide, South Australia 5005, Australia

^c Australian Research Council Centre of Excellence for Nanoscale BioPhotonics (CNBP), Australia

^d School of Biological Sciences, The University of Adelaide, Adelaide, South Australia 5005, Australia

^e Dame Roma Mitchell Cancer Research Laboratories, Adelaide Medical School, Faculty of Health and Medical Sciences, University of Adelaide, Adelaide, South Australia 5005, Australia

† Electronic supplementary information (ESI) available. See DOI: 10.1039/d1cb00113b



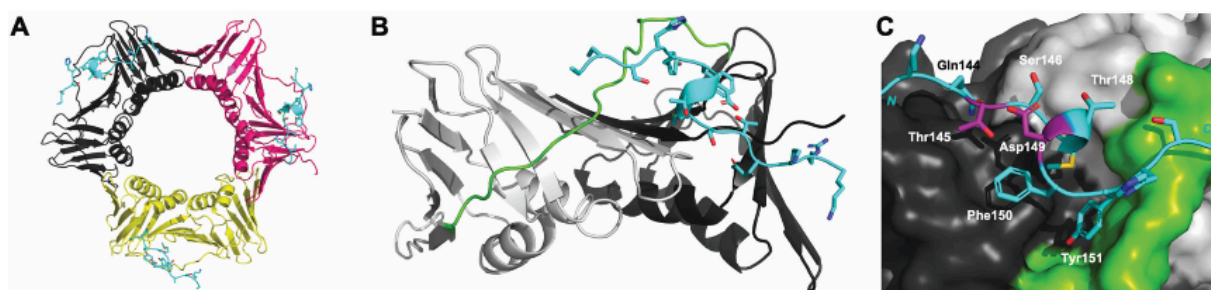


Fig. 1 p21₁₄₁₋₁₅₅ (cyan) bound to PCNA (PDB: 7KQ1). (A) Ring-shaped PCNA with three peptides (cyan) bound to the PIP-box binding site. PCNA monomers (three) shown in shades of grey, pink and yellow. (B) Single PCNA subunit with two domains shown in shades of grey and inter-domain connecting loop (IDCL) in green. p21₁₄₁₋₁₅₅ shown in cyan. (C) PCNA shown as surface representation (grey) with IDCL highlighted (green). p21₁₄₁₋₁₅₅ 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.

with a bifunctional reagent.¹⁵⁻¹⁸ To this end, we previously reported the first macrocyclic peptidomimetic of p21₁₃₉₋₁₅₂, 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 3₁₀-helix in solution, where the 3₁₀-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 3₁₀-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 (p21₁₄₃₋₁₅₄, 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 *bis*-alkylation 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 bimanane-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.

Results/discussion

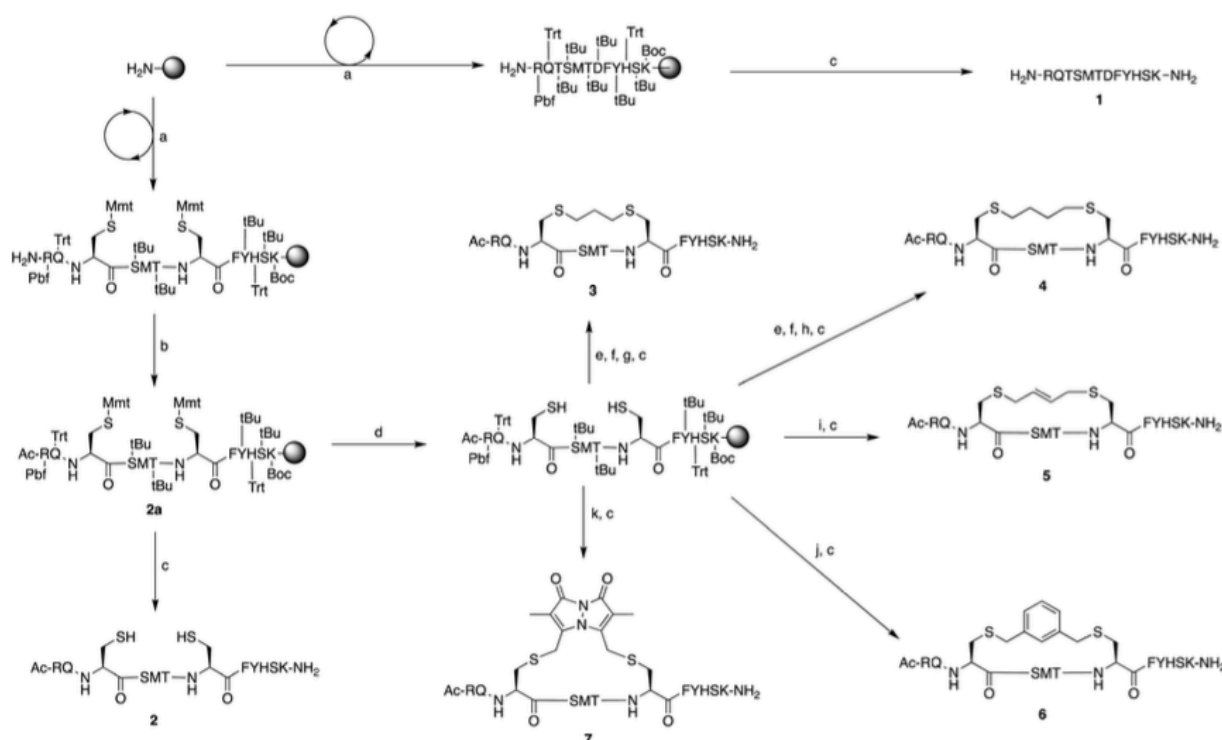
Design & synthesis

The peptide p21₁₃₉₋₁₆₀ is reported to bind PCNA with high specificity and provides an ideal template to develop PCNA inhibitors. It is known that the p21₁₃₉₋₁₆₀ sequence can be shortened without drastically impacting affinity for PCNA^{13,19} and consequently a short variant, p21₁₄₃₋₁₅₄ (**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 p21₁₄₃₋₁₅₄ (Fig. 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 (Fig. 1C), and were successfully linked in a p21 lactam bridge peptide in our earlier study to stabilise the 3₁₀-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 3₁₀-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 bimanane linker²³ 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/*t*Bu solid-phase peptide synthesis (SPPS) protocol detailed in the Experimental section (see ESI†) 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 ESI†).^{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 alkyl iodides *in situ*,²¹ whereas reaction with *trans*-1,4-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





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, 10 min. (b) Acetylation: Ac₂O (50 equiv.), DIPEA (50 equiv.), DMF, 15 min. (c) Cleavage: 92.5 : 2.5 : 2.5 : 2.5 TFA/TIPS/DODT/H₂O, 2 h. (d) Mmt deprotection: 2% TFA in DCM, 1 min × 40. (e) NaI (17.5 equiv.), TCEP (0.5 equiv.), DMF, N₂, 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.

(2.5% of each TIPS, DODT & H₂O) 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 ESI†).

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

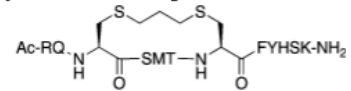
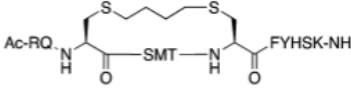
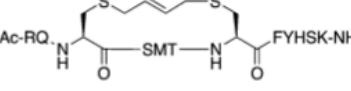
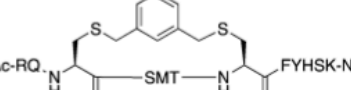
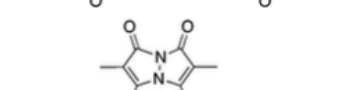
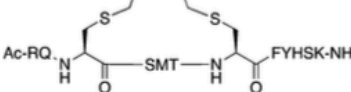
peptide 3, with its propyl linker, gave the second highest affinity for PCNA at 769 nM, whereas the larger butyl linker of peptide 4 (8 atom linker *cf.* 7 atoms of 3) resulted in reduced affinity for PCNA with a K_D value of 1.99 μM. The more rigid *trans*-butenyl-based linker in peptide 5 resulted 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.

Structural analysis

Next, the binding conformation of macrocycles 3–7 bound to PCNA was investigated by X-ray crystallography and



Table 1 SPR binding affinity K_D values. SE – standard error. χ^2 – fitting error. NS represents a non-specific interaction with the sensor chip, consequently the binding affinity for PCNA could not be determined. Ac – Acetyl; β A – Beta-alanine; FITC – fluorescein isothiocyanate

Peptide	Sequence	Linker (atom length)	Affinity $K_D \pm SE$ (μ M)	χ^2
1	H-RQTSMDFYHSK-NH ₂	—	0.102 \pm 0.005	0.070
2	Ac-RQCSMTCFYHSK-NH ₂	—	NS	—
3		Propyl (7)	0.769 \pm 0.078	0.250
4		Butyl (8)	1.99 \pm 0.140	0.160
5		<i>trans</i> -Butenyl (8)	2.82 \pm 0.080	0.106
6		<i>m</i> -Xylene (9)	3.86 \pm 0.035	0.527
7		Bimane (9)	0.570 \pm 0.030	0.118
8	FITC- β A-RQTSMDFYHSK-NH ₂	—	—	—
9		Bimane (9)	25.2 \pm 1.90	0.055

Open Access Article. Published on 21 July 2021. Downloaded on 8/29/2023 12:58:19 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.



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 Fig. 2 and 3, as well as Fig. S2, S4 and S5 (ESI[†]). Data collection and refinement statistics are summarised in Table S2 (ESI[†]). 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 Fig. S1, ESI[†]), 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 (Fig. S2F and G, ESI[†]). Macrocycles 4 and 7 were then modelled onto the PCNA surface and energy

minimised in the same way and are shown in Fig. 2, 3 and Fig. S3, S6 (ESI[†]).

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 Fig. 2A–D). The structures of macrocyclic peptides 3–7 bound to PCNA indicate a percentage buried surface area (%BSA) of 60% for the PIP-box residues collectively (Table S3, ESI[†]), 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 Fig. 2C–H. Interestingly, the main-chain hydrogen bond between residue 146 (*i*) and 149 (*i* + 3), that defines the classic 3₁₀-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 (Fig. S2A–D, ESI[†])

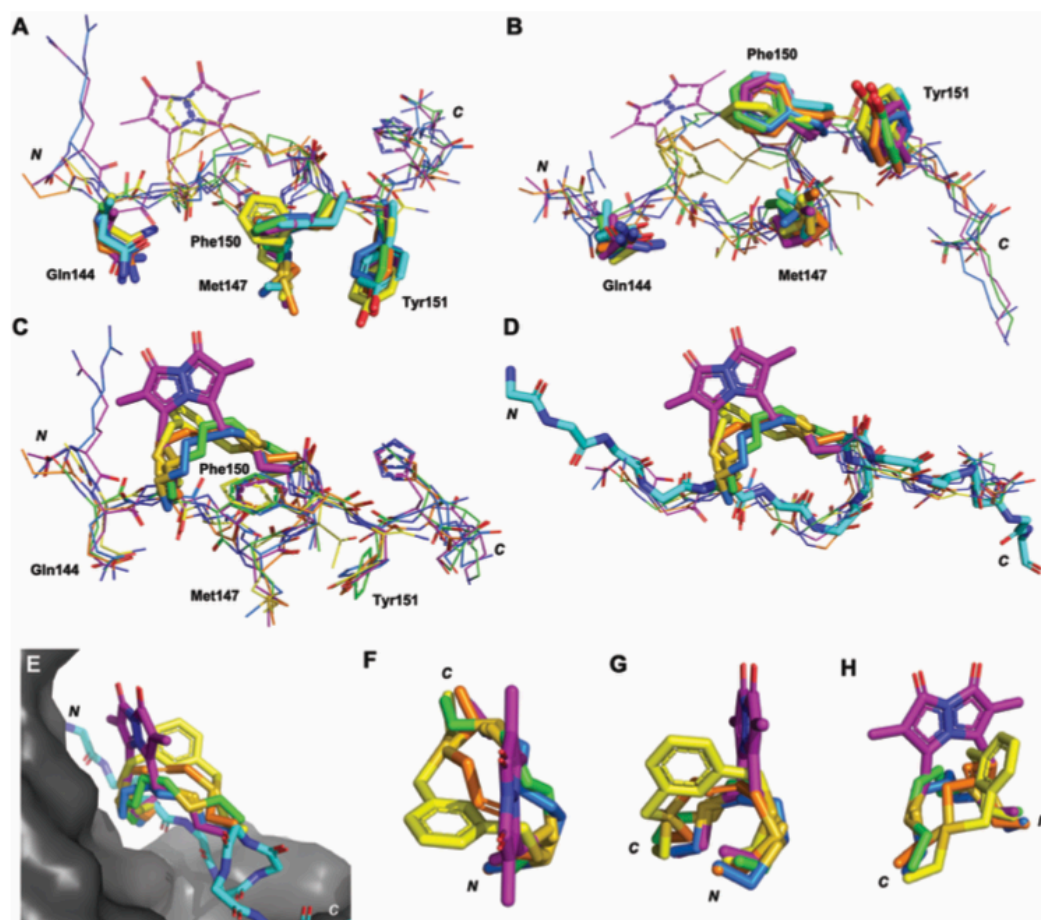


Fig. 2 Differences in the linker position, and impact of the linker on backbone conformation. Peptides p21_{141–155} (light blue, 7KQ1), **3** (green, co-crystal 7M5L), **4** (blue, computationally modelled), **5** (orange, co-crystal 7M5M), **6** (yellow, co-crystal 7M5N) and **7** (purple, computationally modelled). (A and 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–H) Conformation of the peptide linkers relative to one another.

and a 2.9 Å intramolecular hydrogen bond to the Cys145* amide (Fig. 3B, green), where * indicates the modified cysteine residue. 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_{141–155} (Fig. 2), however Tyr151 does not make a hydrogen bond with Gln131 as in p21_{141–155}. 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 (*cf.* 3₁₀-helix from 146–149 for p21_{141–155}). 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 (Fig. 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_{141–155} (Fig. 2); however, the %BSA for Met147 is notably lower at 70% than in p21_{141–155}, or the other macrocycles (100%, Table S3, ESI†). Three intermolecular hydrogen bonds are observed between peptide **5** and PCNA to anchor the macrocycle onto the surface (Fig. S4, ESI†). 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 (Fig. 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



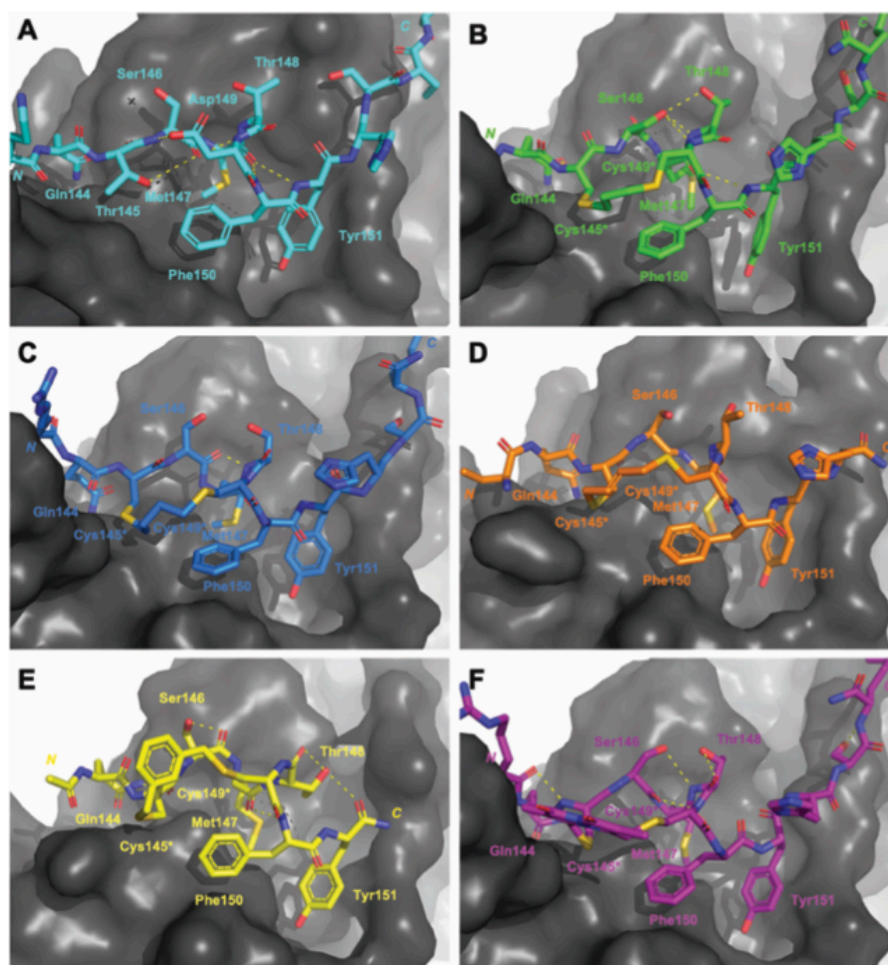


Fig. 3 Structures of peptides (stick representation) bound to PIP-box binding site on PCNA (grey, surface). Intramolecular polar interactions are shown as yellow dashed lines. (A) p21₁₄₁₋₁₅₅ (light blue), PDB: 7KQ1.¹³ (B) Peptide **3** – with propyl linker (green), co-crystal structure (PDB: 7M5L). (C) 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.

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 μ M) 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 p21₁₄₁₋₁₅₅, however the aromatic residues are notably shifted (Fig. 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 (Fig. S5, ESI[†]). 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 (Fig. S5, ESI[†]). A weak intramolecular 3.0 Å hydrogen bond between the Met147 carbonyl and Phe150 amide defines a 3_{10} -helical turn, however this is not at the same location as in native p21₁₄₁₋₁₅₅ (*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 (Fig. 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 μ M), than for peptides 3–5 and 7.

The computational model of peptide 4 bound to PCNA shows the conserved residues are positioned similarly to p21₁₄₁₋₁₅₅, except the Tyr151 %BSA is reduced to 70% compared to 90–100% in p21₁₄₁₋₁₅₅ and peptides 3, 5–7 (Table S3, ESI[†]). 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 (Fig. S3, ESI[†]). One intramolecular 2.0 Å interaction is observed between the Ser146 carbonyl and Thr148 amide (Fig. 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 (Fig. 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, ESI†). 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 (Fig. 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) (Fig. 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 (Fig. S6, ESI†).

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 (H_α), amide proton (NH), α -carbon (C_α) and carbonyl carbon (CO) for each residue (Fig. 4 and Table S6, ESI†). The secondary shift for each signal (H_α , NH, C_α , CO; Fig. 4) was calculated as the difference between the observed resonance in macrocycle **7** (Table S5, ESI†) and the observed resonance in the linear precursor peptide **2** for each residue (Table S4, ESI†). 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 H_α and positive C_α secondary shifts for residues 147 and 148 together suggest helical structure in this region (Fig. 4). Furthermore, the CO secondary shifts of residues 145–147 are positive, which is 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 (Fig. S8, ESI†) which is also

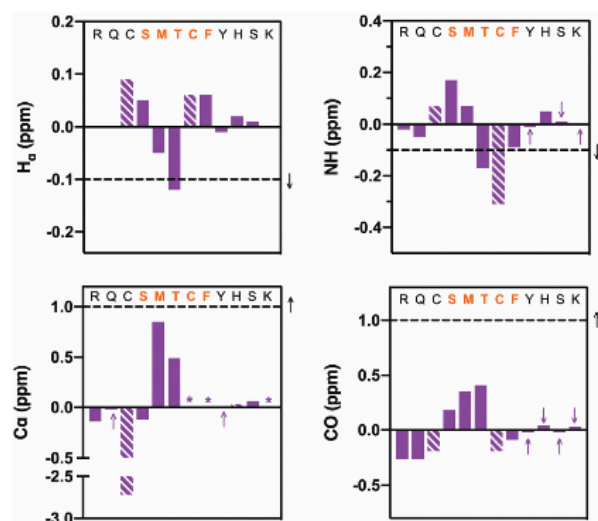


Fig. 4 NMR structural data (10% aq. D₂O, pH ~ 5) for peptide **7**: H_α , NH, C_α and carbonyl (CO) carbon secondary shifts are calculated relative to the corresponding resonances of peptide **2**. Each column represents a consecutive residue. Dashed columns represent bimane-modified cysteine residues. 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 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.

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

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 control 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 the 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 24 h. After treatment, media containing the peptide was removed and cells were washed with ice-cold PBS to remove all traces of



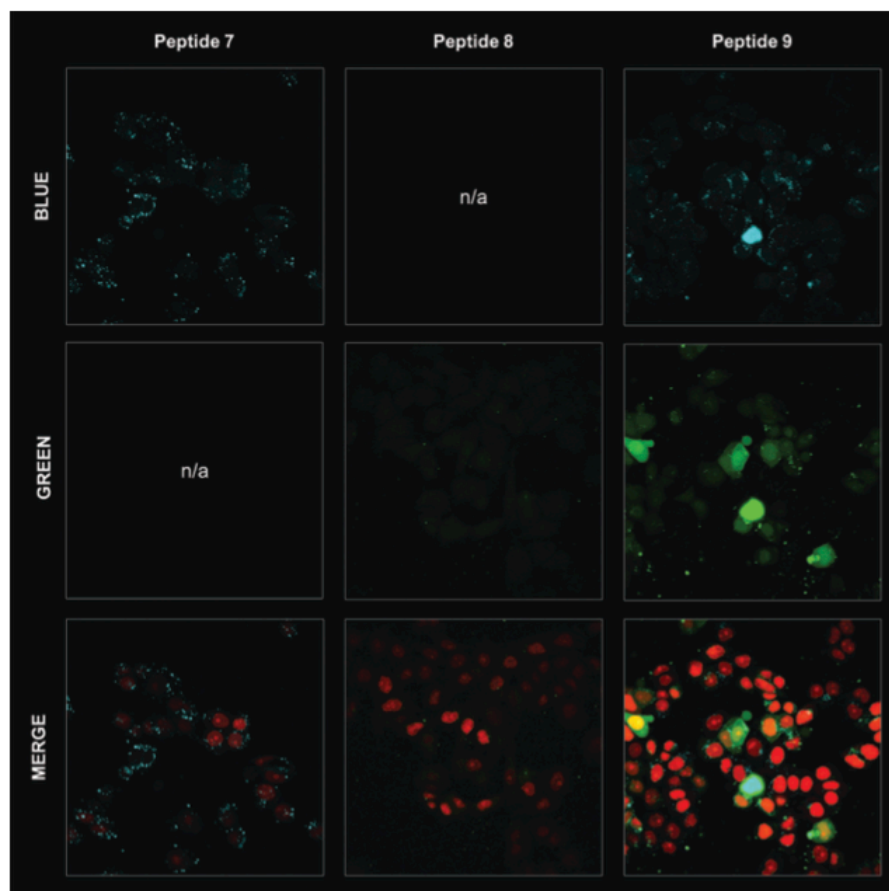


Fig. 5 Breast cancer cells MDA-MB-468 treated with 10 μM of peptide **7**, peptide **8** or peptide **9**, then fixed and imaged by confocal microscopy. Peptide **7** is cell permeable as shown by punctate blue fluorescence throughout the cytoplasm. Peptide **8** is not cell permeable, where no green 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.

Open Access Article. Published on 21 July 2021. Downloaded on 8/29/2023 12:58:19 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.



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 (Fig. 5). Peptide **8** treated cells were imaged at ex. 488 nm, and em. 490–534 nm (GREEN channel in Fig. 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 Fig. 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} Macrocytic 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.

Conclusions

Two cysteine residues were incorporated at positions 145 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 bimeane (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 3.86 μ M. The bimeane-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 bimeane 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 study¹⁹ 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 bimeane 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 bimeane 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.

Conflicts of interest

There are no conflicts of interest to declare.

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 MX1³⁵ and MX2³⁶ beamlines at the Australian Synchrotron, part of ANSTO, and made use of the Australian Cancer Research Foundation (ACRF) detector. 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).

References

- 1 T. Tsurimoto, *Front. Biosci.*, 1999, **4**, d849–858.
- 2 G. Maga and U. Hubscher, *J. Cell Sci.*, 2003, **116**, 3051–3060.
- 3 G. L. Moldovan, B. Pfander and S. Jentsch, *Cell*, 2007, **129**, 665–679.
- 4 S. N. Naryzhny and H. Lee, *FEBS Lett.*, 2007, **581**, 4917–4920.
- 5 I. Stoimenov and T. Helleday, *Biochem. Soc. Trans.*, 2009, **37**, 605–613.
- 6 D. Zhongyun, M. Wortman, Z. Tan and K. Dillehay, WO 2012/033938A2, 2012.
- 7 A. J. Horsfall, A. D. Abell and J. B. Bruning, *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, *Biochemistry*, 2000, **39**, 7388–7397.
- 9 J. B. Bruning and Y. Shamoo, *Structure*, 2004, **12**, 2209–2219.
- 10 E. M. Boehm and M. T. Washington, *BioEssays*, 2016, **38**, 1117–1122.
- 11 J. M. Gulbis, Z. Kelman, J. Hurwitz, M. O'Donnell and J. Kuriyan, *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, *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, *J. Biol. Chem.*, 2021, **296**, 100773.
- 14 E. Warbrick, D. P. Lane, D. M. Glover and L. S. Cox, *Curr. Biol.*, 1995, **5**, 275–282.
- 15 T. A. Hill, N. E. Shepherd, F. Diness and D. P. Fairlie, *Angew. Chem., Int. Ed.*, 2014, **53**, 13020–13041.
- 16 A. P. Higuero, H. Jubb and T. L. Blundell, *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, *MedChem-Comm*, 2015, **6**, 111–119.



- 18 D. P. Fairlie and A. D. de Araujo, *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, *Chem. – Eur. J.*, 2018, **24**, 11325–11331.
- 20 L. Peraro, T. R. Siegert and J. A. Kritzer, *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, *Angew. Chem., Int. Ed.*, 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, *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, *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, *Angew. Chem., Int. Ed.*, 2014, **53**, 6965–6969.
- 25 A. J. Horsfall, D. P. McDougal, D. B. Scanlon, J. B. Bruning and A. D. Abell, *ChemBioChem*, 2021, **22**, DOI: 10.1002/cbic.202100241.
- 26 A. J. Kroker and J. B. Bruning, *Biochemistry*, 2015, **54**, 3483–3493.
- 27 K. N. Choe and G.-L. Moldovan, *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, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 1871–1876.
- 29 T. B. Potocky, A. K. Menon and S. H. Gellman, *J. Biol. Chem.*, 2003, **278**, 50188–50194.
- 30 P. Ramoino, A. Diaspro, M. Fato and C. Usai, *Molecular Regulation of Endocytosis*, 2012, ch. 6, DOI: 10.5772/46061.
- 31 P. M. Cromm, J. Spiegel, P. Kuchler, L. Dietrich, J. Kriegesmann, M. Wendt, R. S. Goody, H. Waldmann and T. N. Grossmann, *ACS Chem. Biol.*, 2016, **11**, 2375–2382.
- 32 Y. Tian, Y. Jiang, J. Li, D. Wang, H. Zhao and Z. Li, *ChemBioChem*, 2017, **18**, 2087–2093.
- 33 S. R. Perry, T. A. Hill, A. D. de Araujo, H. N. Hoang and D. P. Fairlie, *Org. Biomol. Chem.*, 2018, **16**, 367–371.
- 34 A. M. Spokoiny, Y. Zou, J. J. Ling, H. Yu, Y. S. Lin and B. L. Pentelute, *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. Panjekar, J. R. Price, A. Riboldi-Tunncliffe, R. Williamson and T. Caradoc-Davies, *J. Synchrotron Radiat.*, 2015, **22**, 187–190.
- 36 D. Aragao, J. Aishima, H. Cherukuvada, R. Clarken, M. Clift, N. P. Cowieson, D. J. Ericsson, C. L. Gee, S. Macedo, N. Mudie, S. Panjekar, J. R. Price, A. Riboldi-Tunncliffe, R. Rostan, R. Williamson and T. T. Caradoc-Davies, *J. Synchrotron Radiat.*, 2018, **25**, 885–891.



Electronic Supplementary Information

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*} Andrew D. Abell^{1-3*}

Experimental

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) H₂O 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.

Synthesis & 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-NH₂

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 C₆₄H₉₈N₂₀O₂₀S: 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 H₂O 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 & Isolation:

The peptide was cleaved from the resin by addition of 92.5:2.5:2.5:2.5 TFA/TIPS/DODT/H₂O (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/H₂O, 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]

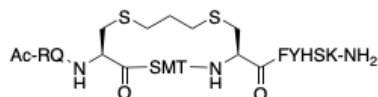
Rink Amide PL resin (0.2 mmol, 644 mg, 0.31 mmol/g, Agilent) was swollen in 1:1 DMF/DCM (15 mL) for 15 min. The Fmoc-protecting group was removed by treatment of the resin with a solution of 20% piperidine and 0.1 M HOBt in DMF (8 mL) for 15 min. The solution was drained and the resin washed with DMF (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 μ L) and DIPEA (870 μ L) in DMF (10 mL) for 15 min. The resin was washed with DMF (3 x 8 mL), and DCM (5 x 8 mL).

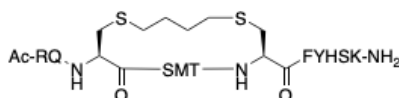
For peptides **3-7**, the cysteine side-chains were selectively deprotected: Mmt groups were removed by repetitive treatment of the resin with 2% TFA in DCM (8 mL) for 1 min, followed by washing with DCM (3 x 8 mL). Treatments were repeated until the solution no longer turned yellow on addition to the 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-NH₂

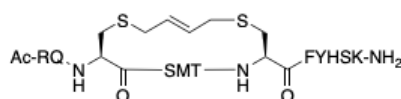
Peptide 2: Following peptide assembly as described in 'Solid-phase Peptide Synthesis of linear on-resin precursor peptide to **2a**' and *N*-terminal acetylation, the peptide was cleaved from the resin by *General Procedure for Cleavage and Isolation* to give peptide **2**. The crude peptide was purified 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 C₆₄H₉₈N₂₀O₁₈S₃: 383.6778, observed: $[M+4H]^{4+}$ 383.6702. ⁵RP-HPLC purity (C18, 215 nm) 88.9%.



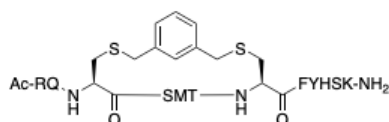
Peptide 3, propyl thioether cyclisation:² Following linear peptide assembly as described in 'Solid-phase Peptide Synthesis of linear on-resin precursor peptide to **2a**' the preswelled resin (0.1 mmol) with resin-bound *N*-terminally acetylated peptide and cysteines freshly deprotected, was transferred 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 μ L) was then added and the vessel sealed and bubbled with N₂ for 15 min. Under a N₂ atmosphere, DIPEA (35 equiv, 274 μ L) was added and the mixture stirred and bubbled with N₂ for a further 20 min. 1,3-Dibromopropane (3.5 equiv, 16 μ L) 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 H₂O (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 C₆₇H₁₀₂N₂₀O₁₈S₃: 524.5692, observed: $[M+3H]^{3+}$ 524.5718. RP-HPLC purity (C18, 215 nm) 91.9%.



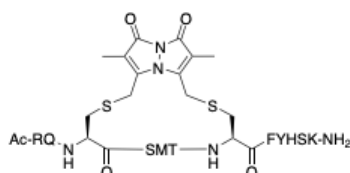
Peptide 4, butyl thioether cyclisation:² Following linear peptide assembly as described in 'Solid-phase Peptide Synthesis of linear on-resin precursor peptide to **2a**' the preswelled resin (0.1 mmol) with resin-bound *N*-terminally acetylated peptide and cysteines freshly deprotected, was transferred 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 μ L) was then added and the vessel sealed and bubbled with N₂ for 15 min. Under a N₂ atmosphere, DIPEA (35 equiv, 274 μ L) was added and the mixture stirred and bubbled with N₂ for a further 20 min. 1,4-Dibromobutane (3.5 equiv, 19 μ L) 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 H₂O (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 C₆₈H₁₀₄N₂₀O₁₈S₃: 529.2416, observed: $[M+3H]^{3+}$ 529.2410. RP-HPLC purity (C18, 215 nm) 91.8%.



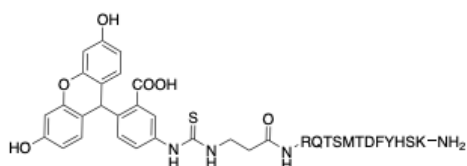
Peptide 5, trans-butene cyclisation: Following linear peptide assembly as described in 'Solid-phase Peptide Synthesis of linear on-resin precursor peptide to **2a**' the preswelled resin (0.1 mmol) with resin-bound *N*-terminally acetylated peptide and cysteines freshly deprotected, was treated with a solution of trans-1,4-dibromo-2-butene (2 equiv, 42.8 mg) and DIPEA (4 equiv, 174 μ L) 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%.



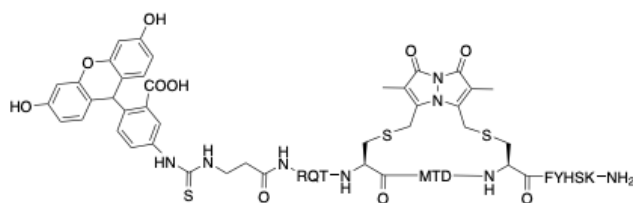
Peptide 6, xylene cyclisation: Following linear peptide assembly as described in 'Solid-phase Peptide Synthesis of linear on-resin precursor peptide to **2a**' the preswelled resin (0.1 mmol) with resin-bound *N*-terminally acetylated peptide and cysteines freshly deprotected, was treated with a solution of dibromo-*m*-xylene (2 equiv, 47.4 mg) and DIPEA (4 equiv, 174 μ L) 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: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 treated with a solution of dibromobimane (2 equiv, 70 mg) and DIPEA (4 equiv, 174 μ L) 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 μ L) 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: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 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 μ L) 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 μ L) 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%.

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 μ L/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 μ L/min).

hPCNA (5 μ L, 12 mg/mL) was diluted into running buffer (245 μ L). Only once the preactivation was complete was the protein further diluted to a final concentration of 25 μ g/mL in 10 mM NaAc (~pH 4.6) by addition of hPCNA/HEPES (50 μ L) to a solution of 100 mM NaAc (50 μ L) and water (400 μ L). This solution was immediately injected over only the second flow cell (10 μ L/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 $\mu\text{L}/\text{min}$). 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 $c = (A_\lambda/\epsilon_\lambda \cdot l) \cdot \text{DF}$ 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 2013,⁴ and ϵ_{380} used for bimeane-containing peptide **7** as reported in Shen 2011,⁵ 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 $\mu\text{L}/\text{min}$, 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.

NMR Analysis

^1H and ^{13}C NMR 1D and 2D spectra were recorded on a Varian Inova 600 MHz instrument in 10% aq. D_2O at 298K, pH \sim 5, and referenced to DSS at 0 ppm. ES suppression sequences were used for all ^1H 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 ^1H 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 2011⁶ 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).

Protein Expression and Purification

A glycerol stock of *E. coli* BL21 (λDE3) cells carrying a hPCNA-pMCSG19 plasmid (with no purification tag) were grown in a 50 mL overnight culture. Eight 1 L baffled flasks of LB with 100 $\mu\text{g}/\text{mL}$ 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. hPCNA 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 \sim 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 x 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 \sim 10 mg/mL using a centrifugal filter unit (50 kDa molecular mass cut off) and stored at -80°C.

Protein-peptide co-crystallisation experiments

hPCNA 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 16,000xg 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 μ L well solution, by mixing 1 μ L protein and peptide with equal volume of well solution.⁷⁻⁹ Diffracting crystals of hPCNA 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 hPCNA 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 hPCNA 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 MX2¹⁰ (peptide **3**) and MX1¹¹ (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.refine¹⁶ in iterative rounds with manual rebuilding in Coot.^{17, 18} Data collection and refinement statistics for hPCNA 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.

Computational modelling

The model of peptide **1** bound to hPCNA was constructed using the solved structure of hPCNA bound with p21₁₄₁₋₁₅₅ 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 hPCNA 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 hPCNA-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 PyMOL²² to validate the model by comparing against the p21₁₄₁₋₁₅₅ 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 server²³ and PoseView²⁴.

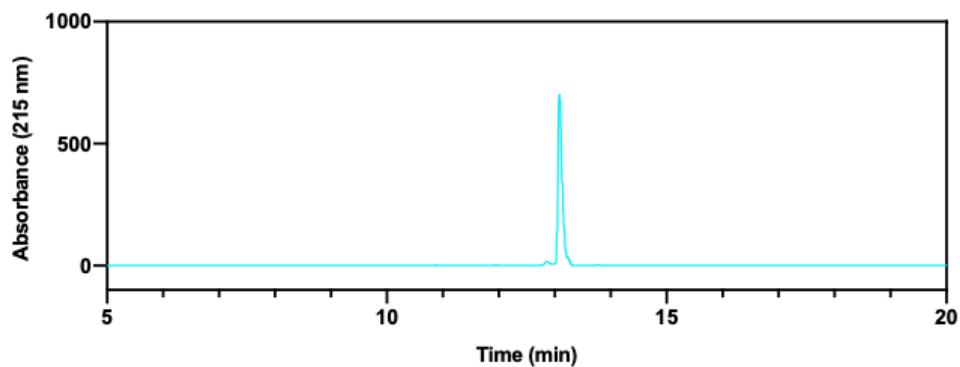
Cell imaging

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% CO₂. 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 μ M peptide **7** or **8** while being protected from light to sustain the fluorescent 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).

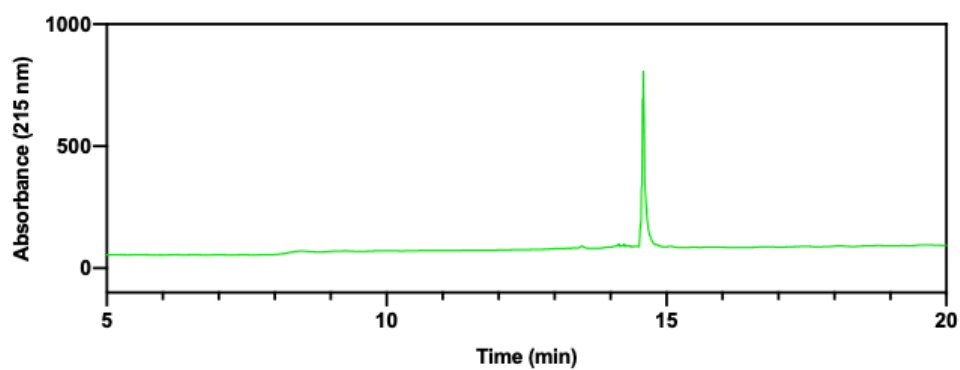
Cell nuclei were visualised using a 594 nm laser with a detection range of 600-700 nm. A 488 nm laser with a detection range of 490-534 nm was used for peptide **8**, which contains a FITC fluorophore. A 405 nm laser with a detection range of 410-485 nm was used for peptide **7**, which contains the 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.

Supplementary Information

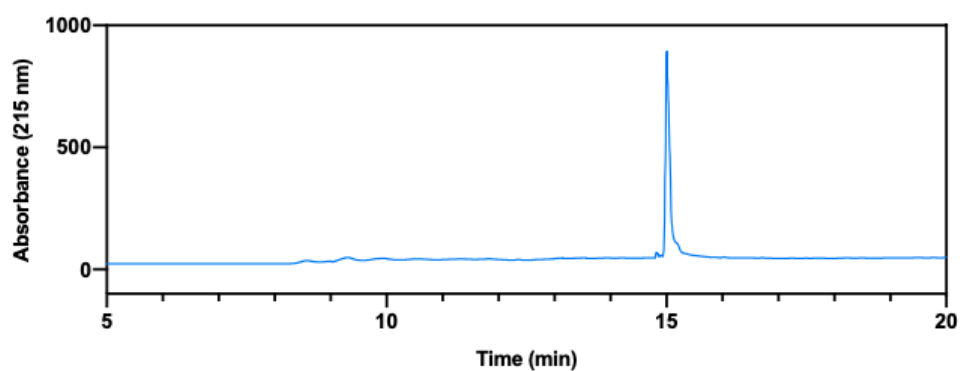
HPLC spectra



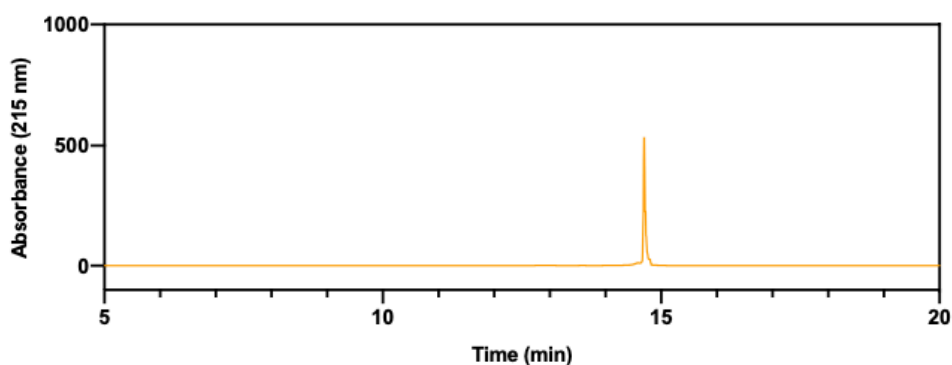
Analytical HPLC spectrum of peptide 1 at 215 nm. Collected over a gradient of 5-50% ACN with 1% TFA, in water with 1% TFA, over 15 minutes (between 5-20 min) at 1.5 mL/min.



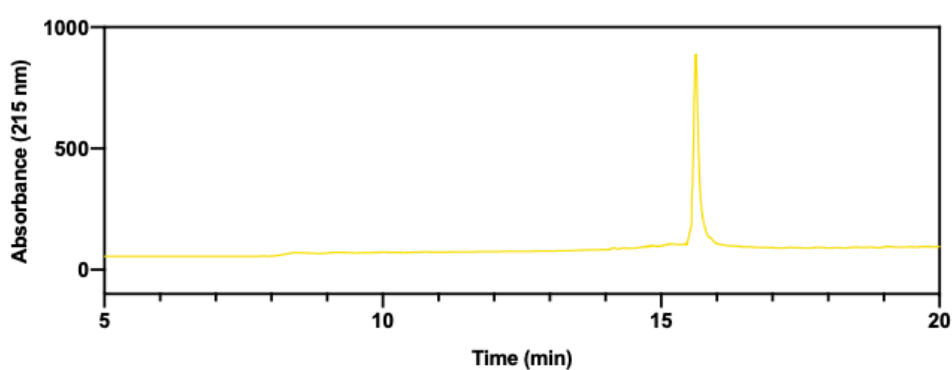
Analytical HPLC spectrum of peptide 3 at 215 nm. Collected over a gradient of 5-50% ACN with 1% TFA, in water with 1% TFA, over 15 minutes (between 5-20 min) at 1.5 mL/min.



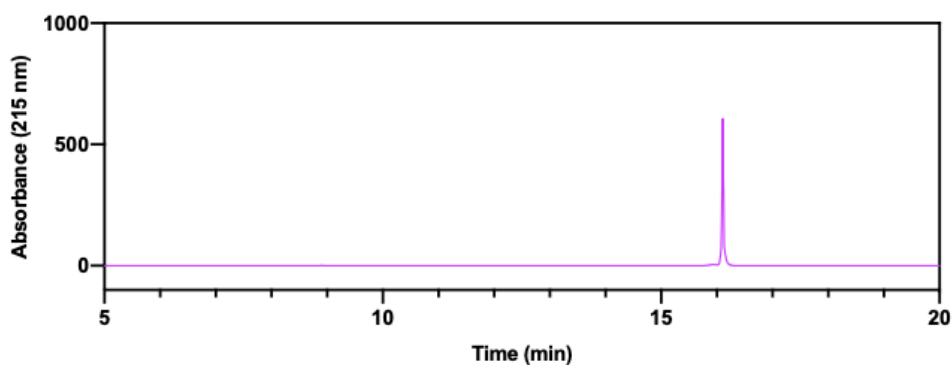
Analytical HPLC spectrum of peptide 3 at 215 nm. Collected over a gradient of 5-50% ACN with 1% TFA, in water with 1% TFA, over 15 minutes (between 5-20 min) at 1.5 mL/min.



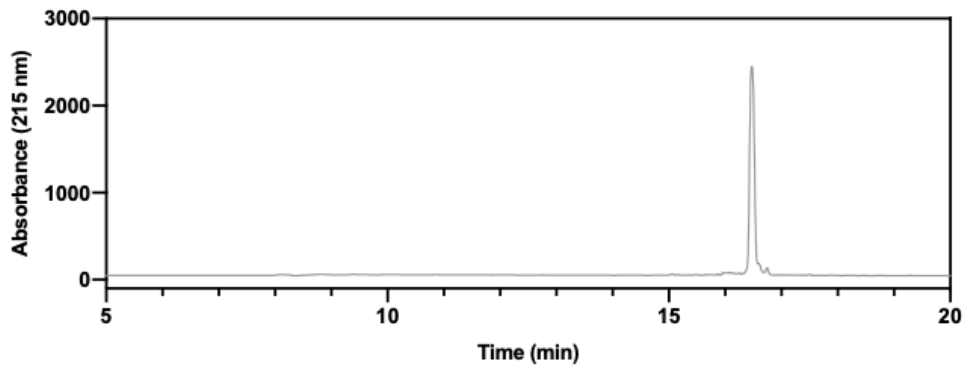
Analytical HPLC spectrum of peptide 5 at 215 nm. Collected over a gradient of 5-50% ACN with 1% TFA, in water with 1% TFA, over 15 minutes (between 5-20 min) at 1.5 mL/min.



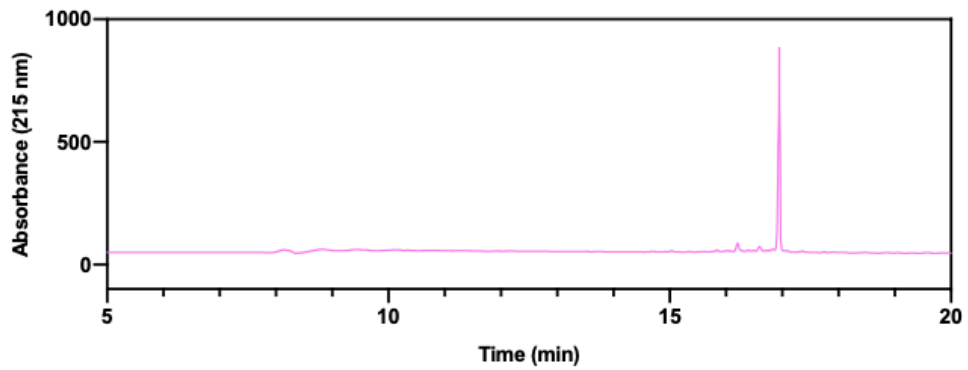
Analytical HPLC spectrum of peptide 6 at 215 nm. Collected over a gradient of 5-50% ACN with 1% TFA, in water with 1% TFA, over 15 minutes (between 5-20 min) at 1.5 mL/min.



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.

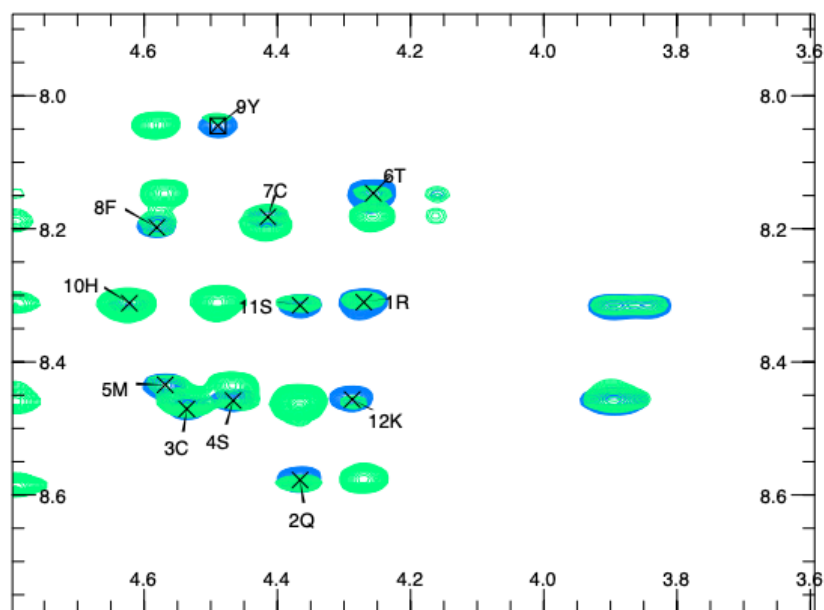
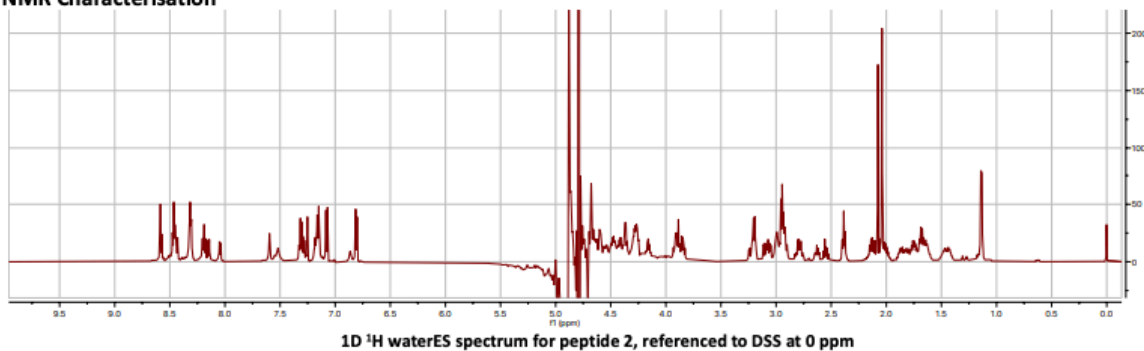


Analytical HPLC spectrum of peptide 8 at 215 nm. Collected over a gradient of 5-50% ACN with 1% TFA, in water with 1% TFA, over 15 minutes (between 5-20 min) at 1.5 mL/min.

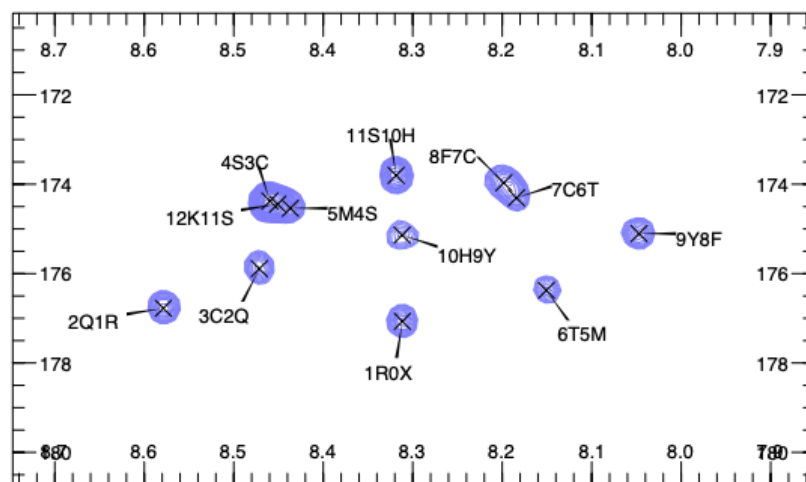


Analytical HPLC spectrum of peptide 9 at 215 nm. Collected over a gradient of 5-50% ACN with 1% TFA, in water with 1% TFA, over 15 minutes (between 5-20 min) at 1.5 mL/min.

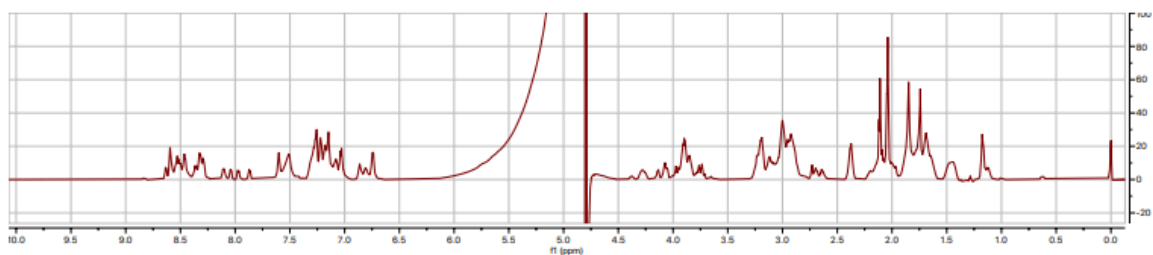
NMR Characterisation



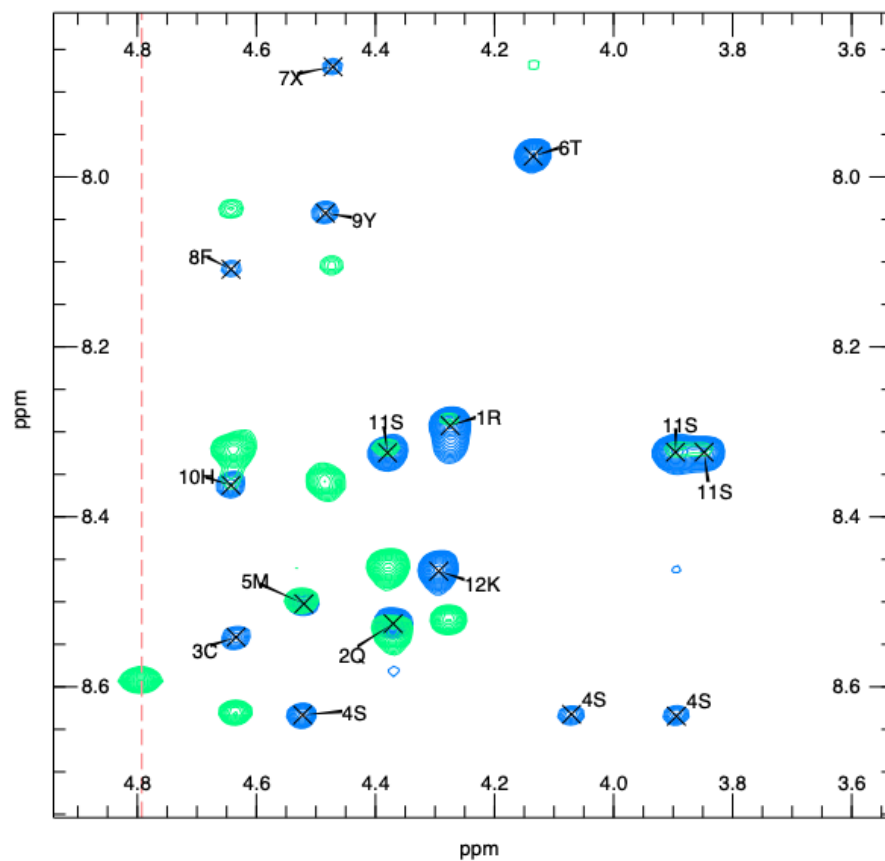
2D ¹H TOCSY (blue) and ROESY (green) for peptide 2 in the mainchain region. Assignments of TOCSY spectrum shown. Green cross peaks at 4.5 ppm (y-axis, F2) correspond to the residual H₂O peak.



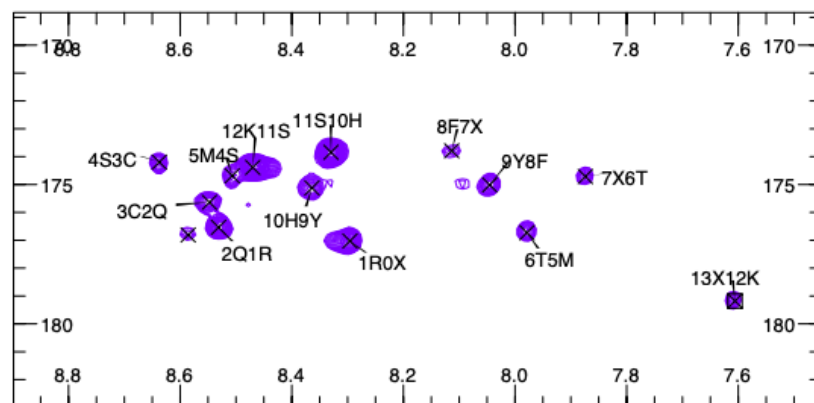
2D heteronuclear HMBC for peptide 2 in the mainchain region. Assignments are NH/CO (f2/f1).



1D ^1H wet1D spectrum for peptide 7, referenced to DSS at 0 ppm



2D ^1H TOCSY (blue) and ROESY (green) for peptide 7 in the main chain region. Red dotted line indicates HDO crosspeaks. Assignments of TOCSY spectrum shown.



2D heteronuclear HMBC for peptide 7 in the main chain region. Assignments are NH/CO (f2/f1). X represents the bimane modified cysteine residue.

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} was calculated with the online calculator detailed in Anthis 2013.⁴ Top Conc is the highest concentration of 8x 1 in 2 dilutions, 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 ^b	380	2000	570.5	30	0.118	40/60
Peptide 9	4694 ^b	380	30000	25190	1900	0.0547	40/60

*The ϵ_{λ} of 2 was used as an approximation of the peptide concentration

Crystallographic Data Statistics

Table S2: Data collection & Refinement statistics. Data collection and refinement statistics of hPCNA bound with peptide **3** (RCSB PDB ID: 7M5L), hPCNA bound with peptide **5** (RCSB PDB ID: 7M5M), and hPCNA bound with peptide **6** (RCSB PDB ID: 7M5N). Statistics for the highest-resolution shell are shown in parentheses.

	Peptide 3 7M5L	Peptide 5 7M5M	Peptide 6 7M5N
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 $R_{merge} = \sum |I - \langle I \rangle| / \sum I$
^b $R_{pim} = \sum h [1 / (n_h - 1)]^{1/2} \sum_i |<I_h> - I_{h,i}| / \sum_h \sum_i I_{h,i}$ (2)
^c $R_{work} = \sum |F_o - F_c| / \sum |F_o|$ for all data excluding data used to calculate Rfree.
^d $R_{free} = \sum |F_o - F_c| / \sum |F_o|$ for all data.

hPCNA-peptide structures supplementary data

Table S3: Percentage buried surface area (BSA, Å² | %) of PIP-box residues (excluding the covalently linked residues Cys145* and Cys149*) and the total BSA of the PIP-box residues including the covalent linker.

Peptide bound to hPCNA	p21 ¹⁴¹⁻¹⁵⁵	Peptide 3	Peptide 3	Peptide 4	Peptide 5	Peptide 6	Peptide 7
Structure type	Cocrystal PDB: 7KQ1	Cocrystal PDB: 7M5L	Computational (confidence) model	Computational model	Cocrystal PDB: 7M5M	Cocrystal PDB: 7M5N	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

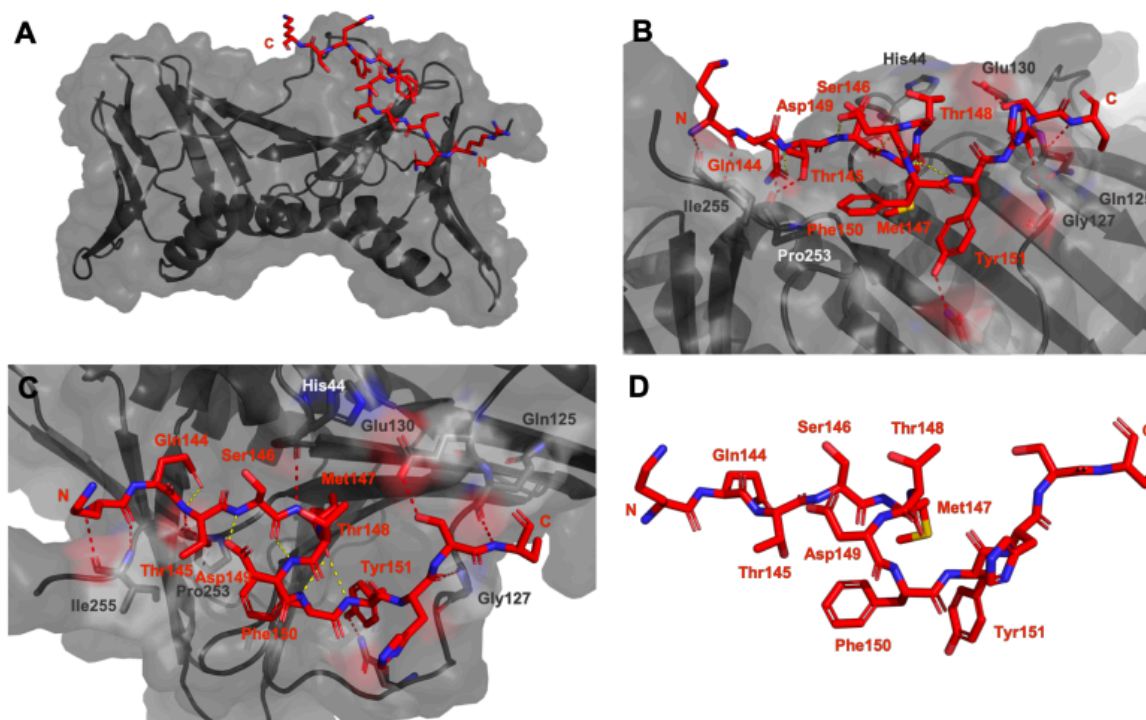


Figure S1: Computational model of peptide 1 bound to hPCNA. Peptide 1 is shown in red as sticks, and hPCNA in grey with cartoon and transparent surface representation. The PIP-box residues of peptide 1 are labelled in red, and PCNA residues in grey. Intermolecular interactions are shown as red dashes, and intramolecular interactions as yellow dashes.

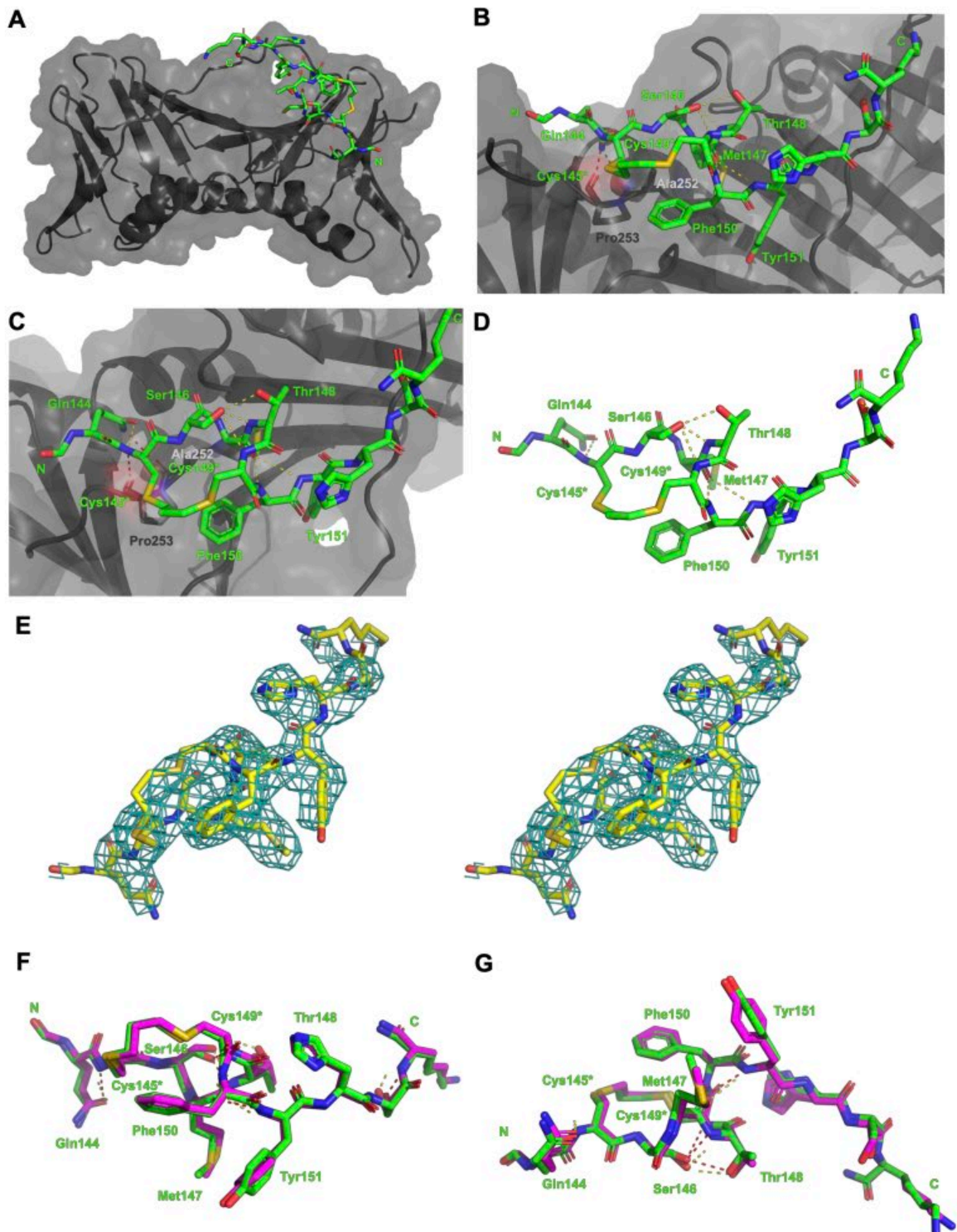


Figure S2: Co-crystal structure of peptide **3** bound to hPCNA (PDB: 7M5L). Peptide **3** is shown in green as sticks, and hPCNA in grey with cartoon and transparent surface representation (A-C). B-D The PIP-box residues of peptide **3** are labelled in green, and PCNA residues in grey. Intermolecular 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 1.0 σ . F & G Overlay of co-crystal structure of peptide **3** from cocrystal structure (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.

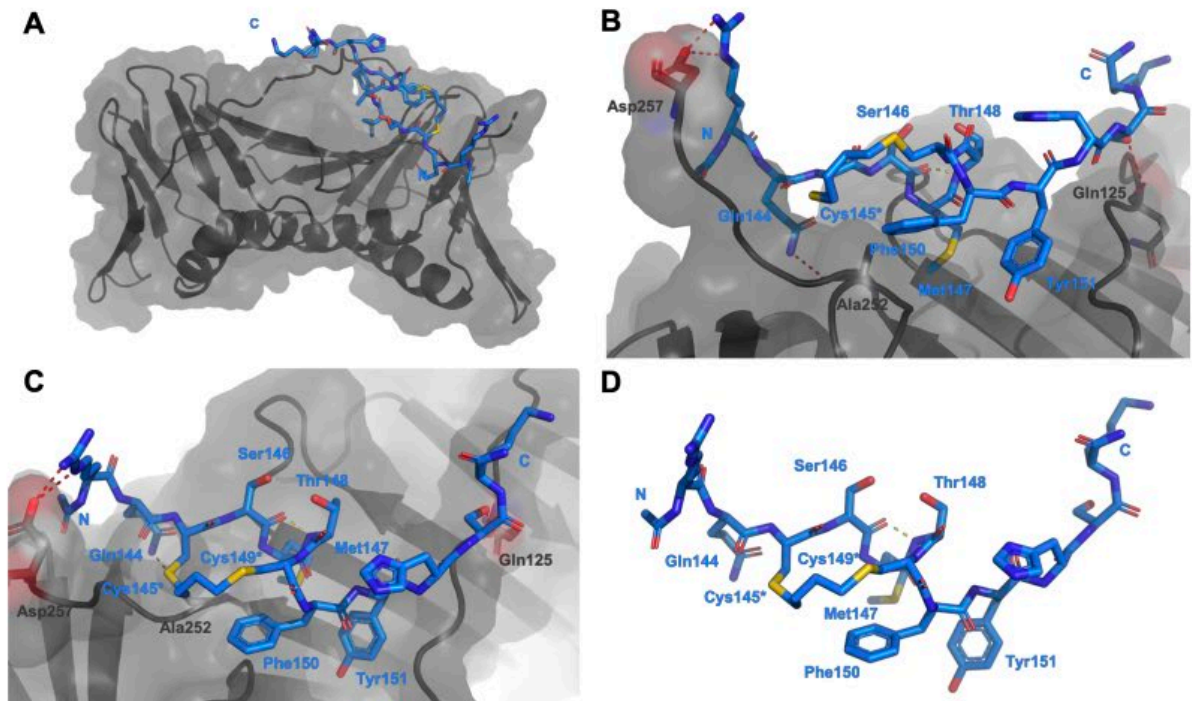


Figure S3: Computational model of peptide 4 bound to hPCNA. Peptide 4 is shown in blue as sticks (A-D), and hPCNA in grey with cartoon and transparent surface representation (A-C). The PIP-box residues of peptide 4 are labelled in blue, and PCNA residues in grey. Intermolecular interactions are shown as red dashes, and intramolecular interactions as yellow dashes.

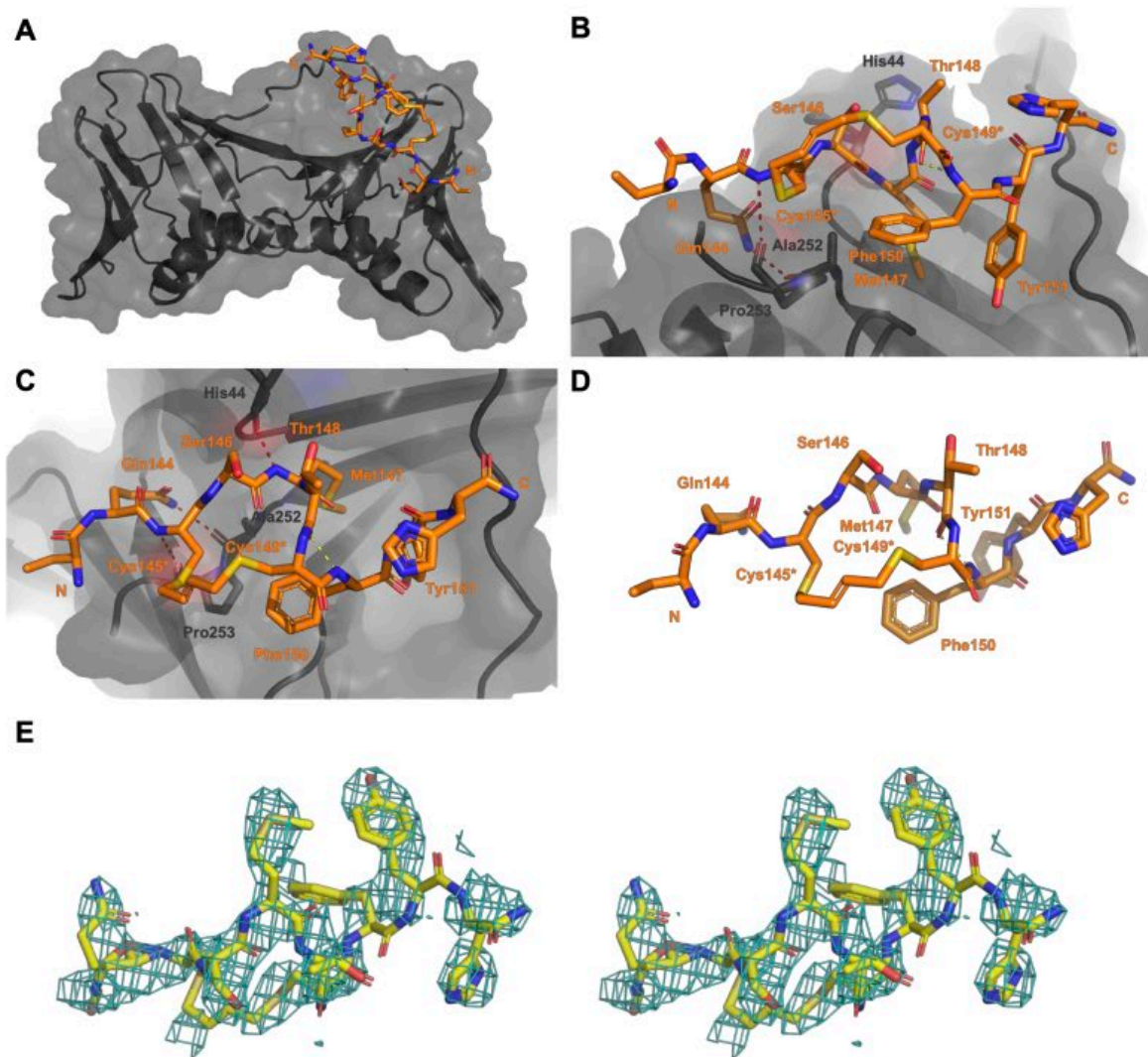


Figure 54: Co-crystal structure of peptide **5** bound to hPCNA (PDB: 7M5M). **A-D** Peptide **5** is shown in orange as sticks, and hPCNA in grey with cartoon and transparent surface representation (**A-C**). The PIP-box residues of peptide **3** are labelled in orange, and PCNA residues in grey. Intermolecular 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 1.0 σ .

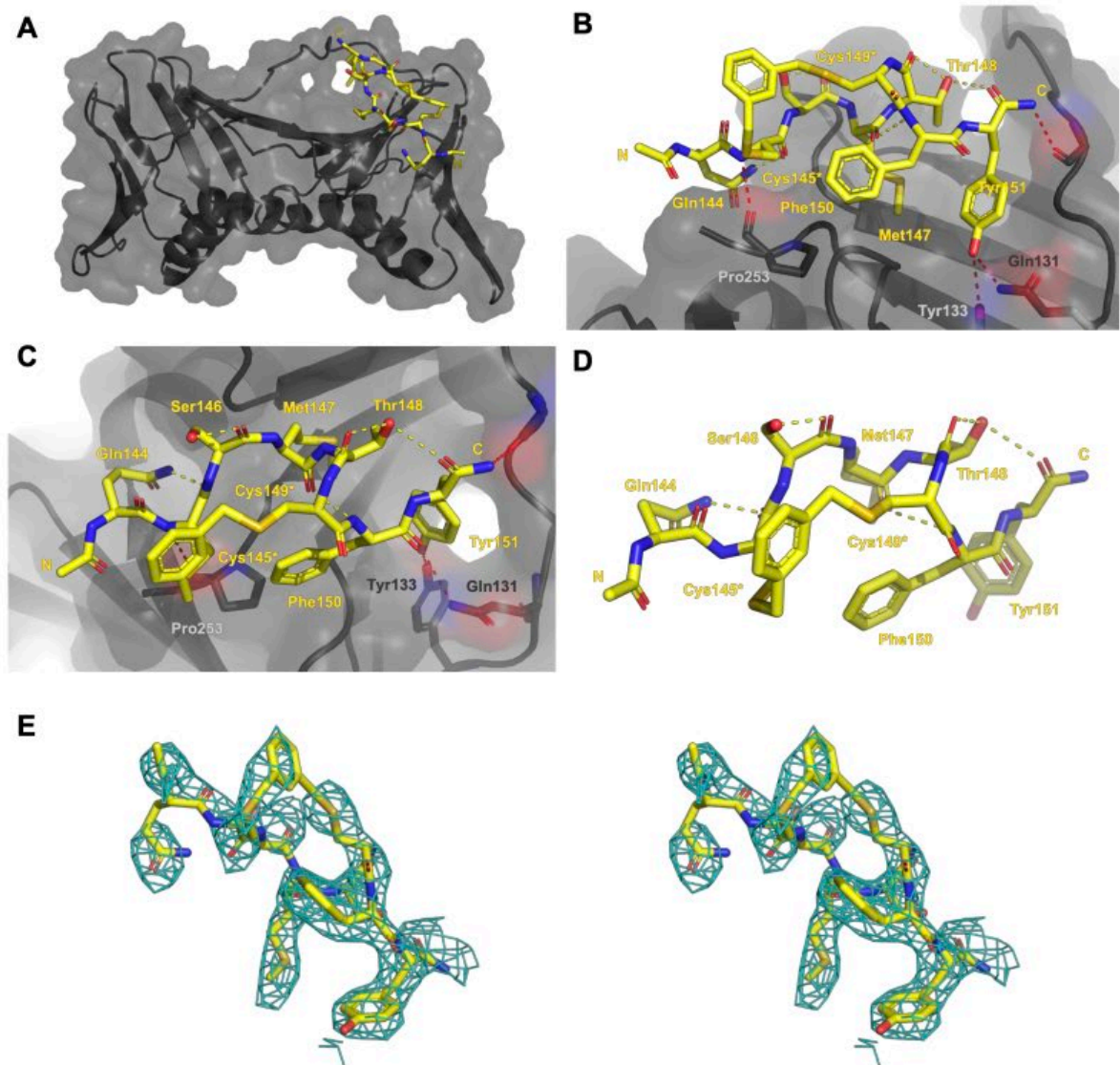


Figure S5: Co-crystal structure of peptide **6** bound to hPCNA (PDB: 7M5N). Peptide **6** is shown in yellow as sticks (**A-E**), and hPCNA in grey with cartoon and transparent surface representation (**A-C**). The PIP-box residues of peptide **6** are labelled in yellow, and PCNA residues in grey. 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 .

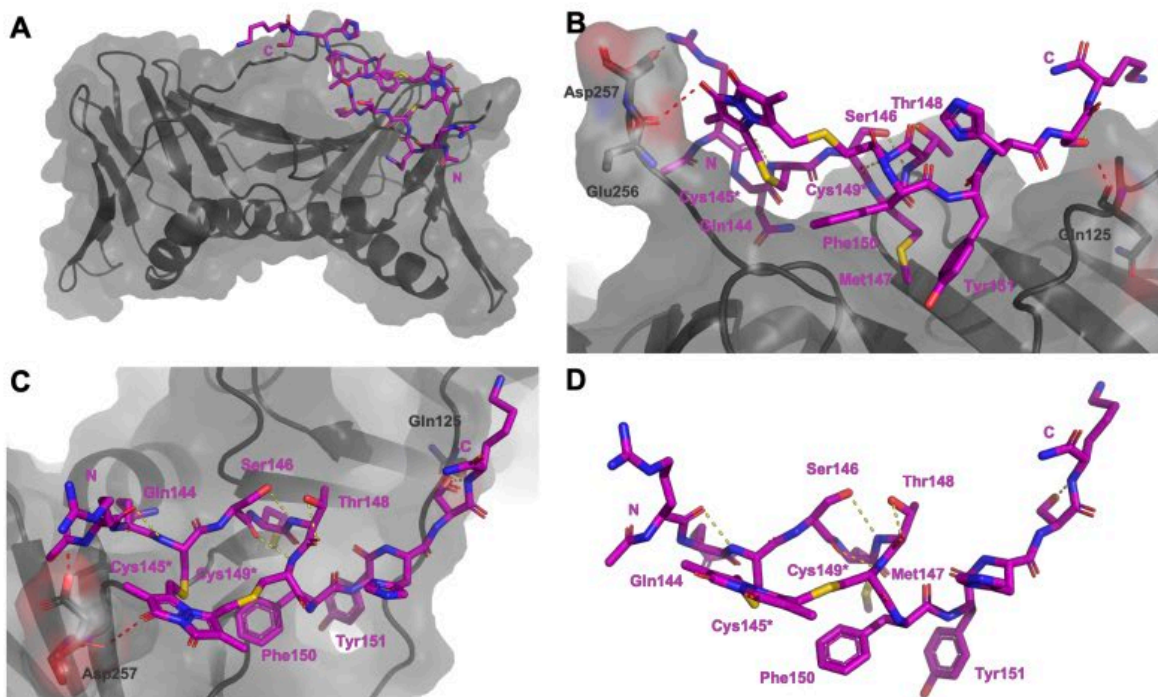


Figure S6: Computational model of peptide **7** bound to hPCNA. Peptide **7** is shown in purple as sticks (**A-D**), and hPCNA in grey with cartoon and transparent surface representation (**A-C**). The PIP-box residues of peptide **7** are labelled in purple, and PCNA residues in grey. Intermolecular interactions are shown as red dashes, and intramolecular interactions as yellow dashes.

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 2011⁶. NN – nearest neighbour corrections applied to random coil value; H α did not require NN corrections due to the sequence used here. $\Delta\delta$ – Secondary shift values calculated as the difference between the observed resonance (H α , C α , CO or NH) and the respective NN value (obs-NN), shown graphically in green in Figure S7.

2	H α	RC	NN	$\Delta\delta$	C α	RC	NN	$\Delta\delta$	CO	RC	NN	$\Delta\delta$	NH	RC	NN	$\Delta\delta$
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 2011⁶. NN – nearest neighbour corrections applied to random coil value; H α did not require NN corrections due to the sequence used here. $\Delta\delta$ – Secondary shift values calculated as the difference between the observed resonance (H α , C α , CO or NH) and the respective NN value, shown graphically in blue in Figure S7.

7	H α	RC	NN	$\Delta\delta$	C α	RC	NN	$\Delta\delta$	CO	RC	NN	$\Delta\delta$	NH	RC	NN	$\Delta\delta$
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. $\Delta\delta$ – Secondary shift values calculated as the difference between the observed resonance (H α , C α , CO or NH) of 7, and the respective value for 2 (7-2), shown graphically in purple in Figure S7.

7-2	7H α	2H α	$\Delta\delta$	7C α	2C α	$\Delta\delta$	7CO	2CO	$\Delta\delta$	7NH	2NH	$\Delta\delta$
1R	4.27	4.27	0	56.15	56.29	-0.14	176.54	176.80	-0.26	8.29	8.31	0.02
2Q	4.37	4.37	0	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	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	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	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	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	-	58.28	-	173.79	173.98	-0.19	7.87	8.18	0.31
8F	4.64	4.58	0.06	-	57.65	-	175.01	175.10	-0.09	8.11	8.20	0.09
9Y	4.48	4.49	-0.01	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	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	58.25	58.19	0.06	174.39	174.41	-0.02	8.32	8.31	-0.01
12K	4.29	4.29	0	-	56.22	-	179.13	179.10	0.03	8.46	8.46	0.00

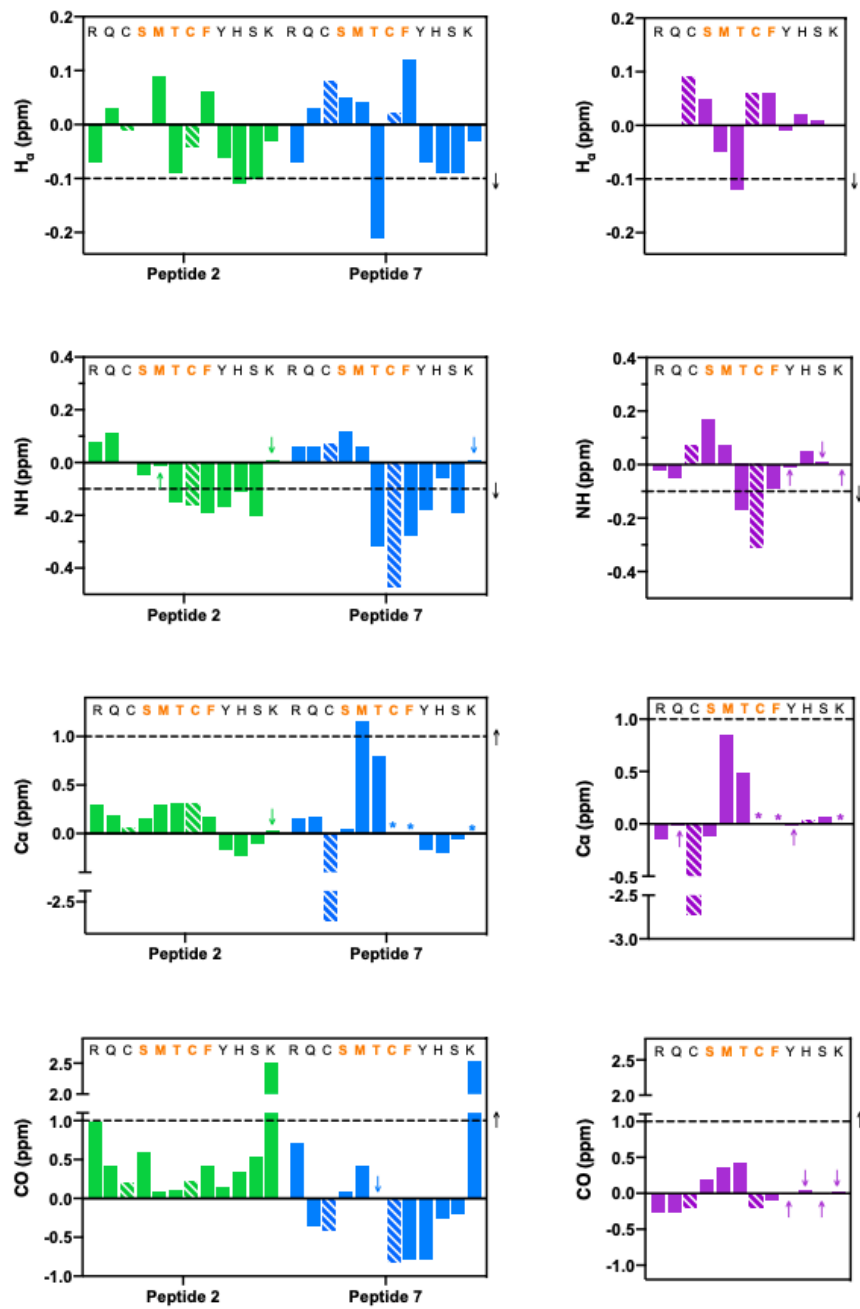


Figure S7: NMR secondary shift of mainchain resonances of peptide 2 (green) and 7 (blue) calculated relative to literature random coil values, with nearest neighbour corrections applied; and for peptide 7 relative to linear precursor peptide 2 (purple). Each bar represents the resonance for each consecutive amino-acid in the sequence, where the sequence is annotated at 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.

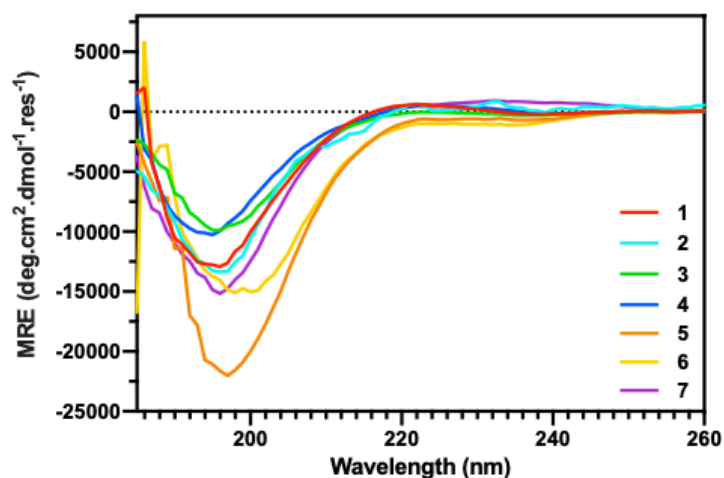


Figure S8: Circular dichroism spectra collected for peptides 1-7 in 10 mM phosphate buffer at pH 7.2. The spectra shown are the average of 8 scans, collected at 50 nm/min with a pitch of 1 mm in a 1 mm cuvette, using a Jasco J-810 (UniSA BioPhysical Characterisation facility); then smoothed using the in-built software with a Savitsky-Golay function and a convolution width of 7.

References:

1. W. S. Hancock and J. E. Battersvy, *Anal. Biochem.*, 1976, **71**, 260-264.
2. 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, *Angew. Chem. Int. Ed. (English)*, 2018, **57**, 17073-17078.
3. A. J. Horsfall, K. R. Dunning, K. L. Keeling, D. B. Scanlon, K. L. Wegener and A. D. Abell, *ChemBioChem*, 2020, **21**, 3423-3432.
4. N. J. Anthis and G. M. Clore, *Protein Sci.*, 2013, **22**, 851-858.
5. X. Shen, C. B. Pattillo, S. Pardue, S. C. Bir, R. Wang and C. G. Kevil, *Free Radical Biol. Med.*, 2011, **50**, 1021-1031.
6. D. S. Wishart, *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, *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, *Nuclear Receptors Research*, 2018, **5**, 101350.
9. J. L. Pederick, A. P. Thompson, S. G. Bell and J. B. Bruning, *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, *Journal of Synchrotron Radiation*, 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, *Journal of Synchrotron Radiation*, 2015, **22**, 187-190.
12. W. Kabsch, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2010, **66**, 125-132.
13. P. Evans, *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. Read, *J. Appl. Crystallogr.*, 2007, **40**, 658-674.
15. J. M. Gulbis, Z. Kelman, J. Hurwitz, M. O'Donnell and J. Kuriyan, *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, *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, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2012, **68**, 352-367.
18. P. Emsley and K. Cowtan, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2004, **60**, 2126-2132.
19. N. W. Moriarty, R. W. Grosse-Kunstleve and P. D. Adams, *Acta Crystallogr D Biol Crystallogr*, 2009, **65**, 1074-1080.
20. R. Abagyan, M. Totrov and D. Kuznetsov, *J. Comput. Chem.*, 1994, **15**, 488-506.
21. R. Abagyan and M. Totrov, *J. Mol. Biol.*, 1994, **235**, 983-1002.
22. Schrodinger, LLC, unpublished work.
23. D. Piovesan, G. Minervini and S. C. Tosatto, *Nucleic Acids Res.*, 2016, **44**, W367-374.
24. K. Stierand, P. C. Maass and M. Rarey, *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, *Acta Crystallogr D Biol Crystallogr*, 2015, **71**, 646-666.

Chapter 5: Publication

*“Towards a High Affinity Peptidomimetic Targeting PCNA
from Aspergillus fumigatus.”*

Statement of Authorship

Title of Paper	Towards a High Affinity Peptidomimetic Targeting PCNA from <i>Aspergillus fumigatus</i>
Publication Status	<input checked="" type="checkbox"/> Published <input type="checkbox"/> Accepted for Publication <input type="checkbox"/> Submitted for Publication <input type="checkbox"/> Unpublished and Unsubmitted work written in manuscript style
Publication Details	Vandborg, B. C., Horsfall, A. J., Pederick, J. L., Abell, A. D., and Bruning, J. B. Towards a High-Affinity Peptidomimetic Targeting Proliferating Cell Nuclear Antigen from <i>Aspergillus fumigatus</i> . <i>Journal of Fungi</i> (2023) 9, 1098,

Principal Author

Name of Principal Author (Candidate)	Bethiney Vandborg
Contribution to the Paper	Protein synthesis and purification, peptide design, experimental design, validation experiments, discussed results and wrote and edited manuscript.
Overall percentage (%)	70%
Certification:	I certify that this work is original research I conducted during the period of my Higher Degree by research 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	<div style="display: flex; justify-content: space-between;"> <div style="border-bottom: 1px solid black; width: 80%;"></div> <div style="border-bottom: 1px solid black; width: 15%; text-align: center;">Date</div> <div style="border-bottom: 1px solid black; width: 5%;"></div> </div>
	29/08/2023

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 to 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.

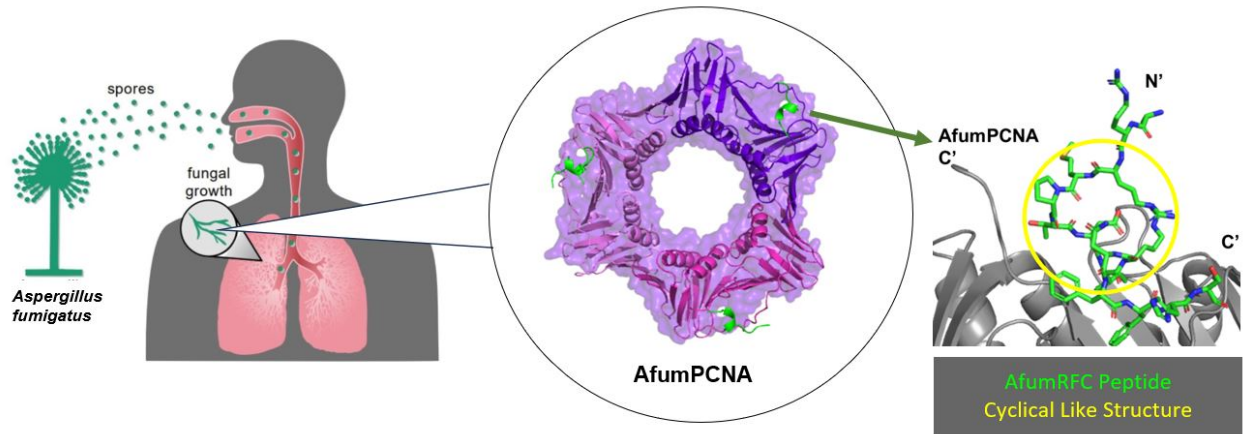
Name of Co-Author	Aimee Horsfall
Contribution to the Paper	Designed and synthesised peptides, edited manuscript.
Signature	<div style="display: flex; justify-content: space-between;"> <div style="border-bottom: 1px solid black; width: 80%;"></div> <div style="border-bottom: 1px solid black; width: 15%; text-align: center;">Date</div> <div style="border-bottom: 1px solid black; width: 5%;"></div> </div>
	29/08/2023

Name of Co-Author	Jordan Pederick
Contribution to the Paper	Assisted with experimental design and writing and editing of the manuscript.
Signature	<div style="display: flex; justify-content: space-between;"> <div style="border-bottom: 1px solid black; width: 80%;"></div> <div style="border-bottom: 1px solid black; width: 15%; text-align: center;">Date</div> <div style="border-bottom: 1px solid black; width: 5%;"></div> </div>
	11/10/2023

Name of Co-Author	Andrew Abell		
Contribution to the Paper	Supervised AJH		
Signature		Date	29/08/2023

Name of Co-Author	John Bruning		
Contribution to the Paper	Supervised BV, edited manuscript.		
Signature			17/9/23

Graphical Abstract



Article

Towards a High-Affinity Peptidomimetic Targeting Proliferating Cell Nuclear Antigen from *Aspergillus fumigatus*

Bethiney C. Vandborg ^{1,2}, Aimee J. Horsfall ^{1,3}, Jordan L. Pederick ^{1,2}, Andrew D. Abell ^{1,3} and John B. Bruning ^{1,2,*}

- ¹ Institute of Photonics and Advanced Sensing (IPAS), The University of Adelaide, Adelaide 5005, Australia; bethiney.vandborg@adelaide.edu.au (B.C.V.); jordan.pederick@adelaide.edu.au (J.L.P.); andrew.abell@adelaide.edu.au (A.D.A.)
² School of Biological Sciences, The University of Adelaide, Adelaide 5005, Australia
³ ARC Centre of Excellence for Nanoscale BioPhotonics, The University of Adelaide, Adelaide 5005, Australia
 * Correspondence: john.bruning@adelaide.edu.au

Abstract: Invasive fungal infections (IFIs) are prevalent in immunocompromised patients. Due to alarming levels of increasing resistance in clinical settings, new drugs targeting the major fungal pathogen *Aspergillus fumigatus* are required. Attractive drug targets are those involved in essential processes like DNA replication, such as proliferating cell nuclear antigens (PCNAs). PCNA has been previously studied in cancer research and presents a viable target for antifungals. Human PCNA interacts with the p21 protein, outcompeting binding proteins to halt DNA replication. The affinity of p21 for hPCNA has been shown to outcompete other associating proteins, presenting an attractive scaffold for peptidomimetic design. p21 has no *A. fumigatus* homolog to our knowledge, yet our group has previously demonstrated that human p21 can interact with *A. fumigatus* PCNA (afumPCNA). This suggests that a p21-based inhibitor could be designed to outcompete the native binding partners of afumPCNA to inhibit fungal growth. Here, we present an investigation of extensive structure–activity relationships between designed p21-based peptides and afumPCNA and the first crystal structure of a p21 peptide bound to afumPCNA, demonstrating that the *A. fumigatus* replication model uses a PIP-box sequence as the method for binding to afumPCNA. These results inform the new optimized secondary structure design of a potential peptidomimetic inhibitor of afumPCNA.

Keywords: PCNA; DNA replication proteins; non-tag purification; peptide characterization; *Aspergillus fumigatus*; X-ray crystallography



Citation: Vandborg, B.C.; Horsfall, A.J.; Pederick, J.L.; Abell, A.D.; Bruning, J.B. Towards a High-Affinity Peptidomimetic Targeting Proliferating Cell Nuclear Antigen from *Aspergillus fumigatus*. *J. Fungi* **2023**, *9*, 1098. <https://doi.org/10.3390/jof9111098>

Academic Editor: Gill Diamond

Received: 11 October 2023

Revised: 4 November 2023

Accepted: 7 November 2023

Published: 10 November 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Invasive fungal infections (IFIs) are a prevalent cause of death in immunocompromised patients [1]. A major fungal pathogen causing such infections is *Aspergillus fumigatus*, a filamentous fungus that is usually present in decaying organic matter [2,3]. The conidia of *A. fumigatus* inhaled from the environment can be cleared from the lungs by a healthy immune system. However, when this fails, the conidia can reach the lower airways and evade host immune cells [4]. This can result in the infection of the bronchi and sinuses and dissemination to the brain and other vital organs through the circulatory system [5,6]. This is known as invasive aspergillosis. When invasive aspergillosis invades the nervous system, it has mortality rates of up to 90% [7].

As the infection rates of *A. fumigatus* increase, more species and therefore differing antifungal resistances arise, which have been associated with negative clinical outcomes [8]. Although many antifungals are available, the mortality rates remain high due to the development of drug resistance in *A. fumigatus* [9]. Current treatments of invasive aspergillosis target the components of the cell membrane: for example, the antifungal amphotericin B [10]. Unfortunately, in addition to increasing resistance rates, amphotericin B is associated with severe side effects, the most notable being kidney and liver toxicity [11]. There

has been increasing resistance to the antifungal triazole reported via unknown resistance mutations [12]. As such, there is an urgent need for improved diagnostic protocols and a broader range of antifungal options [13].

The DNA replication process is a desirable target for an antifungal product; hence, *A. fumigatus* proliferating cell nuclear antigen (afumPCNA) has been hypothesised as a fungal target for the development of new antifungals [14]. The PCNA processivity factor is also known as the sliding clamp, as DNA slides through its central cavity. PCNA functions as a docking platform to allow DNA polymerases and a host of DNA replication and repair machinery to interact at the replication fork [15]. PCNA has been proposed as a target for the treatment of multiple diseases as it is essential for cell replication, and its absence has been shown to cause embryonic lethality [16,17].

The X-ray crystal structure of apo afumPCNA solved at 2.6 Å resolution [14] shows a trimeric tertiary structure that is similar to that of hPCNA by the superimposition of the structures (PDB: 8GJF and PDB: 7KQ1), this revealed a root-square standard deviation (RMSD) of 0.939 Å. Despite this, the proteins only have a sequence similarity of 53% [14]. The PCNA homotrimer comprises two domains, with each containing two alpha helices and nine beta strands (Figure 1b), connected by a motif known as the interdomain connecting loop (IDCL), which forms part of the PCNA surface with which binding domain proteins interact. The negatively charged beta sheets allow interactions with replication and repair proteins, and the positively charged alpha helices allow non-specific interactions with double-stranded DNA on the inside of the sliding clamp (Figure 1a). It has been hypothesised that the difference in amino acid sequences present at the surface of PCNA, to which interacting partners bind, could allow for the specificity of afumPCNA over hPCNA in the design of a fungal replication inhibitor [14]. This illustrates the importance of investigating the structure of afumPCNA to understand how interacting peptides bind in aiding rational drug design.

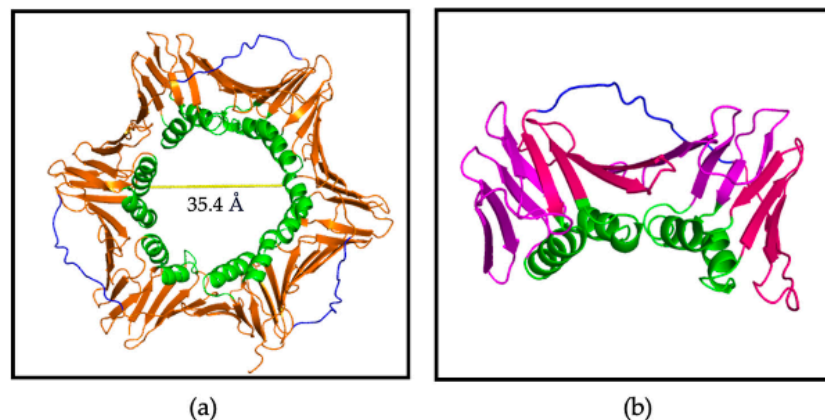


Figure 1. PCNA homotrimer structure (PDB:7KQ0) [18]. (a) Trimer: PCNA surface shown in orange, IDCL shown in blue, and positively charged alpha helices shown in green. Central cavity is 35.4 Å wide. (b) Monomer: beta sheets shown in magenta and pink, alpha helices shown in green, and IDCL shown in blue. Made using Pymol Version 1.2 [19].

In an effort to target PCNA, the most thoroughly characterised peptide inhibitor is derived from the tumour suppressor protein p21. Protein p21 (also known as cyclin-dependent kinase inhibitor 1), a cyclin-dependent kinase inhibitor, binds hPCNA to outcompete binding partners in order to halt DNA replication for repair systems [20], thus regulating the cell cycle during DNA damage. The affinity of peptides derived from p21 binding to hPCNA has been shown to be much higher than other associating proteins [18,21]. PCNA-interacting proteins, including p21, are allowed access to DNA by interacting with PCNA via the PCNA-interacting protein (PIP) box sequence, $Q_{144}X_{145}X_{146}\Phi_{147}X_{148}X_{149}\Psi_{150}\Psi_{151}$,

a consensus sequence in which the glutamine residue (Gln₁₄₄) binds the ‘Q pocket’ of the PCNA hydrophobic patch, Φ_{147} represents a hydrophobic residue, and Ψ_{150} and Ψ_{151} represent aromatic residues. The sequence and affinity of the PIP-box are theorised to correlate with the protein’s function [22]. The combination of the hydrophobic residue and two aromatic residues forms a hydrophobic plug that inserts into the PCNA surface and twists the peptide’s backbone from residues 147 to 151 into a 3_{10} helix that is conserved between binding partners. This secondary structure is critical for high-affinity binding. Other interactions that increase affinity are the ionic charged interactions of the C-terminal flank of the PIP-box with the surface of PCNA, and the N-terminal flank of the PIP-box creates an antiparallel β -sheet with the C-terminus of PCNA.

The fluorescence polarisation experiments of afumPCNA and a 22 amino acid peptide derived from the C-terminus of (human) p21 containing the PIP-box (139–160) have demonstrated their interaction, suggesting that afumPCNA interacts with DNA binding proteins using a similar PIP-box mechanism compared to the human system [14]. Given that *Aspergillus fumigatus* does not have a known p21 equivalent and this p21-derived peptide shows high-affinity interactions, further investigation into human PIP-box sequence interactions with afumPCNA may indicate the characteristics of a high-affinity mimetic. The p21 scaffold may serve as a useful starting point for designing the peptide inhibitors of afumPCNA. This also suggests that an artificial PIP-box could be designed specifically to disrupt the function of afumPCNA and highlights afumPCNA as a potential drug target for treating fungal infections. Fungal PCNA-interacting proteins were investigated as a means of probing these unknown PIP-box sequences; therefore, they can uncover interactions that could be advantageous to a mimetic. These PIP-box candidates were investigated in fungal proteins DNA polymerase (DNAPol), DNA ligase (DNALig), flap endonuclease 1 (FEN1), and replication factor C (RFC).

Here, we present the first structure of a PIP-box peptide bound to afumPCNA and characterise the interactions of afumPCNA with predicted fungal protein PIP-box candidates. Via binding affinity assays and X-ray crystallography studies, the findings support the hypothesis that the fungal replication model uses a PIP-box sequence as the method for binding to fungal PCNA, and a rational design for a potential peptidomimetic is presented.

2. Materials and Methods

2.1. Peptides

The following peptides were obtained and synthesised by Genscript Biotech, Singapore at a purity of >95%, and they were purified via HPLC. Peptides denoted with * were designed by B. Vandborg. The sequences are shown in bold.

5FAM-p21-22mer (5FAM)-**GRKRRQTSMTDFYHSKRRLIFS**

p21 μ p21 μ -15mer **KRRQTSMTDFYHSKR**

p21 μ -afumDNALIG **KRRQRVRSIASFFHSKR***

p21 μ -afumDNAPOL **KRRQKELSRFDHFSK***

p21 μ -afumFEN1 **KRRQSRLEGFFHSKR***

p21 μ -afumRFC **KRRMPTDIRNFFHSKR***

The following peptides were synthesised by Fmoc SPPS, as described below in Section 2.1.1; each has a C terminal carboxyl amide. The sequences are shown in bold.

p21 μ (p21 μ -15mer) **KRRQTSMTDFYHSKR**

p21 μ -RD2 **KRRQTRITEYFHSKR**

p21 μ -Q144M **KRRMTSMTDFYHSKR**

p21 μ -T145K **KRRQKSMTDFYHSKR**

p21 μ -T145D **KRRQDSMTDFYHSKR**

p21 μ -S146R **KRRQTRMTDFYHSKR**

p21 μ -M147L **KRRQTSMTDFYHSKR**

p21 μ -M147I **KRRQTSITDFYHSKR**

p21 μ -D149E **KRRQTSMTDFYHSKR**

p21 μ -F150Y **KRRQTSMTDYHSKR**

p21 μ -Y151F KRRQTSMTDFHFSKR
 p21 μ -FY150151YF KRRQTSMTDYFHSKR

2.1.1. Peptide Synthesis by Fmoc SPPS

All peptides were prepared on Rink Amide functionalized polystyrene resin (Agilent, Santa Clara, CA, USA) and synthesized via Fmoc/*t*Bu solid-phase peptide synthesis (SPPS), as previously described [18]. The peptides were purified via semi-preparatory RP-HPLC, and the purity and identity were confirmed via analytical HPLC and MS, as previously reported [18].

2.2. Expression of Recombinant *afumPCNA*

A. fumigatus PCNA was expressed as described in Vandborg 2023 [23].

A glycerol stock of *E. coli* BL21 (DE3) cells carrying a codon-optimized *afumPCNA*-pMCSG19 plasmid was grown in a 100 mL overnight culture. Two 1 L baffled flasks containing LB with 100 μ g/mL of ampicillin were inoculated with 50 mL of the overnight culture. Cultures were incubated at 37 °C until OD₆₀₀ = 0.7, and expression was induced with a final concentration of 0.5 mM IPTG. Cultures were grown overnight at 16 °C and shaking occurred at 200 rpm. Cultures were pelleted at 5000 \times g for 20 min. After removing the supernatant, pellets were resuspended in 20 mL 20 mM Tris-HCl pH 7.5, 20 mM NaCl, and 2 mM DTT and then lysed via sonication at 70% amplification for 20 s with a 40 s waiting period for 25 cycles. Lysate was clarified via pelleting at 45,000 \times g for 45 min, and the supernatant was collected for purification.

2.3. Purification of Recombinant *afumPCNA*

A. fumigatus PCNA was purified as described in Vandborg 2023 [23].

Buffer solutions were filtered before being used. Clarified lysate containing *afumPCNA* was first purified at 4 °C via fast protein liquid chromatography (FPLC) using anion exchange chromatography and two DEAE columns in series (HiTrap DEAE FF 5 mL column). They were then equilibrated in Buffer A (20 mM Tris-HCl pH 7.5, 20 mM NaCl, 2 mM DTT), and *afumPCNA* was eluted using a linear gradient (0.02 M–0.7 M NaCl). Fractions containing *afumPCNA* were pooled, and ammonium sulphate was added dropwise to a final concentration of 1.5 M from a stock solution of 3 M ammonium sulphate. The sample was allowed to stir gently for 1 h at 4 °C to allow DNA–protein dissociation, and then it was applied to hydrophobic interaction chromatography (HiTrap Phenyl FF (high sub) 5 mL column) and equilibrated in Buffer B (20 mM Tris-HCl pH 7.5, 20 mM NaCl, 2 mM DTT, 0.5 mM EDTA, 1.5 M ammonium sulphate) and eluted in Buffer C (20 mM Tris-HCl pH 7.5, 2 mM DTT, and 0.5 mM EDTA) with a reverse linear gradient (1.5 M–0 M ammonium sulphate). Fractions containing *afumPCNA* were pooled and dialyzed overnight in Buffer A. *afumPCNA* was then applied to a second anion exchange step. The Q Sepharose column (5 mL Q Sepharose FF column (GE)) was equilibrated in Buffer A, and the protein was eluted using a linear gradient (0.02 M–0.7 M NaCl). Fractions containing *afumPCNA* were pooled and dialyzed overnight in 20 mM Tris-HCl pH 7.5, 10% *v/v* glycerol, 2 mM DTT, and 0.5 mM EDTA. The 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.

2.4. Surface Plasmon Resonance Protocol

Surface plasmon resonance (SPR) experiments were performed as previously described [18]. The running buffer used for ligand attachment and analyte-binding experiments was a 10 mM HEPES buffer with 150 mM NaCl, 3 mM EDTA, and 0.05% Tween-20, adjusted to pH 7.4 with 2 M NaOH. A GE CM5 (series S) sensor chip was primed with the running buffer and pre-conditioned with successive injections (2 \times 50 s, 30 μ L/min) of 50 mM NaOH, 10 mM HCl, 0.1% SDS, 0.85% H₃PO₄, and 50 mM glycine pH 9.5, respectively. The surface was then activated with an injection of 0.2 M 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) and 50 mM N-hydroxysuccinimide (NHS) (600 s, 10 μ L/min). *A. fumigatus* PCNA (5 μ L,

12 mg/mL) was diluted into the running buffer (245 μ L). Upon the preactivation of the surface, the protein was further diluted to a final concentration of 25 μ g/mL in 10 mM NaAc (~pH 4.6). This solution was immediately injected over the target flow cell (10 μ L/min) to immobilize ~1500 RU. Both the target and reference flow cells were then blocked with 1.0 M ethanolamine at pH 8.5 (600 s, 10 μ L/min). The chip was left to stabilize before sample injections commenced.

Peptide stock solutions for use in SPR experiments were prepared in MilliQ water. The peptide stock concentration was determined via 205 nm absorbance with NanoDrop2000. The ϵ_{205} for each peptide was calculated using an online calculator (<http://nickanthis.com/tools/a205.html>, accessed on 22 August 2022 [24]); however, additional glycine residue was added to each peptide sequence to account for the terminal amide of the peptides synthesized in-house. The peptide stock solution's concentration was then calculated using Beer's Law.

Steady-state affinity experiments were conducted at a flow rate of 30 μ L/min, with a starting contact time of 40 s and dissociation of 30 s. A 2-fold serial dilution was performed for each peptide, with 8 samples injected sequentially from the lowest to highest concentrations; they were preceded by a buffer-only blank injection. After each injection, the surface was regenerated with 2 M NaCl (2 \times 30 s, 25 μ L/min). All data were analyzed using the GE Biosystems Biacore S200 Evaluation Software, Version 1.0 (Build: 20). All data are summarized in Table 1.

Table 1. p21 Peptide SPR data against afumPCNA in comparison to human PCNA binding affinity as shown in Horsfall 2021 [18]. K_D is the affinity constant. SD, standard deviation. All peptides are C-terminally amidated. Changes to the p21 μ scaffold are indicated in bold. The PIP-box residues are separated from flanking residues with spaces. Conserved PIP positions are underlined. More information can be found in Table S1 and Figure S1.

Name	Sequence	afumPCNA Binding Affinity K_D (nM) \pm SD (nM)	Human PCNA Binding Affinity K_D (nM) \pm SD (nM) [18]
p21(139–160)	¹³⁹ GRKRR <u>Q</u> TSM <u>TDFY</u> HSKRRLIFS ¹⁶⁰	69.7 \pm 20.2	4.3 \pm 1.3
p21 μ	¹⁴¹ KRR <u>Q</u> TSM <u>TDFY</u> HSKR ¹⁵⁵	265.1 \pm 5.9	12.3 \pm 0.5
p21 μ -RD2	¹⁴¹ KRR <u>Q</u> TR <u>ITEYF</u> HSKR ¹⁵⁵	20.3 \pm 6.8	1.1 \pm 0.3
p21 μ -Q144M	¹⁴¹ KRR <u>M</u> TSM <u>TDFY</u> HSKR ¹⁵⁵	41,400 \pm 0.8	1544 \pm 159
p21 μ -T145K	¹⁴¹ KRR <u>Q</u> KSM <u>TDFY</u> HSKR ¹⁵⁵	10,000 \pm 0.7	98 \pm 10.8
p21 μ -T145D	¹⁴¹ KRR <u>Q</u> DSM <u>TDFY</u> HSKR ¹⁵⁵	4100 \pm 0.31	714 \pm 30.4
p21 μ -S146R	¹⁴¹ KRR <u>Q</u> TRM <u>TDFY</u> HSKR ¹⁵⁵	64.4 \pm 18.4	4.3 \pm 1.3
p21 μ -M147L	¹⁴¹ KRR <u>Q</u> TS <u>L</u> TDFY HSKR ¹⁵⁵	382 \pm 51.0	20.5 \pm 1.7
p21 μ -M147I	¹⁴¹ KRR <u>Q</u> TS <u>I</u> TDFY HSKR ¹⁵⁵	37.1 \pm 7.8	11.1 \pm 0.3
p21 μ -D149E	¹⁴¹ KRR <u>Q</u> TSM <u>T</u> E <u>FY</u> HSKR ¹⁵⁵	400.7 \pm 45.6	12.7 \pm 1.4
p21 μ -F150Y	¹⁴¹ KRR <u>Q</u> TSM <u>T</u> D <u>Y</u> Y HSKR ¹⁵⁵	75.2 \pm 18.9	20.2 \pm 0.4
p21 μ -Y151F	¹⁴¹ KRR <u>Q</u> TSM <u>T</u> D <u>F</u> F HSKR ¹⁵⁵	167.2 \pm 19.1	10.6 \pm 1.5
p21 μ -FY150151YF	¹⁴¹ KRR <u>Q</u> TSM <u>T</u> D <u>Y</u> F HSKR ¹⁵⁵	96.4 \pm 19.6	2.2 \pm 0.5

2.5. Protein-Peptide Co-Crystallization Experiments

To form the protein-peptide complex, afumPCNA was mixed with the peptide of interest at a molar ratio of 1:1.2. After incubation on ice for 30 min, the sample was pelleted at 16,000 \times g for 10 min to remove aggregates. Crystals were grown via the hanging drop vapor diffusion method in 24-well plates containing 500 μ L of well solution by mixing 1 μ L of protein and peptide with equal volumes of well solutions. The diffracting crystals of afumPCNA bound to p21 μ grew in 0.2 M Tacsimate pH 4.0 0.1M Na Acetate and 16% PEG 3350 (Hampton Research Aliso Viejo, CA, USA, product code HR2-591) at 16 $^{\circ}$ C after 3 weeks (Figure S2). The diffracting crystals of afumPCNA bound with p21 μ -afumRFC grew in 0.2M Tacsimate pH 4.0 0.1M Na Acetate and 16% PEG 3350 (HR2-591) in a tray

at 16 °C after 3 weeks (Figure S3). Crystals were mounted on cryo-loops, and they were cryoprotected using paratone-N and flash-cooled in liquid nitrogen. Data were collected at 100 K using the MX1 beamline at the Australian Synchrotron (Clayton, VIC, Australia). Diffraction data were indexed and integrated using XDS (X-ray Detector Software), Version January 10, 2022 [25]. Pointless (CCP4i) [26] was used to create an mtz reflection file for scaling. Data were scaled and merged using Aimless (CCP4i) [27] at a resolution of 2.0 Å for afumPCNA bound with p21 μ and 2.30 Å for afumPCNA bound with p21 μ afumRFC. The phase problem was solved via molecular replacements using Phaser MR (CCP4i) [28] and a search model (PDB: 5TUP). Solutions were refined in Phenix Refine [29,30] in iterative rounds with manual rebuilding in Coot [31] (Figures S4 and S5). Data collection and refinement statistics are summarized in Table S2. The final structures of afumPCNA bound with p21 μ and p21 μ RFC are deposited on the RCSB database under accession numbers 8GJF and 8GJ5, respectively.

2.6. Computational Modelling

The models of peptides bound to afumPCNA were constructed using the structure of afumPCNA bound with the p21 μ peptide (PDB: 8GJF) as a starting template, and the necessary, deleted, and unresolved side chains of residues were modelled into the computational structure.

The manual refinement of the computational linker was carried out in Coot [31]. Energy minimisation/annealing ($n = 30$) for refinement was carried out in ICM-Pro Molsoft [32,33]. Refined models were analysed using PyMOL Version 1.2 [19] to validate the model by comparing it against p21 μ (PDB: 8GJF) and assessing side chain interactions. The resulting structures were visualized in PyMOL [19], and they are depicted in Figures S4–S17.

3. Results

3.1. A p21 Peptide Library Interacts with afumPCNA in a Similar Trend Compared to hPCNA

The p21-derived peptide (139–160) (Table 1) [21] was previously shown to bind to afumPCNA via fluorescence polarization with a K_D of 3.1 μ M [14]. Here, we build on this observation and interrogate the binding of afumPCNA with respect to various peptides. Previously, a shorter scaffold of this 22 amino acid p21 peptide was derived and synthesized: p21 μ (141–155). It is 14 amino acids long, and it retained high-affinity binding, which was used to construct a library with variations relative to PIP-box residues [18] as a rational starting point for the investigation of the fungal binding site.

The p21 PIP-box contains a conserved glutamine residue that binds a conserved hydrophobic pocket on hPCNA; this was shown to be valuable in binding afumPCNA. The p21 glutamine residue (Gln₁₄₄), which binds the 'Q pocket' of the hydrophobic patch on human PCNA, forms two hydrogen bonds relative to the carbonyl backbone moieties of residues Ala₂₅₂ and Pro₂₅₃ of the PCNA main chain. These residues are conserved in the afumPCNA sequence (Figure 2); hence, as for the hPCNA investigation, the modification of Gln₁₄₄ into Met, as in p21 μ -Q144M, reduces the binding of peptide p21 μ -Q144M to afumPCNA from 265.1 nM to 41,400 nM.

The importance of residues in the non-conserved position of the PIP-box was previously shown to be important in binding to the PCNA surface in the human system [34]. To probe the effect of altering amino acids in this region of the PIP-box and its affinity to afumPCNA, the peptide was altered from Ser₁₄₆ to an Arg residue. This produced an affinity of K_D 64.4 nM, improved from the p21 μ binding affinity of K_D 265.1 nM. This was also previously observed in hPCNA, which was hypothesised to be due to an increase in side chain length [18]. The Ser to Arg variation changes the distance between residues, strengthening the intramolecular hydrogen bond to the carbonyl of Asp₁₄₉, stabilising the peptide's 3₁₀ helical structure (Figure S10). This suggests that lengthening the side chain would also improve the binding of the p21 μ -D149E peptide to afumPCNA; however, an Asp₁₄₉ modification to Glu showed reduced binding affinity with a K_D of 400.6 nM,

Table 2. Sequences of candidate fungal protein PIP-boxes in comparison to established human protein PIP-boxes. Sequences that fit the model of an eight-residue section with Q at the beginning, a hydrophobic residue in the middle, and aromatic residues at the end are shown in bold. Residues that are found to be identical between human and fungal PIP-boxes are underlined.

Protein Name	<i>Homo Sapiens</i> PIP-Box Sequence	<i>Aspergillus fumigatus</i> PIP-Box Sequence Candidates
FEN1	<u>Q</u> RRLD DF F	<u>Q</u> SRLEG FF
RFC1	<u>M</u> DIR K FF	<u>M</u> PTDIR N FF
DNA Polymerase	<u>Q</u> V S ITG F F	<u>Q</u> KEL S RFD
POLD3	<u>Q</u> R S IM S FF	<u>Q</u> R V RS I AS F F
DNA Ligase	<u>Q</u> R S IM S FF	<u>Q</u> R V RS I AS F F

The major differences between human and fungal PIP-boxes appear at the non-conserved residues of the canonical PIP-box sequence. In particular, in the p21 μ afumDNALig and p21 μ afumRFC sequence, additional residues were interspaced with conserved residues, possibly interfering with the alignment of the canonical structure and the contact with the protein's surface. To elaborate, the RFC1 PIP-box found in humans has the correct amount of non-conserved residues, but the candidate for *A. fumigatus* has two extra residues between the conserved methionine residue and the conserved hydrophobic residue isoleucine (Table 2).

Each *A. fumigatus* PIP-box sequence exhibits a Gln₁₄₄-conserved residue, except the RFC sequence. This significant difference between the human and fungal candidate PIP-box leads to the hypothesis that the p21 μ afumRFC peptide could not bind to afumPCNA with high affinity. However, it is surprisingly bound with <100 nM affinity.

3.3. X-ray Crystallography Study of the p21 μ Peptide Bound to afumPCNA

The first co-crystal structure of afumPCNA bound with the p21 μ scaffold peptide was solved at 2.0 Å resolution (PDB: 8GJF) in order to examine the details of the binding interaction (Figure 3). The structure shows that the overall fold of the p21 μ peptide bound to the surface of afumPCNA is similar to the structure of the p21 μ peptide bound to hPCNA (PDB: 7KQ1), as illustrated by the RMSD value of 0.939 Å. The p21 μ peptide in the afumPCNA structure (PDB: 8GJF) displays a notable charged interaction between Arg₁₄₃ and Glu₁₄₉, a 3.2 Å salt bridge interaction (Figure 4b) that was not previously shown as the extended Arg₁₄₃ side chain was not present in the hPCNA crystal structure (Figure 4a), illuminating a new interaction that also strengthens the 3_{10} helical structure.

Differences in the protein sequence of afumPCNA and hPCNA account for the shift in affinity with the conservation of the secondary structure of the peptide. There is structural conservation around the 3_{10} helical secondary structure and PIP-box, with more variability on the N- and C-terminus (Figure 4c) likely due to the mobility of the ends of the peptide and changes in IDCL residues. Previous literature interpreting the difference in the binding of p21-based peptides to afumPCNA compared to hPCNA used molecular dynamics to illustrate that the weakness of the p21 peptide (139–160) because afumPCNA came from differences in these protein binding domain residues [14]. One prominent example is residue His125, which forms an antiparallel β -sheet with the C-terminal residues of p21 peptide on hPCNA (139–160); however, in the afumPCNA structure, afumPCNA His125 obstructs the formation of a favorable side chain hydrogen bond with the His152 side chain of the p21 peptide (139–160). This has the effect of pushing the C-terminus to be quite distant from the protein's surface while not forming the hydrogen bonds of the β -sheet as observed in the human structure. There is also the loss of 3.4 Å hydrogen bond interactions between hPCNA Gln131 and the Tyr151 phenol of p21 μ in afumPCNA as this residue is Thr131.

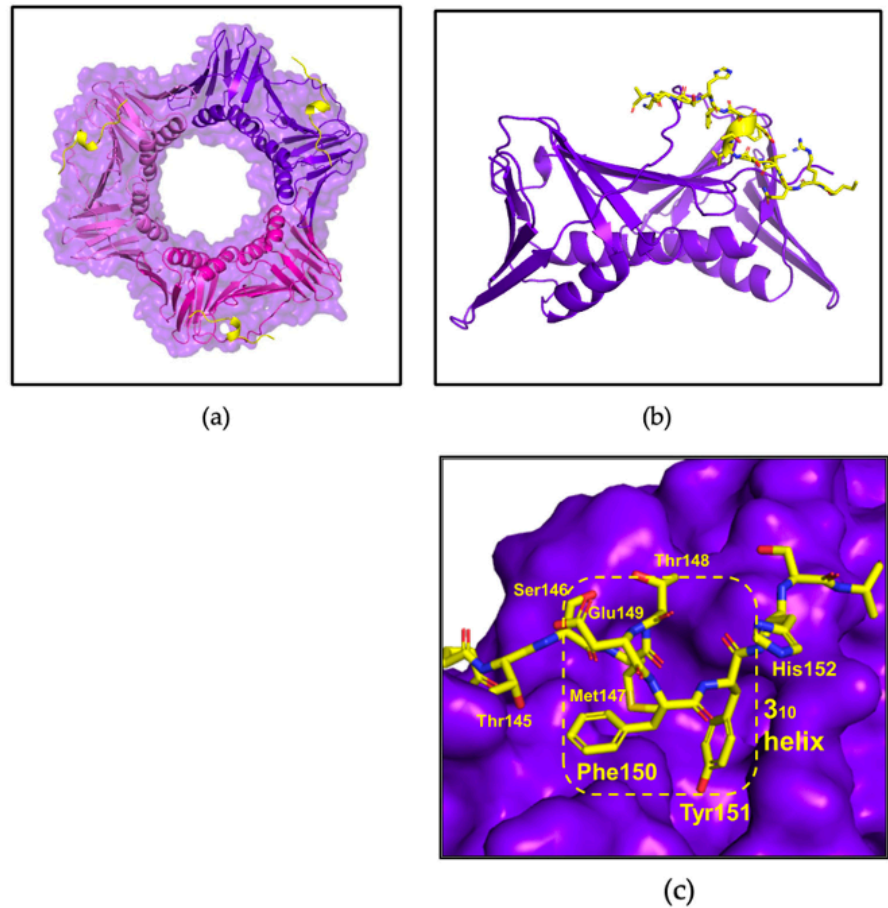


Figure 3. afumPCNA bound with p21 μ structure (PDB: 8GJF). (a) Trimer: afumPCNA surface shown in purple and p21 μ shown in olive green. (b) Monomer: afumPCNA shown in purple and p21 μ shown in olive green. (c) PIP-box binding site: 3_{10} helical structure outlined in yellow and p21 μ peptide shown in olive green. Made using Pymol [19]. Crystallographic information can be found in Table S2 and Figures S2 and S4.

3.4. X-ray Crystallography Study of afumPCNA and p21 μ -afumRFC Reveals a Ring-like Structure

The co-crystal structure of afumPCNA bound with the p21 μ -afumRFC peptide was solved at a resolution of 2.30 Å in order to examine the structural features of an Afum-derived PIP-box (Figure 5a,b).

The KRRMP amino acids of the p21 μ -afumRFC peptide (Table 3) fold over the PIP-box, not interacting with the afumPCNA's surface (Figure 5c). This is caused by a change in the sequence of the PIP-box compared to the human RFC sequence (Table 3); the inclusion of a proline residue causes a kink, and the backbone carbonyl of ProXXX interacts with the Asp₁₄₇ backbone amide and Asn₁₅₀ residue side chain to stabilize the turn in the peptide (Figure 4c). The second arginine, Arg₁₄₃, is located close to the third arginine, Arg₁₄₉, producing a loop structure (143–149). The Met₁₄₄ side chain does not interact with the afumPCNA surface pocket as Gln₁₄₄ in the p21 μ PIP-box does with hPCNA. The Met₁₄₄ backbone amide does interact with the Asn₁₅₀ side chain and in turn also interacts with the Asp₁₄₇ side chain, supporting the 3_{10} helix via extra contacts that hold the compact structure (Figure 4c).

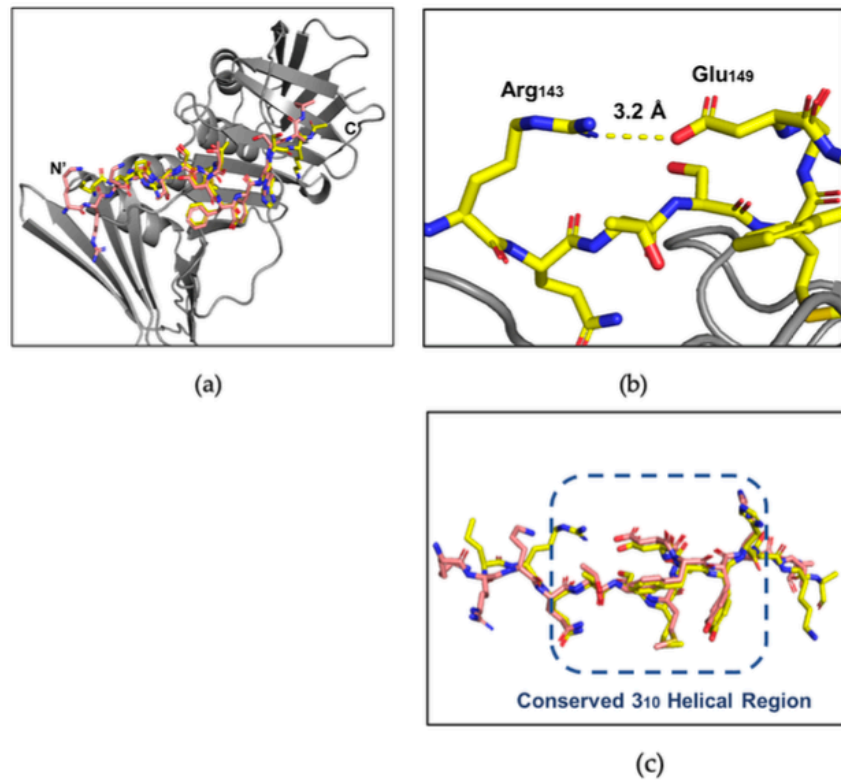


Figure 4. Comparison of the p21μ peptide bound to hPCNA (PDB: 7KQ1) and p21μ peptide bound to afumPCNA (PDB: 8GJF). (a) hPCNA shown in grey, p21μ (PDB: 7KQ1) shown in salmon, and p21μ (PDB: 8GJF) shown in yellow. (b) Arg₁₄₃ and Glu₁₄₉ salt bridge interaction in p21μ (PDB: 8GJF) structure. (c) Conserved 3₁₀ helical and PIP-box region of p21μ peptides, residues 144–152: p21μ (PDB: 7KQ1) shown in salmon and p21μ (PDB: 8GJF) shown in yellow (RMSD: 0.939 Å).

Table 3. Candidate fungal protein Peptide SPR data against afumPCNA. Peptide p21μafumRFC was also tested against human PCNA. Tested in triplicate. KD is the affinity constant. SD, standard deviation. All peptides are C terminally amidated. Changes to the p21μ scaffold are indicated in bold. More information can be found in Table S1.

Name	Sequence	afumPCNA Binding Affinity		Human PCNA Binding Affinity	
		K _D (nM) ± K _D SD (nM)	K _D (nM) ± K _D SD (nM)	K _D (nM) ± K _D SD (nM)	K _D (nM) ± K _D SD (nM)
p21μafumDNALIG	¹⁴¹ KRRQRVRSIASFFHSKR ¹⁵⁷	458 ± 117.77	-	-	-
p21μafumDNAPOL	¹⁴¹ KRRQKELSRFDFHSKR ¹⁵⁶	659.3 ± 105.8	-	-	-
p21μafumFEN1	¹⁴¹ KRRQSRLEGFFHSKR ¹⁵⁵	713 ± 56.9	-	-	-
p21μafumRFC	¹⁴¹ KRRMPTDIRNFFHSKR ¹⁵⁶	94.84 ± 8.76	295 ± 6.9	-	-

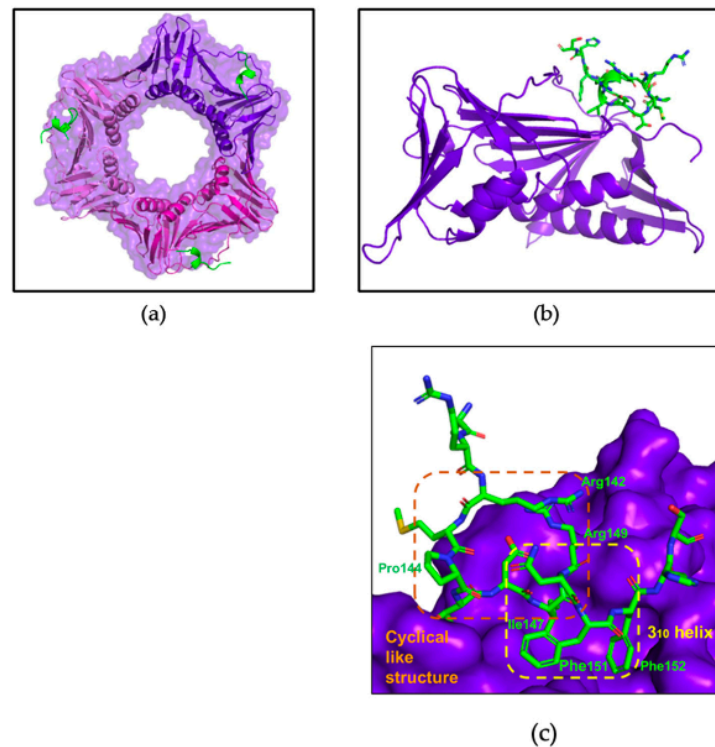


Figure 5. afumPCNA bound with p21 μ -afumRFC structure (PDB:8GJ5). (a) Trimer: afumPCNA surface shown in purple and p21 μ -afumRFC shown in green. (b) Monomer: afumPCNA shown in purple and p21 μ -afumRFC shown in green. (c) PIP-box binding site: 3₁₀ helical structure outlined in yellow and cyclical-like secondary structure outlined in orange. p21 μ -afumRFC peptide shown in green. Made using PyMOL [19]. Crystallographic information can be found in Table S2 and Figures S3 and S5.

4. Discussion

4.1. The p21 Peptide Library Interacts with afumPCNA with Similar Affinity and Structural Trends as hPCNA

The co-crystal structure of afumPCNA bound to the p21 μ peptide supports the hypothesis that *A. fumigatus* adopts a PIP-box sequence as a method for proteins to interact with afumPCNA.

The SPR (Table 1) results indicate that the binding of the p21 peptide library to afumPCNA generally follows the same trends seen in the hPCNA investigation [18]. This includes modifications at similar positions that cause similar changes in binding affinity across the two PCNA species.

It was previously hypothesized from results in molecular dynamics studies [14] that differences in the residues of the IDCL of afumPCNA and hPCNA could provide specificity for afumPCNA over hPCNA for a peptidomimetic inhibitor. However, this does not seem to be supported, as high-affinity peptidomimetic residues have a similar trend in affinity for hPCNA and afumPCNA. This is directly shown in the binding of rational design mutant 2 (RD2) to afumPCNA (Figure S6) [18], which was specifically designed for hPCNA, but it binds to the surface of afumPCNA with the same structure and a K_D of 20.3 nM (Table 1). The two PCNA species cannot be differentiated in specificity through the changes in PIP-box residues via the modifications investigated here.

4.2. Fungal Protein Replication Factor C PIP-Box Candidate Peptide Has a High Affinity for afumPCNA

Previously, a p21 peptide with the human RFC PIP-box, p21 μ -RFC, which has a seven-amino-acid PIP-box, MDIRKFF, was investigated to understand variations in the canonical sequence, and it was found to have a K_D value of 145 nM [18]. It was hypothesised here that this affinity was due to the position of residues Ile₁₄₇, Phe₁₅₀, and Phe₁₅₁, which form a hydrophobic three-pronged plug that inserts into the hydrophobic cleft of hPCNA [18]. Via computational modelling, it was observed that this would result in the extension of the arginine residue of position four over the conserved glutamine pocket in order to interact with hPCNA residue Val₄₅ [18].

The p21 μ -afumRFC PIP-box has an affinity for afumPCNA of less than 100 nM. This may be accounted for solely by the Ile₁₄₇, Phe₁₅₀, and Tyr₁₅₁ residues (Table 1), as previous research has shown these conserved residues to be highly favourable, especially Ile₁₄₇. The p21 μ -afumRFC PIP-box has a lower affinity for hPCNA than afumPCNA (Table 3). This is not only attributed to the lack of Gln₁₄₄, similarly to that of the human RFC peptide, but also the extra residues (Pro₁₄₅, Thr₁₄₆, and Asn₁₅₀) of the PIP-box for which the canonical positions do not exactly fit the conserved motif. It was hypothesised [18,35] that the Gln₁₄₄ residue was essential to the p21 peptide with respect to high-affinity binding; hence, it is present in p21 μ -RD2. The Gln₁₄₄ of p21 is known to contribute significantly to the binding affinity of hPCNA, as a Gln144Ala modification was not able to effectively inhibit DNA replication in vitro [18,35]. Gln₁₄₄ was considered at first to remain important in the Afum binding since the modification of p21 μ -Q144M decreases the binding affinity of the p21 μ -Q144M peptide to afumPCNA from 265.1 nM to 41.4 μ M. This is solely attributed to the single residue change as the secondary structure is maintained (Figure S7). However, here, its importance has still been questioned for the afumPCNA binding domain due to its absence in the candidate fungal RFC PIP-box. The attributes of p21 μ -afumRFC affinity for afumPCNA, although the canonical PIP-box is not followed, appear to be the unique secondary structure that is formed, which is discussed further.

4.3. The p21 μ afumRFC Peptide Has a Unique Structure That Could Be Exploited for an Antifungal Treatment

Peptidomimetic drug pipelines often reach the point of requiring a cell-permeable mechanism; a convenient method of improving cell uptake is via the cyclisation of the peptide. Cyclic peptides have been shown to enter the mammalian cell cytosol via multiple mechanisms, including passive diffusion, which is facilitated predominantly by hydrophobic side chains and small amino acid size (approximately 10 amino acids long), and endocytic uptake and endosomal escape [36].

In a structure such as the p21 PIP-box, which creates a 3₁₀ helix, constraining this structure via cyclisation would allow the preorganization of the backbone and reduce the entropic cost of forming the secondary structure upon binding. Another advantage is that cyclic peptides may have improved cell permeability, which has been investigated in previous studies [37,38]. The investigation studied such macrocycles bound hPCNA with K_D values ranging from 570 nM to 3.86 μ M, with a bimeane-constrained peptide proving to be the most potent. This peptide was also cell-permeable and localized to the cell cytosol of breast cancer cells (MDA-MB-468). The 3₁₀-helical structure was present in the computationally modelled structure. However, the analysis showed the peptide did not have a rigid 3₁₀ helix in the solution when not bound to PCNA as NMR revealed it was not present in the solution [38]. This suggested that the pre-defining of a peptide backbone may not improve PCNA binding affinity. A 'linker' is a covalent tether that connects two distant parts of a peptide sequence to create a bridge and consequently preorganise the peptide backbone into a conformation that is suitable to bind to its target. It has been reported that a linker that affords flexibility in its cyclised structure may be preferable to enable the peptide to adopt its ideal conformation upon binding. This could be provided using the p21 μ afumRFC peptide by constraining the Arg₁₄₃ and Arg₁₄₉ residues as a linker to cyclise the peptide.

The p21 μ -afumRFC secondary structure looks as if it naturally mimics a ring such as that of the bimane structure (Figure 6a). This could be used as a scaffold for a peptidomimetic, which could be improved to be fungal-cell-penetrable, as it has already been shown to not interfere with the 3₁₀ helical turn upon binding. The two Arg residues can be replaced to create a linkage that, based on the X-ray crystallography structure, would not interrupt the 3₁₀-helical conformation and side chain exposure required for binding to afumPCNA, as these are 3.5 Å distance apart in the naturally forming architecture. The ability to outcompete afumPCNA's binding in the cell may be achieved via the incorporation of select p21 μ -RD2 mutant sequence residues in the p21 μ -afumRFC peptide, such as the combination of Tyr₁₅₀ and Phe₁₅₁ aromatic residues, which was shown to be essential in the affinity assay (Table 1).

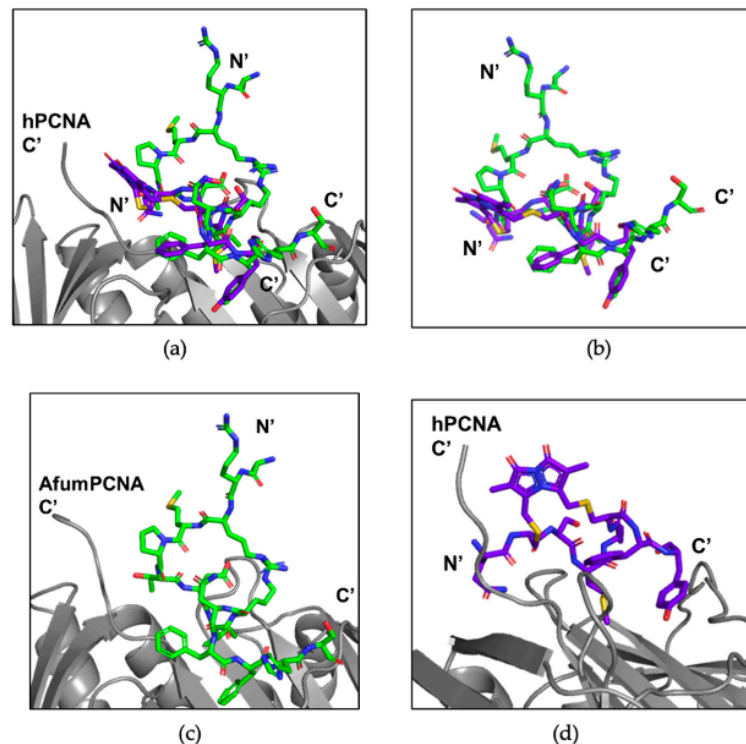


Figure 6. Previous p21-constrained peptides in comparison to the new p21 μ afumRFC peptide crystal structure. (a) p21 μ afumRFC peptide (green) crystal structure bound to the afumPCNA monomer (grey). p21 μ afumRFC peptide (green) produced an affinity of 295 ± 6.9 nM to hPCNA. (b) The Bimane peptide (purple) docked to the hPCNA monomer (grey), as shown in the computational model, produced an affinity of 570 ± 30 nM to hPCNA [38]. (c) p21 μ afumRFC peptide (green) superimposed over the Bimane peptide (purple) and hPCNA (grey) computational model. (d) p21 μ afumRFC peptide (green) superimposed over the Bimane peptide (purple) computational model.

The bimane peptide (Figure 6b,c) has a unique interaction in which the bimane linker interacts with the C-terminal end of PCNA. This is also achieved in the RFC peptide via the interaction of the proline residue interacting with the C-terminal end of afumPCNA. The key difference in the scaffolds is that the RFC peptide also carried out interactions on the other side of the PIP-box binding domain through the Arg₁₄₉ side chain. This is believed to

achieve a more ideal surface packing of the PIP-box onto afumPCNA than previous cyclical peptides have achieved.

Adding a cell-penetrating peptide (CPP) or fungal-specific peptide to the N- or C-terminus of this structure would, based on the X-ray crystallographic structure, not interfere with PIP-box binding and the secondary structure, a problem found for other investigated linkers incorporated into p21 PIP-box peptides. Although high levels of translocation were typically associated with the toxicity of peptides towards fungal cells, SynB, an 18-amino-acid-long peptide, has been found in previous studies of CPPs to be specific for fungal cells with respect to efficiently translocating into the human fungal pathogen *Candida albicans* at concentrations that led to minimal toxicity [39]. Lowered toxicity is vital for experimental studies in assessing the specification of afumPCNA inhibition over fungal cell toxicity.

5. Conclusions

Here, we present the first structure of a p21 PIP-box peptide bound to *A. fumigatus* PCNA, as well as fungal PIP-box candidates, demonstrating the hypothesis that the fungal replication model uses a PIP-box sequence as a method for binding to fungal PCNA.

A high-affinity rational design for a potential cell-permeable peptidomimetic is presented via the combination of a cyclised structure of the p21 μ afumRFC peptide. Via full cyclisation and the incorporation of select p21 μ -RD2 mutant sequence residues, this peptide could be used in the next stages of the drug discovery pipeline as a potential fungal therapeutic. This could be carried out via the addition of a linker to the cyclized secondary structure. Future work will focus on specific fungal cell permeability via the utilisation of the N-terminus of the peptide, which makes no contact with the surface of PCNA and cannot interfere with the helical and cyclized 3₁₀ structure.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/jof9111098/s1>.

Author Contributions: B.C.V. is responsible for conceptualization, methodology, formal analysis, investigation, writing—original draft preparation, and review and editing final. A.J.H. is responsible for resources, investigation, and methodology. J.L.P. is responsible for intellectual support, review and editing. J.B.B. is responsible for resources, writing—review and editing visualization, supervision, project administration, and funding. A.D.A. is responsible for supervision and resources. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article or Supplementary Materials. The data presented in this study are available in Supplementary Materials.

Acknowledgments: B. Vandborg and A. Horsfall and J. Pederick were supported by an Australian Government Research Training Program stipend scholarship. This research was undertaken in part using the MX1 and MX2 beamlines at the Australian Synchrotron and part of ANSTO and made use of the Australian Cancer Research Foundation detector. Special thanks are given to William Walters, Paul Jr Walters, and Polly Walters for their crystallography consultation. Special thanks are given to Jakeb Vandborg for their scientific consultation.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bitar, D.; Lortholary, O.; Le Strat, Y.; Nicolau, J.; Coignard, B.; Tattevin, P.; Che, D.; Dromer, F. Population-based analysis of invasive fungal infections, France, 2001–2010. *Emerg. Infect. Dis.* **2014**, *20*, 1149–1155. [[CrossRef](#)]
2. McNeil, M.; Nash, S.L.; Hajjeh, R.A.; Phelan, M.A.; Conn, L.A.; Plikaytis, B.D.; Warnock, D.W. Trends in mortality due to invasive mycotic diseases in the United States, 1980–1997. *Clin. Infect. Dis.* **2001**, *33*, 641–647. [[CrossRef](#)]
3. Latgé, J.-P.; Chamilos, G. *Aspergillus fumigatus* and Aspergillosis in 2019. *Clin. Microbiol. Rev.* **2019**, *33*, 00140–18. [[CrossRef](#)]

4. Rhodes, J.; Abdolrasouli, A.; Dunne, K.; Sewell, T.R.; Zhang, Y.; Ballard, E.; Brackin, A.P.; van Rhijn, N.; Chown, H.; Tsiotsopoulou, A.; et al. Population genomics confirms acquisition of drug-resistant *Aspergillus fumigatus* infection by humans from the environment. *Nat. Microbiol.* **2022**, *7*, 663–674. [[CrossRef](#)]
5. Yoon, H.; Choi, H.Y.; Kim, Y.K.; Song, Y.J.; Ki, M. Prevalence of fungal infections using National Health Insurance data from 2009–2013, South Korea. *Epidemiol. Health* **2014**, *36*, e2014017. [[CrossRef](#)]
6. Van de Veerdonk, F.L.; Gresnigt, M.S.; Romani, L.; Netea, M.G.; Latgé, J.-P. *Aspergillus fumigatus* morphology and dynamic host interactions. *Nat. Rev. Microbiol.* **2017**, *15*, 661–674. [[CrossRef](#)]
7. Gallien, S.; Fournier, S.; Porcher, R.; Bottero, J.; Ribaud, P.; Sulahian, A.; Molina, J.M. Therapeutic outcome and prognostic factors of invasive aspergillosis in an infectious disease department: A review of 34 cases. *Infection* **2008**, *36*, 533–538. [[CrossRef](#)]
8. Sugui, J.; Kwon-Chung, K.J.; Juvvadi, P.R.; Latge, J.P.; Steinbach, W.J. *Aspergillus fumigatus* and related species. *Cold Spring Harb Perspect. Med.* **2015**, *5*, a019786. [[CrossRef](#)]
9. Steinbach, W. Are we there yet? Recent progress in the molecular diagnosis and novel antifungal targeting of *Aspergillus fumigatus* and invasive aspergillosis. *PLoS Pathog.* **2013**, *9*, e1003642. [[CrossRef](#)]
10. Arastehfar, A.; Carvalho, A.; Houbraken, J.; Lombardi, L.; Garcia-Rubio, R.; Jenks, J.D.; Rivero-Menendez, O.; Aljohani, R.; Jacobsen, I.D.; Berman, J.; et al. *Aspergillus fumigatus* and aspergillosis: From basics to clinics. *Stud. Mycol.* **2021**, *100*, 100115. [[CrossRef](#)]
11. Ben-Ami, R.; Lewis, R.E.; Kontoyiannis, D.P. Enemy of the (immunosuppressed) state: An update on the pathogenesis of *Aspergillus fumigatus* infection. *Br. J. Haematol.* **2010**, *150*, 406–417. [[CrossRef](#)]
12. Resendiz Sharpe, A.; Lagrou, K.; Meis, J.F.; Chowdhary, A.; Lockhart, S.R.; Verweij, P.E.; ISHAM/ECMM *Aspergillus* Resistance Surveillance Working Group. Triazole resistance surveillance in *Aspergillus fumigatus*. *Med. Mycol.* **2018**, *56*, S83–S92. [[CrossRef](#)] [[PubMed](#)]
13. Edouarzin, E.; Horn, C.; Paudyal, A.; Zhang, C.; Lu, J.; Tong, Z.; Giaever, G.; Nislow, C.; Veerapandian, R.; Hua, D.H.; et al. Broad-spectrum antifungal activities and mechanism of drimane sesquiterpenoids. *Microb. Cell* **2020**, *7*, 146–159. [[CrossRef](#)]
14. Marshall, A.C.; Kroker, A.J.; Murray, L.A.; Gronthos, K.; Rajapaksha, H.; Wegener, K.L.; Bruning, J.B. Structure of the sliding clamp from the fungal pathogen *Aspergillus fumigatus* (AfumPCNA) and interactions with human p21. *FEBS J.* **2017**, *284*, 985–1002. [[CrossRef](#)] [[PubMed](#)]
15. Maga, G.; Hubscher, U. Proliferating cell nuclear antigen (PCNA): A dancer with many partners. *J. Cell Sci.* **2003**, *116*, 3051–3060. [[CrossRef](#)]
16. Sakakura, C.; Hagiwara, A.; Tsujimoto, H.; Ozaki, K.; Sakakibara, T.; Oyama, T.; Takahashi, T. The anti-proliferative effect of proliferating cell nuclear antigen-specific antisense oligonucleotides on human gastric cancer cell lines. *Surg. Today* **1995**, *25*, 184–186. [[CrossRef](#)] [[PubMed](#)]
17. Naryzhny, S.; Lee, H. Characterization of proliferating cell nuclear antigen (PCNA) isoforms in normal and cancer cells: There is no cancer-associated form of PCNA. *FEBS Lett.* **2007**, *528*, 4917–4920. [[CrossRef](#)] [[PubMed](#)]
18. Horsfall, A.J.; Vandborg, B.A.; Kowalczyk, W.; Chav, T.; Scanlon, D.B.; Abell, A.D.; Bruning, J.B. Unlocking the PIP-box: A peptide library reveals interactions that drive high-affinity binding to human PCNA. *J. Biol. Chem.* **2021**, *296*, 100773. [[CrossRef](#)]
19. Schrodinger LLC. *The PyMOL Molecular Graphics System*, Version 1.2r3pre; Schrodinger LLC: New York, NY, USA, 2015.
20. Warbrick, E. PCNA binding through a conserved motif. *Bioessays* **1998**, *20*, 195–199. [[CrossRef](#)]
21. Gulbis, J.M.; Kelman, Z.; Hurwitz, J.; O'Donnell, M.; Kuriyan, J. Structure of the C terminal region of p21(WAF1/CIP1) complexed with human PCNA. *Cell* **1996**, *87*, 297–306. [[CrossRef](#)]
22. Prestel, A.; Wichmann, N.; Martins, J.M.; Marabini, R.; Kassem, N.; Broendum, S.S.; Otterlei, M.; Nielsen, O.; Willemoes, M.; Ploug, M.; et al. The PCNA interaction motifs revisited: Thinking outside the PIP-box. *Cell. Mol. Life Sci.* **2019**, *76*, 4923–4943. [[CrossRef](#)] [[PubMed](#)]
23. Vandborg, B.; Holroyd, D.L.; Pukala, T.; Bruning, J.B. Production of recombinant human proliferating cellular nuclear antigen (PCNA) for structural and biophysical characterization. *Protein Expr. Purif.* **2023**, *212*, 106353. [[CrossRef](#)] [[PubMed](#)]
24. Anthis, N.J.; Clore, G.M. Sequence-specific determination of protein and peptide concentrations by absorbance at 205 nm. *Protein Sci.* **2013**, *22*, 851–858. [[CrossRef](#)]
25. Kabsch, W. XDS (X-ray detector software). *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 125–132. [[CrossRef](#)] [[PubMed](#)]
26. Winn, M.D.; Ballard, C.C.; Cowtan, K.D.; Dodson, E.J.; Emsley, P.; Evans, P.R.; Keegan, R.M.; Krissinel, E.B.; Leslie, A.G.W.; McCoy, A.; et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*, 235–242. [[CrossRef](#)]
27. Potterton, E.; Briggs, P.J.; Turkenburg, M.G.W.; Dodson, E. A graphical user interface to the CCP4 program suite. *Acta Crystallogr. D Biol. Crystallogr.* **2003**, *59*, 1131–1137. [[CrossRef](#)]
28. McCoy, A.J.; Grosse-Kunstleve, R.W.; Adams, P.D.; Winn, M.D.; Storoni, L.C.; Read, R.J. Phaser crystallographic software. *J. Appl. Crystallogr.* **2007**, *40*, 658–674. [[CrossRef](#)]
29. Afonine, P.V.; Grosse-Kunstleve, R.W.; Echols, N.; Headd, J.J.; Moriarty, N.W.; Mustyakimov, M.; Terwilliger, T.C.; Urzhumtsev, A.; Zwart, P.H.; Adams, P.D. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **2012**, *68*, 352–367. [[CrossRef](#)]
30. Liebschner, D.; Afonine, P.V.; Baker, M.L.; Bunkóczi, G.; Chen, V.B.; Croll, T.I.; Hintze, B.; Hung, L.W.; Jain, S.; McCoy, A.J.; et al. Phenix: Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in phenix. *Acta Crystallogr. D Biol. Crystallogr.* **2019**, *75*, 861–877. [[CrossRef](#)]

31. Emsley, P.; Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 2126–2132. [[CrossRef](#)]
32. Abagyan, R.; Totrov, M. Biased probability Monte Carlo conformational searches and electrostatic calculations for peptides and proteins. *J. Mol. Biol.* **1994**, *235*, 983–1002. [[CrossRef](#)] [[PubMed](#)]
33. Abagyan, R.; Totrov, M.; Kuznetsov, D. ICM—A 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. [[CrossRef](#)]
34. Horsfall, A.; Abell, A.; Bruning, J. Targeting PCNA with peptide mimetics for therapeutic purposes. *ChemBioChem* **2019**, *21*, 442–450. [[CrossRef](#)]
35. Boehm, E.; Washington, M.T. R.I.P. to the PIP: PCNA-binding motif no longer considered specific. *BioEssays* **2016**, *38*, 1117–1122. [[CrossRef](#)] [[PubMed](#)]
36. Dougherty, P.G.; Ashweta, S.; Dehua, P. Understanding Cell Penetration of Cyclic Peptides. *Chem. Rev.* **2019**, *119*, 10241–10287. [[CrossRef](#)] [[PubMed](#)]
37. Wegener, K.L.; McGrath Amy, E.; Dixon Nicholas, E.; Oakley Aaron, J.; Scanlon Denis, B.; Abell Andrew, D.; Bruning John, B. Rational Design of a 310 -Helical PIP-Box Mimetic Targeting hPCNA, the Human Sliding Clamp. *Eur. J. Chem.* **2018**, *24*, 11325–11331. [[CrossRef](#)]
38. Horsfall, A.J.; Vandborg, B.A.; Kikhtyak, Z.; Scanlon, D.B.; Tilley, W.D.; Hickey, T.E.; Bruning, J.B.; Abell, A.D. A cell permeable bimane-constrained PCNA-interacting peptide. *RSC Chem. Biol.* **2021**, *2*, 1499–1508. [[CrossRef](#)]
39. Gong, Z.; Karlsson, A.J. Translocation of cell-penetrating peptides into *Candida* fungal pathogens. *Protein Sci.* **2017**, *26*, 1714–1725. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Electronic Supplementary Information

Table S1. p21 Peptide Surface Plasmon Resonance (SPR) data against AfumPCNA. The 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. SD, standard deviation; On/Off, indicate times for contact (on) and dissociation phases (off) of each run. All peptides are C-terminally amidated. Peptides notated with * were designed and synthesised by Horsfall 2021 [1]. Changes to p21 μ scaffold are indicated in bold.

Name	Sequence	ϵ_{205}^*	Top Conc (nM)	Affinity K_D (nM)	K_D SD (nM)	On/Off (s)
p21 (139-160)	¹³⁹ GRKRRQTSMTDFYHSK RRLIFS	98620	500	69.71	20.21	40/30
*p21 μ	KRRQTSMTDFYHSKR	67860	1200	265.1	5.97	40/30
*p21 μ -RD2	KRRQ TRITE YFHSKR	67380	150	20.33	6.84	40/30
*p21 μ -Q144M	KRR MT SMTDFYHSKR	69290	32,000	41,400	820	40/30
*p21 μ -T145K	KRR QK SMTDFYHSKR	67860	128,000	10,000	65	40/30
*p21 μ -T145D	KRR QD SMTDFYHSKR	67860	64,000	4110	310	40/30
*p21 μ -S146R	KRR QTR MDFYHSKR	69210	1000	64.4	18.4	40/30
*p21 μ -M147L	KRR QTS LDFYHSKR	66030	5171.2	382	51.04	40/30
*p21 μ -M147I	KRR QTS ITDFYHSKR	66030	500	37.05	7.84	40/30
*p21 μ -D149E	KRR QTS MTEFYHSKR	67860	1000	400.66	45.54	40/30
*p21 μ -F150Y	KRR QTS MTDYFHSKR	65340	646.4	75.20	18.92	40/30
*p21 μ -Y151F	KRR QTS M DF FHSKR	70380	1292.8	167.2	19.07	40/30
*p21 μ -FY150151YF	KRR QTS M DY FHSKR	67860	1292.8	96.4	19.6	40/30
p21 μ -AfumDNALIG	KRR QVR SIASFFHSKR	74030	8000	458	117.77	40/30
p21 μ -AfumDNAPOL	KRR QK ELSRFDFFHSKR	69900	10,000	659.3	105.8	40/30
p21 μ -AfumFEN1	KRR QSR LEGFFHSKR	67120	10,000	713	56.9	40/30
p21 μ -AfumRFC	KRR MP TDIRNFFHSKR	71730	2000	94.84	8.76	40/30

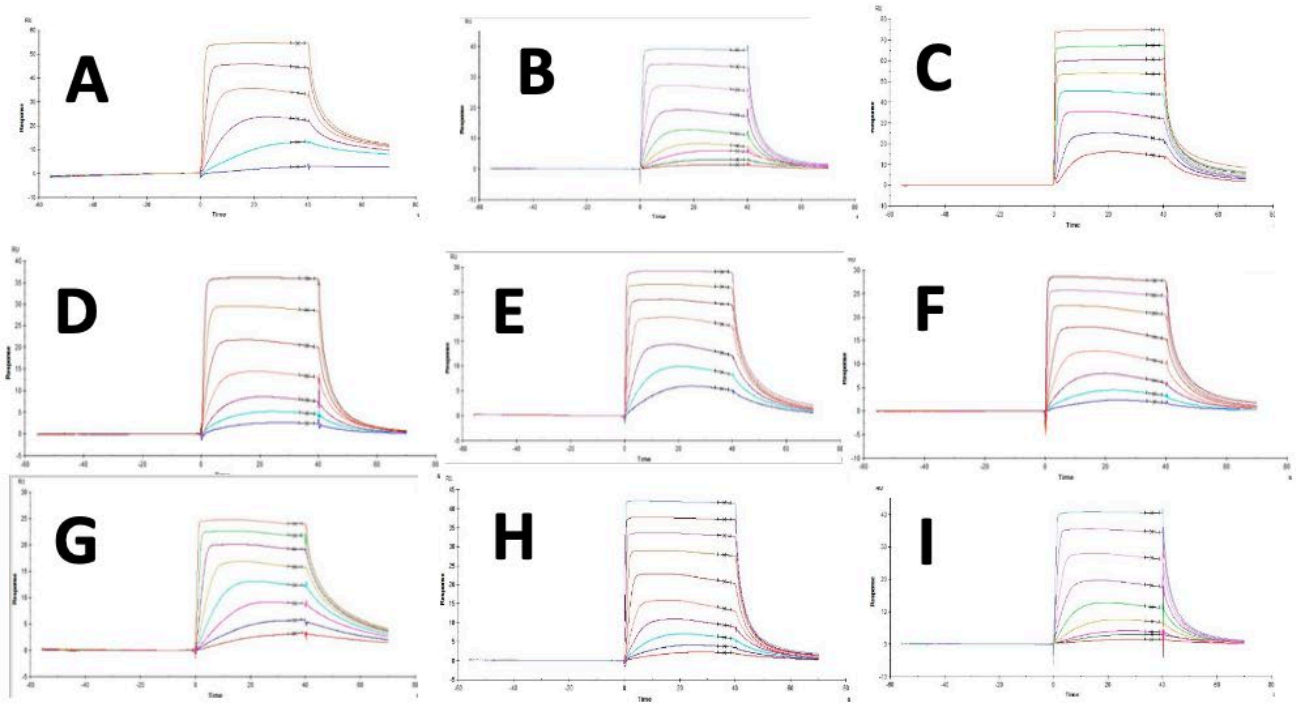


Figure S1. Representative sample of SPR sensorgrams testing peptide library against AfumPCNA.
A) p21 (139-160). **B)** p21 μ . **C)** p21 μ -AfumRFC. **D)** p21 μ -149E. **E)** p21 μ -150Y. **F)** p21 μ -151F. **G)** p21 μ -147I. **H)** p21 μ -147L. **I)** p21 μ -RD2.

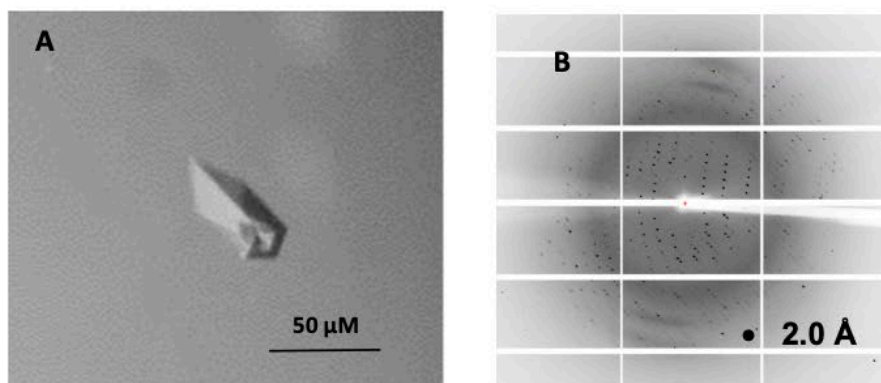


Figure S2. A) Protein crystal of AfumPCNA bound with p21 μ crystal grown in 0.2M Trisacetate pH 4.0 0.1M Na Acetate, 16% PEG 3350 at 16 degrees Celsius after about 3 weeks. **B)** Diffraction image of protein crystal of AfumPCNA bound with p21 μ diffracting at 2.0 Å.

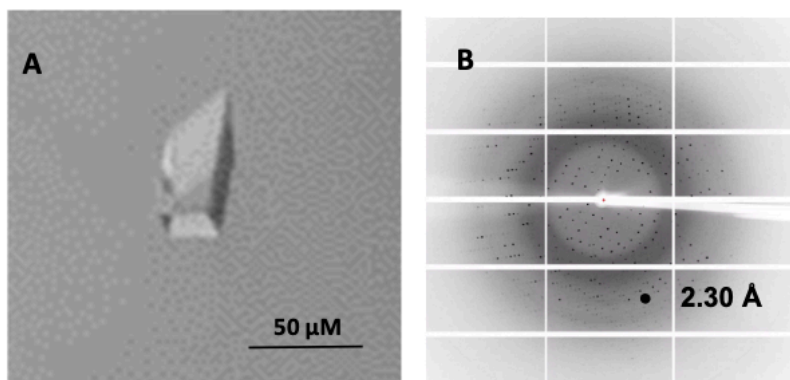


Figure S3. A) Protein crystal of AfumPCNA bound with p21 μ -AfumRFC grown in 0.2M Tacsimate pH 4.0 0.1M Na Acetate, 16% PEG 3350 tray at 16 degrees celcius after 3 weeks. **B)** Diffraction image of protein crystal of AfumPCNA bound with p21 μ -AfumRFC diffracting at 2.30 Å.

Table S2. Data collection and refinement statistics of AfumPCNA bound with p21 μ structure (PDB:8GJF) and AfumPCNA bound with p21 μ -AfumRFC structure (PDB:8GJ5). Statistics for the highest-resolution shell are shown in parentheses. Statistics for the highest-resolution shell are shown in parentheses.

Parameter	8GJF	8GJ5
Wavelength	0.9537	0.9537
Resolution range	36.55 - 2.0 (2.071 - 2.0)	47.2 - 2.3 (2.383 - 2.301)
Space group	C 1 2 1	P 21 21 21
Unit cell	146.0 84.8 70.4 90.0 91.4 90.0	83.9 97.0 108.1 90.0 90.0 90.0
Total reflections	114807 (11427)	79454 (7729)
Unique reflections	57621 (5742)	39748 (3880)
Multiplicity	2.0 (2.0)	2.0 (2.0)
Completeness (%)	99.1 (99.2)	99.8 (99.2)
Mean I/sigma(I)	13.2 (1.6)	21.0 (1.4)
Wilson B-factor	38.4	50.6
R-merge	0.032 (0.49)	0.024 (0.57)
R-meas	0.045 (0.69)	0.034 (0.81)
R-pim	0.031 (0.49)	0.024 (0.57)
CC1/2	0.99 (0.64)	1.00 (0.63)
CC*	1.00 (0.89)	1 (0.88)
Reflections used in refinement	57608 (5742)	39732 (3878)
Reflections used for R-free	2888 (263)	2021 (203)
R-work	0.23 (0.34)	0.23 (0.34)
R-free	0.27 (0.38)	0.25 (0.36)
CC(work)	0.95 (0.71)	0.94 (0.71)
CC(free)	0.92 (0.58)	0.94 (0.67)
Number of non-hydrogen atoms	6392	6018
macromolecules	5971	5874
solvent	421	144
Protein residues	803	797
RMS(bonds)	0.005	0.004
RMS(angles)	1.20	0.93
Ramachandran favored (%)	97.35	94.78
Ramachandran allowed (%)	2.28	4.59
Ramachandran outliers (%)	0.38	0.64
Rotamer outliers (%)	0.32	0.00
Clashscore	6.86	6.66
Average B-factor	50.00	58.67
macromolecules	49.85	58.75
solvent	52.20	55.44

$$^a R_{merge} = \frac{\sum |I - \langle I \rangle|}{\sum I}$$

$$^b R_{pim} = \frac{\sum h [1 / (n_h - 1)]^{1/2} \sum_i |<I_h> - I_{h,i}|}{\sum_h \sum_i I_{h,i}} \quad (2)$$

$$^c R_{work} = \frac{\sum |F_o - F_c|}{\sum |F_o|} \text{ for all data excluding data used to calculate } R_{free}$$

$$^d R_{free} = \frac{\sum |F_o - F_c|}{\sum |F_o|} \text{ for all data.}$$

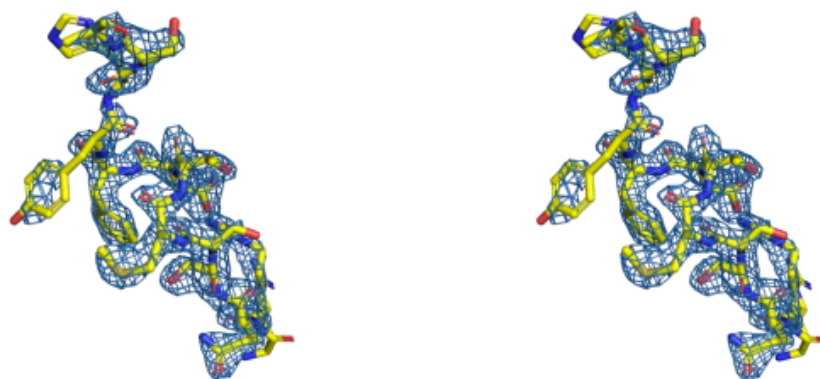


Figure S4. Representative electron density of **p21 μ** (yellow, sticks) shown as a wall-eye stereo image $2F_o-F_c$ composite omit map, view contoured at 1.5σ .

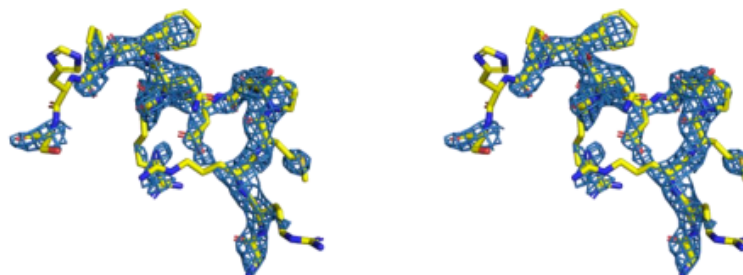


Figure S5. Representative electron density of **p21 μ -AfumRFC** (yellow, sticks) shown as a wall-eye stereo image $2F_o-F_c$ composite omit map, view contoured at 1.5σ .

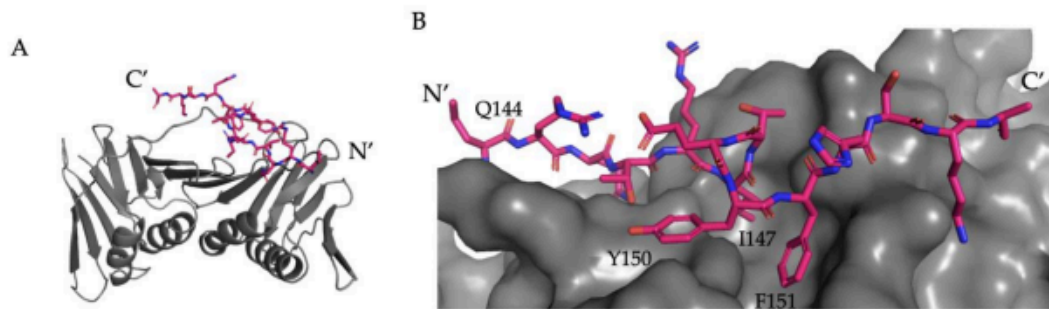


Figure S6. Computationally modelled structure of **p21 μ -RD2** (pink, sticks) on the PIP-box binding site of a AfumPCNA monomer (grey, cartoon). A) Monomer. Peptide N and C terminus labelled. B) Surface, Conserved PIP box residues of peptide are labelled. Peptide N and C terminus labelled.

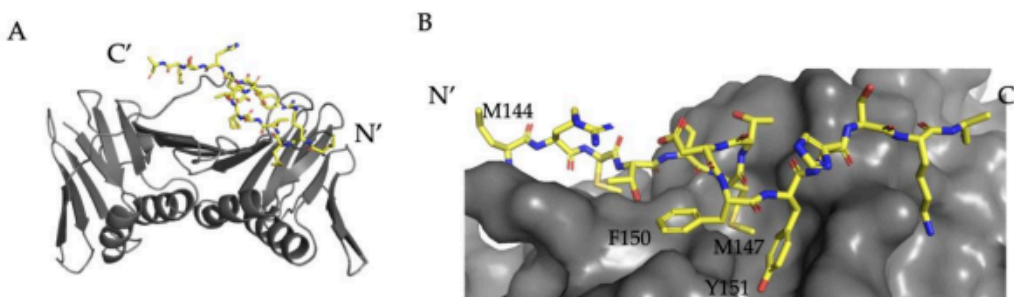


Figure S7. Computationally modelled structure of **p21 μ -Q144M** (yellow, sticks) on the PIP-box binding site of a AfumPCNA monomer (grey, cartoon). A) Monomer. Peptide N and C terminus labelled. B) Surface, Conserved PIP box residues of peptide are labelled. Peptide N and C terminus labelled.

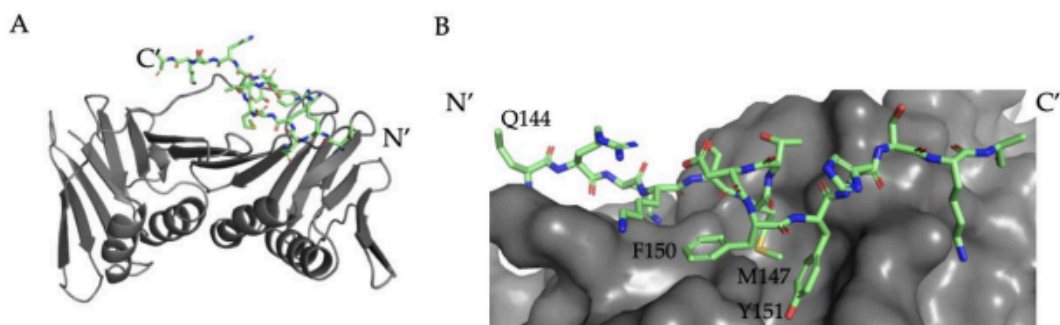


Figure S8. Computationally modelled structure of **p21 μ -T145K** (lime green, sticks) on the PIP-box binding site of a AfumPCNA monomer (grey, cartoon). A) Monomer. Peptide N and C terminus labelled. B) Surface, Conserved PIP box residues of peptide are labelled. Peptide N- and C- terminus labelled.

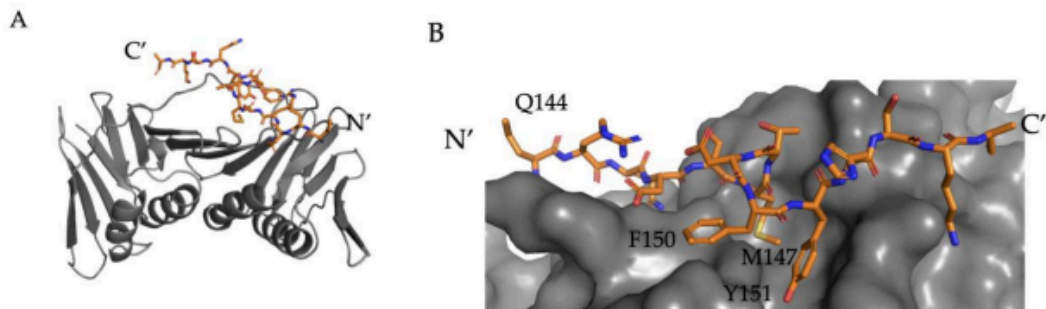


Figure S9. Computationally modelled structure of **p21_μ-T145D** (orange, sticks) on the PIP-box binding site of a AfumPCNA monomer (grey, cartoon). A) Monomer. Peptide N and C terminus labelled. B) Surface, Conserved PIP box residues of peptide are labelled. Peptide N and C terminus labelled.

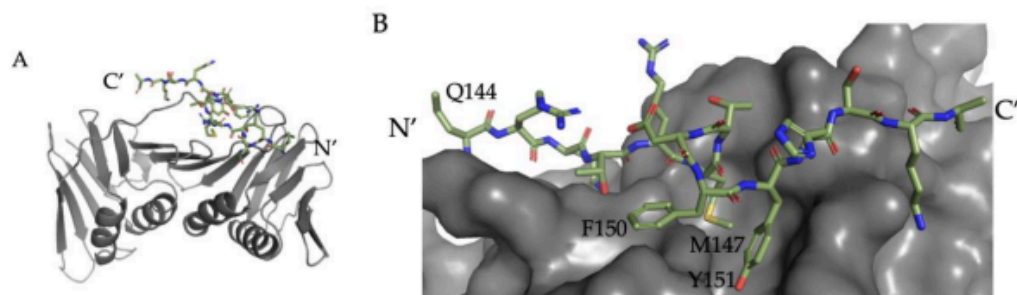


Figure S10. Computationally modelled structure of **p21_μ-S146R** (olive, sticks) on the PIP-box binding site of a AfumPCNA monomer (grey, cartoon). A) Monomer. Peptide N- and C-terminus labelled. B) Surface, Conserved PIP box residues of peptide are labelled. Peptide N- and C-terminus labelled.

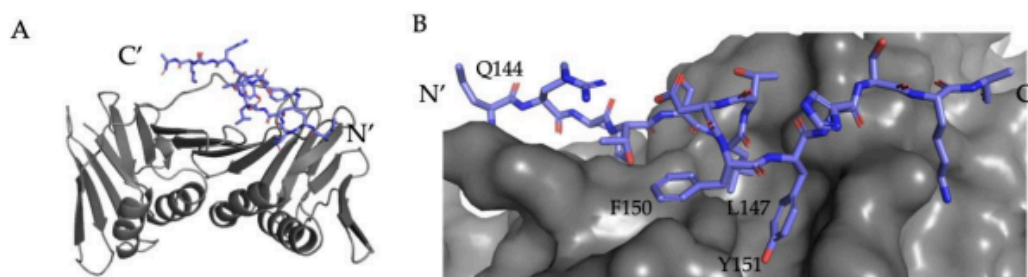


Figure S11. Computationally modelled structure of **p21_μ-M147L** (lilac, sticks) on the PIP-box binding site of a AfumPCNA monomer (grey, cartoon). A) Monomer. Peptide N- and C-terminus labelled. B) Surface, Conserved PIP box residues of peptide are labelled. Peptide N- and C-terminus labelled.

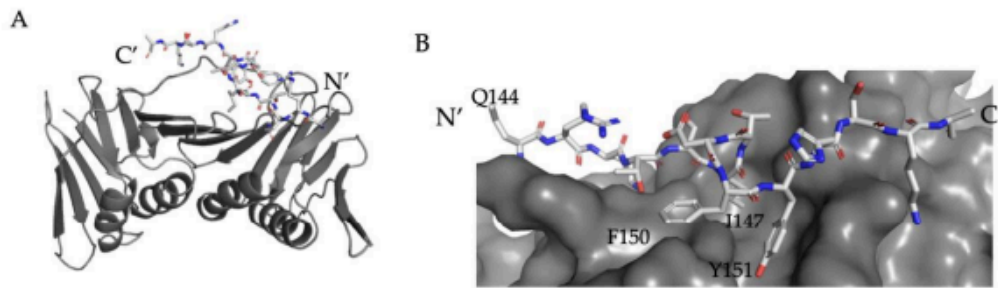


Figure S12. Computationally modelled structure of **p21_μ-M147I** (light grey, sticks) on the PIP-box binding site of a AfumPCNA monomer (grey, cartoon). A) Monomer. Peptide N- and C-terminus labelled. B) Surface, Conserved PIP box residues of peptide are labelled. Peptide N- and C-terminus labelled.

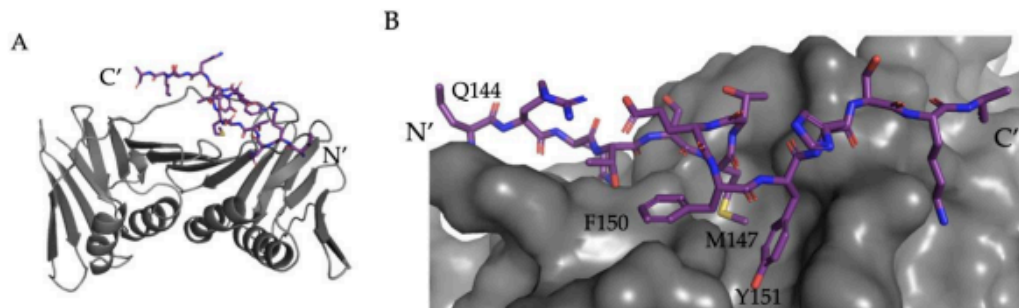


Figure S13. Computationally modelled structure of **p21_μ-D149E** (purple, sticks) on the PIP-box binding site of a AfumPCNA monomer (grey, cartoon). A) Monomer. Peptide N- and C-terminus labelled. B) Surface, Conserved PIP box residues of peptide are labelled. Peptide N- and C-terminus labelled.

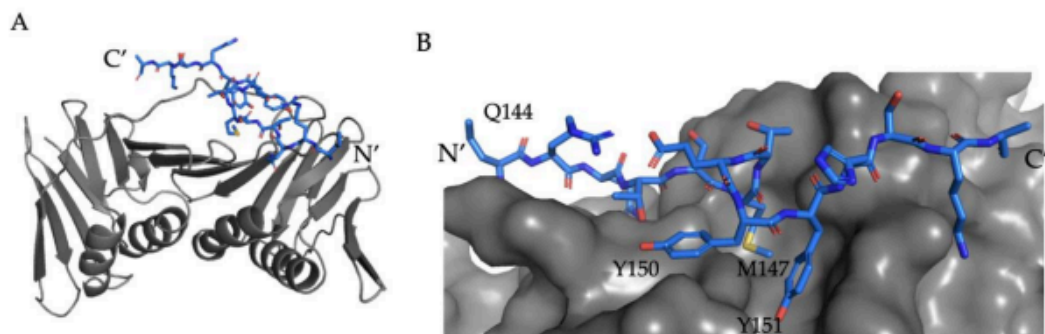


Figure S14. Computationally modelled structure of **p21_μ-F150Y** (blue, sticks) on the PIP-box binding site of a AfumPCNA monomer (grey, cartoon). A) Monomer. Peptide N- and C-terminus labelled. B) Surface, Conserved PIP box residues of peptide are labelled. Peptide N- and C-terminus labelled.

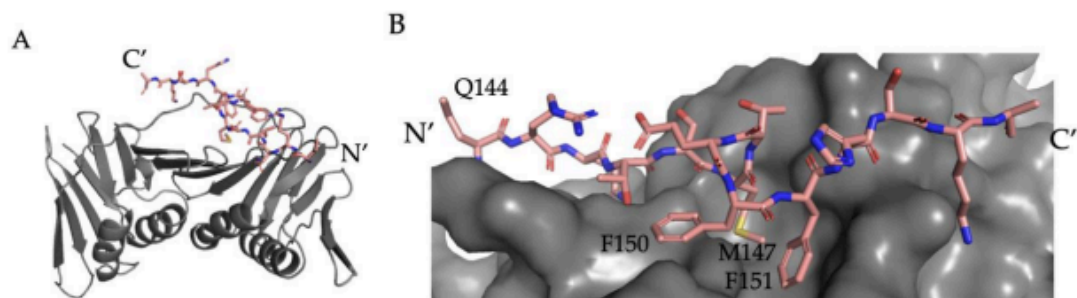


Figure S15. Computationally modelled structure of **p21 μ -Y151F** (salmon, sticks) on the PIP-box binding site of a AfumPCNA monomer (grey, cartoon). A) Monomer. Peptide N- and C-terminus labelled. B) Surface, Conserved PIP box residues of peptide are labelled. Peptide N- and C-terminus labelled.

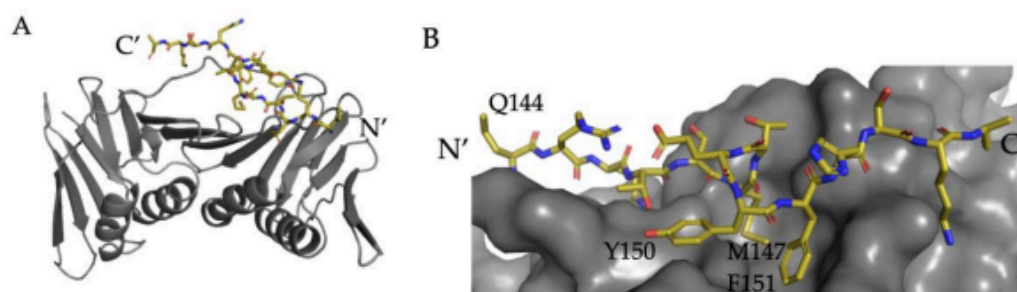


Figure S16. Computationally modelled structure of **p21 μ -FY150151YF** (gold, sticks) on the PIP-box binding site of a AfumPCNA monomer (grey, cartoon). A) Monomer. Peptide N- and C-terminus labelled. B) Surface, Conserved PIP box residues of peptide are labelled. Peptide N- and C-terminus labelled.

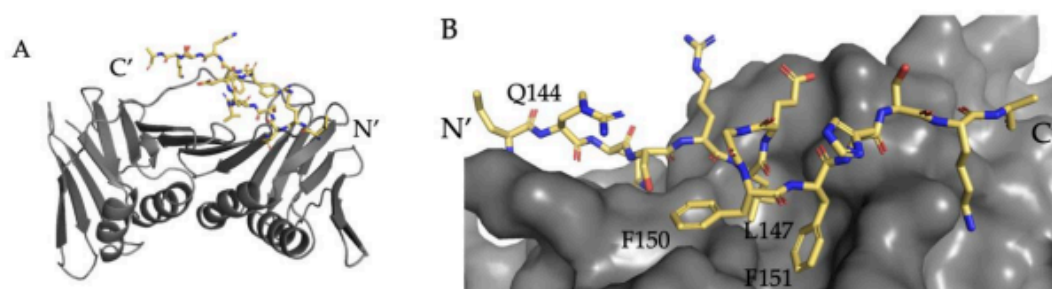


Figure S17. Computationally modelled structure of **p21 μ -AfumFEN1** (yellow, sticks) on the PIP-box binding site of a AfumPCNA monomer (grey, cartoon). A) Monomer. Peptide N- and C-terminus labelled. B) Surface, Conserved PIP box residues of peptide are labelled. Peptide N- and C-terminus labelled.

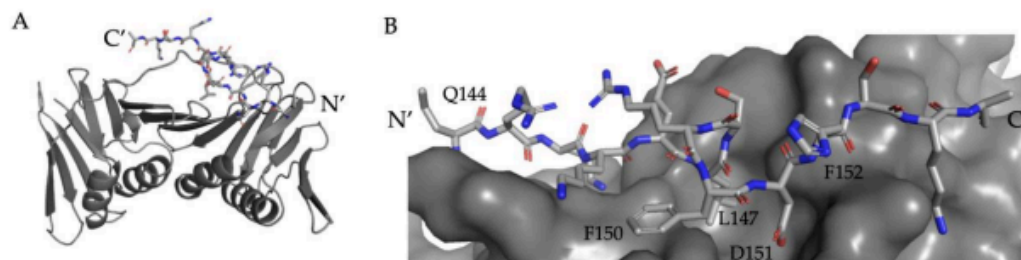


Figure S18. Computationally modelled structure of **p21 μ -AfumDNAPOL** (grey, sticks) on the PIP-box binding site of a AfumPCNA monomer (grey, cartoon). A) Monomer. Peptide N- and C-terminus labelled. B) Surface, Conserved PIP box residues of peptide are labelled. Peptide N- and C-terminus labelled.

References

- [1]Horsfall, A. J., Vandborg, B. A., Kowalczyk, W., Chav, T., Scanlon, D. B., Abell, A. D., and Bruning, J. B. Unlocking the PIP-box: A peptide library reveals interactions that drive high-affinity binding to human PCNA. *J Biol Chem* (2021) 296, 100773, 10.1016/j.jbc.2021.100773.



Full wwPDB X-ray Structure Validation Report i

Mar 15, 2023 – 07:14 AM EDT

PDB ID : 8GJ5
 Title : fungal pcna and peptidomimetic
 Deposited on : 2023-03-14
 Resolution : 2.30 Å (reported)

This wwPDB validation report is for manuscript review

This is a Full wwPDB X-ray Structure Validation Report.

This report is produced by the wwPDB biocuration pipeline after annotation of the structure.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the i symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

The following versions of software and data (see [references](#) i) were used in the production of this report:

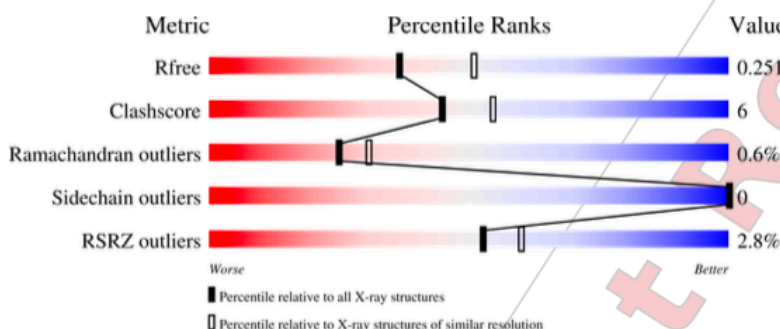
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.32.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.32.1

1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	257	
1	B	257	
1	C	257	
2	D	16	
2	E	16	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	16	 56% 6% 38%

For Manuscript Review

2 Entry composition [\(i\)](#)

There are 3 unique types of molecules in this entry. The entry contains 6017 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Proliferating cell nuclear antigen.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	253	Total 1845	C 1163	N 308	O 362	S 12	0	0	0
1	B	257	Total 1881	C 1185	N 311	O 373	S 12	0	0	0
1	C	254	Total 1861	C 1174	N 306	O 369	S 12	0	0	0

- Molecule 2 is a protein called THR-ASP-ILE-ARG-ASN-PHE-PHE-HIS-SER.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	D	9	Total 75	C 48	N 13	O 14		0	0	0
2	E	14	Total 122	C 76	N 26	O 19	S 1	0	0	0
2	F	10	Total 89	C 57	N 17	O 15		0	0	0

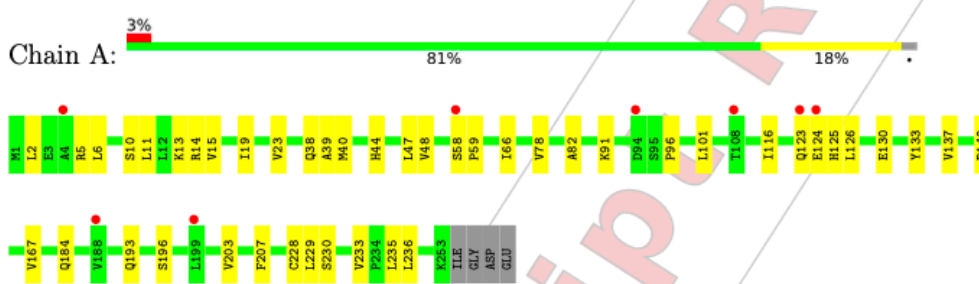
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	33	Total 33	O 33	0	0
3	B	47	Total 47	O 47	0	0
3	C	58	Total 58	O 58	0	0
3	D	2	Total 2	O 2	0	0
3	E	3	Total 3	O 3	0	0
3	F	1	Total 1	O 1	0	0

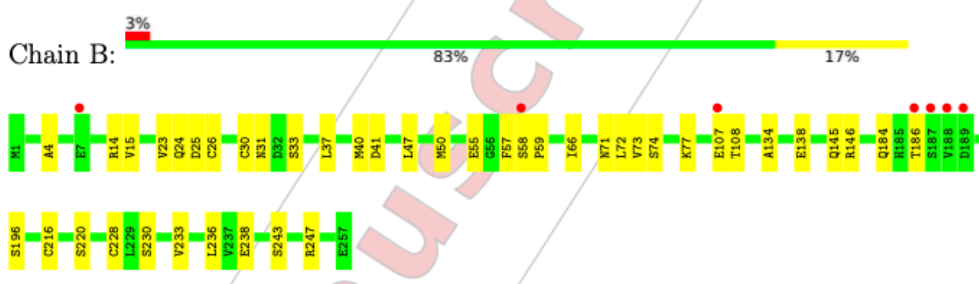
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

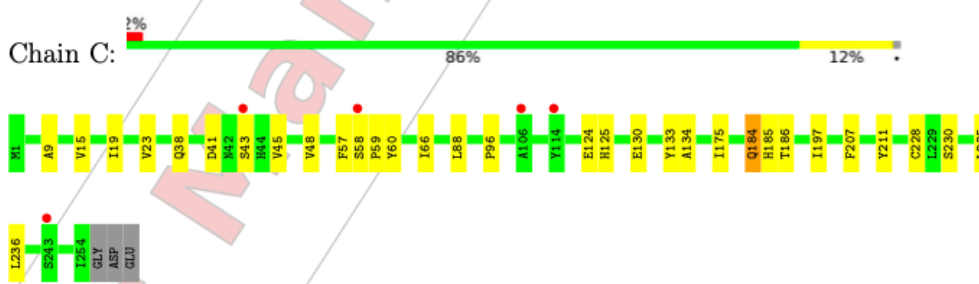
- Molecule 1: Proliferating cell nuclear antigen



- Molecule 1: Proliferating cell nuclear antigen

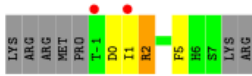


- Molecule 1: Proliferating cell nuclear antigen

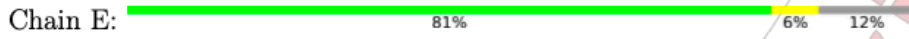


- Molecule 2: THR-ASP-ILE-ARG-ASN-PHE-PHE-HIS-SER

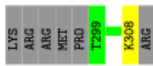




- Molecule 2: THR-ASP-ILE-ARG-ASN-PHE-PHE-HIS-SER



- Molecule 2: THR-ASP-ILE-ARG-ASN-PHE-PHE-HIS-SER



For Manuscript Review

4 Data and refinement statistics 

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	83.93Å 96.98Å 108.05Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.20 – 2.30 47.20 – 2.30	Depositor EDS
% Data completeness (in resolution range)	99.8 (47.20-2.30) 99.8 (47.20-2.30)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.35 (at 2.29Å)	Xtrriage
Refinement program	PHENIX (1.18.2_3874: ???)	Depositor
R, R_{free}	0.239 , 0.251 0.239 , 0.251	Depositor DCC
R_{free} test set	2021 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å ²)	50.6	Xtrriage
Anisotropy	0.435	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 42.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	6017	wwPDB-VP
Average B, all atoms (Å ²)	58.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.55% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z >5	RMSZ	# Z >5
1	A	0.29	0/1869	0.55	0/2544
1	B	0.30	0/1905	0.60	2/2592 (0.1%)
1	C	0.34	0/1885	0.58	1/2566 (0.0%)
2	D	0.29	0/76	0.65	0/101
2	E	0.32	0/125	0.62	0/166
2	F	0.23	0/91	0.43	0/120
All	All	0.31	0/5951	0.58	3/8089 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	184	GLN	CA-CB-CG	7.00	128.80	113.40
1	B	146	ARG	NE-CZ-NH1	-6.07	117.26	120.30
1	C	184	GLN	CA-CB-CG	5.47	125.43	113.40

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1845	0	1788	27	0
1	B	1881	0	1824	26	0
1	C	1861	0	1814	24	0
2	D	75	0	68	3	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	E	122	0	114	2	0
2	F	89	0	83	1	0
3	A	33	0	0	0	0
3	B	47	0	0	0	0
3	C	58	0	0	0	0
3	D	2	0	0	0	0
3	E	3	0	0	0	0
3	F	1	0	0	0	0
All	All	6017	0	5691	75	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 6.

All (75) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:58:SER:HB2	1:C:59:PRO:HD3	1.59	0.84
1:C:43:SER:OG	1:C:211:TYR:OH	1.95	0.84
1:A:40:MET:HE3	1:A:124:GLU:HG2	1.59	0.83
1:C:43:SER:HG	1:C:211:TYR:HH	1.36	0.73
1:A:230:SER:HB2	1:A:233:VAL:HG22	1.75	0.69
1:A:137:VAL:HG21	1:A:167:VAL:HG21	1.78	0.66
1:B:74:SER:HB3	1:C:175:ILE:HG23	1.76	0.65
1:A:228:CYS:HB2	1:A:236:LEU:HB3	1.80	0.64
1:C:58:SER:HB2	1:C:59:PRO:CD	2.29	0.62
1:C:228:CYS:HB2	1:C:236:LEU:HB3	1.82	0.61
1:B:24:GLN:HG2	1:B:25:ASP:OD1	2.00	0.61
1:B:138:GLU:HB2	1:B:196:SER:HB2	1.83	0.61
1:C:134:ALA:H	1:C:230:SER:HB2	1.66	0.59
1:A:5:ARG:HB3	1:A:59:PRO:HD2	1.84	0.59
1:C:66:ILE:HD11	1:C:96:PRO:HG3	1.84	0.58
1:B:108:THR:HG22	1:C:185:HIS:CE1	2.39	0.58
1:A:126:LEU:HD13	2:D:5:PHE:HB2	1.85	0.57
1:A:13:LYS:NZ	1:A:82:ALA:O	2.19	0.56
1:B:228:CYS:HB2	1:B:236:LEU:HB3	1.88	0.56
1:B:238:GLU:HG3	1:B:247:ARG:HG2	1.87	0.56
1:C:133:TYR:CG	1:C:228:CYS:HB3	2.40	0.56
1:B:4:ALA:HB1	1:B:57:PHE:CD2	2.40	0.56
1:A:6:LEU:HD21	1:A:11:LEU:HB3	1.87	0.56
1:A:23:VAL:HG11	1:A:39:ALA:HB1	1.88	0.55
1:B:14:ARG:HD3	1:B:220:SER:HB2	1.88	0.55

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:58:SER:OG	1:B:59:PRO:HD3	2.06	0.55
1:C:19:ILE:HD12	1:C:48:VAL:HG11	1.88	0.55
1:A:66:ILE:HD11	1:A:96:PRO:CD	2.37	0.55
1:C:58:SER:CB	1:C:59:PRO:HD3	2.36	0.54
1:B:24:GLN:HE21	1:B:71:ASN:ND2	2.06	0.53
1:C:184:GLN:HE22	1:C:197:ILE:H	1.56	0.53
1:B:40:MET:HG2	2:E:301:ILE:HD11	1.90	0.53
1:C:207:PHE:CZ	1:C:235:LEU:HB2	2.45	0.52
1:B:108:THR:HG22	1:C:185:HIS:HE1	1.76	0.50
1:C:43:SER:HB3	1:C:45:VAL:HG23	1.94	0.50
1:C:15:VAL:O	1:C:19:ILE:HG12	2.11	0.50
1:A:78:VAL:HG11	1:A:116:ILE:HD11	1.94	0.49
1:A:66:ILE:HD11	1:A:96:PRO:HD3	1.93	0.49
1:A:184:GLN:OE1	1:A:196:SER:HA	2.12	0.48
1:A:207:PHE:CZ	1:A:235:LEU:HB2	2.49	0.48
1:B:145:GLN:HG3	1:B:216:CYS:HB3	1.95	0.48
1:A:44:HIS:O	2:D:0:ASP:HA	2.15	0.46
1:B:15:VAL:HG11	1:B:50:MET:SD	2.55	0.46
1:B:31:ASN:OD1	1:B:33:SER:OG	2.25	0.46
1:B:107:GLU:O	1:B:108:THR:OG1	2.29	0.46
1:B:47:LEU:HB2	2:E:301:ILE:HD12	1.98	0.46
2:D:1:ILE:O	2:D:2:ARG:HG3	2.17	0.45
1:B:55:GLU:HB2	1:B:243:SER:HB2	1.99	0.44
1:B:73:VAL:O	1:B:77:LYS:HG3	2.18	0.44
1:A:133:TYR:CG	1:A:228:CYS:HB3	2.52	0.44
1:A:140:PRO:HG3	1:A:193:GLN:HA	2.00	0.43
1:C:185:HIS:HD2	1:C:186:THR:N	2.16	0.43
1:A:38:GLN:HG2	1:A:47:LEU:HD11	2.01	0.43
1:C:9:ALA:HB2	1:C:88:LEU:HB2	2.01	0.43
1:A:101:LEU:HD13	1:A:116:ILE:HD12	2.00	0.43
1:B:23:VAL:HG12	1:B:41:ASP:HA	2.00	0.42
1:B:26:CYS:SG	1:B:37:LEU:HD11	2.59	0.42
1:C:41:ASP:OD2	1:C:211:TYR:HE1	2.03	0.42
1:B:134:ALA:H	1:B:230:SER:HG	1.65	0.42
1:C:19:ILE:HG22	1:C:23:VAL:CG2	2.50	0.42
1:A:2:LEU:O	1:A:91:LYS:HA	2.19	0.42
1:A:203:VAL:HG11	1:A:229:LEU:HB3	2.01	0.42
1:B:72:LEU:HD23	1:B:72:LEU:HA	1.87	0.41
1:A:123:GLN:HB2	1:A:125:HIS:CD2	2.55	0.41
1:A:40:MET:CE	1:A:124:GLU:HG2	2.42	0.41
1:B:14:ARG:HG2	1:B:220:SER:OG	2.21	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:230:SER:HB2	1:B:233:VAL:CG2	2.50	0.41
1:A:15:VAL:O	1:A:19:ILE:HG13	2.21	0.41
1:B:30:CYS:HB2	1:B:66:ILE:HG23	2.03	0.41
1:C:38:GLN:HA	1:C:48:VAL:O	2.21	0.41
1:C:125:HIS:O	2:F:308:LYS:N	2.49	0.41
1:A:38:GLN:HA	1:A:48:VAL:O	2.21	0.40
1:A:10:SER:O	1:A:14:ARG:HG3	2.21	0.40
1:A:58:SER:OG	1:A:59:PRO:HD3	2.22	0.40
1:C:57:PHE:HB2	1:C:60:TYR:HB2	2.02	0.40

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	251/257 (98%)	241 (96%)	9 (4%)	1 (0%)	34	42
1	B	255/257 (99%)	242 (95%)	12 (5%)	1 (0%)	34	42
1	C	252/257 (98%)	237 (94%)	13 (5%)	2 (1%)	19	23
2	D	7/16 (44%)	5 (71%)	1 (14%)	1 (14%)	0	0
2	E	12/16 (75%)	12 (100%)	0	0	100	100
2	F	8/16 (50%)	8 (100%)	0	0	100	100
All	All	785/819 (96%)	745 (95%)	35 (4%)	5 (1%)	25	31

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	186	THR
2	D	2	ARG
1	A	130	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	124	GLU
1	C	130	GLU

5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	194/223 (87%)	194 (100%)	0	100	100
1	B	198/223 (89%)	198 (100%)	0	100	100
1	C	198/223 (89%)	198 (100%)	0	100	100
2	D	8/16 (50%)	8 (100%)	0	100	100
2	E	13/16 (81%)	13 (100%)	0	100	100
2	F	10/16 (62%)	10 (100%)	0	100	100
All	All	621/717 (87%)	621 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	24	GLN
1	C	24	GLN
1	C	184	GLN
1	C	185	HIS

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

For Manuscript Review

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled '#RSRZ > 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q < 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ > 2	OWAB(Å ²)	Q < 0.9
1	A	253/257 (98%)	0.25	8 (3%) 47 54	44, 60, 83, 94	0
1	B	257/257 (100%)	0.28	7 (2%) 54 62	43, 57, 75, 96	0
1	C	254/257 (98%)	0.19	5 (1%) 65 71	39, 53, 72, 91	0
2	D	9/16 (56%)	0.61	2 (22%) 0 1	65, 69, 76, 79	0
2	E	14/16 (87%)	0.37	0 100 100	53, 57, 74, 78	0
2	F	10/16 (62%)	0.52	0 100 100	61, 68, 77, 81	0
All	All	797/819 (97%)	0.25	22 (2%) 53 60	39, 57, 79, 96	0

All (22) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	123	GLN	4.3
1	A	124	GLU	4.0
1	B	188	VAL	3.8
1	B	58	SER	3.7
2	D	1	ILE	3.2
1	A	58	SER	3.2
1	C	58	SER	3.0
1	B	187	SER	2.9
1	A	94	ASP	2.9
1	B	107	GLU	2.7
1	B	186	THR	2.7
1	A	199	LEU	2.7
1	C	106	ALA	2.7
1	A	108	THR	2.4
1	C	243	SER	2.3
1	B	189	ASP	2.3
1	C	114	TYR	2.2
1	A	4	ALA	2.1
1	A	188	VAL	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	7	GLU	2.0
2	D	-1	THR	2.0
1	C	43	SER	2.0

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.

For Manuscript Review



Full wwPDB X-ray Structure Validation Report i

Oct 16, 2023 – 09:42 AM EDT

PDB ID : 8GJF
Title : afupcna bound with peptide mimetic
Deposited on : 2023-03-15
Resolution : 2.00 Å (reported)

This wwPDB validation report is for manuscript review

This is a Full wwPDB X-ray Structure Validation Report.

This report is produced by the wwPDB biocuration pipeline after annotation of the structure.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the i symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

The following versions of software and data (see [references](#) i) were used in the production of this report:

MolProbity	:	4.02b-467
Xtrriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

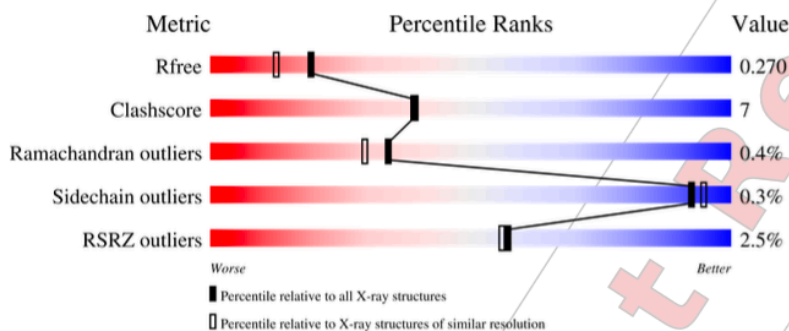
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.




Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	256	<div style="display: flex; align-items: center;"> <div style="width: 2%; height: 10px; background-color: red; margin-right: 2px;"></div> <div style="width: 86%; height: 10px; background-color: green; margin-right: 2px;"></div> <div style="width: 13%; height: 10px; background-color: yellow; margin-right: 2px;"></div> <div style="width: 1%; height: 10px; background-color: orange; margin-right: 2px;"></div> <div style="width: 1%; height: 10px; background-color: grey;"></div> </div> <div style="display: flex; justify-content: space-between; width: 100%;"> 2% 86% 13% . </div>
1	B	256	<div style="display: flex; align-items: center;"> <div style="width: 2%; height: 10px; background-color: red; margin-right: 2px;"></div> <div style="width: 86%; height: 10px; background-color: green; margin-right: 2px;"></div> <div style="width: 13%; height: 10px; background-color: yellow; margin-right: 2px;"></div> <div style="width: 1%; height: 10px; background-color: orange; margin-right: 2px;"></div> <div style="width: 1%; height: 10px; background-color: grey;"></div> </div> <div style="display: flex; justify-content: space-between; width: 100%;"> 2% 86% 13% .. </div>
1	C	256	<div style="display: flex; align-items: center;"> <div style="width: 4%; height: 10px; background-color: red; margin-right: 2px;"></div> <div style="width: 82%; height: 10px; background-color: green; margin-right: 2px;"></div> <div style="width: 18%; height: 10px; background-color: yellow; margin-right: 2px;"></div> </div> <div style="display: flex; justify-content: space-between; width: 100%;"> 4% 82% 18% </div>
2	D	15	<div style="display: flex; align-items: center;"> <div style="width: 87%; height: 10px; background-color: green; margin-right: 2px;"></div> <div style="width: 13%; height: 10px; background-color: grey;"></div> </div> <div style="display: flex; justify-content: space-between; width: 100%;"> 87% 13% </div>
2	E	15	<div style="display: flex; align-items: center;"> <div style="width: 60%; height: 10px; background-color: green; margin-right: 2px;"></div> <div style="width: 13%; height: 10px; background-color: yellow; margin-right: 2px;"></div> <div style="width: 7%; height: 10px; background-color: orange; margin-right: 2px;"></div> <div style="width: 20%; height: 10px; background-color: grey;"></div> </div> <div style="display: flex; justify-content: space-between; width: 100%;"> 60% 13% 7% 20% </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	15	 80% 13% 7%

For Manuscript Review

2 Entry composition i

There are 3 unique types of molecules in this entry. The entry contains 6392 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Proliferating cell nuclear antigen.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	255	Total	C	N	O	S	0	1	0
			1867	1179	311	365	12			
1	B	254	Total	C	N	O	S	0	1	0
			1885	1187	320	365	13			
1	C	255	Total	C	N	O	S	0	0	0
			1892	1195	309	375	13			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	59	ALA	PRO	conflict	UNP A0A229Y5V5
B	59	ALA	PRO	conflict	UNP A0A229Y5V5
C	59	ALA	PRO	conflict	UNP A0A229Y5V5

- Molecule 2 is a protein called LYS-ARG-ARG-GLN-THR-SER-MET-THR-ASP-PHE-TYR-HIS-SER-LYS-ARG.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	13	Total	C	N	O	S	0	0	0
			111	67	22	21	1			
2	E	12	Total	C	N	O	S	0	0	0
			100	61	18	20	1			
2	F	14	Total	C	N	O	S	0	0	0
			116	72	21	22	1			

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	129	Total	O	0	0
			129	129		
3	B	123	Total	O	0	0
			123	123		

Continued on next page...

Continued from previous page...

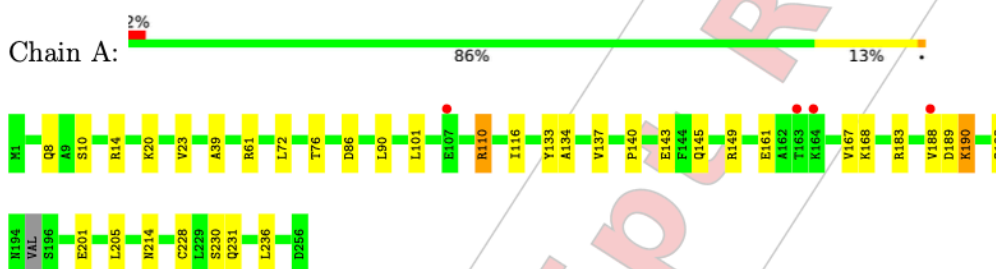
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	C	131	Total O 131 131	0	0
3	D	16	Total O 16 16	0	0
3	E	6	Total O 6 6	0	0
3	F	16	Total O 16 16	0	0

For Manuscript Review

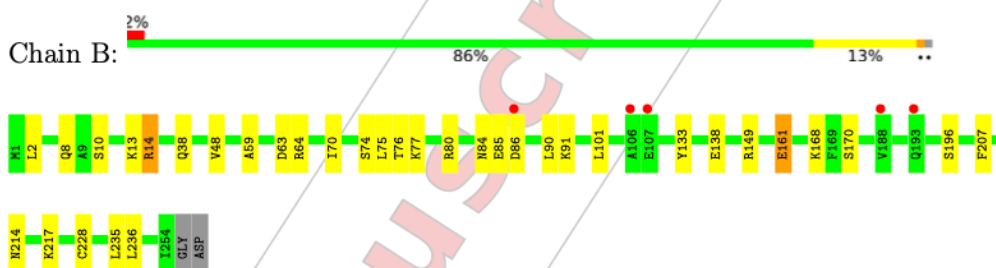
3 Residue-property plots i

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

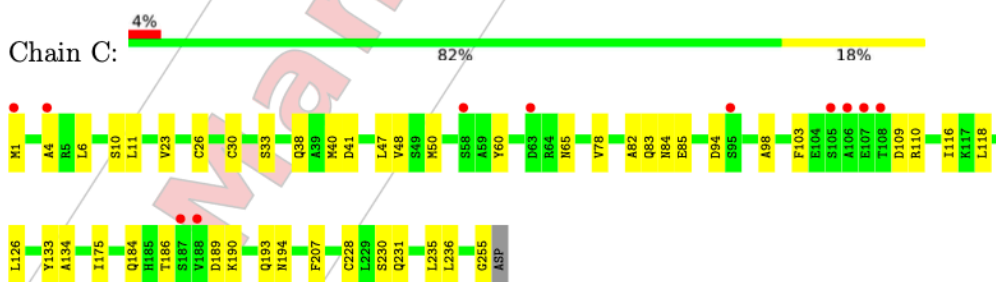
- Molecule 1: Proliferating cell nuclear antigen



- Molecule 1: Proliferating cell nuclear antigen



- Molecule 1: Proliferating cell nuclear antigen

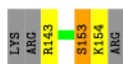


- Molecule 2: LYS-ARG-ARG-GLN-THR-SER-MET-THR-ASP-PHE-TYR-HIS-SER-LYS-ARG





• Molecule 2: LYS-ARG-ARG-GLN-THR-SER-MET-THR-ASP-PHE-TYR-HIS-SER-LYS-ARG



• Molecule 2: LYS-ARG-ARG-GLN-THR-SER-MET-THR-ASP-PHE-TYR-HIS-SER-LYS-ARG



For Manuscript Review

4 Data and refinement statistics i

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	146.00Å 84.83Å 70.45Å 90.00° 91.41° 90.00°	Depositor
Resolution (Å)	42.20 – 2.00 42.20 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.1 (42.20-2.00) 99.1 (42.20-2.00)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.68 (at 2.00Å)	Xtrriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, R_{free}	0.233 , 0.270 0.233 , 0.270	Depositor DCC
R_{free} test set	2888 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å ²)	38.3	Xtrriage
Anisotropy	0.286	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 44.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.000 for -1/2*h+3/2*k,1/2*h+1/2*k,-l 0.000 for -1/2*h-3/2*k,-1/2*h+1/2*k,-l 0.085 for 1/2*h+3/2*k,1/2*h-1/2*k,-l 0.053 for 1/2*h-3/2*k,-1/2*h-1/2*k,-l 0.015 for -h,-k,l	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	6392	wwPDB-VP
Average B, all atoms (Å ²)	49.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.20% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z >5	RMSZ	# Z >5
1	A	0.30	0/1892	1.08	3/2571 (0.1%)
1	B	0.38	1/1911 (0.1%)	0.61	1/2595 (0.0%)
1	C	0.32	0/1916	0.57	1/2600 (0.0%)
2	D	0.25	0/113	0.48	0/150
2	E	2.36	1/102 (1.0%)	0.50	0/136
2	F	0.22	0/118	0.44	0/156
All	All	0.45	2/6052 (0.0%)	0.77	5/8208 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	B	0	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	153	SER	C-N	-23.71	0.79	1.34
1	B	161	GLU	CD-OE2	8.25	1.34	1.25

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	201	GLU	OE1-CD-OE2	-38.20	77.46	123.30
1	A	201	GLU	CG-CD-OE1	21.11	160.52	118.30
1	A	201	GLU	CG-CD-OE2	-17.62	83.06	118.30
1	B	161	GLU	CG-CD-OE2	-6.96	104.39	118.30
1	C	184	GLN	CA-CB-CG	5.01	124.43	113.40

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	161	GLU	Sidechain

5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1867	0	1815	28	1
1	B	1885	0	1863	22	0
1	C	1892	0	1868	34	1
2	D	111	0	97	0	0
2	E	100	0	83	7	0
2	F	116	0	103	2	0
3	A	129	0	0	4	0
3	B	123	0	0	3	0
3	C	131	0	0	10	0
3	D	16	0	0	0	0
3	E	6	0	0	1	0
3	F	16	0	0	0	0
All	All	6392	0	5829	87	1

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 7.

All (87) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:E:153:SER:C	2:E:154:LYS:CA	1.97	1.29
2:E:153:SER:O	2:E:154:LYS:N	1.75	1.19
2:E:153:SER:CA	2:E:154:LYS:N	2.05	1.18
2:E:153:SER:O	2:E:154:LYS:CA	2.10	0.92
2:E:153:SER:O	2:E:154:LYS:HA	1.70	0.91
2:E:153:SER:C	2:E:154:LYS:N	0.79	0.84
1:C:26:CYS:SG	3:C:309:HOH:O	2.40	0.79
1:A:20:LYS:HD3	1:A:76:THR:HG21	1.71	0.73
1:B:86:ASP:N	3:B:301:HOH:O	2.27	0.66

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:76:THR:O	1:B:80:ARG:HG3	1.98	0.64
1:A:161:GLU:HG3	1:A:168:LYS:HB3	1.80	0.63
1:A:149:ARG:NH2	3:A:302:HOH:O	2.32	0.62
1:A:145:GLN:NE2	3:A:302:HOH:O	2.32	0.62
1:B:63:ASP:O	1:B:64:ARG:HD3	2.01	0.61
1:C:60:TYR:OH	3:C:301:HOH:O	2.16	0.60
1:B:14:ARG:NH1	3:B:303:HOH:O	2.33	0.60
1:A:8:GLN:N	1:A:8:GLN:OE1	2.34	0.60
1:B:8:GLN:NE2	1:B:85[A]:GLU:OE1	2.33	0.60
1:C:231:GLN:NE2	3:C:306:HOH:O	2.33	0.59
1:A:134:ALA:H	1:A:230:SER:HB3	1.67	0.59
1:A:8:GLN:HG2	1:A:8:GLN:O	2.05	0.56
1:C:4:ALA:HA	3:C:357:HOH:O	2.06	0.56
1:A:110[B]:ARG:NH2	3:A:306:HOH:O	2.36	0.56
1:C:10:SER:HB3	1:C:84:ASN:HB3	1.88	0.56
1:A:205:LEU:HD21	1:A:231:GLN:HA	1.88	0.55
1:A:86:ASP:OD2	1:A:110[A]:ARG:NH1	2.40	0.55
1:B:228:CYS:HB2	1:B:236:LEU:HB3	1.90	0.54
1:B:138:GLU:HB2	1:B:196:SER:HB3	1.91	0.53
1:C:186:THR:HA	1:C:194:ASN:OD1	2.08	0.52
1:C:6:LEU:HD11	1:C:11:LEU:HD23	1.91	0.52
2:E:143:ARG:N	3:E:202:HOH:O	2.43	0.51
1:C:30:CYS:HB3	3:C:301:HOH:O	2.08	0.51
1:C:65:ASN:ND2	3:C:303:HOH:O	2.28	0.51
1:C:255:GLY:HA3	2:F:142:ARG:HB2	1.94	0.50
1:A:137:VAL:HG21	1:A:167:VAL:HG21	1.94	0.50
1:A:228:CYS:HB2	1:A:236:LEU:HB3	1.94	0.50
1:C:38:GLN:HA	1:C:48:VAL:O	2.12	0.49
1:A:10:SER:O	1:A:14:ARG:HG3	2.13	0.48
1:A:145:GLN:O	1:A:149:ARG:HG3	2.13	0.48
1:C:78:VAL:HG21	1:C:116:ILE:CD1	2.44	0.48
1:A:183:ARG:HB2	1:C:109:ASP:HB3	1.96	0.47
1:A:188:VAL:O	1:A:190:LYS:N	2.48	0.47
1:B:38:GLN:HA	1:B:48:VAL:O	2.14	0.47
1:C:40:MET:HG3	1:C:47:LEU:HD12	1.97	0.47
1:B:214:ASN:OD1	1:B:217:LYS:NZ	2.46	0.47
1:A:110[B]:ARG:HE	1:A:110[B]:ARG:HB3	1.64	0.47
1:B:133:TYR:CG	1:B:228:CYS:HB3	2.50	0.47
1:C:85:GLU:OE1	1:C:85:GLU:N	2.42	0.46
1:C:126:LEU:O	2:F:154:LYS:HE2	2.15	0.46
1:C:134:ALA:H	1:C:230:SER:HB3	1.79	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:13:LYS:HE2	1:B:84:ASN:HD22	1.79	0.46
1:A:90:LEU:CD2	1:A:101:LEU:HD23	2.46	0.46
1:B:207:PHE:CZ	1:B:235:LEU:HB2	2.51	0.46
1:A:140:PRO:HD3	1:A:193:GLN:O	2.15	0.46
1:A:214:ASN:ND2	3:A:313:HOH:O	2.48	0.46
1:C:83:GLN:HB3	1:C:85:GLU:OE1	2.16	0.46
1:B:13:LYS:HE2	1:B:84:ASN:ND2	2.31	0.45
1:B:74:SER:HB3	1:C:175:ILE:HG23	1.99	0.45
1:A:72:LEU:O	1:A:76:THR:HG23	2.17	0.45
1:C:82:ALA:HB2	1:C:103:PHE:CD2	2.52	0.45
1:C:207:PHE:CZ	1:C:235:LEU:HB2	2.52	0.45
1:C:98:ALA:HA	1:C:118:LEU:HG	1.99	0.44
1:C:193:GLN:N	3:C:318:HOH:O	2.50	0.44
1:A:143:GLU:OE1	1:C:110:ARG:NH2	2.44	0.44
1:C:228:CYS:HB2	1:C:236:LEU:HB3	1.98	0.44
1:A:133:TYR:CG	1:A:228:CYS:HB3	2.53	0.44
1:A:23:VAL:HG11	1:A:39:ALA:HB1	1.99	0.43
1:C:10:SER:HB3	1:C:84:ASN:CB	2.47	0.43
1:B:90:LEU:HD22	1:B:101:LEU:HD23	2.00	0.43
1:B:168:LYS:HE2	1:B:170:SER:HB3	2.00	0.43
1:A:161:GLU:HG2	1:A:168:LYS:O	2.19	0.43
1:B:10:SER:O	1:B:14:ARG:HD3	2.19	0.43
1:B:2:LEU:O	1:B:91:LYS:HA	2.18	0.43
1:C:50:MET:HE3	3:C:399:HOH:O	2.18	0.43
1:C:1:MET:HG2	1:C:94:ASP:OD1	2.20	0.42
1:C:30:CYS:SG	3:C:414:HOH:O	2.62	0.42
1:C:189:ASP:O	1:C:190:LYS:HE2	2.20	0.42
1:A:101:LEU:HD11	1:A:116:ILE:HD11	2.01	0.42
1:C:23:VAL:HG12	1:C:41:ASP:HA	2.00	0.42
1:A:161:GLU:CG	1:A:168:LYS:HB3	2.50	0.41
1:A:183:ARG:HD3	1:A:183:ARG:HA	1.77	0.41
1:C:133:TYR:CG	1:C:228:CYS:HB3	2.56	0.41
1:B:77:LYS:HD2	1:C:175:ILE:CD1	2.51	0.41
1:C:60:TYR:HD1	3:C:357:HOH:O	2.05	0.40
1:B:149:ARG:NH2	3:B:311:HOH:O	2.54	0.40
1:B:70:ILE:HG21	1:B:75:LEU:HD22	2.03	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:61:ARG:NH1	1:C:33:SER:OG[1_545]	2.19	0.01

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	254/256 (99%)	248 (98%)	4 (2%)	2 (1%)	19	13
1	B	253/256 (99%)	244 (96%)	8 (3%)	1 (0%)	34	30
1	C	253/256 (99%)	247 (98%)	6 (2%)	0	100	100
2	D	11/15 (73%)	11 (100%)	0	0	100	100
2	E	10/15 (67%)	10 (100%)	0	0	100	100
2	F	12/15 (80%)	12 (100%)	0	0	100	100
All	All	793/813 (98%)	772 (97%)	18 (2%)	3 (0%)	34	30

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	189	ASP
1	A	190	LYS
1	B	59	ALA

5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	192/221 (87%)	190 (99%)	2 (1%)	76	81
1	B	201/221 (91%)	200 (100%)	1 (0%)	88	92
1	C	203/221 (92%)	203 (100%)	0	100	100
2	D	12/15 (80%)	12 (100%)	0	100	100
2	E	11/15 (73%)	11 (100%)	0	100	100
2	F	12/15 (80%)	12 (100%)	0	100	100
All	All	631/708 (89%)	628 (100%)	3 (0%)	92	92

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	110[A]	ARG
1	A	110[B]	ARG
1	B	14	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	84	ASN
1	A	184	GLN
1	B	8	GLN
1	B	84	ASN
1	C	125	HIS
2	D	152	HIS
2	E	144	GLN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
2	E	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	E	153:SER	C	154:LYS	N	0.79

For Manuscript Review

6 Fit of model and data i

6.1 Protein, DNA and RNA chains i

In the following table, the column labelled '#RSRZ > 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q < 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ > 2	OWAB(Å ²)	Q < 0.9
1	A	255/256 (99%)	-0.02	4 (1%) 72 70	27, 44, 82, 123	0
1	B	254/256 (99%)	0.10	5 (1%) 65 63	33, 45, 89, 119	0
1	C	255/256 (99%)	0.24	11 (4%) 35 34	27, 45, 85, 131	0
2	D	13/15 (86%)	-0.01	0 100 100	30, 37, 62, 74	0
2	E	12/15 (80%)	0.27	0 100 100	44, 53, 66, 73	0
2	F	14/15 (93%)	0.08	0 100 100	32, 38, 67, 69	0
All	All	803/813 (98%)	0.11	20 (2%) 57 56	27, 45, 83, 131	0

All (20) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	188	VAL	4.7
1	C	58	SER	3.4
1	C	108	THR	3.4
1	A	164	LYS	3.1
1	C	4	ALA	3.0
1	C	1	MET	3.0
1	B	107	GLU	3.0
1	B	193	GLN	3.0
1	C	106	ALA	2.7
1	B	86	ASP	2.6
1	B	188	VAL	2.6
1	C	187	SER	2.5
1	C	105	SER	2.5
1	C	63	ASP	2.5
1	B	106	ALA	2.5
1	C	95	SER	2.4
1	A	107	GLU	2.2
1	C	107	GLU	2.1
1	A	163	THR	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	188	VAL	2.1

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.

For Manuscript Review

Chapter 6: Manuscript

“AI generated cyclized peptide for sliding clamp peptidomimetics for cancer therapeutic.”

Statement of Authorship

Title of Paper	AI generated cyclised peptide for sliding clamp peptidomimetics for cancer therapeutic.
Publication Status	<input type="checkbox"/> Published <input type="checkbox"/> Accepted for Publication <input type="checkbox"/> Submitted for Publication <input checked="" type="checkbox"/> Unpublished and Unsubmitted work written in manuscript style
Publication Details	Manuscript style work

Principal Author

Name of Principal Author (Candidate)	Bethiney Vandborg		
Contribution to the Paper	Protein expression and purification, experimental design, validation experiments, discussion of results and 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 aim its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	29/08/2023

Co-Author Contributions

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

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

Name of Co-Author	Dayna Holroyd		
Contribution to the Paper		tion, validation and edited	
Signature	Supervisor signature	Date	17/9/23
Name of Co-Author			
Contribution to the Paper		discussed results.	
Signature	Supervisor signature	Date	17/9/23

Please cut and paste additional co-

Name of Co-Author	Amit Michaeli		
Contribution to the Paper	Peptide design, expression and purification, discussed results.		
Signature	Supervisor signature	Date	15/9/23

Name of Co-Author			
Contribution to the Paper	peptide design, discussed results.		
Signature		Date	13/10/2023

Name of Co-Author	John Bruning		
Contribution to the Paper	Supervision of BV and DH, discussed results, edited manuscript.		
Signature		Date	15/9/23

AI generated cyclized peptidomimetics targeting the sliding clamp for a cancer therapeutic.

B. Vandborg, D.L. Holroyd, I. Lerner, A. Michaeli, A.J. Horsfall, J.B. Bruning.

1. Introduction

Human Proliferating Cell Nuclear Antigen (hPCNA)-binding proteins such as p21, the cell cycle regulator, have high affinity interaction with DNA for replication and repair processes. This interaction is enabled by means of the p21 PCNA interacting Protein (PIP) box. The high affinity binding discovered in previous research[27, 55] has shown the advantageous residues selected for its PIP-box create a high affinity structural interaction that has been used in peptidomimetic research. Studies have found further residue substitutions can increase the affinity of synthesized peptides containing the PIP-box for the surface of hPCNA[56], presenting the hypothesis that nature has fine-tuned the native interacting sequence to allow tight interaction, but not permanent[26], further suggesting that the PIP-box can be improved beyond the native p21 sequence.

Current drug screening techniques for new targets require synthesis of compounds to identify hits from molecular libraries. This is a costly system which limits the research into new compounds and therapeutics. Machine learning can be utilized to scan binding domains and calculate peptide combinations that will interact with amino acid residues present. This optimises the drug discovery pipeline by quickly composing new solutions *in silico* for peptide sequences to interact with binding domains. By creating a virtual machine-based method of determining hits to further investigate, costs can be reduced and new routes of research can be carried out without the synthesis of large libraries of molecules. Automated systems of research can be used to allow discovery of novel ligands through the input of the already existing Protein Data Bank[57]. Natural amino acids in small peptides suffer from low stability and fast clearance. Hence, the use of non-natural amino acids may also allow for improvements in synthesised peptidomimetics, with new backbone modifications that may improve these qualities in a new molecule [58] [59].

In peptides, the PIP-box backbone creates the secondary structure of a 3_{10} helix upon binding [27], which is essential for high affinity interaction. Pre-organization of this backbone may reduce the entropic cost of binding to the surface of hPCNA[26]. Previous research into p21 peptides have shown that flexibility in the secondary structure is preferable for the peptide to still have movement to adopt the helical structure upon plugging into the hPCNA binding domain on the surface[55]. This can be done by cyclisation of the peptide, holding a flexible secondary structure in place. Cyclical peptides have been shown to use passive diffusion and endocytic uptake to enter mammalian cells[60], an important aspect in the drug discovery pipeline for a future therapeutic.

Pepticom is a company that uses algorithms and selection through scoring functions to discover binding peptides to protein targets, allowing *in silico* peptide and macrocycle screening for a new peptidomimetic design. Using this method to design a new generation peptides, here is presented the investigation of artificial amino acids in the place of native p21 PIP-box residues, for the use in the drug discovery pipeline to a hPCNA inhibitor.

2. Methods

2.1 Peptide synthesis

The following peptides were synthesised by Genscript Biotech, Singapore at purity >95%, purified by HPLC.

Tracer - **5FAM-p21-22mer** (5FAM)-GRKRRQTSMTDFYHSKRLLIFS

Probe - **p21 μ** p21 μ -15mer KRRQTSMTDFYHSKR

2.1.2 PNA peptide Synthesis

Designed peptides by Pepticom.

Using Pepticom's proprietary software and its *ab initio* peptide design algorithm and based on atomic resolution 3D structures of the targeted protein, Peptide Nucleic Acids (PNA) peptides, based on the p21 μ sequence, were selected for synthesis. Lyophilized peptides were produced as described previously [61] at EMC Microcollections (Tubingen, Germany) and stored at -20°C until use.

2.1.3 Peptide stock preparation

Peptides (approximately 1 mg) were dissolved in MilliQ water and centrifuged (7800 rpm, 10 min) to remove any particulate. The peptide stock concentration was determined by 205 nm absorbance (A₂₀₅), a measurement taken in triplicate with a NanoDrop2000 and baseline referenced to 750 nm absorbance using a 1 in 20 dilution. The ϵ_{205} for each peptide was calculated using an online calculator (<http://nickanthis.com/tools/a205.html>). The peptide stock solution concentration was then calculated per $c = (A_{205}/\epsilon_{205} \times 1) \times DF$, where the concentration is in molar, A₂₀₅ is absorbance at 205 nm calculated as an average of three readings, 1 is the pathlength in centimeter (0.1 cm for Nanodrop), ϵ_{205} is the molar absorptivity at 205 nm, and DF is the dilution factor. The peptides were then diluted into the assay buffer before further dilution as necessary. Except for the fluorescein tagged peptide (5FAM-p21-22mer), which the ϵ_{205} value of 490 was used.

2.1.4 PNA Peptide stock preparation

Peptides (approximately 1 mg) were dissolved in MilliQ water and centrifuged (7800 rpm, 10 min) to remove any particulate. The peptide stock concentration was determined by mass molarity calculations $n=m/M_r$, and further dilutions by $c=n/v$. The peptides were diluted into the assay buffer before further dilution. Those peptides that required more work to achieve solubility were given a higher dilution factor in water.

2.2. Expression of recombinant hPCNA

Human PCNA was expressed as described in Vandborg 2023[62]. A glycerol stock of *E. coli* BL21 (DE3) cells carrying a codon optimized hPCNA-pMCSG19 plasmid were grown in a 100 mL overnight culture. Two 1 L baffled flasks containing LB with 100 µg/mL of ampicillin were inoculated with 50 mL of the overnight culture. Cultures were incubated at 37°C until OD₆₀₀=0.7 and expression induced with a final concentration of 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 20 mL 20 mM Tris pH 7.5, 20 mM NaCl, 2 mM DTT, then lysed by sonication at 70% amplification for 20 seconds with a 40 second waiting period for 25 cycles. Lysate was clarified through pelleting at 45,000xg for 45 min and the supernatant was collected for purification.

2.3. Purification of recombinant hPCNA

Human PCNA was purified as described in Vandborg 2023[62].

Buffer solutions were filtered before being used. Clarified lysate containing hPCNA was first purified at 4 °C by fast protein liquid chromatography (FPLC) via anion exchange chromatography using two DEAE columns in series (HiTrap DEAE FF 5 mL column), equilibrated in Buffer A (20 mM Tris pH 7.5, 20 mM NaCl, 2 mM DTT), and hPCNA eluted using a linear gradient (0.02 M–0.7 M NaCl). Fractions containing hPCNA were pooled and ammonium sulphate added dropwise to a final concentration of 1.5 M from a stock solution of 3 M ammonium sulphate. The sample was allowed to stir gently for 1 hour at 4 °C to allow DNA-protein dissociation, and then applied to hydrophobic interaction chromatography (HiTrap Phenyl FF (high sub) 5 mL column), equilibrated in Buffer B (20 mM Tris pH 7.5, 20 mM NaCl, 2 mM DTT, 0.5 mM EDTA, 1.5 M ammonium sulphate and eluted in Buffer C (20 mM Tris pH 7.5, 2 mM DTT, 0.5 mM EDTA) with a reverse linear gradient (1.5 M – 0 M ammonium sulphate). Fractions containing hPCNA were pooled and dialyzed overnight in Buffer A. hPCNA was then applied to a second anion exchange step. The Q Sepharose column (5 ml Q Sepharose FF column (GE)) was equilibrated in Buffer A and protein was eluted using a linear gradient (0.02 M–0.7 M NaCl). Fractions containing hPCNA were pooled and dialyzed overnight in 20mM Tris pH 7.5, 100mM NaCl, 0.01% Triton X, 2 mM DTT, 10 % glycerol. Protein for fluorescence polarisation was concentrated to ~10 mg/mL using a centrifugal filter unit (50 kDa molecular mass cut off) and stored at –80 °C.

2.4 Fluorescence Polarisation Competition Binding Assay

The fluorescence polarisation (FP) competition binding assay was performed as described in Wegener et al. 2018 [26]. Briefly, the probe and control peptide p21µ was two-fold serially diluted, control peptide p21µ began from 108 µM, into solutions containing a constant concentration of 5FAM-p21-22mer, the tracer (50 nM), and purified hPCNA (175 nM), and interaction tested using fluorescence polarisation competition binding. This was performed using a buffer of 20 mM Tris pH 7.5, 100 mM NaCl, 0.01% Triton x-100, 12 mM DTT, 10% glycerol. Each assay was incubated at 25°C for 30 minutes prior to measurement and carried out in duplicate due to limited materials. each experiment was repeated three times. The fluorescence polarization experiments were conducted on a

Synergy H4 plate reader using black polystyrene 96-well plates with clear bottoms in a total volume of 60 μL per well. The resulting polarization values were background tracer subtracted. Experimental peptides were tested the same way however have varying starting concentrations.

2.4.1 Statistical analysis

All statistics for binding experiments, such as IC_{50} values, was performed using GraphPad Prism version 10.0.0 for Windows (GraphPad Software, Boston USA www.graphpad.com). K_i value determination was done by the following equation determined by Nikolovska-Coleska 2004 [54].

$$K_i = I_{50} / L_{50} / K_d + P_0 / K_d + 1$$

The concentration of the free inhibitor at 50% inhibition (I_{50}) (also known as IC_{50}). The concentration of the free labelled ligand at 50% inhibition (L_{50}). the concentration of the free protein at 0% inhibition (P_0). The dissociation constant of the protein-ligand complex (K_d), defined in previous research [26]. This equation was derived from the basic principles of a competitive binding assay and is not restricted by the concentrations of the protein and the ligand[54].

2.5 Computational Models

Computational models were made using the unbound structure (PDB:6E49) [63]. Protein loop and homology modelling were performed using the Prime suit [64] [65]. Structure comparisons and visualization were performed using the PyMOL Molecular Graphics System [66].

3. Results

3.1 Positive control experiment using p21 μ starting scaffold peptide shows FP assay efficacy.

The truncated p21 peptide designed in Horsfall 2021 was used to validate the fluorescence polarization method chosen from Wegener 2017 to test the AI derived peptides [26, 56]. This peptide was used as the starting scaffold for derivation of the computationally designed library. The K_i value and standard curve achieved indicated the protein-peptide interaction measured to be a reasonable experimental procedure to characterize the AI generated peptides.

Table 1: Fluorescence polarization competition assay of control peptide p21 μ . A top concentration of 108 μM was used before 8x 1 in 2 dilutions used to calculate IC_{50} and K_i value. The concentration of the free inhibitor at 50% inhibition (I_{50}) also known as IC_{50} here represents the affinity value, the average IC_{50} of three experiments performed in duplicate are shown accompanied by a standard deviation from repeat experiments, K_i is the inhibition constant of the given inhibitor.

Peptide	Molecular Weight (g/mol)	Average IC_{50} (μM) \pm SD (μM)	K_i (μM)

p21 μ	1940	1.35 \pm 0.425	0.397
-----------	------	------------------	-------

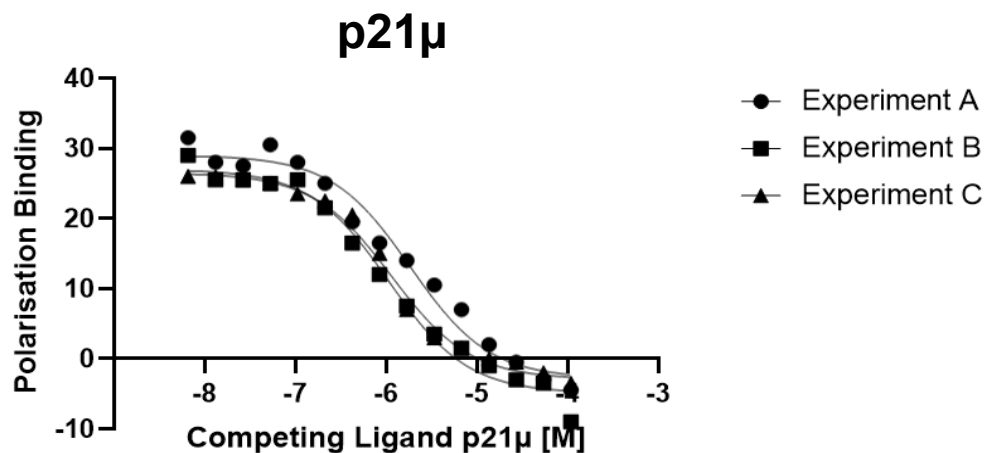


Figure 1. Fluorescence polarization competition assay IC₅₀ curve of control peptide p21 μ . R² of 0.9927. Curves represent single experiments, where each data point is the mean of triplicate samples. Each experiment was repeated three times.

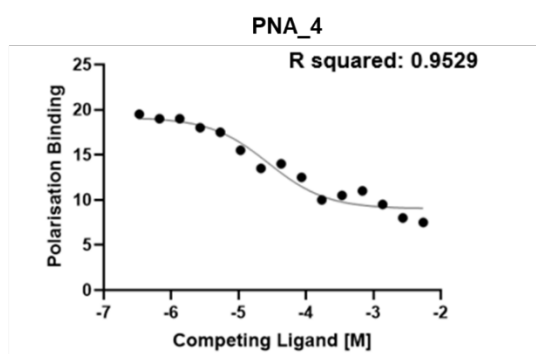
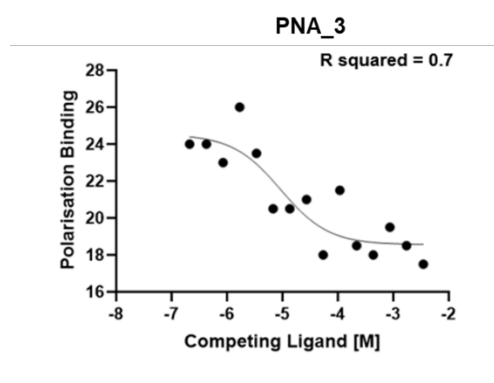
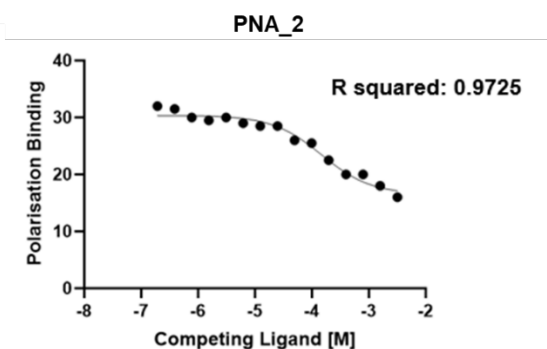
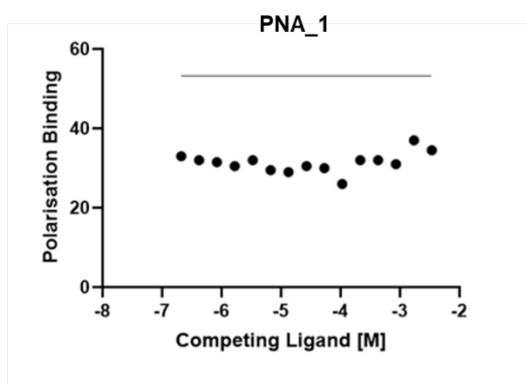
3.1 Investigation of first-generation cyclical peptides generated from p21 μ scaffold yields leads.

The Fluorescence Polarisation (FP) assay revealed IC₅₀ values of multiple generated cyclized peptides interacting with hPCNA (Table 2). The R² value was used to indicate the experiment validity, a value ranging from 0 to 1. It is an indication of how close the data is to the fitted regression line, the closer the value is to 1. The R² value shows the experiment to be valid; however, there was significant differences between IC₅₀ values between repeated experiments, shown by the standard deviation values. This was initially hypothesized to be due to variations in the preparation of the peptide or the composition of the peptides allowing small scale aggregation that was not removed in preparation measures, or the peptide sequence allowing for non-specific binding.

Multiple IC₅₀ experiments were unsuccessful, which may have been due to this possible aggregation or non-specific binding. These include PNA_1, 3, 5, and 6 (Figure 2). However, the peptides PNA_2, 7 and 8 displayed R² values above 0.95 and again in duplicate experiments (Figure 2). PNA_7 and 8 revealed the most fit results, the IC₅₀ values, and optimizations results could give a reliable peptidomimetic starting point (Table 2).

Table 2: Fluorescence polarization competition assay of first-generation peptides. The top concentration is the highest concentration of 8x 1 in 2 dilutions used to calculate IC₅₀ and K_i value. The concentration of the free inhibitor at 50% inhibition (I₅₀) also known as IC50 here represents the affinity value, the average IC50 of three experiments performed in duplicate are shown accompanied by a standard deviation from repeat experiments, K_i is the inhibition constant of the given inhibitor.

Peptide	Molecular Weight (g/mol)	Top (μM)	Conc. Average IC_{50} (μM) \pm SD (μM)	K_i (μM)
PNA_1	1825	1142	N/A	N/A
PNA_2	1973	1056	110 \pm 53	35
PNA_3	1799	1158	N/A	N/A
PNA_4	1138	1831	N/A	N/A
PNA_5	1281	1626	N/A	N/A
PNA_6	1118	1863	N/A	N/A
PNA_7	2070	1006	121 \pm 95	39
PNA_8	2270	918	152 \pm 129	49



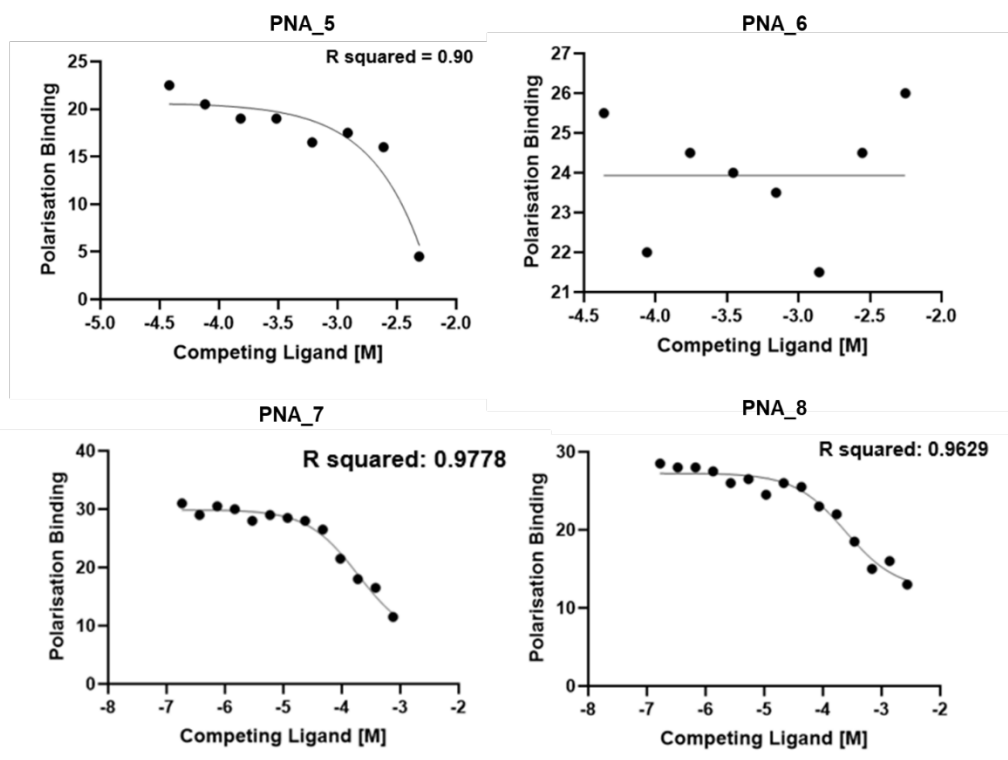


Figure 2. Fluorescence polarization competition assay IC_{50} curve of first-generation peptides. Curves represent averaged data from multiple experiments, where each data point is the mean of duplicate samples. Each experiment was repeated three times.

3.2 Second-generation cyclical peptide investigation provides new artificial sequences which bind to PCNA at a high rate that are repeatable.

PNA8 was used as a new starting point for the design of a second generation of peptides, PNA8_1 to PNA8_7. An improvement was made in the variability between experiments. PNA8_1 has the highest variability of 25% standard deviation between IC_{50} values. The IC_{50} values of PNA8_1, 2 and 5 have an improvement in binding to hPCNA from the starting scaffold of PNA_8 (Table 4), for the increased binding shown in the IC_{50} values difference from PNA_8 of 49 μ M, to PNA8_5 of 7 μ M. Interestingly the new generation peptides also have the least amount of visible aggregation present, as the solution when solubilized was immediately clear and after pelleting no visible pellet was detected. This could be due to an improvement in the peptide manufacturing.

The remaining peptides PNA8_6 and 7 had visible cloudiness and aggregation upon solubilization which led to variable results and no conclusive IC_{50} value. Further measures were employed to allow solubilization of these peptides, such as using a higher volume of water to dissolve solid powder and lowering the initial concentration for the dilution series.

Table 3. In depth analysis of second-generation cyclical peptides sequences at positions along the PIP-box. P21 μ and PNA8 included for reference. P indicates position of amino acid in the peptide, e.g. P1 is the N-terminal amino-acid, P2 is the second amino-acid. P1 and P15 around joined in the round for cyclisation.

Peptide	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15
p21μ	Lys	Arg	Arg	Gln	Thr	Ser	Met	Thr	Asp	Phe	Tyr	His	Ser	Lys	Arg
PNA8	D-Glu	D-Phe(4-CF3)	Trp(5-OH)	D-Lys	Trp	Pro	Lys	Orn	D-2-Nal	Trp	D-Trp	D-Nle	Hse(Me)	Ile	Arg
PNA8_1	D-Glu	D-Phe(4-CF3)	Trp(5-OH)	D-Lys	Trp	Pro	Lys	Nar	D-2-Nal	Trp	D-Trp	D-Nle	Hse(Me)	Ile	Arg
PNA8_2	D-Glu	D-Phe(4-CF3)	Trp(5-OH)	D-hARG	Trp	Pro	Lys	Orn	D-2-Nal	Trp	D-Trp	D-Nle	Hse(Me)	Ile	Arg
PNA8_5	D-Glu	D-Phe(4-CF3)	Trp(5-OH)	D-Lys	Trp	Pro	Lys	Orn	D-2-Nal	Trp	D-Trp	D-hLeu	hLeu	Ile	Arg
PNA8_6	D-Gln	2-Nal	Nar	D-hARG	Trp	Hyp	Lys	Nar	D-Lys	Cys(Mob)	D-Abu(cHex)	D-Nle	Leu	Ser	ADMA
PNA8_7	D-Thr	D-hPhe	Ser(Bn)	D-hARG	Tyr	Pro	D-Arg	Nar	D-Leu	Ser(Me)	D-hTyr(Me)	D-Nle	Nva	Hse	Ala(bthiophen-3-yl)

Table 4: Fluorescence polarization assay of second-generation peptides. A top concentration of 1812 μM was used before 8x 1 in 2 dilutions used to calculate IC_{50} and K_i value. The concentration of the free inhibitor at 50% inhibition (I_{50}) also known as IC_{50} here represents the affinity value, the average IC_{50} of three experiments performed in duplicate are shown accompanied by a standard deviation from repeat experiments, K_i is the inhibition constant of the given inhibitor.

Peptide	Molecular Weight (g/mol)	Average IC_{50} (μM) \pm SD (μM)	K_i (μM)
PNA8_1	2299	106 \pm 21	34
PNA8_2	2313	78 \pm 3	25
PNA8_5	2297	24 \pm 4	7
PNA8_6	2232	N/A	N/A
PNA8_7	2094	N/A	N/A

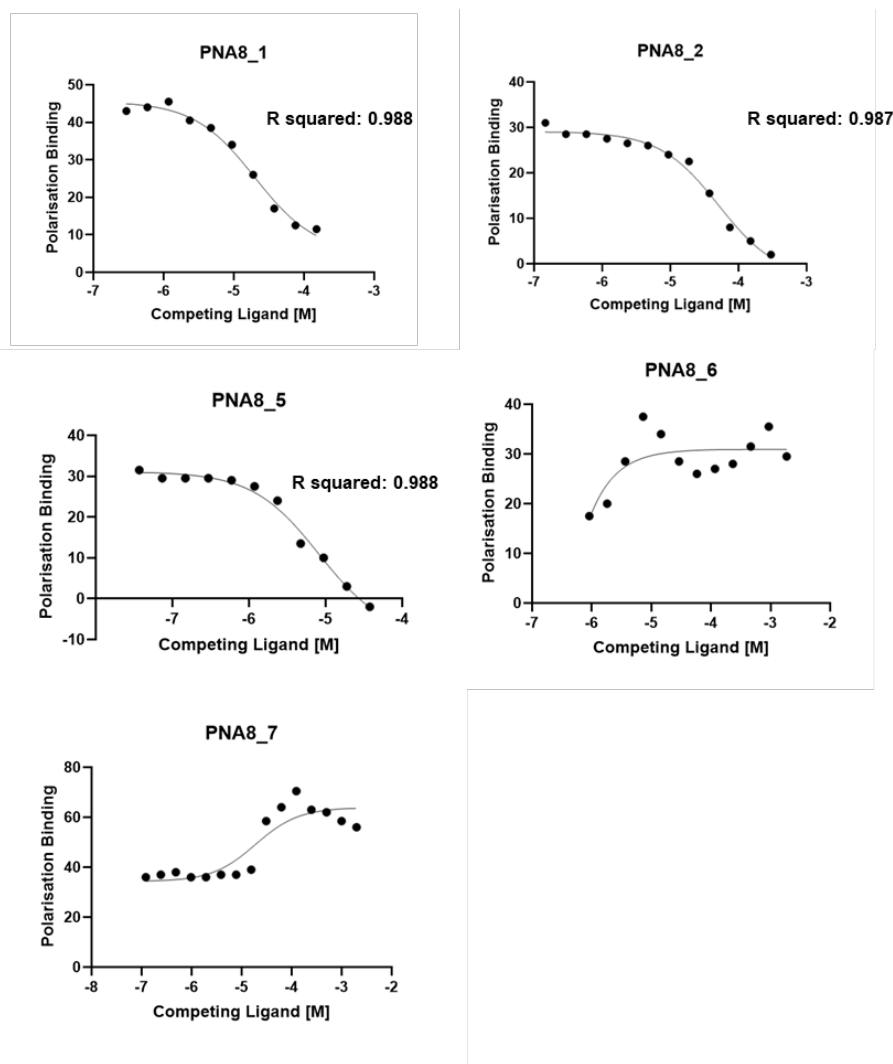


Figure 3. Fluorescence polarization competition assay IC_{50} curve of second-generation peptides. PNA8_1 R^2 of 0.988. PNA8_2 R^2 of 0.987. PNA8_5 R^2 of 0.988. PNA8_6 and PNA8_7 did not provide a valid R^2 value. Curves represent averaged data from multiple experiments, where each data point is the mean of duplicate samples. Each experiment was repeated three times.

4. Discussion and Conclusion

4.1 The newly designed first generation PNA_8 peptide presents new interactions with the PCNA binding surface.

The second generation PNA peptides were used in computational model studies to investigate structurally the impact of the new sequences and understand the binding studies further. The PNA_8 peptide presents new interactions with the PCNA binding surface not seen in previous work from its change in sequence as specific positions (P) (Figure 3B). P8, an ornithine residue at which point the cyclisation of the peptide curves, has a side chain which interacts with residue Glu232 of PCNA, with a 2.2 Å hydrogen bond (Figure 3A). P1 and P15 are joined for cyclisation, and their side chains interact with each other; the negatively charged D-glutamic acid sidechain makes a salt bridge with the positively charged arginine side chain at 2.4 Å. This P15

arginine also interacts with P4 D-Lys through the carbonyl backbone and amide backbone respectively, with a hydrogen bond at 2.7 Å (Figure 3A).

The PCNA residue Asp124 interacts with the sidechains of two PNA_8 residues; the P3 modified Trp(5-OH) residue interacts with the modified hydroxyl group and the side chain of P4 D-Lys. These two interactions form a 2.8 Å hydrogen bond and 3.0 Å salt bridge respectively (Figure 3A). P11 D-Trp and P14 Iso interact between their carbonyl and amide backbone respectively, with a 2.4 Å hydrogen bond (Figure 3A).

PNA_8 was chosen as a scaffold mimetic for the synthesis of a second generation of peptides, based on its measured K_i , lower top concentration in experimentation, and reasonable standard of deviation, to be improved.

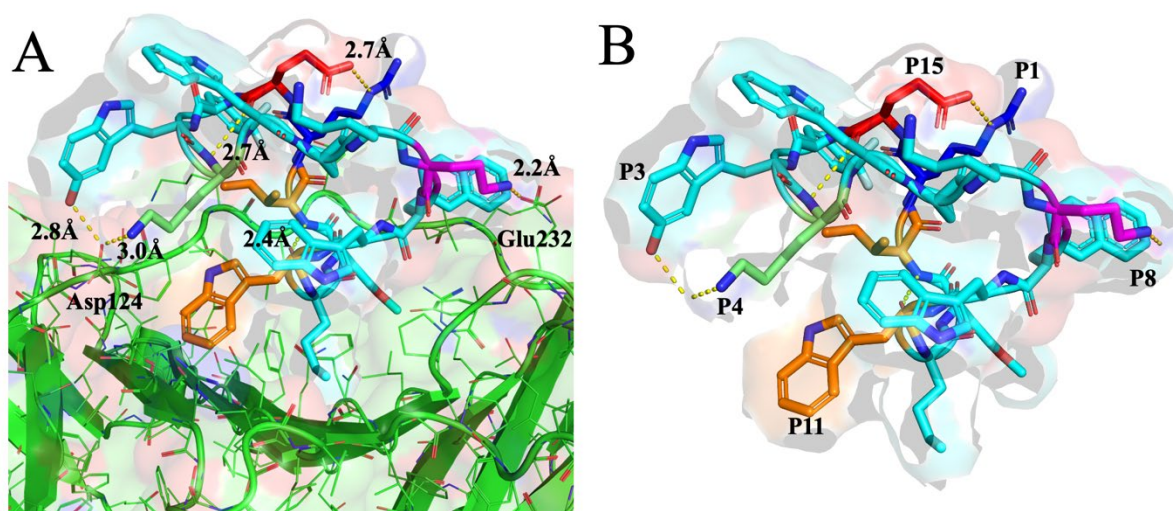


Figure 3. Computational model of PNA_8 peptide docked to surface of PCNA (PDB:6E49). PNA_8 sequence shown in Table 3. A) interactions between PNA8 residues of note and PCNA surface. B) PNA8 residues of note highlighted; P8 Orn (ornithine) shown in magenta, P1 D-Glu shown in blue, P15 Arg shown in red, P4 D-Lys shown in green, P11 and P14 D-Trp and isoleucine shown in orange.

4.2 PNA_8 gives light to a second-generation of peptides which provides a new vastly improved sequence.

The distinguishing factor between PNA8 and PNA8_1 is the substitution of the P8 Orn ($N_5H_{12}N_2O_2$) to a L-nor-arginine (Nar) ($N_5H_{12}N_4O_2$). These are similar in structure, but have a notable exception of the additional N atoms this provides. The difference in K_i value between PNA8 and 8_1 improved from 49 μ M to 34 μ M, this could be strengthening the interactions between P8 and Glu232. The following second-generation peptides do not have this Nar residue, but maintain the ornithine.

The second-generation peptide PNA8_2 has the distinguishing featural change from PNA8 at P4 from the D-Lys to a D-homoarginine (D-hArg) (Table 4). This is reminiscent of the change made in previous research of the p21 μ PIP-box, of position 146 in the peptide which is originally a serine, to an arginine residue, as the longer positive side chain allows for an increased strength salt bridge interaction with negatively charged residues positioned later in the sequence. This is shown in the D-hArg residue computational model as the lengthening of the side chain to decrease the distance of the interaction from 3.0 Å to 2.8 Å, allowing for a

stronger salt bridge interaction with the previously mentioned negatively charged PCNA residue Asp124 (Figure 4A).

The distinguishing factor between PNA8 of the first-generation peptides and of PNA8_5 in its second generation is the replacement of Nle and Hse to two Leu analogues at P12 and P13 (Figure 4C). More specifically, D-hLeu and hLeu (Figure 4C). This improvement in affinity for hPCNA may be due to these leucine residue side chains providing more branched methyl ends, which could be interacting in surface hydrophobic pockets provided from the hydrophobic C-terminal PCNA residues such as Phe₂₄₉, Phe₂₄₉, Ile₂₅₀ and Ala₂₅₁ (Figure 4B). Most interestingly this is no change to the PIP-box residues, but rather a flanking residue; because of its cyclical nature this causes a larger impact on binding affinity as it is closer to the central PIP-box binding domain.

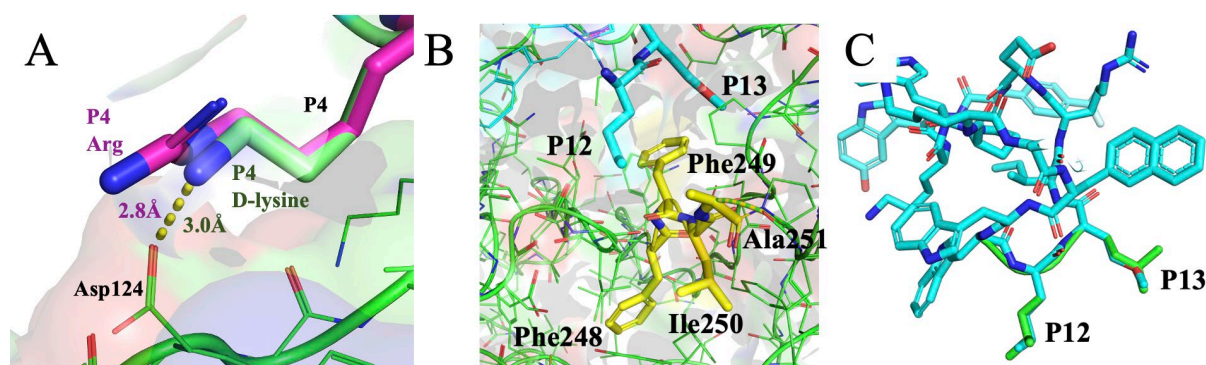


Figure 4. Second generation peptides structural features. A) Position 4 D-Lys shown in PNA8 of computational model shown in green sticks in 3.0 Å salt bridge interaction with PCNA residue Asp124. Superimposed with P4 D-hArg shown in PNA8_2 computational model in magenta sticks, in 2.8 Å salt bridge interaction with PCNA residue Asp124. B) Position 12 and 13 Nle (norleucine) and Hse (homoserine) shown in PNA8 in computational model in cyan sticks. PCNA hydrophobic residues Phe₂₄₉, Phe₂₄₉, Ile₂₅₀ and Ala₂₅₁ shown in yellow sticks. C) P12 and P13 Leu residues shown in PNA8_5 computational model in green sticks, superimposed with PNA8 model in cyan sticks.

The PNA8_5 peptide has been shown to have the best K_i of the second-generation peptides and can be used in further validation experiments. X-ray crystallography will need to be carried out to confirm and validate structural conclusions.

Peptidomimetic drug pipelines often reach the point of requiring a modification to enable cell permeability; a convenient method of improving cell uptake is through cyclisation of the peptide. Cyclic peptides have been shown to enter the mammalian cell cytosol by multiple mechanisms, including passive diffusion, facilitated predominantly by hydrophobic side chains and small size (approximately 10 amino acids long), and endocytic uptake and endosomal escape[60]. Furthermore, previous research has shown that specifically p21 based cyclic peptides may have cell penetration without the use of an auxiliary tag[67]. Previous research has shown the use of a fluorescent sensor peptide to indicate binding to PCNA. This sensor could be applied in a cellular assay to investigate the effectiveness of this peptide library and PNA8_5. This could also be indirectly done through a simple microscale thermophoresis (MST) experiment to determine if the peptide binding to the

designed target PCNA in a cellular setting is achieving the desired outcome. An MST assay could be performed to determine the cell viability and therefore cell proliferation after the peptide has been applied.

The positions of the cyclized peptides which don't interfere with the PIP binding residues could be used for tags and other cell penetrating features. To add a Cell Penetrating Peptide (CPP) or nuclear localizing (NLS) tag to mediate the transport of proteins from the cytoplasm into the nucleus such as SV40 to the N- or C- terminal of this structure would, based on computational model, not interfere with PIP-box binding and secondary structure. Avoiding this interference is crucial as this was found to be a problem found for other investigated linkers incorporated in the p21 PIP-box peptides [26, 68].

These new interactions generated in the PNA second generation library were not seen previously in the p21 peptide and present an interesting new sequence for a possible PCNA inhibiting therapeutic. An investigation into their combination could improve the new peptide further in a third generation for optimization and reach nanomolar concentrations before further cellular testing.

Chapter 7: Discussion

Summary of Findings and Future Directions

One of the main problems to overcome of conventional cancer therapy is the low specificity of chemotherapeutic drugs for carcinogenic cells over healthy cells. Most cancer therapeutics act both on healthy and diseased tissues, generating severe side effects for the patients [6]. This highlights the necessity for effective cancer treatments and therapeutics. To fight this, PCNA has been used as a target in recent years for cancer and mimetic research. This thesis presents the beginning research of the peptide to drug pipeline in the investigation for optimal peptidomimetic cancer and fungal therapeutics. Through the analysis of protein-peptide interactions, the biochemical structures can be used to our biomedical advantage to understand their activity and help treat major diseases. The insights illustrated are crucial to the future design of PCNA inhibitors.

PCNA is a promising target for the inhibition of cancer cell growth as it is critical for cell survival[69]. The trimer structure of PCNA forms a ring-shaped oligomer allowing DNA to pass through the middle and interacting proteins to dock on the outside of the ring. Without this structural formation, there is a loss of critical protein interactions at the replication fork and therefore a loss of DNA replication in the cell. Due to the location of subunit-subunit termini, the addition of a purification tag to the protein can hamper crystallography and biophysical experiments, as the trimer complex folding can be impeded [70]. To avoid these complications, a tag-less, step-wise purification was designed. Chapter 2 presented this method which provides a replicable system to obtain tag-less protein, allowing the further structural research of PCNA for cancer treatments, which was utilised in Chapters 5 and 6 [62]. Using this new method new PCNA targeting mimetic structures have been found.

The known regulator protein p21 was used to investigate new approaches for targeting and inhibiting PCNA for the development of cancer therapeutics. A peptide of the binding domain of p21 has been shown to bind maintaining the affinity of the full protein [27]. The p21 protein has been shown in the human system to block binding to the docking platform and halt cell replication [27]. Using the p21 native structure as a starting scaffold in Chapter 3 this thesis presents current research into structural modifications of the design. Molecular modelling and experimental approaches have been established to develop these novel synthetic peptides. A mutagenesis approach was used to identify what characteristics increase or decrease peptide affinity for PCNA [56]. Differences in PIP-boxes was shown at a residue level, to identify structural points which form a tightly bound confirmation. This research revealed a new peptide sequence with a stronger binding affinity than the native sequence.

Further studies into peptidomimetic characteristics for cell permeability revealed a linker secondary structure that allows uptake into breast cancer cells [68]. A linker structure was used to cyclise the peptide and provide support to the critical secondary structure. The inclusion of this linker structure in Chapter 4 did not interfere with the binding of the peptide to the protein surface, and allowed key confirmation to be achieved upon interaction. It influenced the peptide secondary structure and upon *in vitro* investigation allowed uptake in breast cancer cells. This demonstrates how macrocyclization has improved cell permeability of short peptidomimetics. This is a significant step towards the end of the drug pipeline for this therapeutic.

While this knowledge will be used for the design of future inhibitors for a therapeutic, it requires investigation into the later stages of the pipeline. Investigation of the p21 native sequence led to the synthesis of the rational design peptide presented in Chapter 3, which binds hPCNA with the highest affinity of any p21 peptidomimetic. In Chapter 4, A peptide linker was found in the investigation of macrocyclic peptides which bound to hPCNA with high affinity, and initial research suggests this linker affords cellular permeability. The incorporation of this linker into the rational design peptide could then be tested to determine the drug-like qualities required of the mimetic for further use. A more in-depth investigation of cell permeability must be performed to confirm cellular and nuclear permeability of the new peptide. The hypothesis of the anti-proteolytic nature of the linker must be tested through a proteolysis experiment, using mouse proteases *in vitro* to determine if the linker is unrecognisable to these enzymes, and will not be inevitably broken down. If cleavage occurs, determination of the location is important to overcome this. It can be then tested *in vitro* with cancer cell lines for a number of purposes, such as to test selectivity for cancer lines over healthy human cells. It can also be tested for the interaction with cancer cells, for the peptide's ability to inhibit DNA-replication and therefore cause cell death. This can be indirectly done through assessment by an MTS assay, to determine if the peptide binding to the designed target *in vitro*, and is achieving the desired outcome. This can therefore determine the cell viability and proliferation rate after the peptide has been applied. All of this must be done before moving on to *in vivo* experimentation.

New structural designs are critical to the ongoing problem solving of cancer resistance. Pepticom has designed novel peptide drug candidates, utilizing the *in silico* AI software to optimize the discovery process, accelerate time to market, and reduce development risks and costs. These sequences have been optimised to synthesise peptides for control of specific chemical attributes and risk factors. Through Pepticom's algorithm, the design of new peptide sequences using synthetic amino acids in cyclical peptide structures were tested in Chapter 6, and found high affinity confirmations to be used in further PCNA studies *in vitro*. The resultant candidates of this preliminary research have been identified, and can now progress to lengthened measures, the next step would be to use X-ray crystallography to analyse the structure of the new designed peptides bound to hPCNA, to determine if the co-structure is interacting as hypothesised.

This isn't the full extent of the use of PCNA targeting peptidomimetic research, it can be used on other species, as PCNA has been shown to be present in almost all forms of life [71]. *Aspergillus fumigatus* is abundant in the environment and the most common cause of invasive fungal infection [37]. Central nervous system aspergillosis has mortality rates of ~90% [38]. As PCNA has important roles in replication and cell survival, it is hypothesised that it is a potential target for antifungal treatments [44], illustrating the importance of investigating the crystal structure of *A. fumigatus* PCNA bound with mimetics, to understand how interacting peptides bind to aid rational drug design. This was done in Chapter 5 to reveal a newly designed mimetic with a new secondary structure with high affinity binding which can be utilised in further fungal mimetic research. The peptide sequence and secondary structure can be tested to understand the antifungal effect on fungal cell lines. The antifungal

susceptibility of these synthesised peptides can be tested by using a broth microdilution method. Antifungals such as anidulafungin and amphotericin B can be used for positive control purposes, the designed peptides of interest are diluted into media, using two-fold serial dilution to cover a wide concentration range. *A. fumigatus* isolates can be cultured on dextrose agar with antibiotics until conidiation. After resuspension of isolates at 24 hours to prepare the inoculum, these aspergillus isolates endpoints are read. Antifungal susceptibility is interpreted through the minimum inhibitory concentration (MIC). This *in vitro* experimentation can determine the next steps to be taken in the anti-fungal drug pipeline.

Concluding remarks

The current project presents the biomedical developments in the world of peptidomimetic research in PCNA inhibitors for cancer and fungal therapeutics. The presented findings build on the previous research of this drug target to identify key features of interactions for the progression of the drug discovery pipeline. As this pipeline is ready for development it is expected the next stages of *in vitro* and *in vivo* work will progress the structures to a drug like state. X-ray crystallography needs to be attempted on Pepticom's cyclical peptides, and after analysis such as for those peptides from previous work with crystal structures already examined, proteolysis experiments will be done to inform real world use and weaknesses that may require alteration in the backbone of the peptides, following this antifungal and anticancer susceptibility can be tested in cellular experiments. This *in vitro* data will inform the further adaptation and use of these peptides as an effective drug for a broad and cancer specific treatment to lower the impact of cancer on patients.

Chapter 8:

Bibliography

Bibliography

This includes references for publications cited in Chapters 1, 6 and 7. Bibliographies for all other sections are included in their chapters.

- [1] 'Cancer in Australia statistics'. *Cancer Australia* (2023), <https://www.canceraustralia.gov.au/impacted-cancer/what-cancer/cancer-australia-statistics>.
- [2] Australia High Cancer Survival Rates Attributed To Earlier Detection. (2019) *The Guardian*, <https://www.theguardian.com/australia-news/2019/sep/2011/australia-high-cancer-survival-rates-attributed-to-earlier-detection>,
- [3] Collins, K., Jacks, T., and Pavletich, N. P. The cell cycle and cancer. *Proceedings of the National Academy of Sciences* (1997) 94, 2776-2778, doi:10.1073/pnas.94.7.2776.
- [4] Pappou, E. P., and Ahuja, N. The role of oncogenes in gastrointestinal cancer. *Gastrointest Cancer Res* (2010), S2-s15,
- [5] Dagogo-Jack, I., and Shaw, A. T. Tumour heterogeneity and resistance to cancer therapies. *Nat Rev Clin Oncol* (2018) 15, 81-94, 10.1038/nrclinonc.2017.166.
- [6] Pucci, C., Martinelli, C, & Ciofani, G. Innovative approaches for cancer treatment: current perspectives and new challenges. . *Ecancermedicalscience* (2019) 13, 10.3332/ecancer.2019.961.
- [7] Amjad, M., Chidharla, A & Kasi A. (2023) *Cancer Chemotherapy.*, StatPearls Publishing.
- [8] Forman, A. D., Meyers, C. A., and Levin, V. A. (2009) CHAPTER 33 - Neurotoxic Effects of Pharmaceutical Agents IV: Cancer Chemotherapeutic Agents, In *Clinical Neurotoxicology* (Dobbs, M. R., Ed.), pp 372-381, W.B. Saunders, Philadelphia.
- [9] Weber, G. F. DNA Damaging Drugs. *Molecular Therapies of Cancer* (2014), 9-112, 10.1007/978-3-319-13278-5_2.
- [10] Wang, Y., Probin, V., and Zhou, D. Cancer therapy-induced residual bone marrow injury- Mechanisms of induction and implication for therapy. *Curr Cancer Ther Rev* (2006) 2, 271-279, 10.2174/157339406777934717.
- [11] Baskar, R., Dai, J., Wenlong, N., Yeo, R., and Yeoh, K. W. Biological response of cancer cells to radiation treatment. *Front Mol Biosci* (2014) 1, 24, 10.3389/fmolb.2014.00024.
- [12] Bernier, J., Hall, E. J., and Giaccia, A. Radiation oncology: a century of achievements. *Nat Rev Cancer* (2004) 4, 737-747, 10.1038/nrc1451.
- [13] Majeed, H. G., V. (2023) *Adverse Effects of Radiation Therapy*, StatPearls Publishing.
- [14] Anand, U., Dey, A., Chandel, A. K. S., Sanyal, R., Mishra, A., Pandey, D. K., De Falco, V., Upadhyay, A., Kandimalla, R., Chaudhary, A., Dhanjal, J. K., Dewanjee, S., Vallamkondu, J., and Pérez de la Lastra, J. M. Cancer chemotherapy and beyond: Current status, drug candidates, associated risks and progress in targeted therapeutics. *Genes & Diseases* (2023) 10, 1367-1401, <https://doi.org/10.1016/j.gendis.2022.02.007>.
- [15] Hamdan, S. M., and van Oijen, A. M. Timing, Coordination, and Rhythm: Acrobatics at the DNA Replication Fork*. *Journal of Biological Chemistry* (2010) 285, 18979-18983, <https://doi.org/10.1074/jbc.R109.022939>.
- [16] Choe, K. N., and Moldovan, G. L. Forging Ahead through Darkness: PCNA, Still the Principal Conductor at the Replication Fork. *Mol Cell* (2017) 65, 380-392, 10.1016/j.molcel.2016.12.020.
- [17] Balakrishnan, L., and Bambara, R. A. Okazaki fragment metabolism. *Cold Spring Harb Perspect Biol* (2013) 5, 10.1101/cshperspect.a010173.
- [18] Moldovan, G. L., Pfander, B., and Jentsch, S. PCNA, the maestro of the replication fork. *Cell* (2007) 129, 665-679, 10.1016/j.cell.2007.05.003.
- [19] Kroker, A. J., and Bruning, John B. p21 Exploits Residue Tyr151 as a Tether for High-Affinity PCNA Binding. *Biochemistry* (2015) 54, 3483-3493, 10.1021/acs.biochem.5b00241.
- [20] Roa, S., Avdievich, E., Peled, JU., MacCarthy, T., Werling, U., Kuang, FL., Zhao, C., Berman, A., Cohen, PE., Edelman, W., Schnarff, MD. . Ubiquitylated PCNA plays a role in somatic

- hypermethylation and class-switch recombination and is required for meiotic progression. *Proceedings of the National Academy of Sciences of the United States of America* (2008) 105, 16248–16253,
- [21] Sakakura C., H. A., Tsujimoto H., Ozaki K., Sakakibara T., Oyama T., Ogaki M., and Takahashi T. . The anti-proliferative effect of proliferating cell nuclear antigen-specific antisense oligonucleotides on human gastric cancer cell lines. *Surg Today* (1995) 25, 184-186, 10.1007/BF00311097.
- [22] Mailand, N., Gibbs-Seymour, I., and Bekker-Jensen, S. Regulation of PCNA-protein interactions for genome stability. *Nat Rev Mol Cell Biol* (2013) 14, 269-282, 10.1038/nrm3562.
- [23] Gu, L., Smith, S., Li, C., Hickey, R. J., Stark, J. M., Fields, G. B., Lang, W. H., Sandoval, J. A., and Malkas, L. H. A PCNA-derived cell permeable peptide selectively inhibits neuroblastoma cell growth. *PLoS One* (2014) 9, e94773, 10.1371/journal.pone.0094773.
- [24] Tan, Z., Wortman, M., Dillehay, K. L., Seibel, W. L., Evelyn, C. R., Smith, S. J., Malkas, L. H., Zheng, Y., Lu, S., and Dong, Z. Small-molecule targeting of proliferating cell nuclear antigen chromatin association inhibits tumor cell growth. *Mol Pharmacol* (2012) 81, 811-819, 10.1124/mol.112.077735.
- [25] Abbas, T., and Dutta, A. p21 in cancer: intricate networks and multiple activities. *Nat Rev Cancer* (2009) 9, 400-414, 10.1038/nrc2657.
- [26] Wegener, K. L., McGrath, Amy E., Dixon, Nicholas E., Oakley, Aaron J., Scanlon, Denis B., Abell, Andrew D., and Bruning, John B.. Rational Design of a 3 10 -Helical PIP-Box Mimetic Targeting hPCNA, the Human Sliding Clamp. *European Journal Chemistry* (2018) 24, 11325–11331, 10.1002/chem.201801734.
- [27] Gulbis, J. M., Kelman, Z., Hurwitz, J., O'Donnell, M., & Kuriyan, J. . Structure of the C terminal region of p21(WAF1/CIP1) complexed with human PCNA. *Cell* (1996) 87, 297–306, 10.1016/s0092-8674(00)81347-1.
- [28] Pan, Z. Q., Reardon, J. T., Li, L., Flores-Rozas, H., Legerski, R., Sancar, A., and Hurwitz, J. Inhibition of nucleotide excision repair by the cyclin-dependent kinase inhibitor p21. *J Biol Chem* (1995) 270, 22008-22016, 10.1074/jbc.270.37.22008.
- [29] Tan, Z., Wortman, M., Dillehay, KL., Seibel, WL., Evelyn., CR., Smith, SJ., Malkas, LH., Zheng, Y., Lu, S., Dong, Z. Small-Molecule Targeting of Proliferating Cell Nuclear Antigen Chromatin Association Inhibits Tumor Cell Growth. *Molecular Pharmacology* (2012) 81, 811-819,
- [30] Vagner, J., Qu, H., and Hruby, V. J. Peptidomimetics, a synthetic tool of drug discovery. *Curr Opin Chem Biol* (2008) 12, 292-296, 10.1016/j.cbpa.2008.03.009.
- [31] Bidwell, G. L., 3rd, and Raucher, D. Cell penetrating elastin-like polypeptides for therapeutic peptide delivery. *Adv Drug Deliv Rev* (2010) 62, 1486-1496, 10.1016/j.addr.2010.05.003.
- [32] dos Santos, M. A., Teixeira, F. B., Moreira, H. H., Rodrigues, A. A., Machado, F. C., Clemente, T. M., Brigido, P. C., Silva, R. T., Purcino, C., Gomes, R. G., Bahia, D., Mortara, R. A., Munte, C. E., Horjales, E., and da Silva, C. V. A successful strategy for the recovering of active P21, an insoluble recombinant protein of *Trypanosoma cruzi*. *Sci Rep* (2014) 4, 4259, 10.1038/srep04259.
- [33] Taylor, J. The synthesis and study of side-chain lactam-bridged peptides. *American Peptide Society* (2004) 66, 49-75,
- [34] Bitar , D., Lortholary , O, Le Strat , Y, Nicolau , J, Coignard , B, Tattevin , P, Che , D & Dromer ,F. . Population-based analysis of invasive fungal infections, France, 2001-2010. *Emerg Infect Dis* (2014) 20, 1149–1155, 10.3201/eid2007.140087.
- [35] McNeil, M., Nash, SL, Hajjeh, RA, Phelan, MA, Conn, LA, Plikaytis, BD, & Warnock DW. Trends in mortality due to invasive mycotic diseases in the United States, 1980-1997. *Clin Infect Dis* (2001) 33, 641–647, 10.1086/322606.
- [36] Abruzzo, G. K., Flattery, A. M., Gill, C. J., Kong, L., Smith, J. G., Pikounis, V. B., Balkovec, J. M., Bouffard, A. F., Dropinski, J. F., Rosen, H., Kropp, H., and Bartizal, K. Evaluation of the echinocandin antifungal MK-0991 (L-743,872): efficacies in mouse models of disseminated

- aspergillosis, candidiasis, and cryptococcosis. *Antimicrob Agents Chemother* (1997) 41, 2333-2338, 10.1128/aac.41.11.2333.
- [37]Yoon, H., Choi, HY, Kim, YK, Song, YJ, & Ki, M. Prevalence of fungal infections using National Health Insurance data from 2009-2013, South Korea. *Epidemiol Health* (2014) 36, 10.4178/epih/e2014017.
- [38]Gallien S, F. S., Porcher R, Bottero J, Ribaud P, Sulahian A, Socie G and Molina JM. Therapeutic outcome and prognostic factors of invasive aspergillosis in an infectious disease department: a review of 34 cases. *Infection* (2008) 36, 533–538, 10.1007/s15010-008-7375-x.
- [39]Sugui, J., Kwon-Chung, KJ, Juvvadi, PR, Latge, JP, & Steinbach, WJ. *Aspergillus fumigatus* and related species. *Cold Spring Harb Perspect Med* (2015) 5, a019786, 10.1101/cshperspect.a019786.
- [40]Steinbach, W. Are we there yet? Recent progress in the molecular diagnosis and novel antifungal targeting of *Aspergillus fumigatus* and invasive aspergillosis. *PLoS Pathog* (2013) 9, 10.1371/journal.ppat.1003642.
- [41]Jenks, J. D., and Hoenigl, M. Treatment of Aspergillosis. *J Fungi (Basel)* (2018) 4, 10.3390/jof4030098.
- [42]Ben-Ami, R., Lewis, RE & Kontoyiannis, DP. Enemy of the (immunosuppressed) state: an update on the pathogenesis of *Aspergillus fumigatus* infection. *Br J Haematol* (2010) 150, 406–417, 10.1111/j.1365-2141.2010.08283.
- [43]Stoimenov, I., and Helleday, T. PCNA on the crossroad of cancer. *Biochem. Soc. Trans.* (2009) 37, 605–613, 10.1042/BST0370605.
- [44]Marshall, A. C., Kroker, A.J., Murray, L.A., Gronthos, K., Rajapaksha, H., Wegener, K.L., and Bruning, J.B. Structure of the sliding clamp from the fungal pathogen *Aspergillus fumigatus* (AfumPCNA) and interactions with human p21. *FEBS J.* (2017) 284, 985–1002, 10.1111/febs.14035.
- [45]Cho, I. T., Kim, D. H., Kang, Y. H., Lee, C. H., Amangyelid, T., Nguyen, T. A., Hurwitz, J., and Seo, Y. S. Human replication factor C stimulates flap endonuclease 1. *J Biol Chem* (2009) 284, 10387-10399, 10.1074/jbc.M808893200.
- [46]Elcock, A. H., and McCammon, J. A. Identification of protein oligomerization states by analysis of interface conservation. *Proceedings of the National Academy of Sciences* (2001) 98, 2990-2994, doi:10.1073/pnas.061411798.
- [47]Janin, J., and Rodier, F. Protein-protein interaction at crystal contacts. *Proteins* (1995) 23, 580-587, 10.1002/prot.340230413.
- [48]Hatti, K., Biswas, A., Chaudhary, S., Dadireddy, V., Sekar, K., Srinivasan, N., and Murthy, M. R. N. Structure determination of contaminant proteins using the MarathonMR procedure. *Journal of Structural Biology* (2017) 197, 372-378, <https://doi.org/10.1016/j.jsb.2017.01.005>.
- [49]Karplus, P. A., and Diederichs, K. Linking crystallographic model and data quality. *Science* (2012) 336, 1030-1033, 10.1126/science.1218231.
- [50]Evans, P., and McCoy, A. An introduction to molecular replacement. *Acta Crystallogr D Biol Crystallogr* (2008) 64, 1-10, 10.1107/s0907444907051554.
- [51]Kleywegt, G. J., and Alwyn Jones, T. (1997) [11] Model building and refinement practice, In *Methods in Enzymology*, pp 208-230, Academic Press.
- [52]Terwilliger, T. C. Using prime-and-switch phasing to reduce model bias in molecular replacement. *Acta Crystallogr D Biol Crystallogr* (2004) 60, 2144-2149, 10.1107/s0907444904019535.
- [53]Dieckman, L. M., Freudenthal, B. D., and Washington, M. T. PCNA structure and function: insights from structures of PCNA complexes and post-translationally modified PCNA. *Subcell Biochem* (2012) 62, 281-299, 10.1007/978-94-007-4572-8_15.
- [54]Nikolovska-Coleska, Z., Wang, R., Fang, X., Pan, H., Tomita, Y., Li, P., P. Roller, P., Krajewski, K., G. Saito, N., Jeanne A. Stuckey, J., Wang, S. Development and optimization of a binding assay

- for the XIAP BIR3 domain using fluorescence polarization. *Analytical Biochemistry* (2004) 332, 261-273, <https://doi.org/10.1016/j.ab.2004.05.055>.
- [55]Horsfall, A., Abell, A., Bruning, J. Targeting PCNA with peptide mimetics for therapeutic purposes. *ChemBioChem* (2019) 21, 442-450, 10.1002/cbic.201900275.
- [56]Horsfall, A. J., Vandborg, B. A., Kowalczyk, W., Chav, T., Scanlon, D. B., Abell, A. D., and Bruning, J. B. Unlocking the PIP-box: A peptide library reveals interactions that drive high-affinity binding to human PCNA. *J Biol Chem* (2021) 296, 100773, 10.1016/j.jbc.2021.100773.
- [57]H.M. Berman, J. W., Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne. . The Protein Data Bank *Nucleic Acids Research* (2000) 28, 235-242 <https://doi.org/10.1093/nar/28.1.235>.
- [58]Shapira, L., Lerner, S., Assayag, G., Vardi, A., Haham, D., Bar, G., Kozokaro, VF., and Robicsek, M. E., Lerner, I., Michaeli, A. Discovery of novel spike/ACE2 inhibitory macrocycles using in silico reinforcement learning. *Front. Drug. Discov.* (2022) 2, 10.3389/fddsv.2022.1085701.
- [59]Li Petri, G., Di Martino, S., De Rosa, M. Peptidomimetics: An overview of recent medicinal Chemistry efforts toward the discovery of novel small molecule inhibitors. *J. Med. Chem.* (2022) 65, 7438–7475, 10.1021/acs.jmedchem.2c00123.
- [60]Dougherty, P. G., Ashweta, S, & Dehua, P. Understanding Cell Penetration of Cyclic Peptides. *Chem. Rev.* (2019) 119, 10241–10287, 10.1021/acs.chemrev.9b00008.
- [61]Michaeli, A., Lerner, I., Zatspein, M., Mezan, S., Kilshtain, AV. Discovery of Novel GABAAR Allosteric Modulators Through Reinforcement Learning. *Curr Pharm Des.* (2020) 26, 5713-5719, 10.2174/1381612826666201113104150.
- [62]Vandborg, B., Holroyd, DL, Pukala, T & Bruning, JB. Production of recombinant human proliferating cellular nuclear antigen (PCNA) for structural and biophysical characterization. . *Protein Expr Purif.* (2023) 212, 10.1016/j.pep.2023.106353.
- [63]Buzovetsky, O., Kwon, Y., Tuyet Pham, N., Kim, C., Ira, G., Sung, P., Xiong, Y. Role of the Pif1-PCNA Complex in Pol δ -Dependent Strand Displacement DNA Synthesis and Break-Induced Replication. *Cell Reports* (2017) 21, 1707-1714, <https://doi.org/10.1016/j.celrep.2017.10.079>.
- [64]Jacobson, M. P., R. A. Friesner, Z. Xiang, B. Honig. On the role of the crystal environment in determining protein side-chain conformations. *J. Mol. Biol.* (2002), 597–608,
- [65]Jacobson, M. P., D. L. Pincus, C. S. Rapp, T. J. F. Day, B. Honig, D. E. Shaw, R. A. Friesner. A hierarchical approach to all-atom protein loop prediction. *Proteins* (2004), 351–367,
- [66]Schrodinger, L. The PyMOL Molecular Graphics System, Version 1.2r3pre. (2015),
- [67]Horsfall, A. J., Chav, Theresa, Bruning, John B , Abell, Andrew D. A turn-on fluorescent PCNA sensor. *Bioorganic & Medicinal Chemistry Letters.* (2021) 41, <https://doi.org/10.1016/j.bmcl.2021.128031>.
- [68]Horsfall, A. J., Vandborg, B. A., Kikhtyak, Z., Scanlon, D. B., Tilley, W. D., Hickey, T. E., Bruning, J. B., and Abell, A. D. A cell permeable bimeane-constrained PCNA-interacting peptide. *RSC Chemical Biology* (2021), 1499–1508, 10.1039/D1CB00113B.
- [69]De March, M., Merino, N., Barrera-Vilarmau, S., Crehuet, R., Onesti, S., Blanco, F.J., and De Biasio, A. . Structural basis of human PCNA sliding on DNA. *Nat. Commun.* (2017) 8, 10.1038/ncomms13935.
- [70]Carson, M., Johnson, DH., McDonald, H., Brouillette, C.,and Delucas, LJ. His-tag impact on structure. *Acta Crystallography D Biology Crystallography* (2007) 63, 295-301, 10.1107/S0907444906052024.
- [71]Boehm, E., & Washington, MT. R.I.P. to the PIP: PCNA-binding motif no longer considered specific. *Bioessays.* (2016) 38, 1117–1122, 10.1002/bies.201600116.

Chapter 9: Appendix

Designing Fluorescent Nuclear Permeable Peptidomimetics to Target Proliferating Cell Nuclear Antigen

Aimee J. Horsfall,[▽] Theresa Chav,[▽] Jordan L. Pederick, Zoya Kikhtyak, Bethiney C. Vandborg, Wioleta Kowalczyk, Denis B. Scanlon, Wayne D. Tilley, Theresa E. Hickey, Andrew D. Abell, and John B. Bruning*Cite This: *J. Med. Chem.* 2023, 66, 10354–10363

Read Online

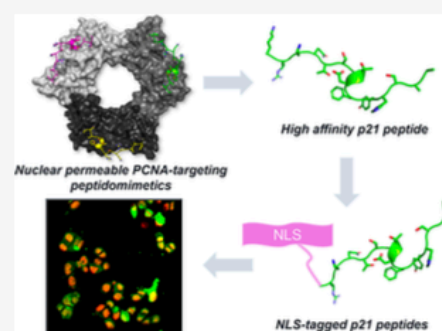
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: Human proliferating cell nuclear antigen (PCNA) is a critical mediator of DNA replication and repair, acting as a docking platform for replication proteins. Disrupting these interactions with a peptidomimetic agent presents as a promising avenue to limit proliferation of cancerous cells. Here, a p21-derived peptide was employed as a starting scaffold to design a modular peptidomimetic that interacts with PCNA and is cellular and nuclear permeable. Ultimately, a peptidomimetic was produced which met these criteria, consisting of a fluorescein tag and SV40 nuclear localization signal conjugated to the N-terminus of a p21 macrocycle derivative. Attachment of the fluorescein tag was found to directly affect cellular uptake of the peptidomimetic, with fluorescein being requisite for nuclear permeability. This work provides an important step forward in the development of PCNA targeting peptidomimetics for use as anti-cancer agents or as cancer diagnostics.



INTRODUCTION

Human proliferating cell nuclear antigen (PCNA) is a member of the sliding clamp family of proteins and is an essential processivity factor and mediator of DNA replication, DNA repair, and cell cycle processes.^{1–8} PCNA is primarily located in the nucleus and is known to recruit over 200 proteins that carry out DNA-replication and repair. The ring-shaped structure of PCNA allows it to be loaded onto template-primer junctions to encircle double-stranded DNA. Once docked onto dsDNA, PCNA acts as a mobile docking platform for protein machinery to first associate and subsequently interact with the DNA.^{2,9–11} This function of PCNA is crucial for DNA replication and repair. As PCNA is upregulated in a large number of cancers, in response to an increased demand for DNA replication, it presents as a key target to develop anti-cancer therapeutics.^{7,12–16}

PCNA interacting proteins do so through one of the three degenerate binding sites on the trimeric PCNA (Figure 1).^{3,5,7,11,12,15,17} Each of the three monomers is comprised of two almost symmetrical domains that are connected through an unstructured loop, termed the interdomain connecting loop (IDCL). Many PCNA interacting proteins bind a hydrophobic cleft, the PCNA-interacting protein (PIP) binding site, which is located between these two domains on the PCNA surface.^{1,18,19} Proteins bind at the PIP binding site through a sequence known as the PIP-box motif, defined as Qxxϕxxj/ψ, where Q is glutamine, “x” is any amino-acid, ϕ is a hydrophobic residue, and ψ an aromatic residue (commonly

phenylalanine or tyrosine).^{19–25} For example, the cell cycle regulator protein, p21, contains the PIP-box sequence QTSMTDFY that binds PCNA with the highest known affinity (Figure 1).^{1,19,26–31} Consequently, p21 competitively blocks the interaction of other proteins with PCNA and stalls DNA replication to provide a necessary checkpoint for cell cycle regulation. A 22-mer peptide derived from p21 (residues 139–160) has been shown to inhibit polymerase δ from binding PCNA and thus shut down SV40 DNA replication.³⁰ The peptide p21_{139–160} is known to bind human PCNA with high affinity (~5 nM).^{1,15,19,30} As such, p21_{139–160} provides an ideal template to design a nuclear-permeable peptidomimetic that can act as a cancer therapeutic through blocking these PCNA-mediated interactions. Hence, p21_{139–160} could be used as a template to develop a peptide that targets PCNA for use as a diagnostic, or as an inhibitor of PCNA for application as a cancer therapeutic. However, this peptide is not nuclear permeable and an investigation to develop a nuclear permeable p21-derived scaffold was undertaken.

Efforts to further develop a p21-derived peptide into a cancer therapeutic have been limited by several factors

Received: March 16, 2023

Published: July 25, 2023



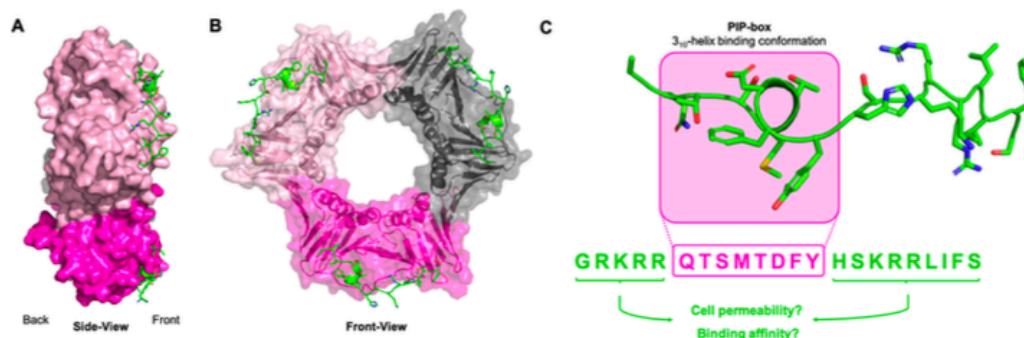


Figure 1. Structure of human PCNA complexed with p21_{139–160} (PDB: 1AXC) (A) surface representation of PCNA (light and dark pink) as a side view with p21_{139–160} bound to the front face in green. (B) Front view of PCNA as a transparent surface representation, and cartoon representation that shows the three monomers (light pink, dark pink, and grey) and p21_{139–160} in green cartoon representation with side chains as sticks, bound to each monomer. (C) p21_{139–160} in green bound in cartoon representation with side chains as sticks, and PIP-box highlighted in pink.

including lack of nuclear penetration and also the length of the peptides used. The nuclear permeability of p21 protein has been explored in detail with endogenously expressed p21 protein being translocated from the cytoplasm to the nucleus.¹ Large p21 fragments (residues 76–164) ectopically expressed also translocate to the nucleus, and residues 140–142 (RKR) of p21 were identified as essential to this nuclear permeability.^{32,33} Shorter p21-derived peptides (141–160 or 139–164), containing these “essential” residues (140–142), are not nuclear permeable. Such peptides were conjugated to long nuclear localization signal (NLS) peptides, such as Penetratin, to improve nuclear permeability and allowed the activity of the resulting peptide-conjugates to be investigated in cells (e.g. DNA replication inhibition, PCNA interaction, etc.).^{34–37} All such peptides are large constructs with greater than 30 amino acids, creating challenges for optimal pharmacokinetics.

This study sets out to improve the drug-like properties of such constructs, by identifying the optimal p21 peptide length while considering the effect on binding affinity (choosing the shortest sequence that maintains high affinity), identifying an optimal NLS sequence to maintain affinity for PCNA and determining which of the sequences are nuclear penetrant. First, to determine the importance of each section of the p21 sequence on binding affinity for PCNA, one amino-acid was sequentially truncated from either the *N*- or *C*-terminus of p21_{139–160}. The binding affinity of the resulting peptides for PCNA was characterized by surface plasmon resonance (SPR). This then informed the design of short p21 peptides that maintain high affinity for PCNA as an important step toward determining the shortest functional peptide scaffold that targets PCNA. The cell and nuclear permeability of fluorescein-tagged variants of these short peptides was then assessed in a breast cancer cell line. Macrocyclic peptides generally display improved stability and cell permeability, where recently a cell permeable fluorescent peptidomimetic targeting PCNA was reported.³⁸ A series of NLS sequences was conjugated to a short linear p21 peptide, and this short macrocyclic p21 peptide, to improve nuclear permeability in breast cancer cells. This study not only characterized the role of the sequences that flank the p21 PIP-box on PCNA binding affinity and cell permeability, but importantly, successfully identified a short p21-derived peptidomimetic sequence that is nuclear permeable.

RESULTS

Shortened p21 Scaffold Maintains High-Affinity Binding.

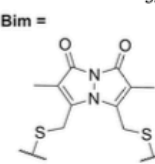
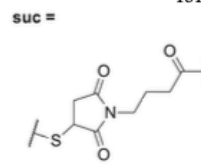
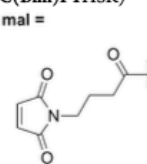
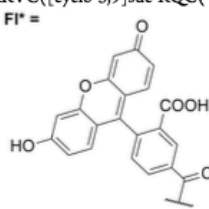
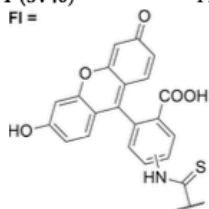
Initial efforts focused on characterizing the minimum p21 sequence required to maintain high-affinity binding to PCNA, as determined by SPR, utilizing p21_{139–160} as a starting scaffold. The upper limit of “high-affinity binding” was defined to be ~100 nM, based on the rationally designed Pogo–DNA Ligase I hybrid peptide (PL), which binds PCNA with 100 nM affinity (ITC).³⁰ PL is reported to inhibit SV40 DNA replication in HeLa cell extracts, providing a benchmark for a level of PCNA affinity that affects function.^{30,39}

A series of 15 p21 peptides were prepared, each with an additional amino acid truncated from either the *N*- or *C*-terminus of the p21_{139–160} sequence (1-0), known to bind PCNA, up until the PIP-box (144–151) (Table 1). The binding affinity of each peptide for PCNA was determined by SPR to identify the shortest sequence able to maintain high-affinity (<100 nM) binding. The binding affinity of 1-0 for PCNA was determined to be 5.14 nM, which is in agreement with literature reports.^{1,15,19,30} The binding affinities of *C*-terminally truncated peptides 1-1 and 1-2 for PCNA were similarly determined to be 14.1 and 15.9 nM. Thus, the two most *C*-terminal residues of 1-0 were not required to maintain affinity for PCNA. The affinity for PCNA was observed to gradually diminish from peptide 1-1 at 14.1 nM, through to 1-6 at 37.4 nM, where each peptide possessed one less *C*-terminal residue. The affinity of peptide 1-7 for PCNA, which lacks the seven *C*-terminal residues of 1-0, was notably lower (210 nM), and further loss of affinity was not observed with additional *C*-terminal truncations (1-8, 150 nM and 1-9, 308 nM). These data together suggest that the three residues *C*-terminal of the PIP-box (residues 152–154, as in 1-6) are required for p21 peptides to maintain high affinity for PCNA.

Peptides 1-10, 1-11, and 1-12 which lack 1, 2, or 3 *N*-terminal residues of 1-0, respectively, maintain low nanomolar binding affinity to PCNA (10.7, 8.16, and 7.57 nM, respectively). The peptide 1-13 lacks the four *N*-terminal residues of 1-0 and gave decreased affinity for PCNA (20.4 nM). Notably, removal of all five residues *N*-terminal to the PIP-box (1-14) greatly decreased affinity for PCNA to 694 nM. This indicates that at least a single residue *N*-terminal to the PIP-box (residue 143) must be included for p21 peptides to maintain high-affinity binding to PCNA.

Table 1. Peptide Sequences^a and SPR Binding Affinity,^b SD is Standard Deviation

peptide	sequence	affinity K_D (nM)	K_D SD (nM)
1-0	GRKRRQTSMTDFYHSKRRLIFS	5.14	0.82
1-1	GRKRRQTSMTDFYHSKRRLIF	14.1	1.8
1-2	GRKRRQTSMTDFYHSKRRLI	15.9	5.4
1-3	GRKRRQTSMTDFYHSKRRL	20.0	1.1
1-4	GRKRRQTSMTDFYHSKR	22.2	4.1
1-5	GRKRRQTSMTDFYHSKR	23.9	4.1
1-6	GRKRRQTSMTDFYHSK	37.4	4.7
1-7	GRKRRQTSMTDFYHS	210	95
1-8	GRKRRQTSMTDFYH	151	52
1-9	GRKRRQTSMTDFY	308	34
1-10	RKRRQTSMTDFYHSKRRLIFS	10.7	2.2
1-11	KRRQTSMTDFYHSKRRLIFS	8.16	1.09
1-12	RRQTSMTDFYHSKRRLIFS	7.57	0.48
1-13	RQTSMTDFYHSKRRLIFS	20.4	11.1
1-14	QTSMTDFYHSKRRLIFS	694	133
2-0	RQTSMTDFYHSK	97.6	4.5
1-0F	Fl*-GRKRRQTSMTDFYHSKRRLIFS		
2-0F	Fl-A _β RQTSMTDFYHSK		
3-1	RKRRQTSMTDFYHSK	55.6	10.8
3-2	RQTSMTDFYHSKR	68.2	22.4
3-3	RKRRQTSMTDFYHSKR	41.1	3.1
3-1F	Fl-A _β RKRRQTSMTDFYHSK	163	39
3-2F	Fl-A _β RQTSMTDFYHSKR	258	68
3-3F	Fl-A _β RKRRQTSMTDFYHSKR	162	44
2-0a	mal-RQTSMTDFYHSK		
4-1F (SV40)	Fl-A _β PKKKRKVC		
4-2F (cMyc)	Fl-A _β PAAKRVKLD		
4-3F (TAT)	Fl-A _β GRKKRRQRRC		
4-4F (R6W3)	Fl-A _β RRWRRRWRRC		
4-1a	PKKKRKVC		
5-1F (SV40)	Fl-A _β PKKKRKVC(suc-RQTSMTDFYHSK)	429	6.5
5-2F (cMyc)	Fl-A _β PAAKRVKLD(suc-RQTSMTDFYHSK)	406	66
5-3F (TAT)	Fl-A _β GRKKRRQRRC(suc-RQTSMTDFYHSK)	116	15
5-4F (R6W3)	Fl-A _β RRWRRRWRRC(suc-RQTSMTDFYHSK)	397	34
6-0	[cyclo-5,9]-Ac-RQC(-)SMTC(Bim-)FYHSK	728	129
6-0a	[cyclo-5,9]-mal-RQC(-)SMTC(Bim-)FYHSK		
6-0F	[cyclo-6,10]-FITC-A _β RQC(-)SMTC(Bim-)FYHSK	24550	650
6-1F (SV40)	Fl-A _β PKKKRKVC([cyclo-5,9]suc-RQC(-)SMTC(Bim)FYHSK)	176	27
6-2F (cMyc)	Fl-A _β PAAKRVKLD([cyclo-5,9]suc-RQC(-)SMTC(Bim)FYHSK)		
6-3F (TAT)	Fl-A _β GRKKRRQRRC([cyclo-5,9]suc-RQC(-)SMTC(Bim)FYHSK)		
6-4F (R6W3)	Fl-A _β RRWRRRWRRC([cyclo-5,9]suc-RQC(-)SMTC(Bim)FYHSK)		
6-1 (SV40)	PKKKRKVC([cyclo-5,9]suc-RQC(-)SMTC(Bim)FYHSK)	181	5.0



^aA large non-specific component was observed in the SPR sensorgrams for the interaction of these peptides with PCNA and affinity could not be confidently determined. ^bThe PIP-box binding motif is indicated in bold. Fl, 5/6-fluorescein thiourea; Fl*, 5-carboxyamido fluorescein; mal, N-maleimidobutyramide; suc, N-succinimidyl butyric thioether; Bim, bimanine; A_β represents beta-alanine.

A short p21 peptide with both an N-terminal and C-terminal truncation was prepared (2-0, residues 143–154), where only one amino acid N-terminal to the PIP-box (residue 143, see 1-13) was included; and three residues C-terminal of the PIP-box were included (residues 152–154, see 1-6). Peptide 2-0 spans residues 143–154 of p21, and the binding affinity for PCNA

was determined to be 97.6 nM. Hence, a short p21 peptide that maintains high-affinity (<100 nM) binding to PCNA was rationally designed.

The Influence of Positive Charge on Cell Uptake of p21 Peptides. Second, the role of the p21 sequence that flanks the PIP-box on cell permeability was investigated using

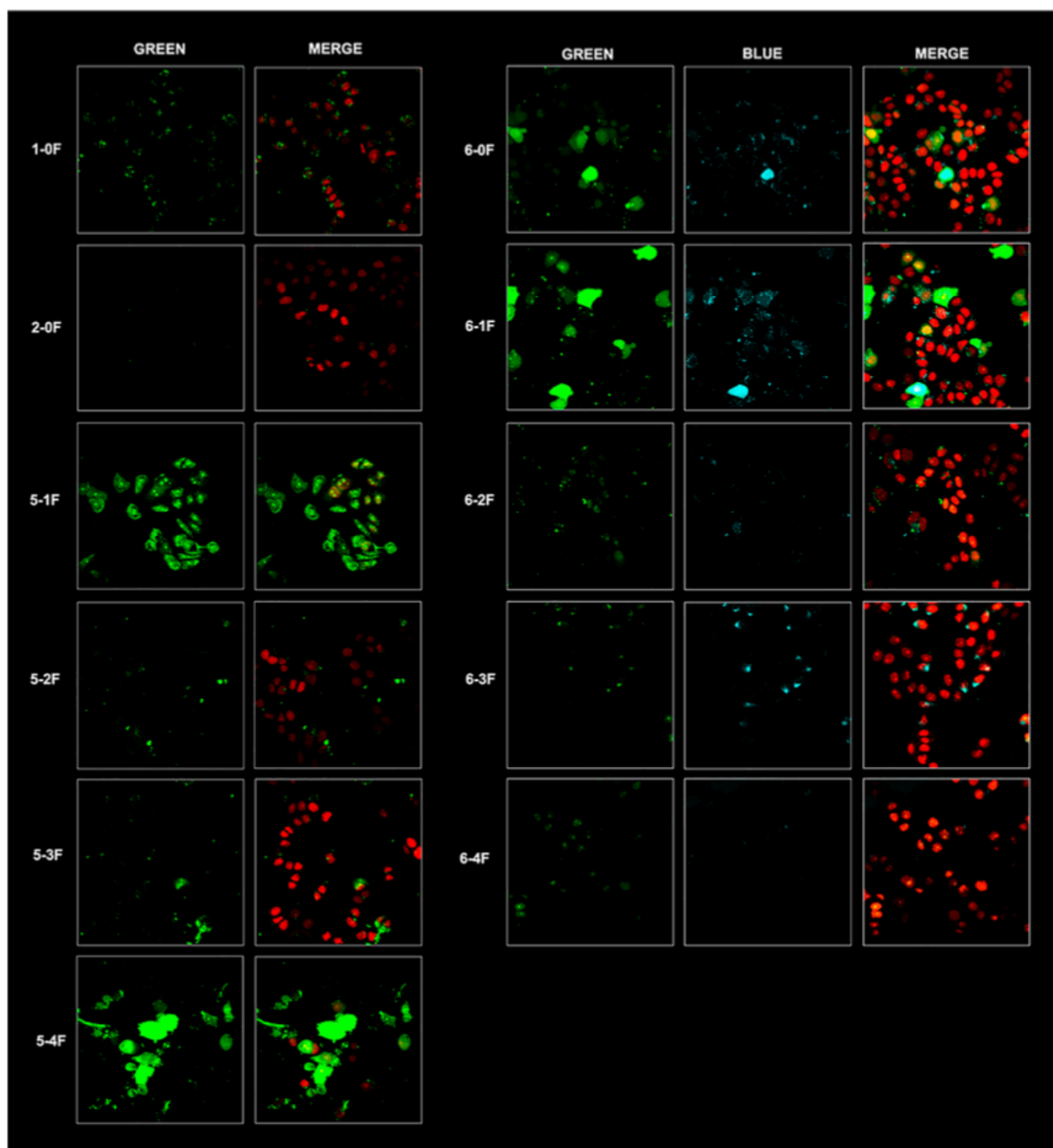


Figure 2. MDA-MB-468 breast cancer cells treated with 10 μ M of peptide for 24 h. Green represents fluorescein, blue indicates the bimane fluorophore, and red indicates the mKate fluorophore expressed at the cell nucleus. Merged images: all relevant channels overlaid.

breast cancer cells. *N* α -fluorescein-tagged peptides **1-0F** and the optimized short peptide **2-0F** (derivatives of **1-0** and **2-0**) were prepared and subjected to a cell uptake assay in order to determine cell permeability. Breast cancer cells (T-47D) were treated with 5 μ M of **1-0F** or **2-0F** and incubated for 48 h. The cells were then fixed, stained with phalloidin, and imaged by fluorescence microscopy (Figure S3). Bright green fluorescence, corresponding to the fluorescein-tagged peptide, was observed in the cytoplasm of **1-0F** treated cells, indicating this

peptide was cell permeable. This fluorescence was punctate, which suggests the peptide is likely endosomally trapped.^{40–42} In contrast, green fluorescence was not observed in the **2-0F**-treated cells, indicating this peptide was not cell permeable. The sequence of short peptide **2-0F** lacks a number of positively charged amino acids, relative to **1-0F**, possibly explaining poor uptake because it is known that positively charged residues such as arginine and lysine often improve cell uptake of peptides.^{43–45} A second series of three peptides was

prepared (3-1, 3-2 and 3-3, Table 1) which systematically reintroduces the arginine and lysine residues that flank the PIP-box sequence, to determine whether this added positive charge could reinstate cell permeability, as seen for the longer 1-0F. The binding affinity of these peptides for PCNA was determined by SPR and revealed a K_D value of between 41 and 68 nM for all three peptides (Table 1). *N* α -Fluorescein-tagged analogs 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 3-2F-treated cells lacked discernible green fluorescence within the cells corresponding to the fluorescein-tagged peptides, consistent with the peptide being impermeable (Figure S3). A modest amount of green fluorescence was observed in the cell cytoplasm of the 3-1F- or 3-3F-treated cells, suggesting the *N*-terminal charged residues may help promote cell entry (Figure S3). The improvement in uptake by reintroduction of positively charged residues was minimal and suggests that the cell uptake of 1-0F is not solely reliant on the positively charged residues.

SV40 NLS Confers Cellular and Nuclear Permeability.

Next, an NLS sequence was appended to the short optimized p21 peptide 2-0 in an attempt to confer both cell and nuclear permeability. The shorter peptide (2-0) was chosen as the longer 1-0F was also not nuclear permeable, and the goal was to identify the shortest nuclear permeable scaffold. Four peptide sequences derived from SV40_{126–132},^{32,46,47} cMyc_{320–328},^{48,49} and TAT_{46–57},^{46,50} and a synthetic sequence termed R6W3^{51,52} are reported to be capable of permitting cell and nuclear entry to a covalently tethered cargo. These sequences vary in length, origin and overall character (*i.e.* charge and hydrophobicity), thereby providing a diverse set of characteristics to trial. Peptides SV40_{126–132}, cMyc_{320–328}, TAT_{46–57}, and R6W3, each with an additional cysteine included at the C-terminus of the sequence, and an *N* α -fluorescein tag, were prepared to give peptides 4-1F, 4-2F, 4-3F, and 4-4F, respectively (Table 1). An analog (2-0a) was prepared with an *N*-alkyl maleimide covalently attached to the *N*-terminus of 2-0, and 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 (Table 1, Figure S1).

Modified MDA-MB-468 breast cancer cells, which express a nuclear mKate fluorophore, were treated with 10 μ M of 1-0F, 2-0F, 5-1F, 5-2F, 5-3F, or 5-4F, to determine whether these peptides were cell and nuclear permeable. The treated cells were fixed and imaged by confocal microscopy (Figure 2). As before, green punctate fluorescence corresponding to the fluorescein-tagged peptide was evident in 1-0F-treated cells, confirming 1-0F is cell permeable. Green fluorescence was not observed for the 2-0F-treated cells, indicating the short p21 peptide was not cell permeable, consistent with the T-47D cell line. The 5-1F-treated cells displayed bright green fluorescence throughout the cytoplasm, which also appears to colocalize with the red mKate fluorescence indicating the cell 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, a lack of red fluorescence corresponding to mKate expression in the nucleus is indicative of lysed or dead cells, suggesting 5-4F at 10 μ M is cytotoxic. No notable cytotoxicity was noted during incubation of the cells with other tested peptides.

Macrocyclic NLS-p21 Derivatives Retain Nuclear and Cellular Permeability. Macrocylic peptides are known to have increased proteolytic stability, and often improved cell permeability, compared to their linear counterparts. We recently reported a short macrocyclic p21 peptide, 6-0, that binds PCNA and is inherently fluorescent and cell permeable in a breast cancer cell line.³⁸ We chose to investigate if appending an NLS to this peptidomimetic could confer a nuclear permeability.

MDA-MB-468 cells were treated with 10 μ M of 6-0 and imaged by confocal microscopy. The 6-0-treated cells displayed blue punctate fluorescence within the cell cytoplasm (Figure S4) owing to the bimane fluorophore embedded in the macrocycle, consistent with our previous report that 6-0 is cell permeable.³⁸ Given this peptide is already cell permeable, it was envisaged that tethering 6-0 to an NLS sequence should further enhance cell and nuclear uptake of the conjugate, compared to the 2-0 conjugates, 5-1F–5-4F. An analog 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 sequences (4-1F, 4-2F, 4-3F, or 4-4F) as before, to give 6-1F, 6-2F, 6-3F, and 6-4F, respectively. An *N* α -fluorescein-tagged derivative of the p21 macrocycle, without an NLS, 6-0F was also prepared.³⁸ MDA-MB-468 cells were treated with 10 μ M of 6-0F, 6-1F, 6-2F, 6-3F, or 6-4F, incubated for 24 h, then fixed and imaged by confocal microscopy (Figure 2).

Cells treated with 6-0F displayed blue punctate fluorescence of the macrocycle in the cell cytoplasm, which largely colocalized with the green fluorescein signal. The green fluorescence was more diffuse, suggesting that some of the fluorescein tag may have been liberated or cleaved from the peptide. Thus, peptide 6-0F was cell permeable, but not nuclear permeable.³⁸ The 6-1F-treated cells, and to a lesser extent the 6-2F-treated cells, displayed blue fluorescence within the cell cytoplasm that colocalized with the green fluorescence of the fluorescein-tag appended to the NLS sequence. The blue and green fluorescence in 6-1F-treated cells was a combination of diffuse and punctate fluorescence, suggesting some of the peptide is endosomally trapped. Some of this signal overlapped with the red fluorescence of the nuclear mKate signal, indicating some of the peptide was translocated to the nucleus. Small pockets of well colocalized green and blue fluorescence were apparent in the 6-3F-treated cells. The 6-4F-treated cells were administered at a lower concentration, than the 6-0F–6-3F cells, owing to the apparent cytotoxicity observed for the 5-4F-treated cells. Only faint green fluorescence was notable in the 6-4F-treated cells, indicating these peptides are not significantly cell permeable at this concentration.

Peptide 6-1, an analog of the cell and nuclear permeable peptide (6-1F), was then prepared by the reaction of 4-1a and 2-1a, to investigate whether the fluorescein-tag influences the cell and nuclear permeability profile, where such effects have been reported for a range of auxiliary tags.^{53–55} Peptide 6-1 contains a bimane-constrained p21 peptide appended to an SV40 NLS sequence, but fluorescein was omitted. 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 positively influence cell uptake and distribution.

DISCUSSION & CONCLUSIONS

A viable pre-clinical peptide-based cancer therapeutic that inhibits PCNA and hence DNA-replication and repair, must have high affinity for PCNA and 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_{139–160} (1-0) peptide, that truncation of the 6 *C*-terminal and 4 *N*-terminal residues (residues 155–160 and 139–142, respectively) did not significantly impact PCNA binding affinity. This is further supported by peptide 2-0, which consists of residues 143–154 of the p21 sequence that binds PCNA with high affinity (97.6 nM). As the PL peptide is reported to bind PCNA with 100 nM affinity and inhibits SV40 replication in cell extracts,³⁰ we expect the affinity of 2-0 to be sufficient to allow this peptide to target and interact with PCNA in a cellular environment.

It is interesting that reintroduction of the positively charged residues to the short p21 variant 2-0F (Arg/Lys in 3-1F–3-3F) did not recover the cellular uptake of the peptides, as is seen for the cell permeable 1-0F. It is known that peptides with large positive charge character, in particular those with poly-Arg combinations, display improved cellular uptake.^{52,56} Additionally, residues 140–142 are reported to be important for the nuclear localization of the expressed p21 protein.³² However, from the combination of peptides investigated herein, it is clear that the cell permeability of 1-0F cannot be attributed solely to the presence of the positively charged amino acids, as 3-3F is not notably cell permeable. This suggests that the hydrophobic tail (¹⁵⁷LIFS) of 1-0F is at least partially responsible for the cell uptake of this peptide, where many amphipathic peptides display good cell uptake.^{44,57,58} Additionally, hydrophobic and lipid groups are reported to associate with the cell membrane and help facilitate passive diffusion.

Inclusion of a SV40-derived sequence improved cell and nuclear uptake of both the linear (5-1F) and macrocyclic (6-1F) p21 peptides. This suggests that 5-1F and 6-1F are the optimal choice for further development of a p21-derived peptidomimetic to target PCNA. Impressively, these peptide conjugates are only 20 amino acids in length. The binding affinity for these key compounds was characterized by SPR and indicates that the interaction with PCNA was retained on conjugation of the NLS tags (429 and 176 nM, respectively, Table 1). The cMyc (5-2F and 6-2F)- and TAT (5-3F and 6-3F)-tagged peptides did not display significant uptake in MDA-MB-468 breast cancer cells, 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 proved to be nuclear permeable; however, an analog which lacks the fluorescein-tag (6-1) was not cell permeable. This suggests that the fluorescein-tag is important for both cell and nuclear permeability. This observation highlights the impact that auxiliary fluorescent tags (such as fluorescein) can have on the biophysical properties and behavior of a peptide,^{53–55} and emphasizes the need for bifunctional peptide modifications, or label-free technologies, that allow direct analysis of a compound of interest. There is a clear benefit in designing therapeutics

where the imaging tags are inherently part of the peptide structure (e.g. the bimeane modification) and not auxiliary tags (e.g. fluorescein) that are appended only for imaging studies. Such designs ensure that the results of biological assays are reflective of the true function of the compound that will go on to further testing, rather than that of an analogue. Additional techniques, such as flow cytometry and analysis of the cell lysate, can be used to quantitate the cell uptake of peptides by fluorescence or HPLC.^{59–63} Future work will determine the ability of peptides 5-1F 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 variation.²⁵

EXPERIMENTAL SECTION

Peptide Synthesis. All peptides were synthesized using Fmoc/*t*-Bu Solid-Phase Peptide Synthesis as previously described.⁵⁵ Briefly, peptides were synthesized manually with Rink Amide PL resin (0.1 mmol, 322 mg, 0.31 mmol/g, Agilent), or on a Liberty Blue peptide synthesizer (CEM Corp) using 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). 5(6)-carboxyfluorescein was coupled in the same manner as Fmoc-amino-acids, and of 4-maleimidobutyric acid was coupled with 3 equiv of reagent for an extended reaction time of 2 h. Other peptide modifications are described below. Synthesized peptides were purified using semi-preparative RP-HPLC equipped with a Phenomenex Luna C18(2) or Phenomenex Aeris Peptide XB-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 high-resolution mass spectrometry using an Agilent 6230 ESI-TOF MS. Peptide purity was confirmed to be >95% using an Agilent 1260 Infinity analytical RP-HPLC equipped with a Phenomenex Luna C18(2) column (4.6 mm × 250 mm, 5 μm) over a linear gradient of 0 to 50% acetonitrile, with 0.1% TFA, at 1 mL/min and UV detection at 220 nm. Characterization data are listed in Table S1.

Bimeane Cyclization. Peptides were cyclized with the bimeane crosslinker, while on resin, as previously described.⁴² Briefly, the cysteine Mmt side chains were selectively deprotected by repetitive treatment of the resin with 2% TFA in DCM (5 mL) for 1 min and then treated with a solution of dibromobimane (2 equiv) and DIPEA (4 equiv) in DMF, and reacted for 3 h at rt with intermittent stirring. The solution was then removed and the resin was washed with DMF (5 × 5 mL) and DCM (5 × 5 mL).

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 were combined. The solution was shaken on an orbital rocker for 24 h, and the reaction progress was monitored by HPLC and MS. The solution was then lyophilized to give the crude peptide conjugate. See Figure S1 for the combination of reagent peptides.

Fluorescein Isothiocyanate Attachment. Following linear peptide assembly, and *N*-terminal deprotection, a solution of fluorescein-5/6-isothiocyanate (5 equiv) and DIPEA (10 equiv) in DMF (5 mL) was added to the resin-bound peptide and stirred intermittently at rt for 1 h. The solution was then removed and the resin was washed with DMF (5 × 5 mL) and DCM (5 × 5 mL), and then dried with diethyl ether (3 × 5 mL).

PCNA Expression & Purification. Buffer solutions were filtered and carefully degassed. Clarified lysate containing hPCNA was first purified at 4 °C by fast protein liquid chromatography (FPLC), using anion exchange chromatography using two DEAE columns in series (HiTrap DEAE FF 5 mL column), equilibrated in Buffer A (20 mM Tris pH 7.5, 20 mM NaCl, 2 mM DTT), and hPCNA eluted using a linear gradient (0.02 M to 0.7 M NaCl). Fractions containing hPCNA were pooled and ammonium sulfate was added to a final concentration of 1.5 M from a stock solution of 3 M ammonium

sulphate. The sample was treated for 1 h at 4° with stirring to allow hydrophobic disruption of bonds between DNA and Protein, and then applied to hydrophobic exchange chromatography (HiTrap Phenyl FF [high sub] 5 mL column), equilibrated in 20 mM Tris pH 7.5, 20 mM NaCl, 2 mM DTT, 1.5 M ammonium sulfate and eluted with a reverse linear gradient (1.5 M to 0 M ammonium sulphate). Fractions containing hPCNA were pooled and dialyzed overnight in Buffer A. hPCNA was applied to a second anion exchange step, the Q Sepharose column (5 mL Q Sepharose FF column (GE)) was equilibrated in Buffer A and protein was eluted using a linear gradient (0.02 to 0.7 M NaCl). Fractions containing hPCNA were pooled and dialyzed overnight in 20 mM Tris pH 7.5, 10% glycerol, 2 mM DTT, and 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.

General SPR Protocol. SPR experiments were performed as previously described.²⁵ Briefly, experiments were performed on a Biacore S200 instrument at 25 °C, using a running buffer containing 10 mM HEPES buffer containing 150 or 300 mM NaCl, 3 mM EDTA, and 0.05% Tween20, adjusted to pH 7.4 with 2 M NaOH, as specified in Table S2. PCNA was immobilized on a CMS Series S sensor chip (GE) using an EDC/NHS coupling protocol. The peptide stock concentration was determined by UV absorbance (A_{λ}) with a Nanodrop2000. The ϵ_{λ} and λ used for each peptide are specified in Table S3. The steady-state affinity experiments were performed at 30 μ L/min using a contact and dissociation time as specified in Table S3, followed by regeneration of 2 M NaCl (2 \times 30 s). Each experiment consisted of eight peptide concentrations, prepared as a 1 in 2 serial dilution. The reported K_D values are an average of a minimum of two independent experiments. The data were analyzed using the provided Biacore S200 Evaluation software. Representative SPR sensorgram curves are provided in the Supporting Information.

T-47D Cell Line with Fluorescence Microscopy. T-47D cells were cultured at 37 °C and 5% CO₂, on glass coverslips in a 6-well plate in 2 mL 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% confluency prior to treatment with 1 or 5 μ M 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 min. 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 permeabilized in 5% TritonX100 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 min to detect the cytoskeleton. Followed by two 5 min 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 two 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 overnight. The following day, the samples were imaged with an Olympus IX73 Inverted Fluorescence Microscope (Adelaide Microscopy). The nuclei of the cells were visualized using a standard Ultra-Violet long-pass filter. For peptides that were tagged with a fluorescein fluorophore (1-0F, 2-0F, 3-1F-3-3F), a standard intermediate-blue long-pass filter was used. Visualization of cellular F-actin (phalloidin staining) was performed using a standard Cys long-pass filter. The images were taken with a 40 \times and 100 \times (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 modified via lentiviral technology to stably express nuclear fluorescent (ex 588 nm; em 635 nm) 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% CO₂. 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 and then treated 24 h with 10 μ M 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 fluorescent 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 visualized 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 fluorescein 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 contain the bimane fluorophore. Images were taken with a 30 \times silicon oil objective, with a 2 \times zoom setting with imaging software (Olympus, Cell Sens), bringing the total magnification of images to 60 \times .

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jmedchem.3c00471>.

Schematic for combination of peptides to give thioether conjugates, peptide characterization data, SPR protocol conditions and characterization data and representative SPR sensorgrams, and supplementary cell images (PDF) Molecular formula strings (CSV)

■ AUTHOR INFORMATION

Corresponding Author

John B. Bruning – *Institute of Photonics and Advanced Sensing (IPAS), The University of Adelaide, Adelaide, South Australia 5005, Australia; School of Biological Sciences, The University of Adelaide, Adelaide, South Australia 5005, Australia; orcid.org/0000-0002-6919-1824; Email: john.bruning@adelaide.edu.au*

Authors

Aimee J. Horsfall – *Institute of Photonics and Advanced Sensing (IPAS), The University of Adelaide, Adelaide, South Australia 5005, Australia; School of Physical Sciences, The University of Adelaide, Adelaide, South Australia 5005, Australia; Australian Research Council Centre of Excellence for Nanoscale BioPhotonics (CNBP), Adelaide, South Australia 5005, Australia; Present Address: AdvanCell, Level 7 Macquarie House, 167 Macquarie Street, Sydney, New South Wales 2000, Australia; orcid.org/0000-0003-1276-2742*

Theresa Chav – *Institute of Photonics and Advanced Sensing (IPAS), The University of Adelaide, Adelaide, South Australia 5005, Australia; School of Physical Sciences, The University of Adelaide, Adelaide, South Australia 5005, Australia; Australian Research Council Centre of Excellence for Nanoscale BioPhotonics (CNBP), Adelaide, South Australia 5005, Australia*

Jordan L. Pederick – *Institute of Photonics and Advanced Sensing (IPAS), The University of Adelaide, Adelaide, South Australia 5005, Australia; School of Biological Sciences, The University of Adelaide, Adelaide, South Australia 5005, Australia; orcid.org/0000-0003-4048-9771*

Zoya Kikhtyak – Dame Roma Mitchell Cancer Research Laboratories, Adelaide Medical School, Faculty of Health and Medical Sciences, University of Adelaide, Adelaide, South Australia 5005, Australia

Bethiney C. Vandborg – Institute of Photonics and Advanced Sensing (IPAS), The University of Adelaide, Adelaide, South Australia 5005, Australia; School of Biological Sciences, The University of Adelaide, Adelaide, South Australia 5005, Australia; orcid.org/0000-0002-3234-3089

Wioleta Kowalczyk – CSIRO Manufacturing, Clayton, Victoria 3168, Australia

Denis B. Scanlon – Institute of Photonics and Advanced Sensing (IPAS), The University of Adelaide, Adelaide, South Australia 5005, Australia; School of Physical Sciences, The University of Adelaide, Adelaide, South Australia 5005, Australia

Wayne D. Tilley – Dame Roma Mitchell Cancer Research Laboratories, Adelaide Medical School, Faculty of Health and Medical Sciences, University of Adelaide, Adelaide, South Australia 5005, Australia

Theresa E. Hickey – Dame Roma Mitchell Cancer Research Laboratories, Adelaide Medical School, Faculty of Health and Medical Sciences, University of Adelaide, Adelaide, South Australia 5005, Australia; orcid.org/0000-0002-2752-730X

Andrew D. Abell – Institute of Photonics and Advanced Sensing (IPAS), The University of Adelaide, Adelaide, South Australia 5005, Australia; School of Physical Sciences, The University of Adelaide, Adelaide, South Australia 5005, Australia; Australian Research Council Centre of Excellence for Nanoscale BioPhotonics (CNBP), Adelaide, South Australia 5005, Australia; orcid.org/0000-0002-0604-2629

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acs.jmedchem.3c00471>

Author Contributions

[‡]A.J.H. and T.C. contributed equally to this work.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The research was supported by the Australian Research Council Centre of Excellence in Nanoscale BioPhotonics (CNBP) (CE140100003). A.J.H. and J.L.P. were 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).

ABBREVIATIONS

NLS, nuclear localization signal; PCNA, proliferating cell nuclear antigen; PIP, PCNA-interacting protein/peptide; PL, Pogo—DNA Ligase I; SPR, surface plasmon resonance

REFERENCES

- (1) Gulbis, J. M.; Kelman, Z.; Hurwitz, J.; O'Donnell, M.; Kuriyan, J. Structure of the C-Terminal Region of p21 WAF1/CIP1 Complexed with Human PCNA. *Cell* **1996**, *87*, 297–306.
- (2) Oku, T.; Ikeda, S.; Sasaki, H.; Fukuda, K.; Morioka, H.; Ohtsuka, E.; Yoshikawa, H.; Tsurimoto, T. 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) Maga, G.; Hübscher, U. Proliferating cell nuclear antigen (PCNA): A dancer with many partners. *J. Cell Sci.* **2003**, *116*, 3051–3060.
- (4) Gilljam, K. M.; Feyzi, E.; Aas, P. A.; Sousa, M. M.; Muller, R.; Vågbo, C. B.; Catterall, T. C.; Liabakk, N. B.; Slupphaug, G.; Drablos, F.; Krokan, H. E.; Otterlei, M. Identification of a novel, widespread, and functionally important PCNA-binding motif. *J. Cell Biol.* **2009**, *186*, 645–654.
- (5) Boehm, E. M.; Gildenberg, M. S.; Washington, M. T. The Many Roles of PCNA in Eukaryotic DNA Replication. *Enzymes* **2016**, *39*, 231–254.
- (6) Powers, K. T.; Washington, M. T. Eukaryotic translesion synthesis: Choosing the right tool for the job. *DNA Repair* **2018**, *71*, 127–134.
- (7) Gonzalez-Magana, A.; Blanco, F. J. Human PCNA Structure, Function and Interactions. *Biomolecules* **2020**, *10*, 570.
- (8) Pohler, J. R.; Otterlei, M.; Warbrick, E. An in vivo analysis of the localisation and interactions of human p66 DNA polymerase delta subunit. *BMC Mol. Biol.* **2005**, *6*, 17.
- (9) Majka, J.; Burgers, P. M. J. The PCNA–RFC Families of DNA Clamps and Clamp Loaders. *Prog. Nucleic Acid Res. Mol. Biol.* **2004**, *78*, 227–260.
- (10) De March, M.; Merino, N.; Barrera-Vilmarau, S.; Crehuet, R.; Onesti, S.; Blanco, F. J.; De Biasio, A. Structural basis of human PCNA sliding on DNA. *Nat. Commun.* **2017**, *8*, 13935.
- (11) Moldovan, G. L.; Pfander, B.; Jentsch, S. PCNA, the maestro of the replication fork. *Cell* **2007**, *129*, 665–679.
- (12) Stoimenov, I.; Helleday, T. PCNA on the crossroad of cancer. *Biochem. Soc. Trans.* **2009**, *37*, 605–613.
- (13) Zhongyun, D.; Wortman, M.; Tan, Z.; Dillehay, K.; The University of Cincinnati Identification of PCNA Targeting Compounds for Cancer Therapy and PCNA Function Regulation. *WO 2012033938 A2*, 2012.
- (14) Park, S. Y.; Jeong, M. S.; Han, C. W.; Yu, H. S.; Jang, S. B. Structural and Functional Insight into Proliferating Cell Nuclear Antigen. *J. Microbiol. Biotechnol.* **2016**, *26*, 637–647.
- (15) Horsfall, A. J.; Abell, A. D.; Bruning, J. B. Targeting PCNA with Peptide Mimetics for Therapeutic Purposes. *ChemBioChem* **2019**, *21*, 442–450.
- (16) Søgaard, C. K.; Nepal, A.; Petrovic, V.; Sharma, A.; Liabakk, N.-B.; Steigedal, T. S.; Otterlei, M. Targeting the non-canonical roles of PCNA modifies and increases the response to targeted anti-cancer therapy. *Oncotarget* **2019**, *10*, 7185–7197.
- (17) De Biasio, A.; Blanco, F. J. Proliferating cell nuclear antigen structure and interactions: too many partners for one dancer? *Adv. Protein Chem. Struct. Biol.* **2013**, *91*, 1–36.
- (18) Krishna, T. S. R.; Kong, X.-P.; Gary, S.; Burgers, P. M.; Kuriyan, J. Crystal Structure of the Eukaryotic DNA Polymerase Processivity Factor PCNA. *Cell* **1994**, *79*, 1233–1243.
- (19) Bruning, J. B.; Shamoo, Y. 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) Warbrick, E. PCNA binding through a conserved motif. *BioEssays* **1998**, *20*, 195–199.
- (21) Hishiki, A.; Hashimoto, H.; Hanafusa, T.; Kamei, K.; Ohashi, E.; Shimizu, T.; Ohmori, H.; Sato, M. Structural basis for novel interactions between human translesion synthesis polymerases and proliferating cell nuclear antigen. *J. Biol. Chem.* **2009**, *284*, 10552–10560.

- (22) Punchihewa, C.; Inoue, A.; Hishiki, A.; Fujikawa, Y.; Connelly, M.; Evison, B.; Shao, Y.; Heath, R.; Kuraoka, I.; Rodrigues, P.; Hashimoto, H.; Kawanishi, M.; Sato, M.; Yagi, T.; Fujii, N. 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) Boehm, E. M.; Washington, M. T. R.I.P. to the PIP: PCNA-binding motif no longer considered specific: PIP motifs and other related sequences are not distinct entities and can bind multiple proteins involved in genome maintenance. *BioEssays* **2016**, *38*, 1117–1122.
- (24) Prestel, A.; Wichmann, N.; Martins, J. M.; Marabini, R.; Kassem, N.; Broendum, S. S.; Otterlei, M.; Nielsen, O.; Willemoes, M.; Ploug, M.; Boomsma, W.; Kragelund, B. B. The PCNA interaction motifs revisited: thinking outside the PIP-box. *Cell. Mol. Life Sci.* **2019**, *76*, 4923–4943.
- (25) Horsfall, A. J.; Vandborg, B. A.; Kowalczyk, W.; Chav, T.; Scanlon, D. B.; Abell, A. D.; Bruning, J. B. Unlocking the PIP-box: A peptide library reveals interactions that drive high affinity binding to human PCNA. *J. Biol. Chem.* **2021**, *296*, 296 100773.
- (26) Kroker, A. J.; Bruning, J. B. p21 Exploits Residue Tyr151 as a Tether for High-Affinity PCNA Binding. *Biochemistry* **2015**, *54*, 3483–3493.
- (27) Marshall, A. C.; Kroker, A. J.; Murray, L. A.; Gronthos, K.; Rajapaksha, H.; Wegener, K. L.; Bruning, J. B. Structure of the sliding clamp from the fungal pathogen *Aspergillus fumigatus* (AfumPCNA) and interactions with Human p21. *FEBS J.* **2017**, *284*, 985–1002.
- (28) Waga, S.; Hannon, G. J.; Beach, D.; Stillman, B. The p21 inhibitor of cyclin-dependent kinases controls DNA replication by interaction with PCNA. *Nature* **1994**, *369*, 574–578.
- (29) Chen, J.; Jackson, P. K.; Kirschner, M. W.; Dutta, A. Separate domains of p21 involved in the inhibition of Cdk kinase and PCNA. *Nature* **1995**, *374*, 386–388.
- (30) Zheleva, D. I.; Zhelev, N. Z.; Fischer, P. M.; Duff, S. V.; Warbrick, E.; Blake, D. G.; Lane, D. P. 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) Prives, C.; Gottifredi, V. The p21 and PCNA partnership: a new twist for an old plot. *Cell Cycle* **2008**, *7*, 3840–3846.
- (32) Luo, Y.; Hurwitz, J.; Massague, J. Cell-cycle inhibition by independent CDK and PCNA binding domains in p21Cip1. *Nature* **1995**, *375*, 159–161.
- (33) Rodríguez-Vilarrupla, A.; Díaz, C.; Canela, N.; Rahn, H.-P.; Bachs, O.; Agell, N. Identification of the nuclear localization signal of p21(cip1) and consequences of its mutation on cell proliferation. *FEBS J.* **2002**, *531*, 319–323.
- (34) Cayrol, C.; Knibiehler, M.; Ducommun, B. p21 binding to PCNA causes G1 and G2 cell cycle arrest in p53-deficient cells. *Oncogene* **1998**, *16*, 311–320.
- (35) Massodi, I.; Bidwell, G. L.; Raucher, D. Evaluation of cell penetrating peptides fused to elastin-like polypeptide for drug delivery. *J. Controlled Release* **2005**, *108*, 396–408.
- (36) Baker, R. D.; Howl, J.; Nicholl, I. D. A synchological cell penetrating peptide mimic of p21 WAF1/CIP1 is pro-apoptogenic. *Peptides* **2007**, *28*, 731–740.
- (37) Ball, K. L.; Lain, S.; Fahraeus, R.; Smythe, C.; Lane, D. P. Cell-cycle arrest and inhibition of Cdk4 activity by small peptides based on the carboxy-terminal domain of p21WAF1. *Curr. Biol.* **1997**, *7*, 71–80.
- (38) Horsfall, A. J.; Vandborg, B. A.; Kikhtyak, Z.; Scanlon, D. B.; Tilley, W. D.; Hickey, T. E.; Bruning, J. B.; Abell, A. D. A short, cell permeable bimane-constrained PCNA-interacting peptide. *RSC Chem. Biol.* **2021**, *2*, 1499–1508.
- (39) Kontopidis, G.; Wu, S.-Y.; Zheleva, D. I.; Taylor, P.; McInnes, C.; Lane, D. P.; Fischer, P. M.; Walkinshaw, M. D. 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.
- (40) Potocky, T. B.; Menon, A. K.; Gellman, S. H. 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.
- (41) Ramoino, P.; Diaspro, A.; Fato, M.; Usai, C. Imaging of Endocytosis in Paramecium by Confocal Microscopy. *Molecular Regulation of Endocytosis*; IntechOpen, 2012.
- (42) Horsfall, A. J.; Dunning, K. R.; Keeling, K. L.; Scanlon, D. B.; Wegener, K. L.; Abell, A. D. A bimane-based peptide staple for combined helical induction and fluorescent imaging. *ChemBioChem* **2020**, *21*, 3423–3432.
- (43) Futaki, S.; Suzuki, T.; Ohashi, W.; Yagami, T.; Tanaka, S.; Ueda, K.; Sugiura, Y. Arginine-rich Peptides. *J. Biol. Chem.* **2001**, *276*, 5836–5840.
- (44) Jones, A. T.; Sayers, E. J. Cell entry of cell penetrating peptides: tales of tails wagging dogs. *J. Controlled Release* **2012**, *161*, 582–591.
- (45) Schmidt, N.; Mishra, A.; Lai, G. H.; Wong, G. C. Arginine-rich cell-penetrating peptides. *FEBS Lett.* **2010**, *584*, 1806–1813.
- (46) Qin, X.; Zhao, H.; Jiang, Y.; Yin, F.; Tian, Y.; Xie, M.; Ye, X.; Xu, N.; Li, Z. Development of a potent peptide inhibitor of estrogen receptor α . *Chin. Chem. Lett.* **2018**, *29*, 1160–1162.
- (47) Ragin, A. D.; Morgan, R. A.; Chmielewski, J. Cellular Import Mediated by Nuclear Localization Signal Peptide Sequences. *Chem. Biol.* **2002**, *9*, 943–948.
- (48) Hodel, M. R.; Corbett, A. H.; Hodel, A. E. Dissection of a nuclear localization signal. *J. Biol. Chem.* **2001**, *276*, 1317–1325.
- (49) Dang, C. V.; Lee, W. M. F. Identification of the Human c-myc Protein Nuclear Translocation Signal. *Mol. Cell. Biol.* **1988**, *8*, 4048–4054.
- (50) Smith, K. M.; Himiari, Z.; Tsimbalyuk, S.; Forwood, J. K. Structural Basis for Importin-alpha Binding of the Human Immunodeficiency Virus Tat. *Sci. Rep.* **2017**, *7*, 1650.
- (51) Oehlke, J.; Scheller, A.; Wiesner, B.; Krause, E.; Beyermann, M.; Klauschen, E.; Melzig, M.; Bienert, M. Cellular uptake of an alpha-helical amphipathic model peptide with the potential to deliver polar compounds into the cell interior non-endocytically. *Biochim. Biophys. Acta* **1998**, *1414*, 127–139.
- (52) Habault, J.; Poyet, J.-L. Recent Advances in Cell Penetrating Peptide-Based Anticancer Therapies. *Molecules* **2019**, *24*, 927.
- (53) Birch, D.; Christensen, M. V.; Staerk, D.; Franzyk, H.; Nielsen, H. M. 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.
- (54) Szeto, H. H.; Schiller, P. W.; Zhao, K.; Luo, G. Fluorescent dyes alter intracellular targeting and function of cell-penetrating tetrapeptides. *FASEB J.* **2005**, *19*, 118–120.
- (55) Luitz, M. P.; Barth, A.; Crevenna, A. H.; Bombliès, R.; Lamb, D. C.; Zacharias, M. Covalent dye attachment influences the dynamics and conformational properties of flexible peptides. *PLoS One* **2017**, *12*, No. e0177139.
- (56) Herce, H. D.; Garcia, A. E.; Litt, J.; Kane, R. S.; Martin, P.; Enrique, N.; Rebolledo, A.; Milesi, V. Arginine-rich peptides destabilize the plasma membrane, consistent with a pore formation translocation mechanism of cell-penetrating peptides. *Biophys. J.* **2009**, *97*, 1917–1925.
- (57) Perry, S. R.; Hill, T. A.; de Araujo, A. D.; Hoang, H. N.; Fairlie, D. P. Contiguous hydrophobic and charged surface patches in short helix-constrained peptides drive cell permeability. *Org. Biomol. Chem.* **2018**, *16*, 367–371.
- (58) Kalafatovic, D.; Giralt, E. Cell-Penetrating Peptides: Design Strategies beyond Primary Structure and Amphipathicity. *Molecules* **2017**, *22*, 1929.
- (59) Richard, J. P.; Melikov, K.; Vives, E.; Ramos, C.; Verbeure, B.; Gait, M. J.; Chernomordik, L. V.; Lebleu, B. Cell-penetrating Peptides. A reevaluation of the mechanism of cellular uptake. *J. Biol. Chem.* **2003**, *278*, 585–590.
- (60) Illien, F.; Rodriguez, N.; Amoura, M.; Joliot, A.; Pallerla, M.; Cribier, S.; Burlina, F.; Sagan, S. Quantitative fluorescence spectroscopy and flow cytometry analyses of cell-penetrating peptides

internalization pathways: optimization, pitfalls, comparison with mass spectrometry quantification. *Sci. Rep.* **2016**, *6*, 36938.

(61) Liu, J.; Afshar, S. In vitro Assays: Friends or Foes of Cell-Penetrating Peptides. *Int. J. Mol. Sci.* **2020**, *21*, 4719.

(62) Kohl, F.; Schmitz, J.; Furtmann, N.; Schulz-Fincke, A.-C.; Mertens, M. D.; Küppers, J.; Benkhoff, M.; Tobiasch, E.; Bartz, U.; Bajorath, J.; Stirnberg, M.; Gütschow, M. Design, characterization and cellular uptake studies of fluorescence-labeled prototypic cathepsin inhibitors. *Org. Biomol. Chem.* **2015**, *13*, 10310–10323.

(63) Seisel, Q.; Pelletier, F.; Deshayes, S.; Boisguerin, P. How to evaluate the cellular uptake of CPPs with fluorescence techniques: Dissecting methodological pitfalls associated to tryptophan-rich peptides. *Biochim. Biophys. Acta, Biomembr.* **2019**, *1861*, 1533–1545.

Recommended by ACS

Characterization of PGua₄, a Guanidinium-Rich Peptoid that Delivers IgGs to the Cytosol via Macropinocytosis

Andréanne Laniel, Christine Lavoie, *et al.*

FEBRUARY 13, 2023
MOLECULAR PHARMACEUTICS

READ 

Small Molecules Targeting DNA Polymerase Theta (POLθ) as Promising Synthetic Lethal Agents for Precision Cancer Therapy

Maria Chiara Pismataro, Serena Massari, *et al.*

MAY 03, 2023
JOURNAL OF MEDICINAL CHEMISTRY

READ 

Fragment-Based Discovery of Azocyclic Alkyl Naphthalenesulfonamides as Keap1-Nrf2 Inhibitors for Acute Lung Injury Treatment

Jianyu Yan, Chunlin Zhuang, *et al.*

MAY 31, 2023
JOURNAL OF MEDICINAL CHEMISTRY

READ 

Design and Discovery of Novel Cyclic Peptides as EDPs–EBP Interaction Inhibitors for the Treatment of Liver Fibrosis

Nazi Song, Xianxing Jiang, *et al.*

MARCH 20, 2023
JOURNAL OF MEDICINAL CHEMISTRY

READ 

Get More Suggestions >

Supporting Information

Designing Fluorescent Nuclear Permeable Peptidomimetics to Target Proliferating Cell Nuclear Antigen

Aimee J. Horsfall,¹⁻³ Theresa Chav,¹⁻³ Jordan L. Pederick,^{1,5} Zoya Kikhtyak,⁴ Bethiney C. Vandborg,^{1,5} Wioleta Kowalczyk,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), Adelaide, South Australia 5005, Australia

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, john.bruning@adelaide.edu.au

Table of contents

Peptide data

Figure S1: Thioether conjugation of NLS peptide and p21 peptides.....	S2
Table S1: Peptide characterisation data.	S3
Representative sample of HPLC traces.....	S4

SPR data

Table S2: SPR protocol conditions.	S6
Table S3: SPR characterisation data.	S6
Figure S2: SPR sensorgram curves.....	S7

Supplementary cell images

Figure S3: Cell images in T47D breast cancer cells.....	S8
Figure S4: MDA-MB-468 breast cancer cells treated with 10 μ M of macrocyclic bimeane peptides.....	S8

References.....	S9
-----------------	----

Peptide data

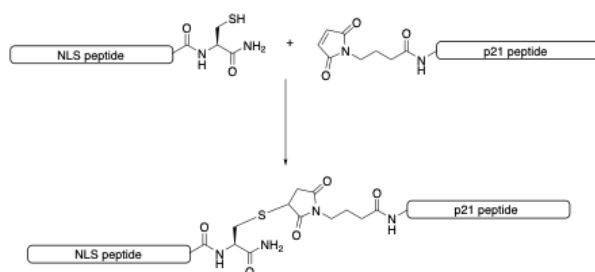


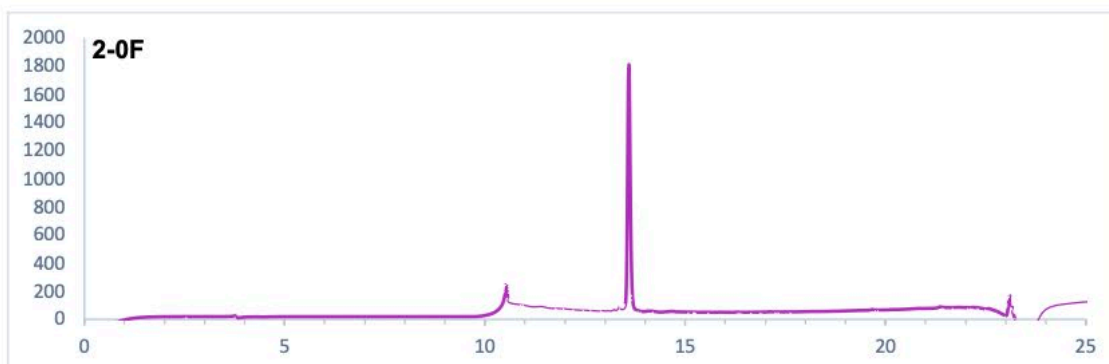
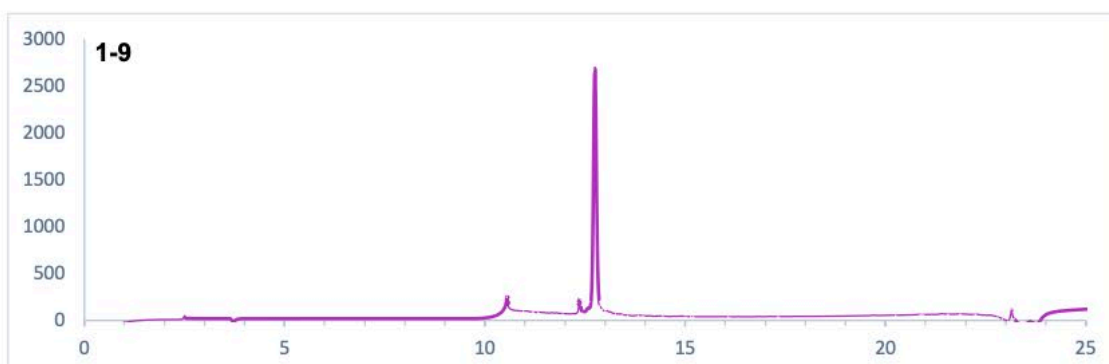
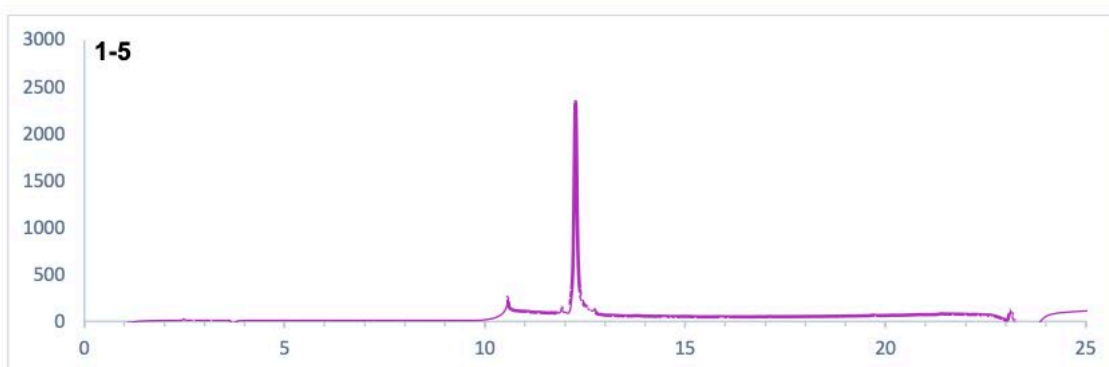
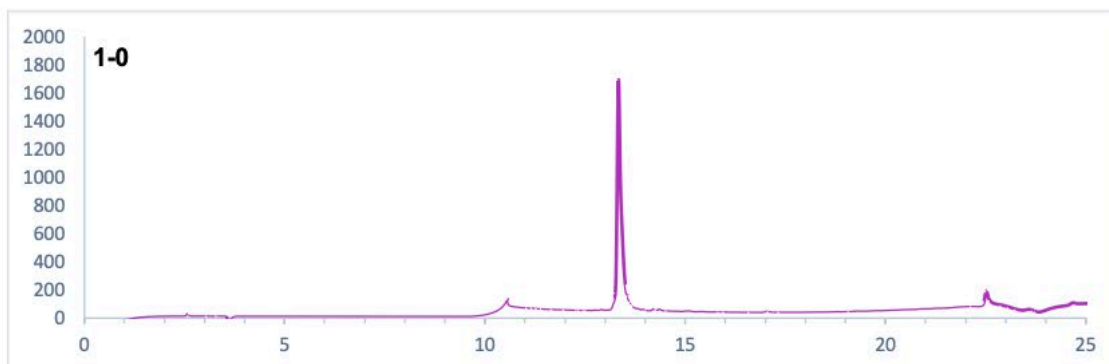
Figure S1: Thioether conjugation of NLS peptide and p21 peptides

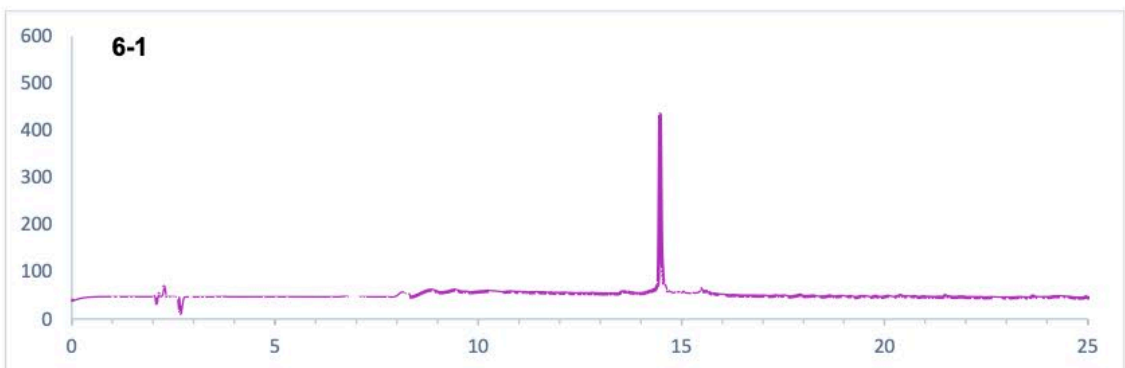
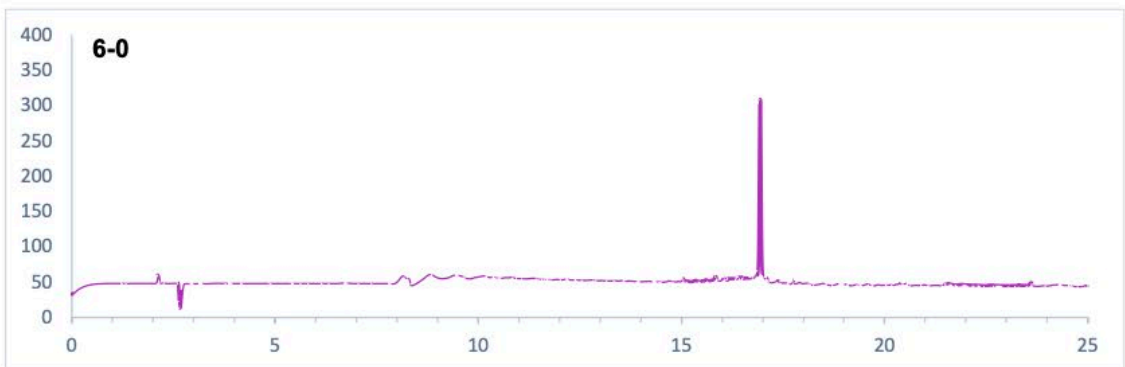
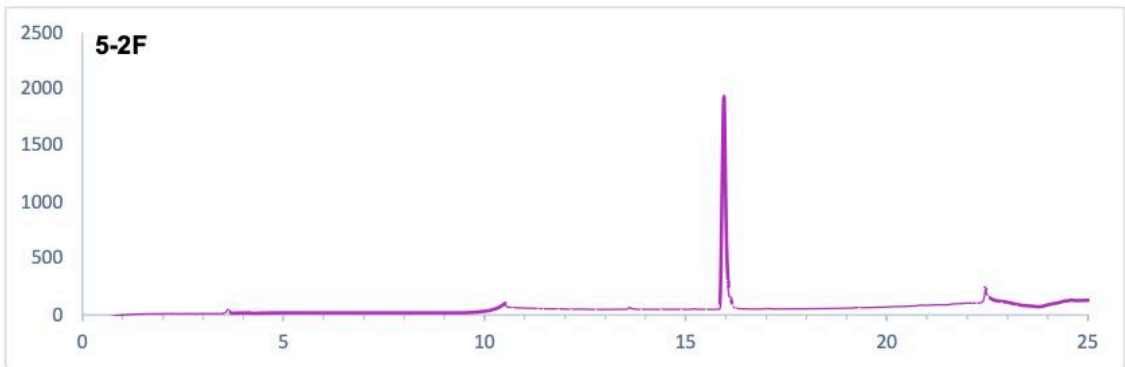
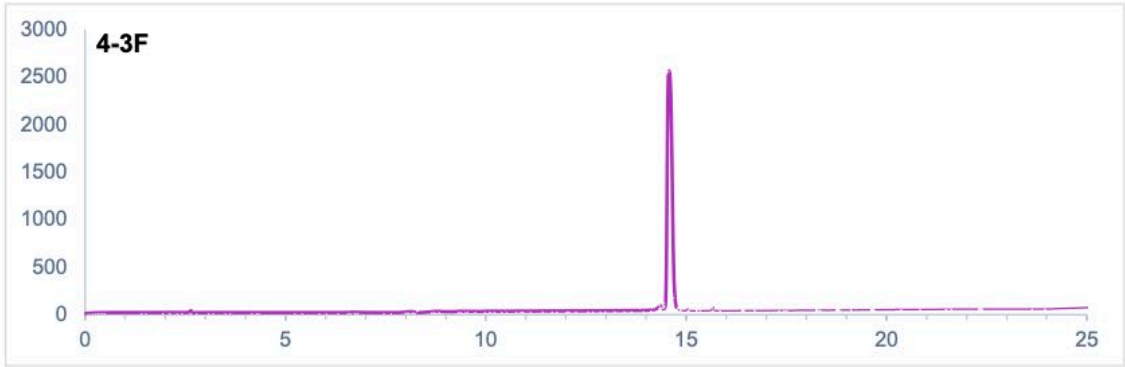
Product	NLS peptide	p21 peptide
5-1F	4-1F	2-0a
5-2F	4-2F	2-0a
5-3F	4-3F	2-0a
5-4F	4-4F	2-0a
6-1F	4-1F	6-0a
6-2F	4-2F	6-0a
6-3F	4-3F	6-0a
6-4F	4-4F	6-0a
6-1	4-1a	6-0a

Table S1: 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+ Calc.	[M+X] ^{z+}	ESI+ [M+X] ^{z+} Found
1-0	GRKRRQTSMTDFYHSKRRLIFS	2770.20	C ₁₂₀ H ₁₉₇ N ₄₃ O ₃₁ S ₁	2769.4722	4 ⁺ : 693.3760		693.3793
1-1	GRKRRQTSMTDFYHSKRRLIF	2683.11	C ₁₁₇ H ₁₉₂ N ₄₂ O ₂₉ S ₁	2681.4561	4 ⁺ : 671.3720		671.3718
1-2	GRKRRQTSMTDFYHSKRRLI	2535.94	C ₁₀₈ H ₁₈₃ N ₄₁ O ₂₈ S ₁	2534.3877	4 ⁺ : 634.6049		634.6042
1-3	GRKRRQTSMTDFYHSKRRL	2422.78	C ₁₀₂ H ₁₇₂ N ₄₀ O ₂₇ S ₁	2421.3036	4 ⁺ : 606.3339		606.3338
1-4	GRKRRQTSMTDFYHSKR	2309.62	C ₉₈ H ₁₆₁ N ₃₉ O ₂₆ S ₁	2308.2196	4 ⁺ : 578.0629		578.0604
1-5	GRKRRQTSMTDFYHSKR	2153.43	C ₉₀ H ₁₄₉ N ₃₅ O ₂₅ S ₁	2152.1185	5 ⁺ : 431.4317		431.4316
1-6	GRKRRQTSMTDFYHSK	1997.25	C ₈₄ H ₁₃₇ N ₃₁ O ₂₄ S ₁	1996.0174	4 ⁺ : 500.0123		500.0123
1-7	GRKRRQTSMTDFYHS	1869.08	C ₇₈ H ₁₂₅ N ₂₉ O ₂₃ S ₁	1867.9224	4 ⁺ : 467.9886		467.9874
1-8	GRKRRQTSMTDFYH	1782.00	C ₇₃ H ₁₂₂ N ₂₈ O ₂₂ S ₁	1780.8904	4 ⁺ : 446.2306		446.2305
1-9	GRKRRQTSMTDFY	1644.86	C ₆₉ H ₁₁₃ N ₂₅ O ₂₀ S ₁	1643.8314	4 ⁺ : 411.9659		411.9653
1-10	RKRRQTSMTDFYHSKRRLIFS	2713.14	C ₁₁₈ H ₁₉₄ N ₄₂ O ₃₀ S ₁	2711.4667	4 ⁺ : 678.8747		678.8744
1-11	KRRQTSMTDFYHSKRRLIFS	2556.96	C ₁₁₂ H ₁₈₂ N ₄₀ O ₂₉ S ₁	2555.3656	5 ⁺ : 512.0811		512.0798
1-12	RRQTSMTDFYHSKRRLIFS	2428.79	C ₁₀₆ H ₁₇₀ N ₃₉ O ₂₈ S ₁	2427.2706	4 ⁺ : 607.8257		607.8257
1-13	RQTSMTDFYHSKRRLIFS	2272.60	C ₁₀₀ H ₁₅₈ N ₃₇ O ₂₇ S ₁	2271.1695	4 ⁺ : 568.8004		568.7994
1-14	QTSMTDFYHSKRRLIFS	2116.41	C ₉₄ H ₁₄₆ N ₃₅ O ₂₆ S ₁	2115.0684	4 ⁺ : 529.7751		529.7752
2-0	RQTSMTDFYHSK	1499.67	C ₆₄ H ₉₈ N ₂₀ O ₂₀ S ₁	1498.6987	4 ⁺ : 500.5740		500.5774
1-0F	FI*-GRKRRQTSMTDFYHSKRRLIFS-OH	3128.53	C ₁₄₁ H ₂₀₈ N ₄₂ O ₃₆ S ₁	3127.5199	4 ⁺ : 782.8878		782.9110
2-0F	FI-A ₆ RQTSMTDFYHSK	1960.13	C ₈₈ H ₁₁₄ N ₂₂ O ₂₈ S ₂	1958.7716	4 ⁺ : 490.7007		490.6728
3-1	RKRRQTSMTDFYHSK	1940.20	C ₈₂ H ₁₂₄ N ₃₀ O ₂₃ S ₁	1938.9959	2 ⁺ : 970.5057		970.5058
3-2	RQTSMTDFYHSKR	1812.05	C ₇₈ H ₁₂₂ N ₂₈ O ₂₂ S ₁	1810.9009	4 ⁺ : 453.7330		453.7330
3-3	RKRRQTSMTDFYHSKR	2252.60	C ₉₄ H ₁₅₈ N ₃₅ O ₂₈ S ₁	2251.1981	4 ⁺ : 563.8073		563.8068
3-1F	FI-A ₆ GRKRRQTSMTDFYHSK	2402.70	C ₁₀₈ H ₁₅₂ N ₄₀ O ₂₉ S ₂	2401.0844	4 ⁺ : 600.7750		600.7727
3-2F	FI-A ₆ RQTSMTDFYHSKR	2272.51	C ₁₀₀ H ₁₄₀ N ₃₉ O ₂₈ S ₂	2270.9738	6 ⁺ : 370.5040		370.5032
3-2F	FI-A ₆ RKRRQTSMTDFYHSKR	2713.06	C ₁₁₈ H ₁₇₄ N ₄₀ O ₃₁ S ₂	2711.2710	3 ⁺ : 904.7683		904.7681
2-0a	mal-RQTSMTDFYHSK	1664.82	C ₇₂ H ₁₀₆ N ₂₁ O ₂₃ S ₁	1663.7413	4 ⁺ : 416.9430		416.9403
4-1F	FI-A ₆ PKKKRKVC	1445.76	C ₆₇ H ₁₀₀ N ₁₉ O ₁₄ S ₂	1444.7108	4 ⁺ : 362.1855		362.1915
4-2F	FI-A ₆ PAAKRVKLDLC	1559.82	C ₇₁ H ₁₀₂ N ₁₉ O ₁₆ S ₂	1558.7061	4 ⁺ : 390.6843		390.6866
4-3F	FI-A ₆ GRKKRRQRRC	1959.29	C ₈₂ H ₁₃₁ N ₃₅ O ₁₆ S ₂	1957.9853	3 ⁺ : 653.6696		653.6609
4-4F	FI-A ₆ RRWRRRRC	2076.40	C ₈₆ H ₁₂₈ N ₃₄ O ₁₆ S ₂	2074.9532	3 ⁺ : 692.6589		692.6472
4-1a	PKKKRKVC	985.30	C ₄₃ H ₆₄ N ₁₆ O ₆ S	984.6379	2 ⁺ : 493.32675		493.3270
5-1F	FI-A ₆ PKKKRKVC(suc-RQTSMTDFYHSK)	3110.58	C ₁₃₉ H ₂₀₅ N ₃₉ O ₃₇ S ₃	3108.4521	5 ⁺ : 622.6982		622.6968
5-2F	FI-A ₆ PAAKRVKLDLC(suc-RQTSMTDFYHSK)	3224.64	C ₁₄₃ H ₂₀₇ N ₃₉ O ₄₁ S ₃	3222.4474	5 ⁺ : 645.4973		645.4969
5-3F	FI-A ₆ GRKKRRQRRC(suc-RQTSMTDFYHSK)	3642.11	C ₁₅₄ H ₂₃₈ N ₅₀ O ₄₁ S ₃	3621.7266	6 ⁺ : 604.6289		604.6287
5-4F	FI-A ₆ RRWRRRRC(suc-RQTSMTDFYHSK)	3741.22	C ₁₆₈ H ₂₃₁ N ₅₅ O ₃₉ S ₃	3738.6945	6 ⁺ : 624.1236		624.1239
6-0	[cyclo-3,7]-Ac-RQC(-)SMTC(Bim-)FYHSK		C ₇₄ H ₁₀₈ N ₂₂ O ₂₅ S ₃		3 ⁺ : 573.9119		573.9140
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
6-0F	[cyclo-4,8]-FI-A ₆ RQC(-)SMTC(Bim-)FYHSK	2138.39	C ₉₇ H ₁₂₂ N ₂₄ O ₂₄ S ₄	2134.7942	3 ⁺ : 712.6059		712.5945
6-1F	FI-A ₆ 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
6-2F	FI-A ₆ PAAKRVKLDLC[mal-RQC(-)SMTC(Bim)FYHSK]	3402.91	C ₁₅₁ H ₂₁₃ N ₄₁ O ₄₀ S ₅	3400.4497	6 ⁺ : 567.7494		567.7497
6-3F	FL-A ₆ GRKKRRQRRC[mal-RQC(-)SMTC(Bim)FYHSK]	3802.38	C ₁₆₂ H ₂₄₂ N ₅₀ O ₄₀ S ₅	3799.7289	5 ⁺ : 760.95358		760.9534
6-4F	FI-A ₆ RRWRRRRC[mal-RQC(-)SMTC(Bim)FYHSK]	3919.49	C ₁₇₀ H ₂₃₇ N ₅₇ O ₃₆ S ₅	3916.6969	6 ⁺ : 653.7906		653.7907
6-1	H-PKKKKRKVC[mal-RQC(-)SMTC(Bim)FYHSK]	2828.39	C ₁₂₃ H ₁₉₅ N ₃₉ O ₃₀ S ₄	2826.3815	4 ⁺ : 707.6032		707.6026

A representative sample has been supplied, visualised at 220 nm.





SPR data

Table S2: SPR protocol conditions. Initial SPR experiments for the truncation series of peptides (1-0 – 1-14) were carried out with a high salt (300 mM) buffer to minimise the non-specific interactions, and improve dissociation times, seen for these peptides containing a large number of positively-charged residues.

Chip #	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 S3: SPR characterisation data. ϵ_λ is the absorptivity factor calculated for the wavelength λ . 'Top Conc' is the highest concentration of peptide in each experiment; each experiment consists of 8 concentrations of peptide prepared as a 1 in 2 serial 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 S2.

Peptide	λ	ϵ_λ	Top Conc (nM)	Ass/Diss time	Chip #
1-0	205	98620	300	60/90	1
			50	60/90	2
1-1	205	95840	300	60/40	1
1-2	205	84460	200	80/40	1
1-3	205	81680	500	60/40	1
1-4	205	78900	500	60/40	1
1-5	205	74770	500	60/40	1
1-6	205	70640	500	60/40	1
1-7	205	67860	500	60/40	1
1-8	205	65080	500	60/40	1
1-9	205	57100	500	60/40	1
1-10	205	95840	200	80/40	1
1-11	205	91710	200	80/40	1
1-12	205	88930	200	80/40	1
1-13	205	84800	500	80/40	1
1-14	205	80670	3000	80/40	1
2-0	205	56820	500	40/60	2
3-1	205	67867	500	40/60	2
3-2	205	65080	1000	40/60	2
3-3	205	76120	500	60/90	2
3-1F	490	86983	1000	40/60	2
3-2F	490	86983	2000	40/60	2
3-2F	490	86983	500	40/60	2
5-1F	490	86983	250	40/60	2
5-2F	490	86983	2000	40/60	2
5-3F	490	86983	1000	40/60	2
5-4F	490	86983	1000	40/60	2
6-0	380	4694 ¹	10000	40/60	2
6-1	380	4694 ¹	1000	40/40	2
6-0F	380	4694 ¹	30000	40/60	2
6-1F	380	4694 ¹	3000	40/60	2

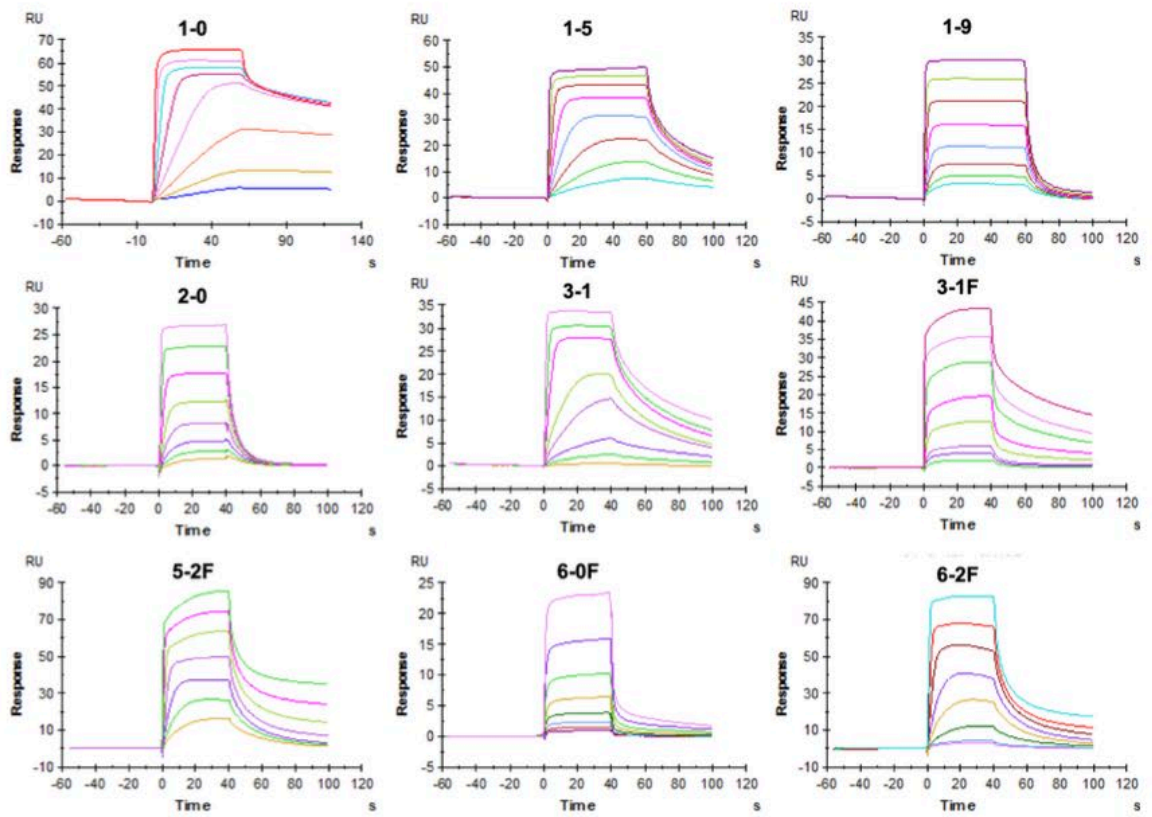


Figure S2: SPR sensorgram curves. A representative selection have been provided. Each trace represents an different concentration of the stated peptide, the top concentration is detailed in Table S3.

Supplementary cell images

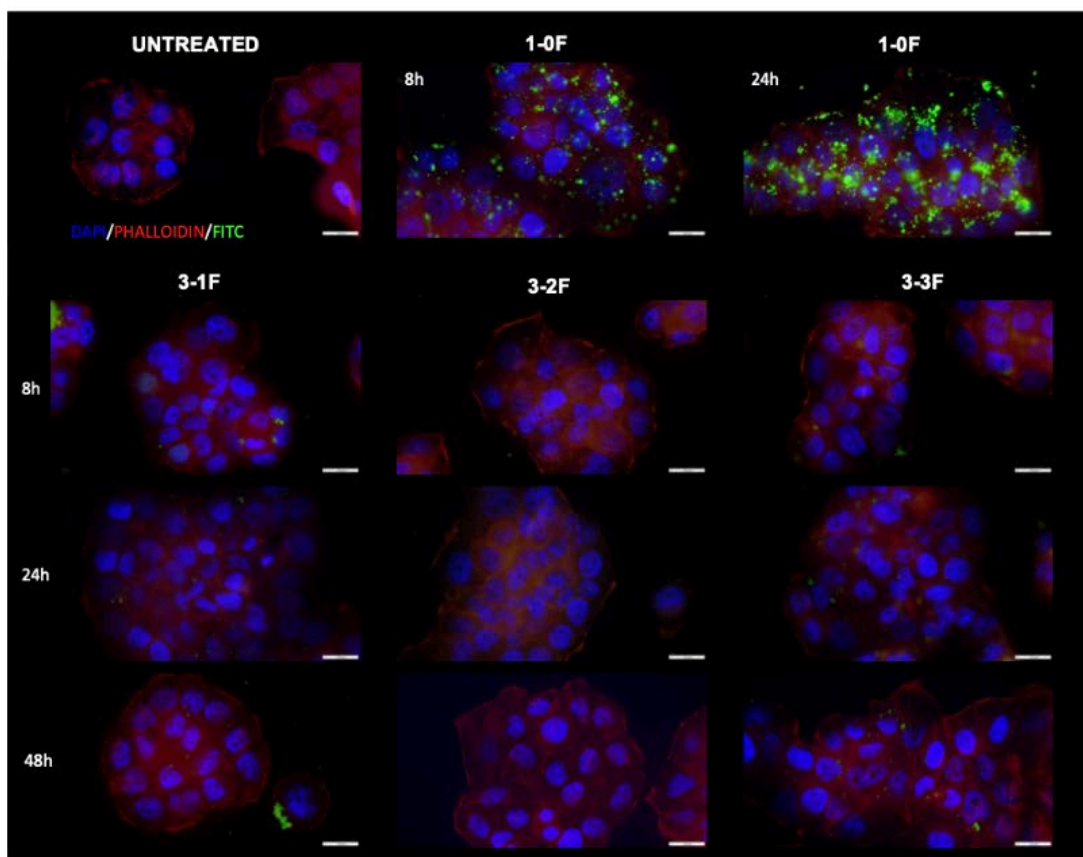


Figure S3: Cell images in T47D breast cancer cells, where blue DAPI stain marks the cell nucleus, red phalloidin marks F-actin indicating the cell cytosol, and green Fluorescein marks the peptide. The cells were incubated for 8, 24 or 28 h with 5 μ M of peptide. The scale bar indicates 50 μ M.

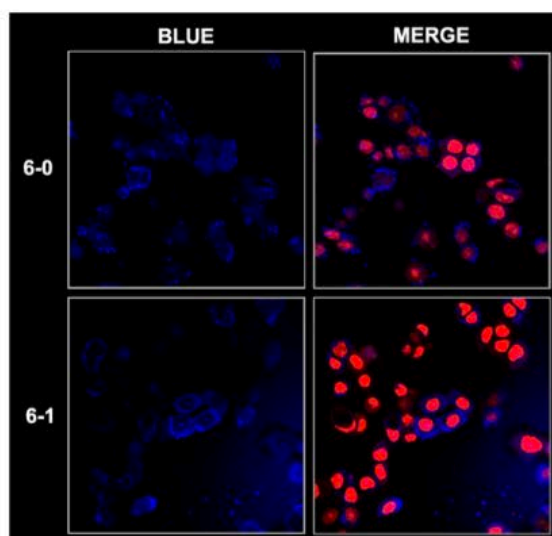


Figure S4: MDA-MB-468 breast cancer cells treated with 10 μ M of macrocyclic bimeane peptides for 24 h

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.