Biophysical and Structural Characterisation of Protein-Peptide Interactions

A Thesis

Submitted for the Degree of

Doctor of Philosophy

by

Peter N. Brown, B.Sc.

Structural Biochemistry Group

The Institute of Structural and Molecular Biology

University of Edinburgh

October 2009

Abstract

Proliferating cell nuclear antigen (PCNA) is an essential protein in the cell. It is involved in transcription and many types of DNA repair and replication. Homologues of this protein are found in all orders of life. The high level of conservation and essential nature of PCNA infers that it may be a potential drug target for anti-caner drugs in humans and also a potential anti-parasitic target. X-ray structures of PCNA from Homo sapiens (Hs), Schizosaccharomyces pombe (Sp) and Leishmania major (Lm) are now available and can be used as a template for structure based drug design.

In this work PCNA from these three species have been prepared in milligram quantities for biochemical and biophysical studies. The previously unknown structure of LmPCNA has been solved in an uncomplexed form and also complexed with a dodecapeptide to a resolution of 3.0Å. A comparison of PCNA structures and their peptide complexes for the three species identifies structural differences which may be relevant in analysing thermodynamic contributions of binding.

All eukaryotic PCNA molecules exist as ring shaped trimers which form around DNA. In this work the oligomeric state of LmPCNA has been determined to be hexameric both in solution and in the crystal. It has also been hypothesised that HsPCNA is hexameric however these would seem to form hexamers in which the trimeric rings associate “back-to-back” while LmPCNA trimers would seem to associate “face-to-face”.

The binding affinities for these three PCNAs have been determined with a selection of peptides derived from the Hs p21 protein. This work has shown, using a selection of different techniques including Surface Plasmon Resonance (SPR),

Isothermal Titration Calorimetry (ITC) and Dynamic Scanning Fluorimetry (DSF); that HsPCNA and SpPCNA have similar affinities for a 12mer peptide (K_d of ~1μM) however LmPCNA shows significantly weaker interactions (K_d of ~10μM). This is most likely due to divergence in the sequence and structure of LmPCNA.

A systematic investigation by SPR on the effect of peptide linker length on binding has been carried out using a series of synthesised peptides with different lengths of chemical spacer. The series of streptavidin immobilised peptides show that longer spacers are required for the recovery of the PCNA peptide binding affinity. The results presented in this work indicate that a linker length of at least 20Å is required for measurable protein binding activity. This interaction is improved with longer peptide spacers.

Declaration

The work presented in this thesis is the original work of the author. This thesis has been composed by the author and has not been submitted in whole or in part for any other degree.

Peter N. Brown

Acknowledgements

I would like to thank my supervisor, Professor Malcolm D. Walkinshaw for providing me with the opportunity to study for the degree of Doctor of Philosophy and his ongoing encouragement and support in me throughout this period of study.

A special note of gratitude to Dr. Cornelia Ludwig and Dr. Jeremy C. Mottram for PCNA constructs of SpPCNA ans LmPCNA respectively. Also to Mr. Anrew Jones who performed DLS on both LmPCNA and SpPCNA.

In addition I would like to thank the following members of the Structural Biochemistry Group for their help and assistance throughout my period of study: Prof. Lindsay Sawyer, Sandra Bruce, Dr. Martin Wear, Dr. Iain McNae, Dr. Hugh Morgan, Dr. Matthew Nowicki, Dr. Paul Taylor, Dr. Julia Richardson, Dr. Elizabeth Blackburn and Dr. Steven Shave.

In addition I would like to thank everyone on the third floor of the Michael Swann building for their help in a variety of tasks during the course of my studies.

Lastly I would like to thank my GCSE science teacher Mr. Coombes and my family for their support and encouragement during my studies.

Abbreviation list

2×TY	2× Tryptone Yeast Extract Media
Å	Ångström
A_{280}	Absorbance at 280 nm
AU	Absorbance Units
BER	Base Excision Repair
CCP4	Collaborative Computational Project Number 4
CD	Circular Dichroism
cDNA	Complementary DNA
Da	Dalton
DLS	Dynamic Light Scattering
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
DSB	Double-Stranded Break Repair
DSF	Dynamic Scanning Flourimetry
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
FPLC	Fast protein liquid chromatography
HEPES	N-(2-Hydroxyethyl)piperazine-N´-(2-ethanesulphonic acid)
HIC	Hydrophobic Interaction Chromatography
HIV	Human Imunodeficiency Virus
Hs	Homo sapiens
IPTG	Isopropyl β-D-thiogalactopyranoside
ITC	Isothermal Titration Calorimetry
K_a	Equilibrium association constant
Kb	Kilobase
K_d	Equilibrium dissociation constant
kDa	Kilodalton
k_{off}	Dissociation rate constant
k_{on}	Association rate constant
k_t	Mass transport coefficient
LB	Luria-Bertani broth
Lm	Leishmania major
MALDI-Tof	Matrix Assisted Laser Desorption Ionisation Time-of-Flight
MMR	DNA Mismatch Repair
NER	Nucleotide Excision Repair
NTA	Nitrolotriacetic acid
OD_{xxx}	Optical density at a wavelength of xxx nm
PCNA	Proliferating Cell Nuclear Antigen

PCR	Polymerase chain reaction
PDB	Protein Data Bank
PEG	Polyethylene glycol
pI	Isoelectric point
PIP-box	PCNA Interacting Protein Box
PISA	Protein Interfaces, Surfaces and Assemblies
RMSD	Root mean square deviation
SA	Streptavidin
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SEC	Size Exclusion Chromatography
Sp	Schizosaccharomyces pombe
SPR	Surface Plasmon Resonance
Tris	Trishydroxymethylaminomethane
T_m	Transition temperature
UV	Ultraviolet
V_m	Matthew's coefficient

Table of contents

Table of Figures

CHAPTER 1. The interactions of PCNA in the cell cycle

Proliferating Cell Nuclear Antigen (PCNA) is an essential and highly conserved protein. This protein is pivotal to the mechanisms of DNA replication and repair. The process of DNA replication can be inhibited by the direct interaction of proteins involved in cell cycle regulation. The stalling of DNA replication can allow DNA damage to be repaired.

The successful inheritance of genetic information requires accurate replication of the genome and epigenetic information. Many of these processes are facilitated by PCNA which binds co-factors to orchestrate these activities (Tsurimoto, T. 1998). Selection between co-factors is mediated by a variety of mechanisms including PCNA modifications, competition, phosphorylation and ubiquitination driven degradation (Moldovan, G. L. *et al* 2007).

PCNA forms a ring which surrounds DNA and coordinates other proteins. Interaction of PCNA with co-factors is mediated by a common binding interface (described in-depth in section 1.2.1). This interface is the most conserved area of the protein (Warbrick, E. *et al* 1998). Binding studies of peptides with three homologous PCNAs from human, yeast and Leishmania will be presented.

1.1 Protein mediated biochemistry of DNA repair and replication

PCNA is essential in the mechanisms of DNA replication, repair and chromatin conservation. DNA replication is a complex mechanism involving hundreds of proteins . DNA repair has been shown to require over 700 associated proteins in the DNA damage response pathways (Matsuoka, S. *et al* 2007). In this chapter. only those which interact with PCNA shall be discussed. The main forms of DNA repair; nucleotide excision, DNA mismatch and double strand break repair all

involve PCNA. As mentioned above, the mechanism of chromatin conservation is also facilitated by PCNA. These critical processes are all dependent on the presence of PCNA, this gives clear evidence as to the importance of PCNA within the cell cycle, as to why null mutations are lethal and this protein is a potential drug target.

1.1.1 Protein mediated biochemistry of DNA replication

The classical view of DNA replication involves a few proteins directly interacting with DNA to facilitate unwinding and replication (Stryer, L. 1995). To copy DNA, the double helix must be unwound by DNA helicase to reveal two complementary strands. The single strands of DNA are stabilised by single stranded DNA binding proteins e.g. Replication Protein A (RPA), preventing the re-annealing of the complementary base sequences (Liu, Y. Y. *et al* 2005). DNA polymerase replicates DNA molecules in a 5'=>3' direction; the lagging strand requires an RNA primer to initiate replication and create an Okasaki fragment (Huang, L. *et al* 1994). The 3' end of a newly synthesised section of one Okasaki fragment anneals to another one by the action of DNA ligase. This is a very simplified view of the proteins required for the replication of DNA as DNA replication is mediated by at least 130 proteins (Wood, R. D. *et al* 2001), many of which interact with those to be mentioned in tables 1.1 and 1.3, or have ancillary activities.

In the course of replication the DNA helix becomes supercoiled. The enzymatic activity of DNA gyrase relieves supercoiling but can impose negative supercoiling aiding DNA helicase activity of the dissociation of DNA strands from one another (Zylicz, M. *et al* 1989). DNA replication is performed by DNA polymerase of which there are seven families, four in humans (table 1.1), polymerase α, β, δ, ε, γ, λ, σ, μ, ζ, terminal deoxynucleotidyl transferase (TDT) and translesion

Family	Human protein	Function
A	γ	A mitochondria DNA polymerase
B	α	A priming polymerase, producing RNA primer followed by DNA nucleotides until removed by RFC
	δ	Is the main DNA replication factor in Eukaryotes
	ε	Has been demonstrated to substitute for polymers δ in the lagging strand
	ζ	Promotes replication through DNA lesions (Prakash, S. *et al* 2005)
X	β	Common repair polymerase
	σ	To date, has no known function
	λ & μ	Involved in non-homolougous end joining to aid in repair of double strand breaks
	TdT	Expressed in lymphoid tissue to allow greater immunological diversity
Y	TLS (ι, κ, λ & η)	Repair polymerase for double stranded breaks

Table 1-1. **Human polymerases grouped by family and activity**

This table shows the four different families of DNA or RNA polymerases in humans. Within each family, there are differing numbers of polymerases, and each polymerase function is detailed here. Some functions are fulfilled by many polymerases like that of double strand break repair due to the complexity of this mechanism. The TLS polymerase complex exist for this one mechanism alone. Others are completed by the actions of a single polymerase e.g. the formation of a 20 base RNA primer prior to the formation of an Okasaki fragment by RNA polymerase α. The A family of polymerases exists in order to replicate mitochondrial DNA and contains only a single polymerase. The B family of polymerases are primarily involved in the stable replication of DNA. The X family of polymerases are involved in common DNA repair mechanisms.The Y family of polymerases were determined to be a seperate family from X (Ohmori, H. *et al* 2001) and comprises of proteins essential for translesional synthesis (TLS), existing in all orders of life. Most of these polymerases have been shown to bind to PCNA.

Modified from (Burgers, P. M. J. *et al* 2001)

sythesis (TLS) polymerases (Burgers, P. M. J. *et al* 2001). The polymerases detailed here are present as homologues in other species, and in some speices novel polymerases also exist (Burgers, P. M. J. *et al* 2001). Replication Factor C (RFC) is also present in the DNA replicating complex, it is capable of dissociating DNA polymerase α, and facilitates the asssembly of the PCNA complex (Ellison, V. and Stillman, B. 2003; Gulbis, J. M. *et al* 1996; Stryer, L. 1995). RFC has been shown to be essential for the loading of PCNA onto DNA by the observation of DNA replication in circular DNA (Burgers, P. M. J. and Yoder, B. L. 1993).

1.1.2 The role of PCNA in DNA replication

PCNA is a DNA sliding clamp, holding DNA polymerase on the double stranded DNA molecule. During DNA replication, supercoiling occurs. This needs to be subsequently relieved by the activity of DNA gyrase, requiring DNA polymerase to dissociate from the DNA molecule. The presence of PCNA allows DNA polymerase to dissociate relieving the effect of supercoiling without the loss of position of the polymerase (Krishna, T. S. R. *et al* 1994). This enables rapid re-annealing of the polymerase to the correct position on the DNA molecule in order to facilitate DNA replication at a high rate (Waga, S. and Stillman, B. 1998). DNA replication occurs at average rates of 750-1000bp sec^{-1} for prokaryotes and 50-100bp sec^{-1} for eukaryotes. PCNA is involved in the replication of the Okazaki fragment from the RNA primer and proceeds to the next. FEN1 RTH complex associates to DNA in a PCNA-dependent manner and removes a 5' phosphate from the Okazaki fragment. DNA ligase can then anneal two Okazaki fragments together (figure 1.1 modified from (Bambara, R. A. *et al* 1997). PCNA has also been shown to bind other polymerases for DNA repair and cell-cycle regulatory purposes (section 1.2.3).

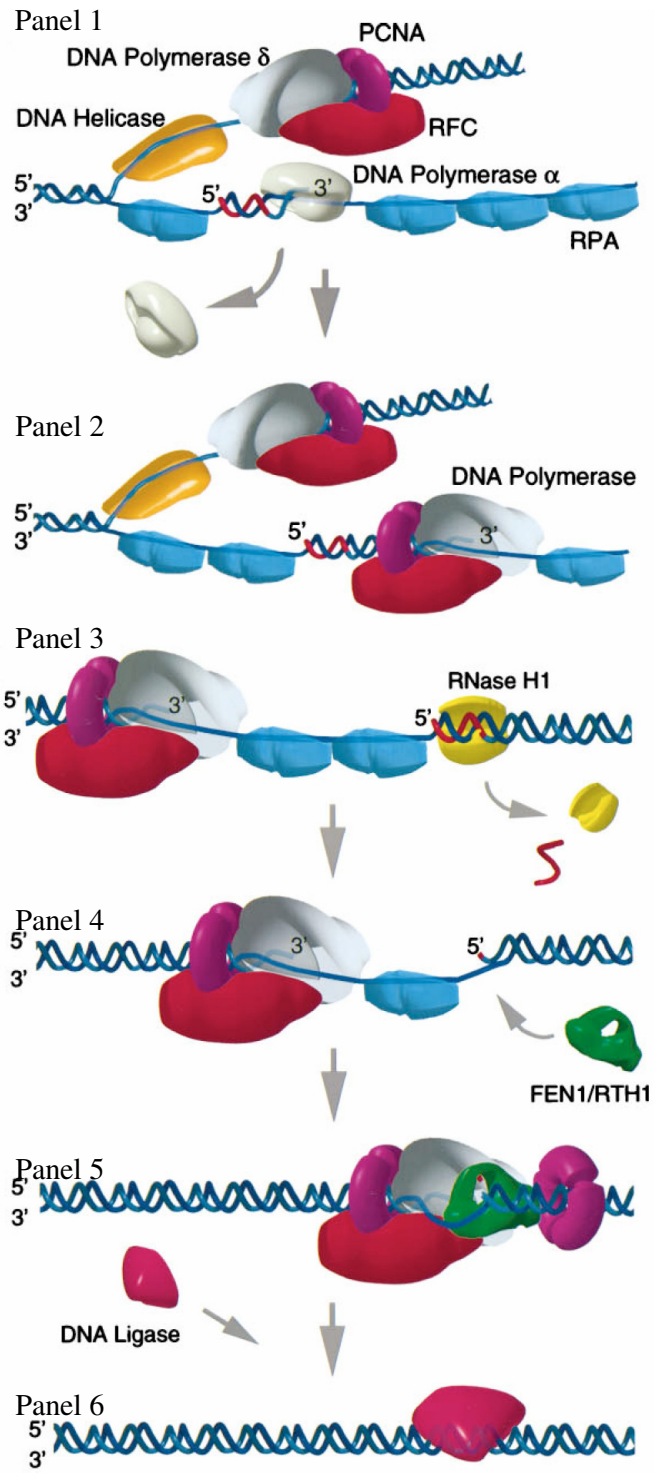


Figure 1-1. **The DNA replication fork**

Panel 1, The replication fork, showing DNA helicase (yellow), DNA polymerase α & δ (white), RFC (red), PCNA (purple) and RPAs (blue). DNA polymerase α has created an RNA primer (scarlet) and some subsequient DNA (blue).

Panel 2, DNA polymerase α has dissociated from the DNA strand and has been replaced with a DNA polymerase δ, RFC and PCNA complex to start creation of an Okazaki fragment.

Panel 3, Okazaki fragment has extended to the previous fragment and RNaseH1 (yellow) is removing the previous RNA primer.

Panel 4, RNase H1 dissociates from the RNA primer and FEN1, RTH1 (green) enters the replication fork

Panel 5, FEN1-RTH1 binds the DNA in a PCNA-dependent manner and cleaves the 5' phosphate from the previous Okazaki fragment.

Panel 6, FEN1, RTH1, DNA polymerase δ, RFC, and PCNA complexes dissociate from the DNA strand and DNA ligase anneals the two new fragments.

Modified from (Bambara, R. A. *et al* 1997)

1.1.3 Protein mediated biochemistry of DNA proof-reading and repair

There are four main forms of DNA repair. These are base excision repair (BER), nucleotide excision repair (NER), DNA mismatch repair (MMR) and double strand break repair (DSB) (Beckman, R. A. and Loeb, L. A. 1993). Base excision repair is most important during replication of the genome as bases can be deaminated or alkylated resulting in their improper incorporation. For example, deamination of cytosine will result in a uracil and alkylation of cytosine will result in 5'-methylcytosine. In contrast deamination of 5'-methylcytosine will result in thymine and can not be repaired by this mechanism (Yoon, J. H. *et al* 2003). Deaminated bases will result in improper base pairing, and at this point, the DNA strand can be bent and allow the mutated base to be "flipped" out of the main body of the strand and excised by DNA glycosylase. The human DNA glycosylases include uracil-DNA glycosylase (UNG), selective monofunctional uracil-DNA glycosylase (SMUG1), thymine DNA glycosylase (TDG) and methyl-CpG-binding protein 4 (MBD4) (Beckman, R. A. and Loeb, L. A. 1993). This excision leaves the sugar-phosphate backbone in place, so it doesn't allow a repair by DNA polymerases. Due to lyase activity by APlyases e.g. Ape1 the sugar phosphate backbone can be cleaved (Wood, R. D. *et al* 2001). This proceedure will leave a gap that can typically be filled by DNA polymerase β with the aid of PCNA, and the new 3' end can be ligated to the pre-existing 5' by the action of DNA ligase (Wood, R. D. *et al* 2001).

1.1.4 Protein-mediated biochemistry of nucleotide excision repair

Nucleotide excision repair (NER) exists in two different forms, global genome (GGNER) and transcription coupled (TCNER) (Sancar, A. *et al* 2004). GGNER involves a continual monitoring for helix distortions throughout the whole

genome at any time, TCNER only occurs on DNA strands being transcribed at that time (Sancar, A. *et al* 2004). These two pathways involve nine core proteins, XPA, XPB, XPC, XPD, XPE, XPF, XPG, CSA and CSB. Deficiencies in these proteins result in the diseases Xeroderma pigmentosum (XPA-G) and Cockayne syndrome (CSA-B). Other proteins involved in NER are ERCCI , RPA and Rad23. GGNER recognises damage by the XPC-Rad23B complex or DNA damage binding proteins (DDB1 or DDB2). TCNER uses RNA polymerase, stopping at e.g. thymine dimers, as well as CSA and CSB, then bind the DNA instead of the XPC-Rad23B complex. Once the damage is recognised, the pathways invoved in NER are the same for both GGNER and TCNER, and both XPB and XPD are subunits of the transcription factor 2H (TFIIH). XPB and XPD act as helicases, unwinding DNA at the point of damage, and XPG cleaves at the 3' side of DNA damage, while XPF-ERCC1 cleave the DNA on the 5' side of DNA damage (Wood, R. D. *et al* 2001). The gap between these two areas is typically of 25-30 bases, allowing the damaged section of DNA to be removed. Once this is accomplished, DNA polymerase δ or ε repairs the DNA strand with the aid of PCNA. RPA protects the template strand from cleavage as the damage is removed, and DNA ligase attaches the newly synthesised 3' end to the pre-existing 5'.

1.1.5 Protein mediated biochemistry of DNA mismatch repair

DNA Mismatch Repair (MMR) repairs DNA strands where complementarity has been lost and can be referred to as long or short type. Long type MMR is primarly implemented in replication errors, but it can be used in all types of mismatches and can possibly replace up to a few kilobases (Sancar, A. *et al* 2004). Short patch MMR is used in third party damage, e.g. damage incured from UV

radiation, free radicals or carcinogens, and will usually only replace up to approximately 10 nucleotides. In this mechanism, the cell needs to determine the site of damage by detecting one of the eight types of mismatch, then determine the incorrect strand for excision, and finally repair the damaged section. In MMR, only the daughter strand will be excised, a process which is best understood in E.coli systems and is thought to be highly conserved (Sancar, A. *et al* 2004). In E.coli systems, there are three main proteins required for MMR: these are MutS, MutH and MutL. MutS is required to detect base-base mismatches along with short regions of deletions or insertions, MutH is used to determine the daughter from the parent DNA strand and to only cleave the daughter, MutL enhances the activity of MutS, MutH and DNA helicase (Mechanic, L. E. *et al* 2000). The section is cleaved by MutH, and if this cleavage were on the 5' side of damage, then RecJ or ExoVIII remove the DNA; if the cleavage is on the 3' side of the damage, then Exo1 is used. Finally the section of DNA is replaced using the DNA polymerase III, β subunit of which is a PCNA homologue. Once the DNA section is replaced it is rapidly methylated by Dam methylase.

The human proteins currently known to be in this mechanism are MSH (a MutS homolog) and MLH (a MutL homolog), but yet no MutH homolog has been found. There are multiple forms of MSH and MLH (Gu, Y. *et al* 2002), MSH2-MSH6 (MutSα) complex indentifies base-base mismatch, insertion and deletion mutations, while MSH2-MSH3 (MutSβ) complex recognises long insertion or deletion mutations (Jiricny, J. 2006). MLH has two main forms, MLH and PMS. MutLα is comprised of MLH1-PMS2 and coordinates MMR. MutLβ is a complex of

MLH1 and PMS1, MutLγ is a complex of MLH1 and MLH3; the functions of these complexes in humans are not yet well understood.

1.1.6 Protein-mediated biochemistry of double stranded break repair

Double strand break (DSB) repair also exists in two main forms Non-Homologus End Joining (NHEJ) and Homologus Recombinational Repair (HRR) (Sancar, A. *et al* 2004). This process is most important prior to replication of the genome, and during cellular differentiation (Boulton, S. J. and Jackson, S. P. 1996). DSBs are detected by the activity of the Mre11-Rad50-Nbs1 complex (Lee, J. H. and Paull, T. T. 2005) and in NHEJ, the two free ends are rejoined by DNA ligase IV in combination with XRCC4 (Moore, J. K. and Haber, J. E. 1996; Wilson, T. E. *et al* 1997). NHEJ requires DNA overhangs known as sticky ends to aid in accurate re-alignment. HRR uses homologous sequences of DNA to find the appropriate end for re-attachment, for example a sister chromatid.

1.1.7 Protein-mediated biochemistry of chromatin conservation

The basic chromatin repeat unit consists of a nucleosome. This is an arrangement of 146 base pairs wrapped around a protein histone core (H2A, H2B, H3 and H4). These nucleosomes form higher order chromatin structures which regulate epigenetic control of genes and maintain chromosome integrity at the telomeres and centromere. For the chromosome to be replicated, the higher order chromatin structures and nucleosomes are required to be modified and removed. Once the DNA is replicated, nucleosomes are required to be replaced and the chromatin remodelled. The mechanisms for this are unknown, however there are some known classes of proteins which mediate these activities, some of which have demonstrated affinity for PCNA (Moldovan, G. L. *et al* 2007).

Prior to DNA replication, the chromatin is required to be destabilised from higher order structures to remove the histone core and linearise the DNA. The complete mechanism of this is currently unknown, however the imitation switch (ISWI) class of chromatin remodelling factors have been shown to assemble chromatin in-vitro (Becker, P. B. and Horz, W. 2002). ISWI has also been previously shown to mobilise nucleosomes and be present in many protein complexes, which can enhance or disrupt the formation of chromatin e.g. nucleosome remodelling factor (NURF), chromatin accessibility complexes (CHRAC) and ATP-utilising chromatin assembly and remodelling factor (ACF) (Bozhenok, L. *et al* 2002). Williams syndrome transcription factor (WSTF) has been shown to bind ISWI, and this WSTF-ISWI complex is known to bind PCNA (Poot, R. A. *et al* 2004).

Once replication has taken place, the sister chromatids are refolded into the new chromatin forms. The H3-H4 dimers, stripped from the DNA strand, rebind to the sister chromatids rapidly after DNA synthesis (Polo, S. E. and Almouzni, G. 2006). This is facilitated by different co-factors e.g. chromatin assembly factor 1 (CAF I), which is known to bind PCNA (Maga, G. and Hubscher, U. 2003), and histone regulator A (HIRA) both of which complex to Asf1, a histone chaperone. CAF I, a known PCNA binder, has been shown to be involved in the delivery of histones to replicated DNA in drosophila embryos (Tyler, J. K. *et al* 1999; Tyler, J. K. *et al* 2001) and human cell lines (Smith, S. and Stillman, B. 1989). The CAF I-Asf1 complex can assist the formation of nucleosomes immediately after DNA replication, during normal replication, or during DNA damage repair mechanisms.

The epigenetic control of a gene is not only dependent on the presence of the nucleosome, but on nucleosome modification and cytosine methylation. The

methylation of cytosines can be achieved by DNA cytosine methyltransferase 1 (DNMT1) (Hermann, A. *et al* 2004). DNMT1 methylates cytosine at positions of hemi-methylation, where one strand of DNA contains methylated cytosines. This protein is directed to the replication fork and stimulates activity by the binding of PCNA via an N-terminal PIP box (Iida, T. *et al* 2002). This recruitment of DNMT1 has also been observed for DNA repair mechanisms (Mortusewicz, O. *et al* 2005). Epigenetic control of DNA is also mediated by nucleosome modifications including histone 3- lysine 9 (H3K9) methylation and histone acetylation. The methylation of H3K9 promotes a repressed, higher ordered chromatin state and can be performed by the histone methyltransferases e.g. hetrochromatin associated protein 1 (HP1), G9a and SETDB1. G9a can associate with PCNA via interactions with DNMT1, which binds PCNA (Esteve, P. O. *et al* 2006). SETDB1 binds methyl-CpG binding protein MBD1 in the presence of methyl-cytosine. The SETDB1-MBD1 histone methyltransferase protein complex can associate to PCNA via CAF I. This association allows histones deposited after DNA replication to be rapidly modified (Sarraf, S. A. and Stancheva, I. 2004), which increases the conservation of epigenetic control. CAF I is associated to PCNA at the replication fork as mentioned above.

The rapid re-formations of chromatin structures are due to the interaction with PCNA via the PIP box, present on some of these proteins (Sarraf, S. A. and Stancheva, I. 2004). This demonstrates the importance of PCNA for chromatin conservation.

1.2 The structure of PCNA

The structure of PCNA has been solved for a diverse number of species, from *H. sapiens* (Gulbis, J. M. *et al* 1996) to bacteriophage RB69 (Shamoo, Y. and Steitz,

T. A. 1999) (table 1.2). The PCNA sliding clamp proteins from these two sources are structurally similar, however sequence identity is only 16% as determined by LALIGN version 2.2 (Myers, E. W. and Miller, W. 1988). The monomer PCNA chains can vary in length from 218 amino acids (RB69) to 356 amino acids (E.coli). The alignment of the PCNA sequences from humans, mice, *S. pombe*, *T. cruzi*, *T. brucei*, *L. major* and bacteriophage RB69 is shown here (figure 1.2). Sequence identity between mammalian PCNA is very high e.g. humans and mice have a sequence identity of ~96%. Sequence identity between PCNAs from mammals and lower eukaryotes is significantly reduced e.g. humans and *S. pombe* PCNAs have a sequence identity of ~50%. The areas of greatest variance within eukaryotic PCNA are at a loop region (185-205) and at the C-terminus (figure 1.2). Sequence identity between eukaryotic and protozoan PCNA is lower e.g. humans and *L. major* have a sequence identity of ~39%. This lower sequence identity is due to an elongated loop region (185-205) which is present in protozoan PCNA (figure1.2). There is negligible sequence identity between eukaryotic and prokaryotic PCNA. This lack of sequence identity is a result of a dimeric rather than trimeric structural architecture (discussed below and in figure 1.5).

The monomer of PCNA has a conserved topology. The predicted secondary structure (figure 1.2) agrees with that determined crystallographically (Gulbis, J. M. *et al* 1996). The PCNA monomer exists as two globular domains (figure 1.5) (Kelman, Z. 1997; Krishna, T. S. R. *et al* 1994). The globular domains are linked by the Inter-Domain Connector Loop (IDCL) (figure 1.2 and 1.3). These two domains each contain a repetition of two βαββββ folds (figure 1.2 and 1.3). This makes a total of four βαββββ folds in eukaryotic PCNA (Krishna, T. S. R. *et al* 1994). Between the

Species	PDB code	Complex state
Homo sapiens	1VYM	Native
	1W60	Native
	1AXC	p21 (143-160)
	1VYJ	Pogo-Ligase peptide (1-16)
	1U7B	Flap endonuclease 1 (331-350)
	1UL1	Flap endonuclease 1 (2-356)
	2ZVK	DNA polymerase η (693-712)
	1U76	DNA polymerase δ, p66 (452-466)
	2ZVL	DNA polymerase κ (862-871)
	2ZVM	DNA polymerase ι (420-432)
Saccharomyces cerevisiae	1PLR	Native
	1PLQ	Native
	1PLR	Native
	3GPM	E113G mutation
	3GPN	E113G mutation
	2OD8	Cdc 9 (35-45)
	3F1W	G178S mutation
	1SXJ	Replication factor C
Arabidopsis thaliana	2ZVV	p21 (139-160)
	2ZVW	p21 (139-160)
Archaeoglobus fulgidus	1RWZ	Native
	1RXZ	Flap endonuclease 1 (11mer)
	1RXM	Flap endonuclease 1 (12mer)
Pyrococcus furiosus	1ISQ	PIP-box 12mer
	1GE8	M73L mutation
	1IZ4	M73L, D143A mutation
	1IZ5	M73L, D143A D147A mutation
Sulfolobus solfataricus	2IX2	Native
	2HII	Native
	2IJX	Native
	2IO4	F2V mutation
	2HIK	Native heterotrimer
	2NTI	Native heterotrimer
	2IZO	Flap endonuclease 1 (5-349)
	3FDS	DNA polymerase IV (1-352)
Sulfolobus tokodaii	1UD9	Native
Human Herpes virus	1YYP	DNA polymerase (1223-1242)
	1DML	DNA polymerase (1200-1235)
Enterobacteria phage RB69	1B77	Native
	1B8H	DNA polymerase (893-903)
Escherichia coli	2POL	Native
	1MMI	Native
	3D1E	DNA polymerase II (505-510)
	3D1F	DNA polymerase III (2-10)
	1OK7	DNA polymerase IV (1-16)
	1UNN	DNA polymerase IV (243-351)
	3BEP	DNA-bound
	3D1G	Bound small molecule inhibitor
Streptococcus pyogenes	2AVT	Native

Table 1-2. **Crystal structures of PCNA deposited in the PDB as of 20/5/09'**
The 49 solved structures of the PCNA sliding clamp from various species. These structures illustrate the native forms of PCNA and the binding interaction with different peptides. 3BEP shows the interaction of PCNA with DNA within the ring and 3D1G shows the interaction of E.coli PCNA binding interface with an inhibitor.

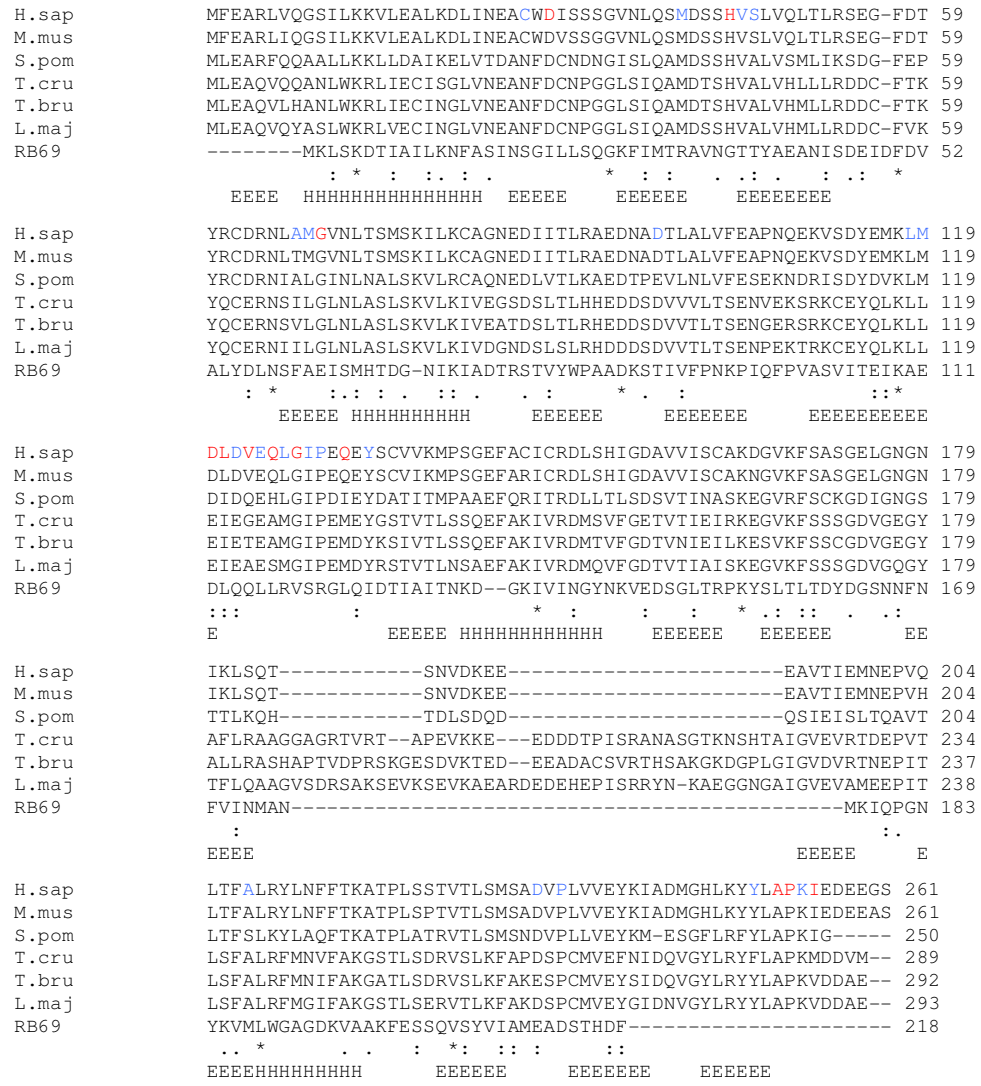


Figure 1-2. **Sequence alignment of PCNAs**

This figure shows an alignment of PCNA from seven different sources *H. sapiens*, M.musculus, *S. pombe*, *T. cruzi*, *T. brucei*, *L. major* and Bacteriophage RB69. The (*) denotes areas of conservation between proteins (:) denotes strong similarity and (.) denotes similar residues. Strong areas of conservation are common at areas incorporated at the active site or important structural bodies in the protein.

Coloured residues in the human sequence are shown to be involved in the PCNA peptide interaction. Those coloured red are determined to be involved in the hydrogen bonding of PCNA to p21 peptide, and residues coloured blue are in contact with this peptide, as determined from the pdb file 1AXC, from (Gulbis, J. M. *et al* 1996). This figure also shows secondary structure as determined by PredictProtein (Rost, B. *et al* 2004). E = β-strand, H = helical.

This alignment also shows a large area of the protein (186-223 L.maj) which is conserved between the protazoan PCNAs, but not present in humans, mice and even in the distant relative RB69. This elongated loop region and has no know function. Determined using http://www.ebi.ac.uk/clustalw/ (Aiyar, A. 2000)

first and second βαββββ folds and the third and fourth βαββββ folds, there are short β-strands (Krishna, T. S. R. *et al* 1994). The second and third βαββββ folds are joined by the IDCL (figure 1.2 and 1.3). Phylogenetic analysis has shown that this βαββββ fold exists in most orders of life (Iwai, T. *et al* 2000; Tsurimoto, T. 1998). The E.coli PCNA monomer contains three globular domains. The three repeating globular domains of a single E.coli PCNA monomer have an average sequence identity of 14% (Kelman, Z. and ODonnell, M. 1995). The low sequence identity of the E.coli PCNA folds suggests that the repetition of this fold occurred early in evolutionary terms.

The eighteen β-strands of each PCNA monomer form three main β-sheets and provide structural rigidity (figure 1.4A). The helices of PCNA are rich in basic residues and bundle together (figure 1.4B). As will be discussed later, the basic helix bundles associate with DNA. The β-sheets and α-helix bundle are present on the opposite side of the PCNA monomer.

The monomers associate forming a ring structure (figure 1.5). Inter-monomer interactions forming the trimer are strong, showing a K_d of ~21nM in *H. sapiens* PCNA (Yao, N. *et al* 1996). The high affinity of this interaction is the result of core hydrophobic interactions, and the hydrogen bonding of two anti-parallel β-sheets from adjacent monomers (figure 1.6). Differences in the stability of homologous proteins are believed to be linked to the number of hydrogen bond interactions at this interface (Yao, N. *et al* 1996). Human PCNA is determined from crystallographic analysis to have eight hydrogen bonds (figure 1.6A) at the binding interface (Gulbis, J. M. *et al* 1996) and RB69 has only four hydrogen bonds at the binding interface

(figure 1.6B) (Shamoo, Y. and Steitz, T. A. 1999). RB69 has a less stable trimeric complex (Yao, N. *et al* 1996) due to the loss of these hydrogen bonds.

The ring structure of PCNA is present as a trimer (eukaryotes) and dimer (prokaryotes) (figure 1.5). Despite the negligible sequence homology between prokaryotes and eukaryotes, the structural architecture of PCNA is conserved (figure 1.5). At the centre of the PCNA ring, there is a circular opening of ~35Å diameter of sufficient size to surround a DNA helix (Kelman, Z. 1997). The PCNA helix bundles (figure 1.4) associate with DNA perpendicular to the phosphate backbone (Tsurimoto, T. 1998). The ring structure in both eukaryotes and prokaryotes both contain twelve βαβββ folds which account for the conserved structural similarities (figure 1.5) (Bruck, I. and O'Donnell, M. 2001).

The first evidence of the ring structure of PCNA and its mode of DNA interaction was provided by size exclusion studies (Stukenberg, P. T. *et al* 1991). PCNA in the presence of RFC and ATP was loaded onto circular and linear DNA. The PCNA-DNA complexes were incubated and passed over a size exclusion column. It was shown that PCNA would co-elute with circular DNA however not linear DNA (Stukenberg, P. T. *et al* 1991). The theory that PCNA formed a ring structure was later confirmed by electron microscopy studies of the PCNA-DNA interaction (Gogol, E. P. *et al* 1992). This ring structure has been hypothesised to potentially form a dimer of rings (Henderson, D. S. *et al* 2000; Naryzhny, S. N. *et al* 2005) . This will be discussed further in chapter 4.

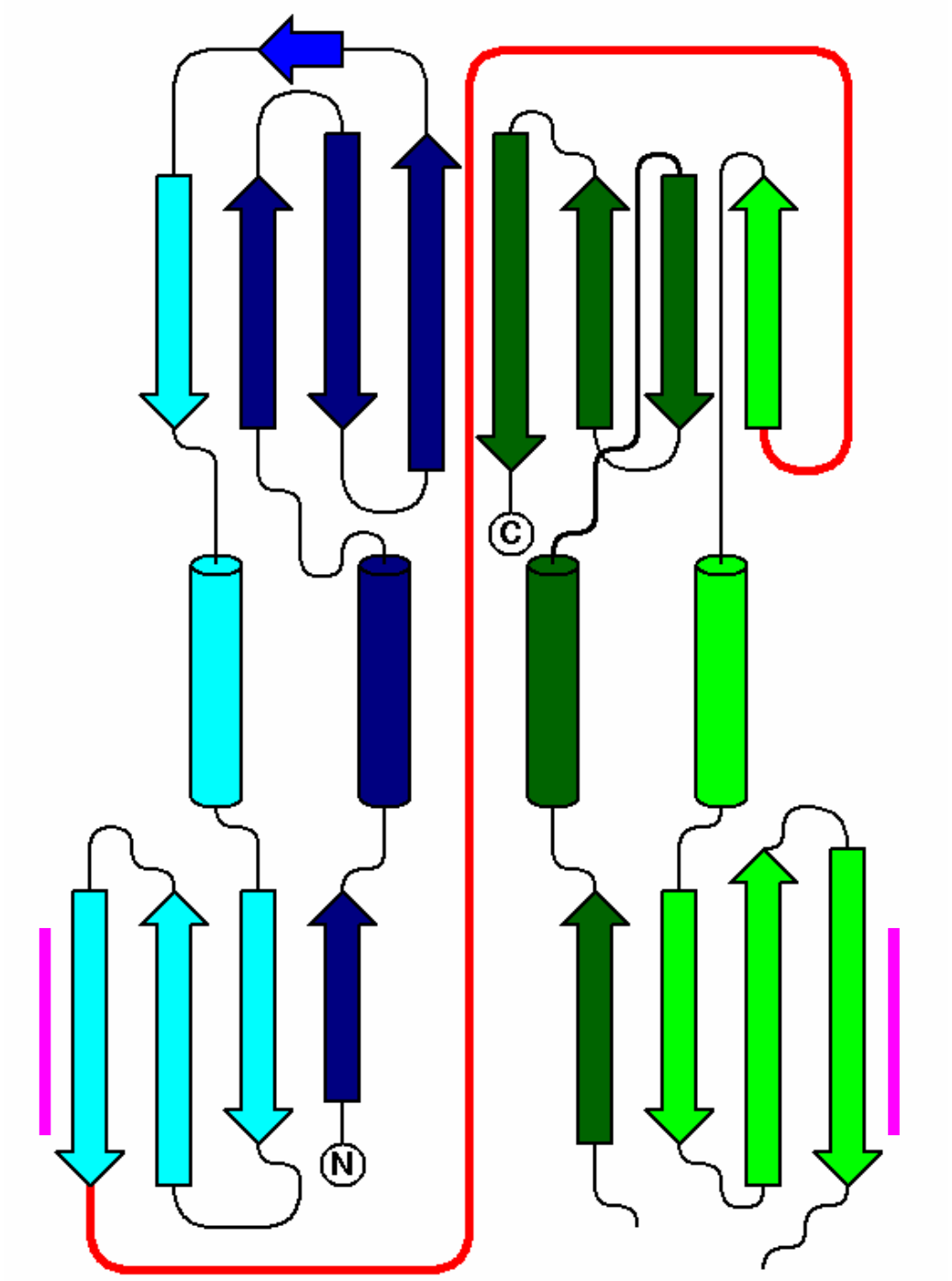


Figure 1-3. **The topology diagram of HsPCNA**
Figure showing the topology of the HsPCNA monomer as shown in the structure 1AXC (Gulbis, J. M. *et al* 1996). The four βαβββ are represented in dark blue, light blue, light green and dark green. The IDCL is shown in red and demonstrates the clear separation in the two domains. The lower β-sheets form the monomer-monomer binding interfaces at the magenta line. Generated using TopDraw (Bond, C. S. 2003).

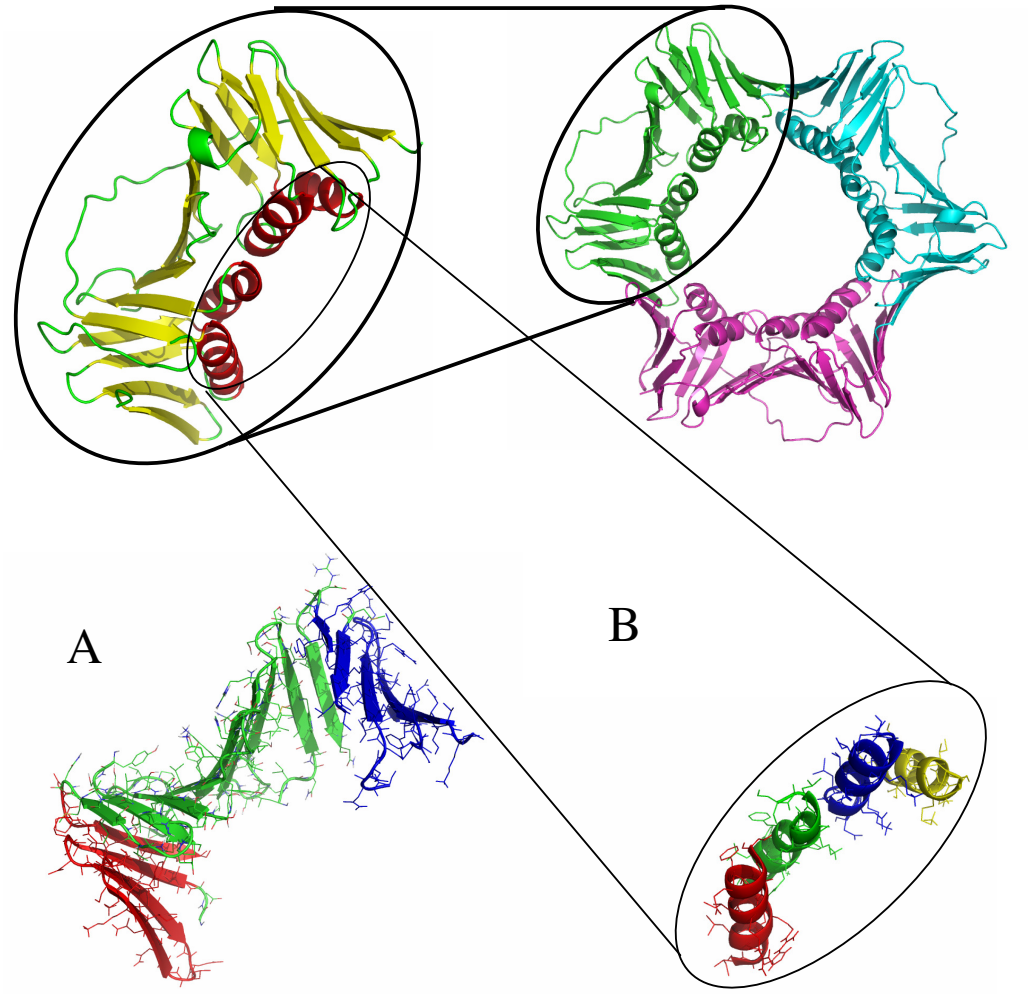


Figure 1-4. **The exploded view of PCNA showing the secondary structural motifs** Figure showing the secondary structural motifs of *H. sapiens* PCNA in an exploded form. A) The arrangement of the β-strands into β-sheets in PCNA. This arrangement shows three main β-sheets coloured red, green and blue and provide structural rigidity to PCNA. Red and blue sheets aid in the inter-monomer binding interaction (detailed further in figure 1.6). The green β-sheet is involved in the PCNA-PIP box interaction (detailed further in figure 1.7). B) The arrangement of the four α-helices, coloured red, green, blue and yellow, into the helix bundle. This image shows hydrophobic residues toward β-sheet region and basic residues toward DNA. These four helices are required for DNA interaction.

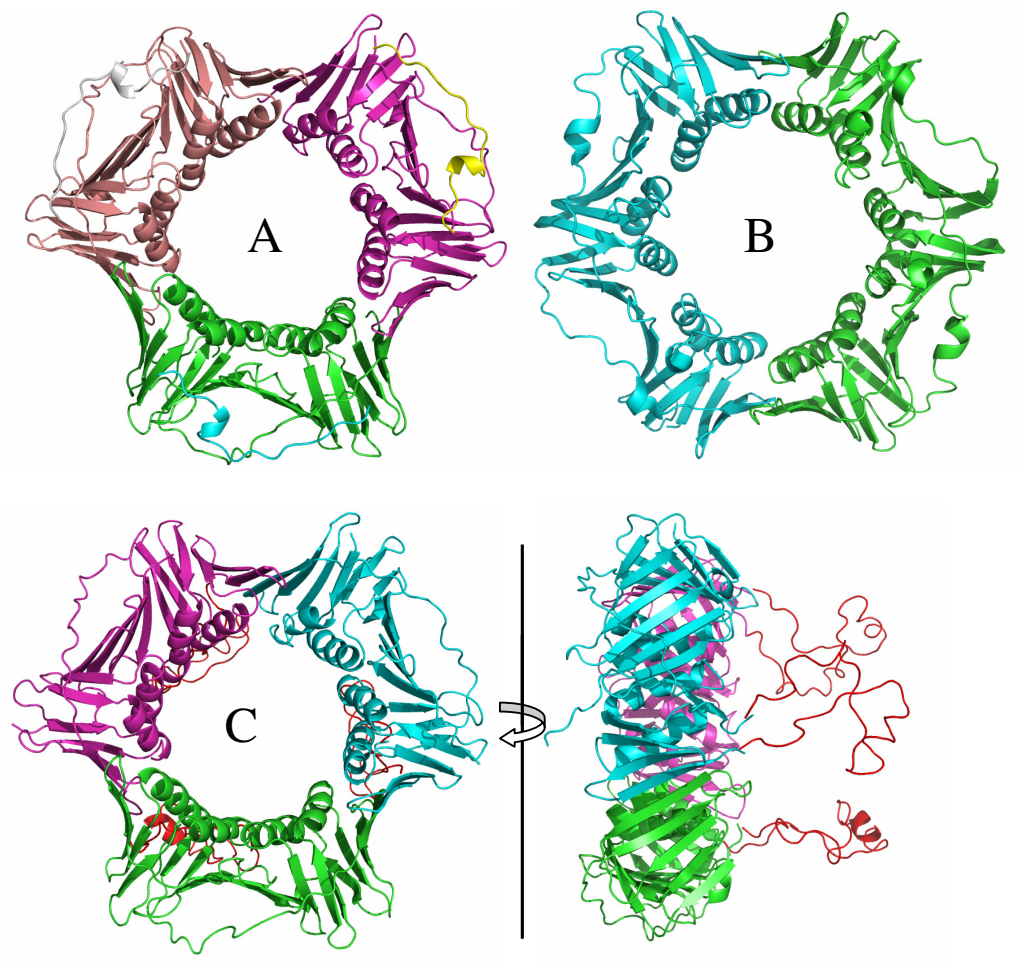


Figure 1-5. **Structural views of PCNA**
Figure showing the conserved structure of PCNA in different orders of life. These protein structures all demonstrate 12 βαββββ folds resulting in structural similarities A) A eukaryotic (*H. sapiens*) PCNA showing a trimeric form with peptide binding sites shown with bound peptide, from pdb file 1AXC (Gulbis, J. M. *et al* 1996).
B) A prokaryotic (E.coli) PCNA showing conserved structural features however present in a dimeric form, from the pdb file 2POL (Kong, X. P. *et al* 1992), as stated above E.coli PCNA shares only 16% identity with *H. sapiens* PCNA.
C :left) The structure of *L. major* PCNA demonstrating conservation of the structural features of PCNA. This protein is present as a trimeric form. This structure has been determined as in chapter 3. *L. major* PCNA shares 38% sequence homology with *H. sapiens* PCNA.
C :right) As C :right rotated 90° through the vertical. Demonstrating elongated loop regions (red) discussed in figure 1.2 and not present in *H. sapiens* PCNA. These regions have not been determined crystallographically and are modelled into the structure.
Images generated using PyMOL (DeLano, W. L. 2002)

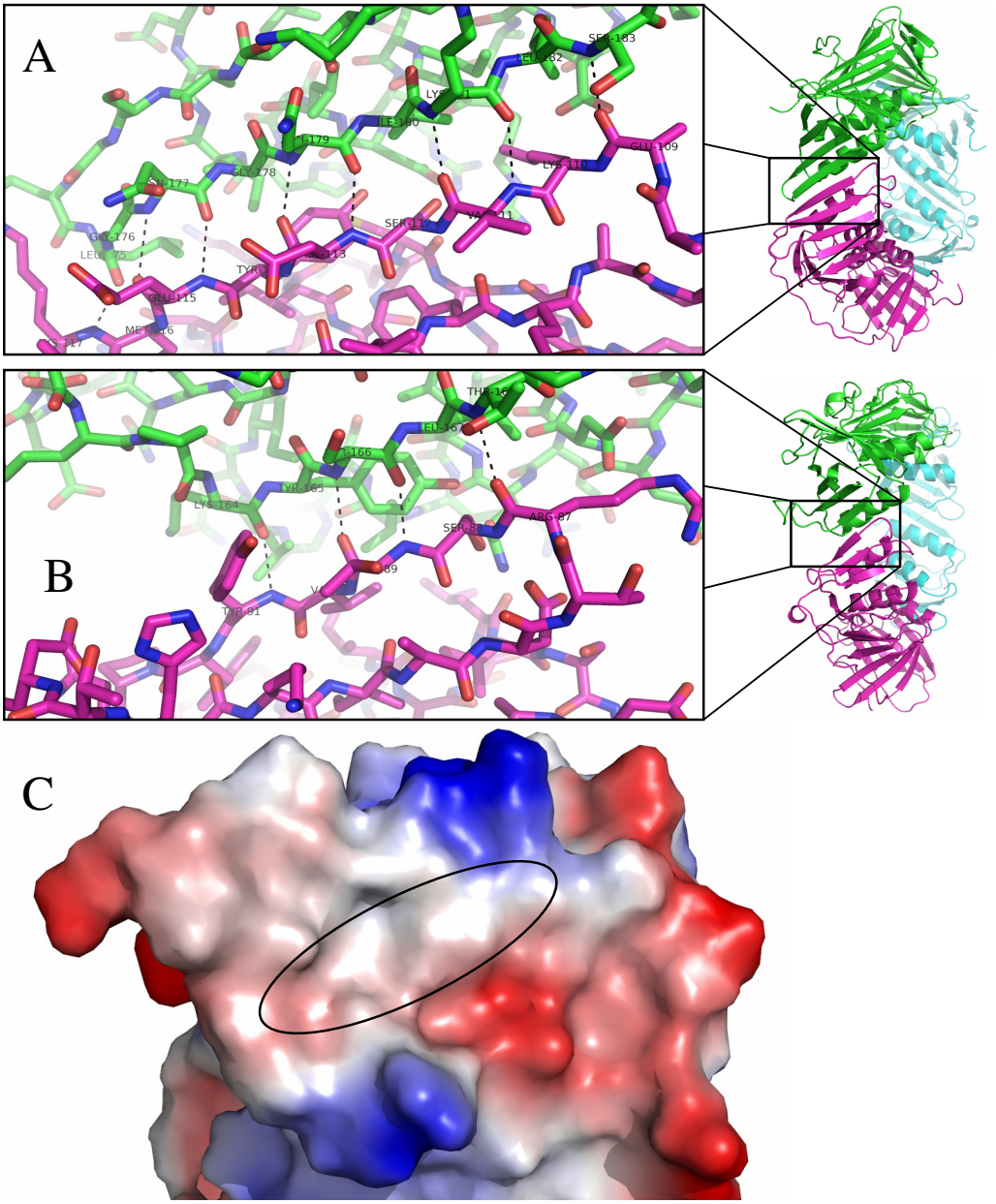


Figure 1-6. **The monomer-monomer interface of *H. sapiens* PCNA**
Figure demonstrating the inter-monomer interactions in a PCNA ring. A) The hydrogen bonding network at the *H. sapiens* PCNA monomer-monomer interaction. There are a total of eight hydrogen bonds (black) resulting in the high affinity interaction. These hydrogen bonds occur between β-strands E107-K117 and L175-S183 (from the pdb file 1AXC). B) The binding interface of RB69 PCNA monomer-monomer interface. This figure shows four hydrogen bonds (black) between β-strands R87-Y91 and K164-T169 (from the pdb file 1B77). C) The vacuum electrostatic representation of the binding interface of the *H. sapiens* PCNA monomer. The oval represents the hydrophobic core of the binding interface. Above the oval is the β-strand, required for hydrogen bonding, and below the positively charged α-helix for DNA interaction

1.2.1 The PCNA PIP box interaction

The majority of binding partners to PCNA (discussed in section 1.1) are essential for successful DNA replication and repair. Most of the interacting proteins which bind PCNA were found to contain a consensus sequence. This consensus sequence has been termed PCNA interacting protein (PIP) box. Some proteins containing this consensus sequence are presented here and grouped by cellular function (table 1.3).

The sequence of the PIP box was demonstrated in peptide studies (Warbrick, E. *et al* 1995) and verified in D. melanogaster (Warbrick, E. *et al* 1998) and present in other organisms also. The PIP box is a conserved sequence of residues with the Q-x-x-h-x-x-a-a motif, where (x) represents any amino acid, (h) any mildly hydrophobic amino acid and (a) represents an aromatic amino acid (Kontopidis, G. *et al* 2005).

The PIP box has demonstrated specificity for a PCNA binding interface. The binding region for the PIP box is formed by an anti-parallel β-sheet and a loop region, termed the inter-domain connector loop (IDCL) (figure 1.2 and 1.6). A hydrophobic pocket is formed between the anti-parallel β-sheet and the C-terminal side of the IDCL (figure 1.7). The conservation of the residues which make up the PIP box binding surface suggests their importance (figure 1.2). These interactions have been demonstrated in crystallographic studies between HsPCNA and p21 peptide (141-160) (Gulbis, J. M. *et al* 1996) (figure1.6) and through mutation analysis (Oku, T. *et al* 1998). In-vitro analysis by ITC of HsPCNA binding to p21 (141-160) peptide has given a K_d of interaction of 87.7nM (Zheleva, D. I. *et al* 2000). 27 hydrogen bonds are formed between the peptide and HsPCNA upon interaction (Gulbis, J. M. *et al* 1996).

Activity	Protein	Binding sequence	Note
DNA replication	pol δ - p12	GRKRLITDSYPV	
	- p66	NRQVSITGFFQR	
	- p125	ATQCQLEADV	
	FEN-1	STQGRLDDFFKV	
	DNA ligase I	MQRSIMSFFHP	
	RFC 1	MDIRKFFGV	
	PDIP 1	KKQTKVEFPEAR	
Trans-lesion polymerases	Pol η	GMQTLESFFKP	
	Pol ι	AKKGLIDYYLM	
	Pol κ	PKHTLDIFFK	
	Pol λ	ESVPVLELFSN	
MMR	EXO 1	GLQIKLNELWKN	Not Investigated
BER	UNG	IGQKTLYSFFSP	
	MPG	MYFCMNISSQGD	
	NTH 1		Not Found
	hMYH	MGQQVLDNFFRS	
	APE 1		Not Investigated
	APE 2	RGQKNLKSYFQP	
	Pol β	VEQLQKVHFIT	
	XRCC 1		Not Found
	XPG	QTQLRIDSFFRL	
DNA damage tolerance	TdT		Not Found
	PARP-1	QDLIKMIF	
	WRN	QWKLLRD	
	hBLM	QQRVKDFF	
	hRECQ 5	QNLIRHFF	
	WSTF	QDEIAEDY	
		QDIIHSIH	
		QASVIKKF	
		QCLVALLH	
Chromatin control	DNMT 1	QTTITSHF	
	HDAC 1		Not Investigated
	p300		Not Investigated
	ESCO 1/2	QLIIDAGQ	
Chromatid cohesion	Ctf 18		Not Investigated
	CDK 2		Not Found
Cell cycle control	EGFR		Not Investigated
	p21	QTSMTDFY	
	p57	SGPLISDFF	
	Cyclin D1		Not Investigated
	MCL 1	QRNHETAF	
	p15	QKGIGEFF	
	CDC 25C		Not Investigated
	ING 1b	QLHLVNYV	
	Gadd 45		Not Investigated
	MyD 118		Not Investigated
	CR 6		Not Investigated
	p53	QETFSDLW	
	MDM 2	QMIVLTYF	
Transposase	RAR α		Not Found

Table 1-3. **The interacting partners of PCNA in humans**
Table of known PCNA binding partners in *H. sapiens* modified from (Moldovan, G. L. *et al* 2007) supplementary data (S1). Showing the activity and interacting sequence of the binding protein.

PCNA binding partners not only bind the main PIP box interface (figure 1.7) but also other regions of PCNA have been determined to be important in protein association. One of the most characterised regions required for PCNA protein interactions is the PCNA C-terminus. Deletion mutations of the PCNA C-terminus have demonstrated this section to be required for the loading of PCNA onto DNA by RFC in eukaryotes (Fukuda, K. *et al* 1995). E.coli DNA polymerase III, a homologue for eukaryotic DNA polymerase ε, interacts with the C-terminal region of PCNA (subunit β) (Naktinis, V. *et al* 1996). A series of point mutations were introduced to the C-terminus of E.coli PCNA and were analysed by SPR. These mutations resulted in a reduced binding affinity for immobilised δ subunit of DNA polymerase III and DNA polymerase III core. This result was confirmed by phosphorylation analysis. A kinase phosphorylation motif was placed on the C-terminus of E.coli PCNA. The recombinant kinase target was protected from phosphorylation by the δ subunit of DNA polymerase III (Naktinis, V. *et al* 1996). These experiments demonstrate the importance of the prokaryotic PCNA C-terminus. The C-terminus of PCNA in both prokaryotes and eukaryotes is present at the hydrophobic pocket adjacent to the IDCL (figure 1.7), which may be the reason for its involvement in PCNA-protein interactions. The 27 hydrogen bonds facilitate the formation of the PCNA-peptide complex.

Away from the PIP box interface of eukaryotic PCNA are two specific residues (aspartic acid 41 and aspartic acid 97) which are also highly conserved in PCNA (figure 1.2 and 1.6). These two residues have been determined to be important in the interaction with RFC. Mutation studies of these two aspartic acid residues to alanine reduce the ATPase activity of RFC (Fukuda, K. *et al* 1995). This was also

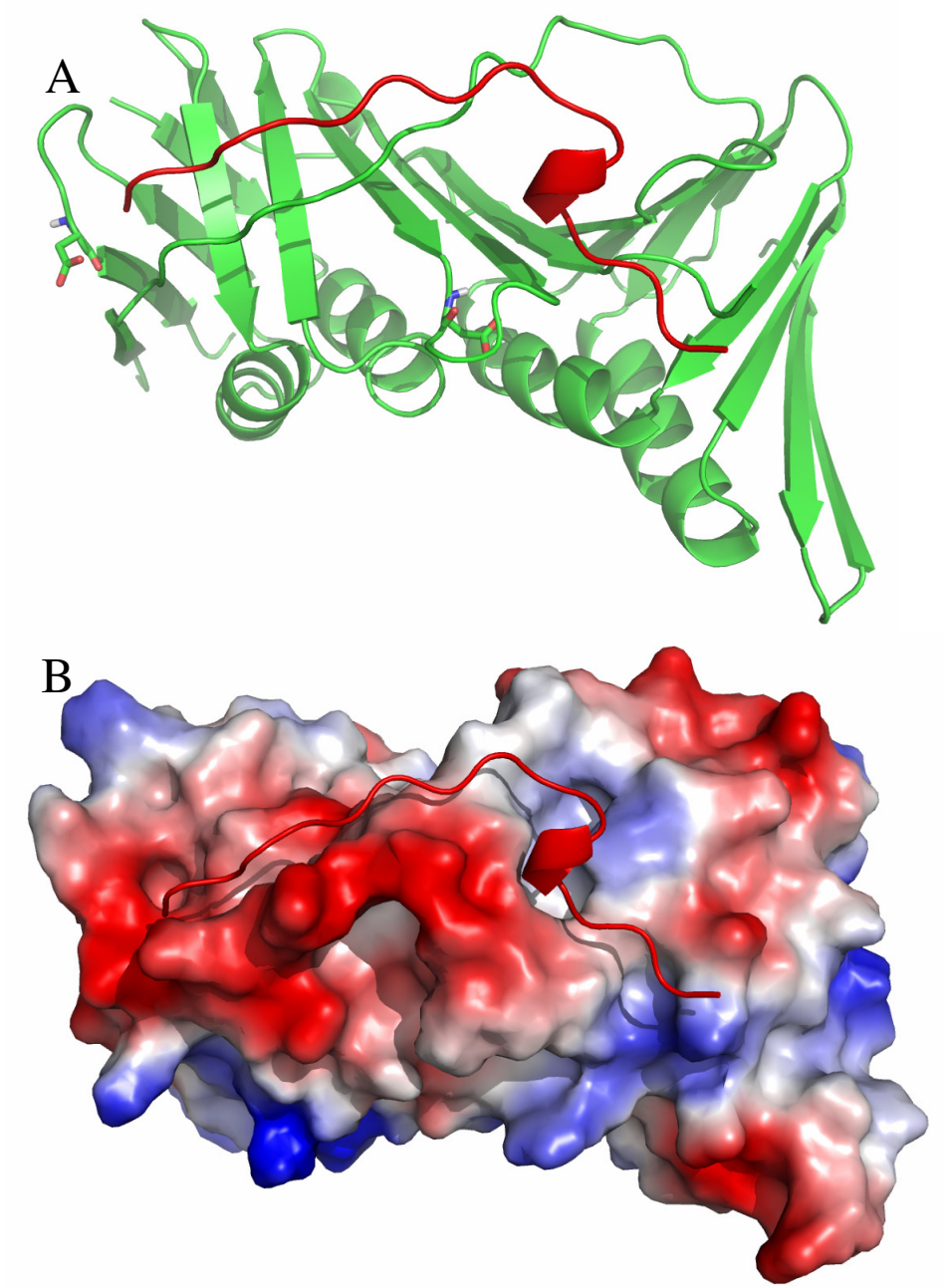


Figure 1-7. **The PIP box binding interface of *H. sapiens* PCNA-p21 (141-160)** The structure of the *H. sapiens* PCNA binding interface bound to p21 (141-160) peptide from the pdb file 1AXC (Gulbis, J. M. *et al* 1996). A) Cartoon form showing PCNA in green and p21 in red. The position of the helical region of p21 represents the hydrophobic pocket and is present between the anti-parallel β-sheet and the C-terminal side of the IDCL. The IDCL presents the potential for hydrogen bonding contacts. Residues D41 and D97 are shown in stick form. D41 is located near the helical region of p21. D97 is located at the N-terminal side of the IDCL. B) The electrostatic representation of PCNA showing p21 in red, and a clear hydrophobic pocket at the helical region in white. Generated using PyMOL (DeLano, W. L. 2002).

shown for aspartic acid 41 in mutation sensitivity studies (Ayyagari, R. *et al* 1995). D41 is located on a loop region near the hydrophobic pocket, suggesting that this residue may be involved in modifying the shape of this loop (figure 1.7A). D97 is shown to be at the N-terminal side of the IDCL and is potentially required for positioning the IDCL, or is in direct contact with the interacting partner (figure 1.7A).

It has been demonstrated that sections of PCNA not including the IDCL, 1-68 and 195-261, can separately bind cyclin D (Matsuoka, S. *et al* 1994). This demonstrates that residues within these regions are involved in PCNA-cyclin interactions.

1.3 PCNA as a potential drug target

The essential nature of PCNA in DNA replication and repair (discussed in section 1.1) suggests that the modulation of PCNA activity may be a lead as a drug target. Structure based drug design can use the PCNA hydrophobic pocket as a template to create a small molecule ligand. The creation of a small molecule ligand of PCNA has been previously demonstrated in E.coli PCNA, with a molecule termed RU7, which presented a K_i of ~10μM (figure 1.8) (Georgescu, R. E. *et al* 2008).

In normal cell lines, PCNA is up-regulated by the tumour suppressor protein (p53) (Xu, J. and Morris, G. F. 1999). When PCNA is present with low levels of p53, DNA replication takes place (Paunesku, T. *et al* 2001). High cellular concentrations of p53 stimulates the up-regulation of other proteins e.g. p21 (Radhakrishnan, S. K. *et al* 2006). p21 results in the stalling of normal replication and targets PCNA for DNA repair mechanisms (Paunesku, T. *et al* 2001). At high cellular concentrations of

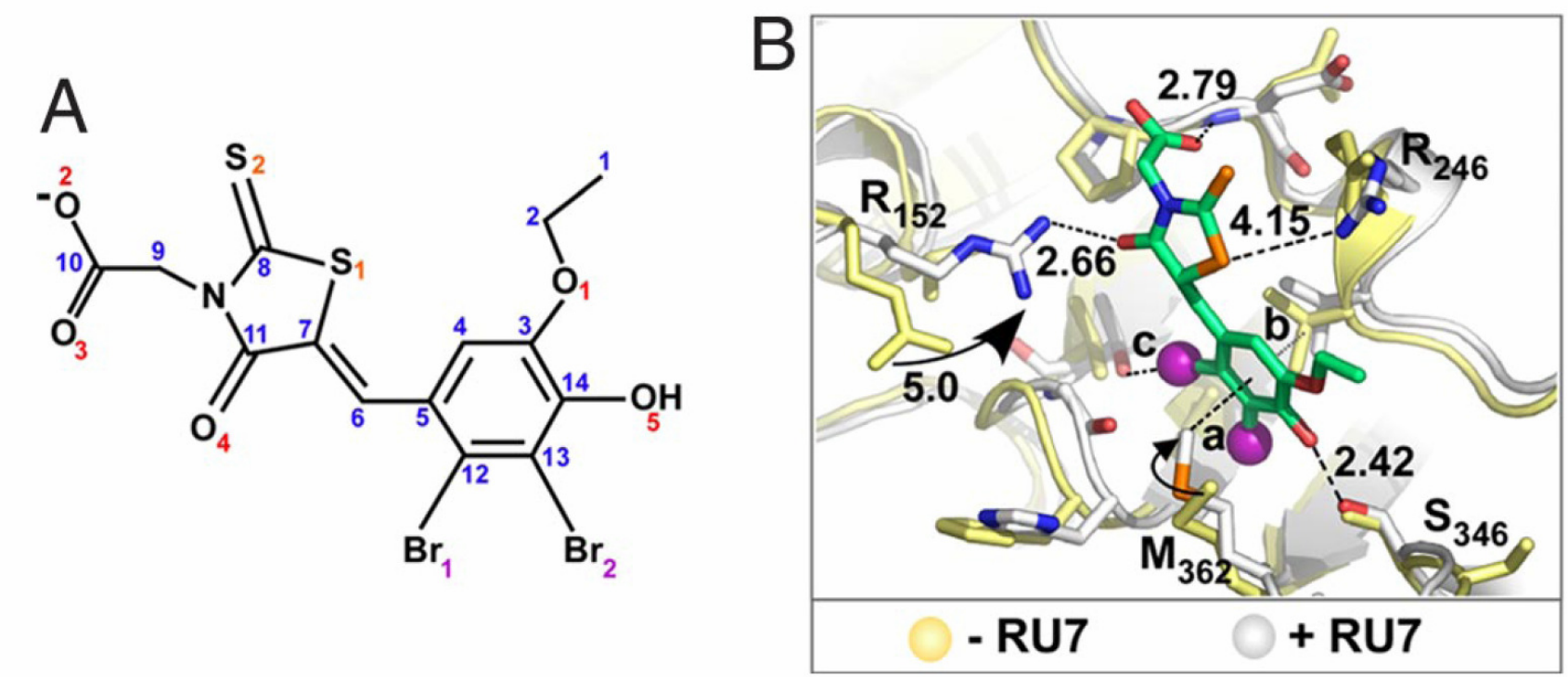


Figure 1-8. **A small molecule inhibitor represented in the binding pocket of E.coli PCNA**
Figure representing the first, effective small molecule inhibitor of E.coli PCNA. A) The schematic representation of the RU7 small molecule inhibitor. B) The small molecule inhibitor bound into the E.coli PCNA PIP box binding interface. The native structure of E.coli PCNA is represented in yellow and the complexed structure of E.coli PCNA is shown in white showing a shift in residue side chains. Distances are given in Ångströms. Image is taken directly from (Georgescu, R. E. *et al* 2008)

p53, high p21 and normal PCNA, the cell undergoes apoptosis (Chen, J. J. *et al* 1996b; Paunesku, T. *et al* 2001).

The up-regulation of HsPCNA is a common marker of proliferating cancerous cell lines (Paunesku, T. *et al* 2001). High levels of PCNA are linked to malignant tumours and the inhibition of PCNA within these cell lines results in the induction of apoptosis (Paunesku, T. *et al* 2001). This has been demonstrated by treatment of cells with anti-sense oligonucleotides resulting in inhibition of cancerous cells with only slight effect in normal cells (Sakakura, C. *et al* 1994). This data suggests the HsPCNA can potentially be used as an anti-cancer target.

The inhibition of PCNA by p21 is a suggested method of cancer suppression (Mattock, H. *et al* 2001). Mimicking this interaction by peptidomimetics could provide a route to a potential anti-cancer therapy (Kontopidis, G. *et al* 2005). A small molecule inhibitor, like that demonstrated in E.coli (Georgescu, R. E. *et al* 2008), targeted for HsPCNA, may provide a potential route to an anti-cancer therapy.

1.3.1 Fungal related diseases

S. pombe is a form of fission yeast used as a model organism for molecular biology. Yeasts are a type of fungi which are responsible for many ailments in humans e.g. Cryptococcus and Candidiasis. Not only are yeasts a form of pathogenic organism, but also a cause for food spoilage. Food spoilage occurs as a result of some yeasts (Zygosaccharomyces) ability to survive high concentrations of sucrose, ethanol, acetic acid, sorbic acid and benzoic acid; all of which are commonly used in the preservation of food (Corry, J. E. L. 1978).

These lower eukaryotes have a cellular use of PCNA as a control system as in mammalian cells (Ayyagari, R. *et al* 1995). This similarity in mechanism implies that

the inhibition of *S. pombe* (Sp)PCNA results in the stalling of cellular division (Warbrick, E. 2006). The PCNA PIP box binding interface may present structural differences between HsPCNA and SpPCNA. The determination of small differences in the PCNA PIP box binding interface of PCNA between HsPCNA and SpPCNA may allow the creation of a species specific inhibitor. The goal is to allow the inhibitor to target non-human cell lines, thus targeting infection with little side effects. This suggests that a small molecule inhibitor for SpPCNA could present a potential lead for an anti-fungal treatment.

1.3.2 Leishmania related diseases

Leishmania is a protozoan pathogen spread by an insect vector. Infection by this parasite is referred to as leishmaniasis. Approximately 30 species of the phlebotomine sand fly can act as the insect vector. These insects live most commonly in tropical regions (World Health Organisation (WHO) 2009). This disease is most problematic in the developing world with 90% of all visceral leishmaniasis cases occurring in Bangladesh, Brazil, India, Nepal and Sudan; 90% of mucocutaneous leishmaniasis occuring in Bolivia, Brazil and Peru; and 90% of cutaneous leishmaniasis cases occuring in Afghanistan, Brazil, Iran, Peru, Saudi Arabia and Syria (World Health Organisation (WHO) 2009). As a result of this disease being most prevalent in poorer countries, many who require treatment can not afford it. This disease can present in three main forms, visceral, mucocutaneous and cutaneous. These three different forms of leishmaniasis arise from different strains of Leishmania. Leishmania major results in cutaneous leishmaniasis (Gonzalez, Urba *et al* 2008).

The life cycle of Leishmania major has two main stages (figure 1.9). These stages are the vector stage (when present in the sand fly) and the infections (when present in mammals). Leishmania are transferred to the sand fly during feeding and amasitgotes are released into the midgut. The amastigotes grow into promastigotes and can freely divide. These promastigotes then migrate to the proboscis (Laskay, T. *et al* 2003). When the sand fly feeds again, the promastigotes are transferred to the mammal. Mammalian macrophages phagocytose the Leishmania promastigotes. These promastigotes revert to amastigotes and divide further. These amastigotes can lyse infected cells, and continue to infect other cells (Laskay, T. *et al* 2003). These amastigotes can then be transferred to another sand fly during feeding.

Current drugs available to combat leishmaniasis are toxic to human hosts. These drugs are also losing efficacy due to increasing resistance of the parasites. These factors stress the requirement for more effective drugs to combat this disease at a lower financial cost (World Health Organisation (WHO) 2009).

A link has been shown between the occurrence of leishmaniasis and severity of HIV infections. Leishmaniasis can also be spread by the sharing of needles by intravenous drug users; moreover, an infection with each of these diseases can elevate the symptoms of both. This is because leishmaniasis will suppress the immune system causing an accumulative immuno-suppression with HIV; and the parasitic infection will stimulate replication of the virus. The symptoms of leishmaniasis are amplified in HIV carriers this is as co-infection rapidly activates disease in parasite carriers also (Olivier, M. *et al* 2003).

The protozoan cell cycle biochemistry has been shown to be similar to that of the eukaryotic cells, demonstrating many homologous kinases. PCNA is likely an

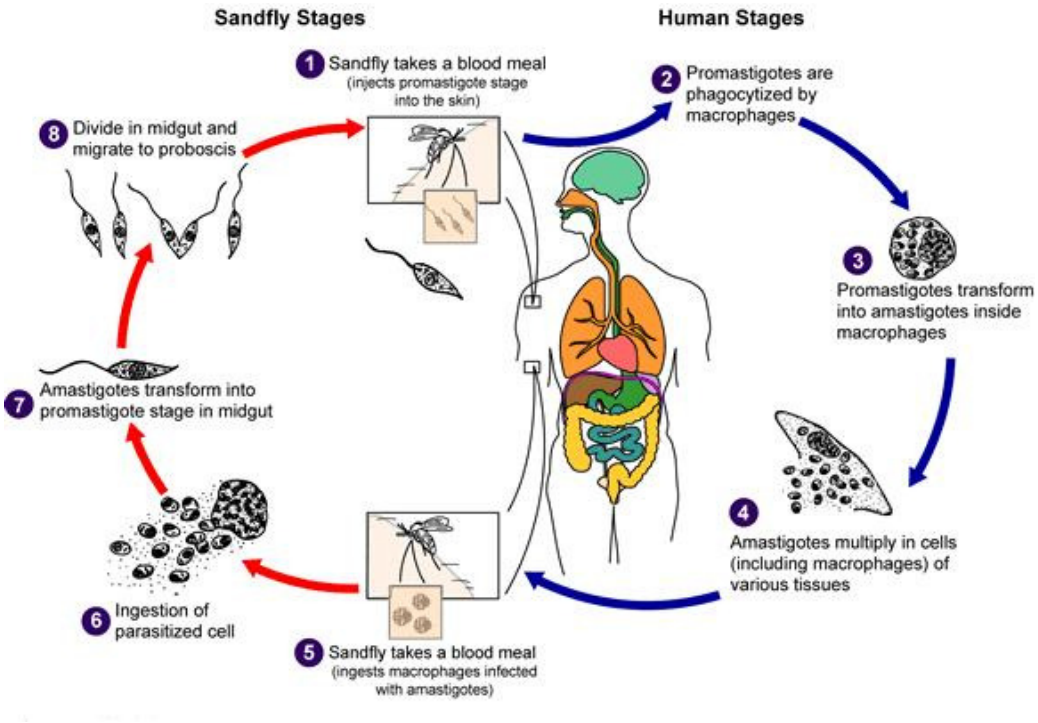


Figure 1-9. **The life cycle of Leishmania major**
Figure showing the two stage life cycle of Leishmania major. In red, to the left, stages present in the sand fly and in blue, to the right, stages present in the mammalian host. 1) the infection of the human host during the feeding of the sandfly. 2) macrophages are invaded by promastigotes. 3) formation of amastigotes from promastigotes in macrophages. 4) the multiplication of amastigotes and lysis of macrophages resulting in the release of amasigotes. 5-6) the infection of the sand fly with amastigotes during feeding. 7) the growth of amastigotes into promastigotes in the midgut of the sand fly. 8) the replication of the promastigotes and migration to the proboscis

Modified from the Centre for Disease Control website accessed 27/07/09'

important protein for the cell cycle in Leishmania similar to that of humans (Mottram, J. C. *et al* 1993). As stated above (section 1.3.1) if the structure of *L. major* (Lm)PCNA were solved, differences in the surface of the PIP box binding interface could be determined. Surface differences may potentially be exploited for the creation of species-specific PCNA inhibitors. This may create drugs capable of specifically targeting and destroying *L. major* cells and leaving *H. sapiens* cells intact. This infers that the creation of a small molecule inhibitor of LmPCNA could act as a potential anti-parasitic lead. The successful creation of a cheap anti-parasitic therapy may also aid in the alleviation of AIDS symptoms in developing countries.

The alignment and structures of LmPCNA (figures 1.2 and 1.3) show a loop region of protozoa PCNA, residues 194-228, which is not present in eukaryotic or prokaryotic PCNA. This loop region has an unknown function and may present a new target for ligand-binding. The elongated loop is present on the opposite side of the PCNA ring from the binding interface, implying that it is not required for the PCNA PIP box interactions (figure 1.4).

1.4 Project outline

This project is based on the biophysical and structural investigation of the PCNA-p21 peptide interaction. X-ray crystal structures of LmPCNA have been solved and compared to the known structures of HsPCNA and SpPCNA. The biochemical interactions of HsPCNA, SpPCNA and LmPCNA with the p21 peptides will be determined by different biochemical techniques.

Using the well defined PCNA-p21 peptide interactions, the effect of different peptides can be analysed. These peptides can allow the thorough determination of the effect of a p21 point mutation. The effect of different biotin tags can be analysed to

better understand the effect of steric hinderance in peptide protein interactions analysed by SPR.

1.4.1 Crystal structure solution of LmPCNA

The solution of the structure of LmPCNA required the purification and crystallisation of the protein. The solution of the structure demonstrated the level of structural similarity of LmPCNA with HsPCNA and SpPCNA. The structure solution failed to provide any information regarding the form of the elongated loop region and no evidence for its activity was determined. A co-crystal structure of LmPCNA with the p21 peptide has also been solved to demonstrate the PCNA peptide interactions present in the crystal. This solution has been used to explain possible differences in strength of interaction determined from the biochemical experiments. This structure is now available to allow the design of LmPCNA inhibitors by structure-based drug design. These novel inhibitors may lead to novel anti-parasitic drugs.

1.4.2 Characterisation of the PCNA peptide interaction

The binding characteristics of HsPCNA, SpPCNA and LmPCNA to peptides can be determined by different biochemical methods. The binding affinity of these three PCNAs for the p21 peptide (141-152) has been determined by SPR, thermal shift assays and ITC. The interaction of the three PCNAs to p21 (141-152, F150A) mutant peptide has also been determined. The wild type and point mutant peptides' binding characteristics have been compared. This data provides information regarding the thermodynamic contribution to the binding interaction of the six carbon ring of F150. The interaction of a variety of peptides with differing biotin tags to PCNA from these three species has also been studied (discussed in greater detail in chapter 7).

In this work, the association of the seven different peptides has been analysed for the affinity for the three PCNAs; HsPCNA, SpPCNA and LmPCNA. These seven peptides (figure 1.10) are shown here:

Peptide0 – KKRQTSMTDFYHSKRRLIFS
Peptide1 – KRRQTSMTDFYH
Peptide2 – Bio-KRRQTSMTDFYH
Peptide3 – Bio-PEG-KRRQTSMTDFYH
Peptide4 – Bio-PEG-PEG-KRRQTSMTDFYH
Peptide5 – Bio-PEG-PEG-Ahx-KRRQTSMTDFYH
Peptide6 – KRRQTSMTDAYH

1.4.3 Effect of spacer lengths in SPR

A common method of affinity determination by SPR is to immobilise the ligand to a sensor surface by the Streptavidin-Biotin interaction. The ligand (peptide) is immobilised to the surface. Immobilisation to StreptAvidin can affect the interaction due to steric hinderance. PEG spacers can be used to relieve this steric hinderance, by separating the ligand from the biotin. The difference in spacer length results in changes in binding affinity between PCNA and immobilised peptide. This difference was confirmed to be due to the relief of steric hinderance, and not due to modification of the PCNA PIP box interaction. The affinities of these peptides with PCNA from three different sources, HsPCNA, SpPCNA and LmPCNA, have been determined with PCNA immobilised to an SPR sensor surface. These experiments determined the level of steric hinderance encountered due to linker size. The affinity of these PCNA peptide interactions have been verified by thermal denaturation assays and ITC.

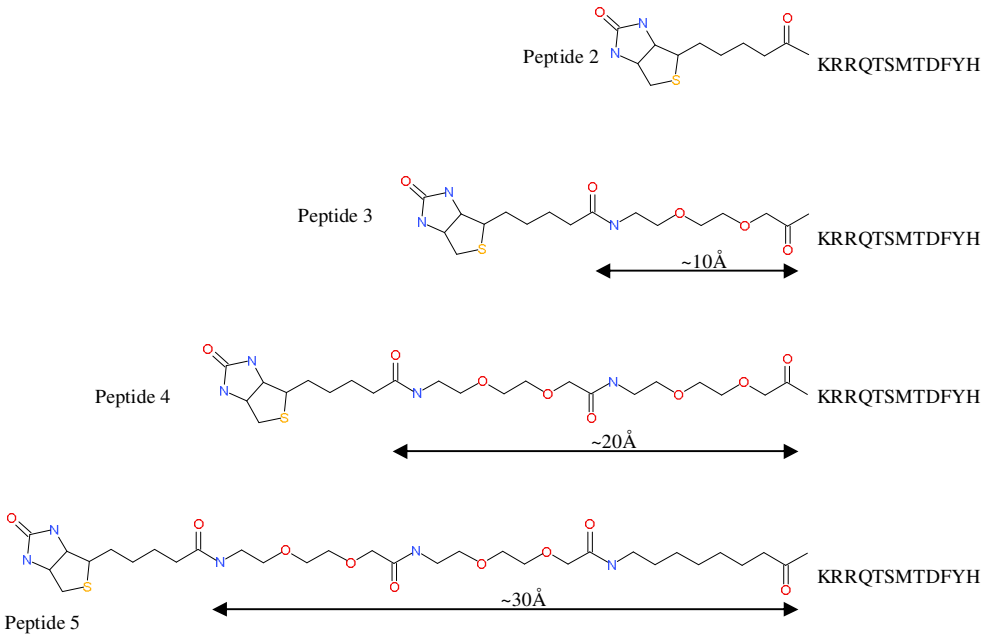


Figure 1-10. **The peptides investigated by SPR in this work**
Figure showing the peptides used in SPR experiments carried out in this study. Peptides 2-5 showing the biotin tags of the peptides for use with spacer investigation. The full chemical sequence of the linkers showing the biotin, PEG and hydrophobic (octanoic acid) spacers are represented. The additional spacers have their approximate length shown below in Ångströms. The biotin tags are an approximately 10Å length.

CHAPTER 2. The molecular biology and purification of three PCNAs

In this chapter, the cloning, expression and purification of PCNA from three organisms are described. A plasmid vector is required to be created containing a gene encoding the protein of interest and appropriate antibiotic selection marker. This plasmid vector is created using molecular cloning techniques. This vector is used in conjunction with bacterial expression systems which are commonly used for protein production for crystallography. These expression systems are acceptable, provided no post-translational modifications are required.

PCNA from all three sources are required to be pure enough for crystallographic analysis (>97%) and with high enough yield (~2mg per purification). For each of these three PCNAs, the protein is required to be folded and capable of binding a peptide with a PIP box present.

2.1 Principles of molecular cloning

Cloning a target gene requires large quantities of the gene of interest to be amplified and have "sticky ends". Sticky ends are a term for DNA overhang, a small section of single stranded DNA at the end of the DNA strand. These sticky ends can then allow ligation to required vector DNA with complementary sticky ends and form a circular plasmid. The final circular plasmid can then be transformed into an expression cell-line for protein production.

Molecular cloning requires a large quantity of DNA. DNA concentration for ligation can be between 250µg ml^{-1} and 0.2µg ml^{-1} (Chang, A. C. Y. and Cohen, S. N. 1978; Shore, D. *et al* 1981). Low DNA concentrations can result in product levels too low for subsequent transformation and high DNA concentrations can result in DNA to be annealed in linear form. The optimal DNA concentration of ligation to

obtain sufficient levels of circular DNA for transformation is required to be between 10µg ml^{-1} and 5µg ml^{-1} (Remaut, E. *et al* 1983; Sambrook, J. *et al* 1989). Insert DNA to vector DNA should have a ratio of ~3:1 respectively for optimal formation of circular plasmids with only single inserts. This quantity of insert DNA can be obtained by Polymerse Chain Reaction (PCR) using Thermus aquaticus (Taq) DNA polymerse (Saiki, R. K. *et al* 1988).

2.1.1 Methods of molecular cloning

Original methods for ligation would be to generate the insert as above and then subsequently cleave the short linear DNA to form sticky ends (figure 2.1). The vector plasmids are linearised by cleavage with the same restriction enzymes. The two fragments can be gel extracted to isolate only the insert and plasmids. These fragments would then be ligated as above (figure 2.1). This method has significant problems as each step results in the loss of insert DNA and a potential for the DNA to become damaged, reducing the probability of the final ligation event being successful. The cleavage of short strands of DNA also results in reduced efficiency as it is known that restriction enzymes have a reduced functionality when cleaving at the end of a DNA strand (Moreira, R. F. and Noren, C. J. 1995). This will result in a reduced quantity of insert DNA being capable of ligation.

2.1.1.1 Modern methods for simple cloning procedures

Modern techniques involve the use of a cloning vector. A cloning vector is a preformed linearised plasmid capable of annealing to PCR products immediately after the PCR reaction. Two of the most common cloning vectors in use are the pCR®2.1-TOPO® from the TOPO® TA cloning system (Invitrogen) and the pJET1.2/blunt cloning vector from the CloneJet™ PCR cloning kit (Fermentas).

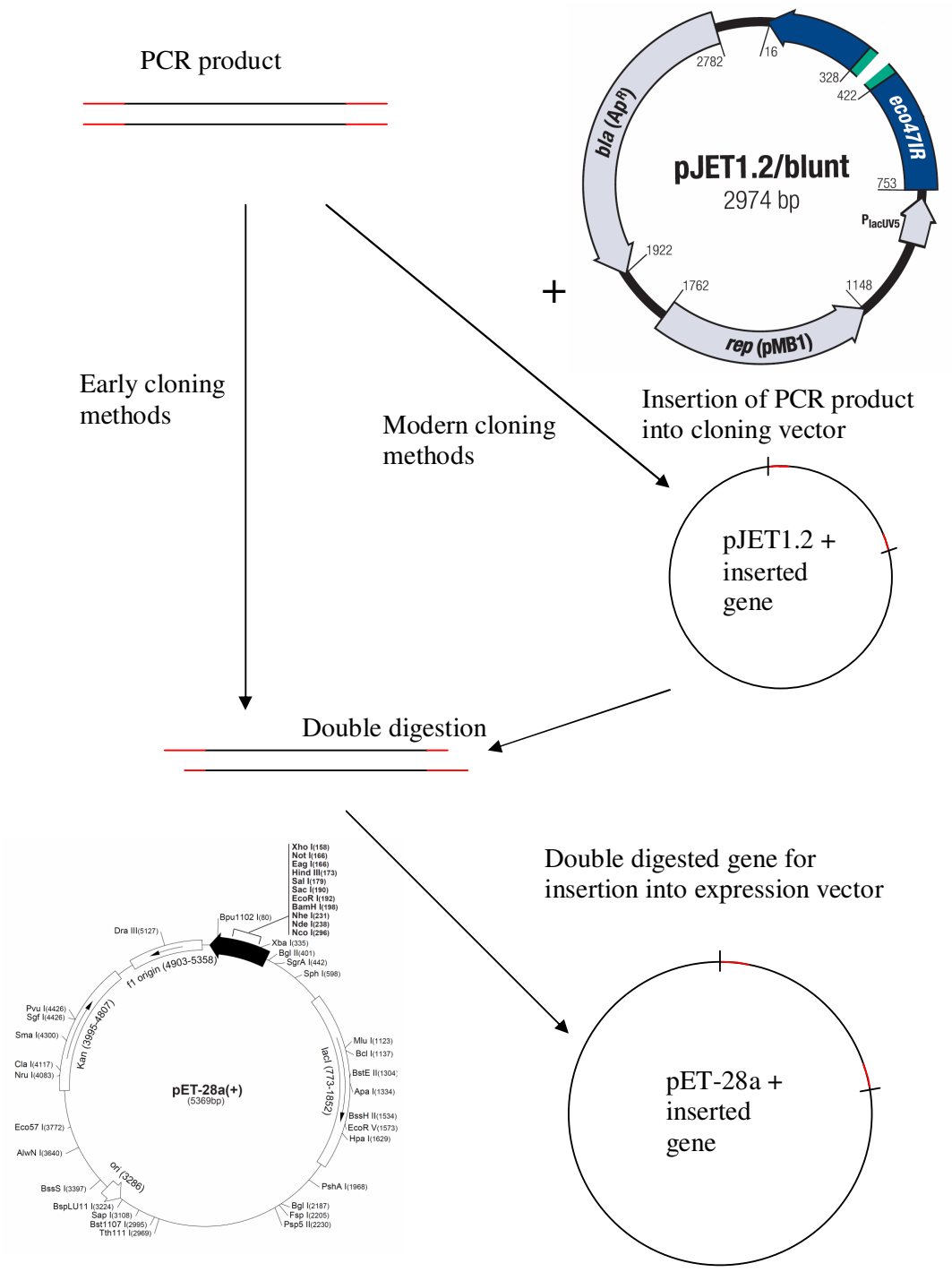


Figure 2-1. **The general principles of the cloning of protein encoding genes**
This figure demonstrates the general steps involved in the cloning of a gene into a bacterial expression vector. The initial step requires the production of a PCR product with appropriate restriction sites (red). Modern cloning techniques allow this gene to be inserted into a cloning vector which can be freely produced for a large quantity of cleaved PCR insert. This doubly digested PCR product can then be inserted into a cleaved expression vector to give the final construct.

These can then be used to produce large quantities of cleaved insert readily.

The TOPO® TA vector is a self ligating system with Topoisomerase covalently bound at each end. This plasmid contains dual antibiotic resistance (Ampicillin and Kanamycin) and also the ability for blue/white selection (Messing, J. *et al* 1977; Norrander, J. *et al* 1983; Vieira, J. and Messing, J. 1982). Dual antibiotic resistance allows for the selection of cells with successful transformants. Blue/white selection allows for the selection of successful ligations due to disruption of the lacZ gene. For successful ligation, a 3'-A overhang is required which is added to the PCR product by the Taq polymerase. Proof reading DNA Taq such as VentR® or PfuTurbo® should ideally be used for cloning in order to reduce mutation rates; these polymerases however remove the 3' overhang reducing the efficiency of TOPO® TA.

The pJET1.2/blunt is a blunt end cloning vector which requires the further addition of T4 DNA ligase to the reaction for the ligation to occur. This vector only confers ampicillin resistance to allow selection for positive transformants. The Eco47I gene is also present in this cloning vector which is disrupted when the insert is present (figure 2.1). Should a cloning vector self-ligate, this gene will be active in the transformed cell and produce a restriction enzyme which is lethal to E.coli host cells. Due to the blunt end ligation system, the CloneJet vector requires the 3'-A overhang to be removed prior to ligation if other forms of Taq DNA polymerase are used.

2.1.2 Materials and methods of the cloning of HsPCNA

A PCR product of the HsPCNA gene was formed using Expand high fidelity Taq plus (Roche). A PCR mixture of high fidelity Taq 2.5U, 5µM each primer, 50µM each deoxyribonucleotide, 1× high fidelity Taq reaction buffer, 1µg template

plasmid DNA, 1mM $MgCl_2$ made up to 50µl with ddH_2O was used to amplify the HsPCNA gene. The PCR was performed with an annealing temperature of 48°C, elongation temperature of 72°C, melting temperature of 92°C and was repeated 28 times, in using a T1 thermal cycler (Biometra). The PCR product was gel extracted from a 1% agarose gel using the QIAquick gel extraction kit (QIAGEN) and the insert was eluted in 30µl ddH_2O. 2µl of the gel extracted PCR product was used and the 3'-A overhang was removed as described in the in the CloneJet reaction for blunt end cloning handbook. The pJET1.2/blunt-HsPCNA plasmid was prepared from a DH5α culture using the QIAprep Spin miniprep kit (QIAGEN). pJET1.2/blunt-HsPCNA and pET-28a plasmids were cleaved using 10 units of BamHI and 10 units of XhoI, 1× BamHI cleavage buffer (NEB), 1× BSA and made up to 50µl with ddH_2O at 37°C for 3hours. Cleaved HsPCNA and pET-28a were ligated using the protocol from the T4 rapid ligation kit (Invitrogen). Clones were verified by colony PCR and double digestion, and the final product was sequenced at the University of Dundee sequencing service.

2.1.3 Cloning of PCNA from three species

A PCR product of the coding region was created using the forward primer -

3'-GCTCGAGATGTTCGAGGCGCGCCTG-5'

and reverse primer -

3'-GGGATCCTAGCCATATGTACTTAGAGGTAC-5'

with high fidelity Taq. Fresh PCR product was used with CloneJet cloning kit. Colony PCR of the resulting colonies verified the presence of insert. Plasmids were isolated from resulting colonies and analysed by double digestion with BamHI and XhoI and sequenced. The cleaved insert was ligated into cleaved pET-15b and

transformed into commercially competent DH5α cells. Resulting colonies were analysed by colony PCR and plasmids were double digested to verify the presence of insert. These purified plasmids were sequenced and verified to be accurate (figure 2.2, top). This formed a recombinant N-terminal 6×HIS tagged HsPCNA.

The SpPCNA plasmid was a gift from Conny Ludwig. This plasmid was verified, and the SpPCNA gene was present on the pQE-32 plasmid between the BamHI and SalI cleavage sites. This generates a recombinant N-terminal 6×HIS tagged SpPCNA.

The LmPCNA construct was created at the University of Glasgow and was a gift from Jeremy C. Mottram, this was assigned pGL1583 from the *L. major* putative PCNA gene lmjF15.1450. PCR from genomic DNA was performed using the forward primer- 5'-CCATATGCTGGAGGCTCAGGTCCAGTACG-3'
and reverse primer- 5'-GCTCGAGTCACTCCGCATCGTCCACC-3'

The PCR product was inserted into pET-28a using NdeI and XhoI cleavage sites and the plasmid product was verified to be accurate by sequencing (appendix 1.1). This formed a recombinant N-terminal 6×HIS tagged LmPCNA.

2.2 Principles of protein expression

Different methods of protein expression exist e.g. bacterial, yeast, baculovirus; each with different applications. In this work I will only discuss principles of recombinant protein expression in bacteria. Bacteria, commonly E.coli are primarily used for large scale protein expression for crystallisation. Different bacterial cell lines, expression vectors, expression media, expression conditions and expression time can be modified to increase the protein yield.

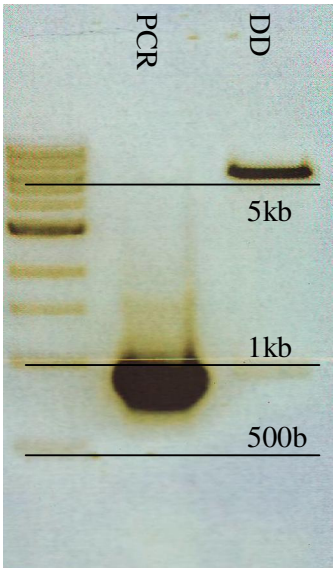
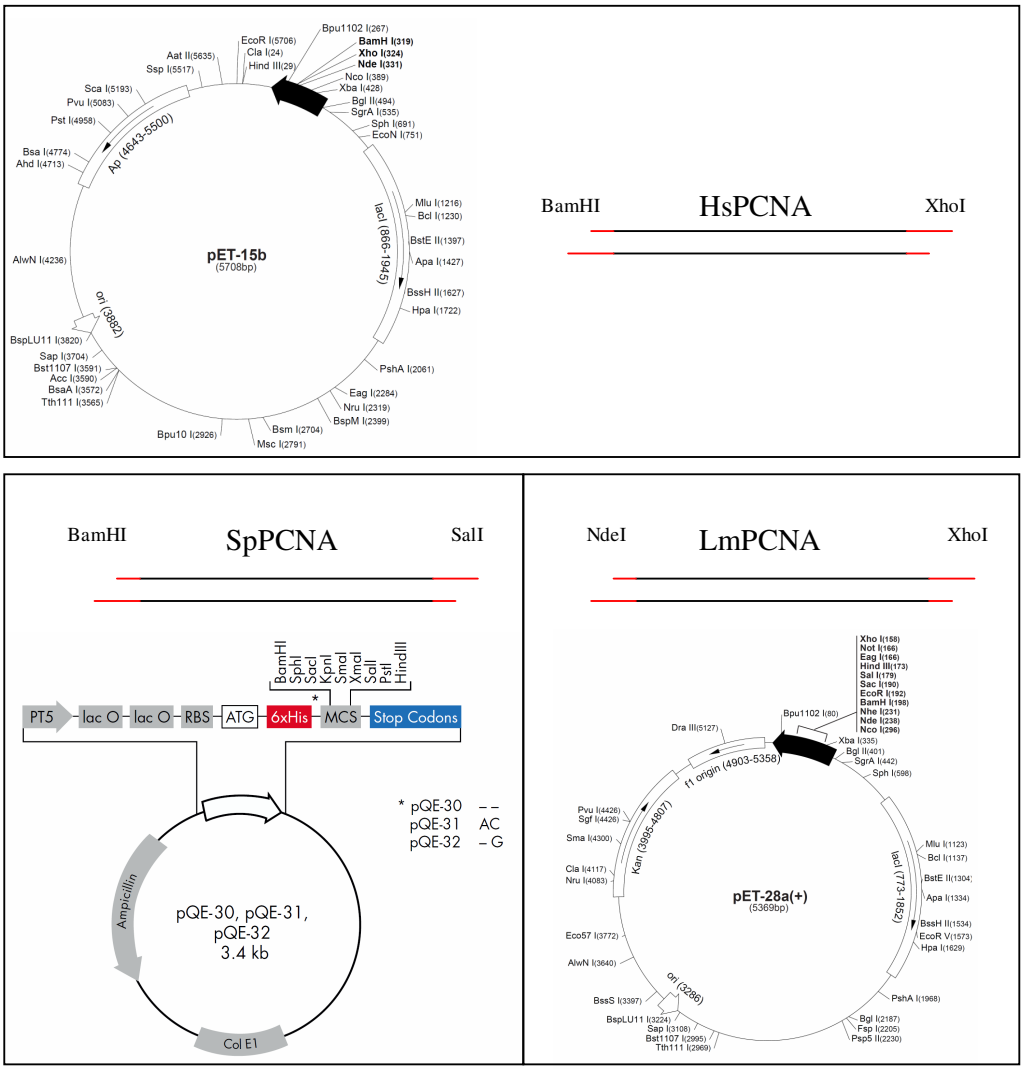


Figure 2-2. **The cloning and expression of HsPCNA-pET28a construct**
Top) The PCR and double digestion of the HsPCNA-pET28a construct, verifying the presence of the HsPCNA insert with BamHI and XhoI restriction. Double digestion demonstrates an insert and plasmid of appropriate size, 800bp and 5.4kb respectively.

Bottom) The constructs of the expression vectors for HIS-tagged PCNA from three species. HsPCNA was cloned into pET-28a using the BamHI and XhoI sites. SpPCNA has been inserted in pQE-32 vector between the BamHI and SalI restriction sites. The LmPCNA gene has been inserted into the pET-28a expression vector between the NdeI and XhoI restriction sites.



2.2.1 E.coli cell lines for protein expression

Different E.coli cell lines can be used in order to maximise expression by overcoming different problems encountered in the expression of different proteins. The (DE3) cell lines contain the bacteriophage λ DE3 lysogen in its genomic DNA, which encodes for the T7 RNA polymerase. This gene is regulated by the lacUV5 promoter, which allows high recombinant transcription levels in the presence of Isopropyl β-D-1-thiogalactopyranoside (IPTG), an allolactose mimic. Cell lines are briefly represented in table 2.1. BL21 (DE3) cell lines are deficient of the Lon and OmpT proteases, reducing the level of protein degradation by protease activity (Studier, F. W. and Moffatt, B. A. 1986). BL21 (DE3) Star® are similar to BL21 (DE3), the Star cell line refers to a mutated rne131 gene resulting in a defective RNase E protein. This mutation allows greater cellular levels of mRNA to accumulate within the cell in order to increase expression (Lopez, P. J. *et al* 1999). Rosetta (DE3) cell lines contain increased expression of tRNA required for translation. This cell line allows higher levels of all tRNA including those for rare codons. This allows proteins encoding greater numbers of rare codons to be produced freely (Baca, A. M. and Hol, W. G. J. 2000; Ikemura, T. 1981). The above cell lines can also contain a PLysS plasmid. This plasmid allows the weak repression of the T7 promotor. The PLysS plasmid encodes for the T7 lysozyme which is capable of binding and inhibiting the T7 RNA polymerase (Moffatt, B. A. and Studier, F. W. 1987). This can allow tighter control of recombinant protein production and is employed primarily to prevent the leaky expression of toxic proteins.

Strain	Genotype	Features
BL21 (DE3)	F– *ompT hsd*SB(rB–, mB–) *gal dcm* (DE3)	Defective in Lon and OmpT protease
BL21 (DE3) Star	F– *ompT hsd*SB(rB–, mB–) *gal dcm rne*131 (DE3)	As BL21 (DE3) and defective in rne131 resulting in ineffective RNase E
Rossetta (DE3) PLysS	F^- ompT $hsdS_B(R_B^- m_B^-)$ gal dcm λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5]) pLysSRARE (Cam^R)	Additional tRNA for all codons including those rare in bacterial cell lines
PLysS plasmid	pLysS(cm^R)	Produces a weak repressor for the T7 promoter region preventing "leaky" expression

Table 2-1. **Table showing a selection of common bacterial expression cell lines** Table showing a small selection of the major cell lines used for recombinant protein expression. The genotypes and main features of these cell lines are represented. The bottom of these show the PLysS plasmid which can be used with most expression cell lines.

Vector	Antibiotic resistance	Promoter	Affinity tag
pT7-7	Ampicillin	T7	None
pQE	Ampicillin	T5	6×HIS
pGEX	Ampicillin	T7	GST
pMAL	Ampicillin	T7	MBD

Table 2-2. **Table showing a selection of common bacterial expression vectors** Table showing a small selection of some expression vectors used for recombinant protein expression. The resistance, promoter and affinity tags of these expression vectors cell lines are represented.

Media	Ingredients	Maximal Growth
Luria-Bertani (LB)	5 g yeast extract, 10 g tryptone and 10 g NaCl	~2Au
2×TY	10 g yeast extract, 16 g tryptone and 5 g NaCl	~3Au
Overnight Express	Proprietary information	Greater than photo spectrometer range however 6 times biomass

Table 2-3. **Table showing a selection of common bacterial expression media** Common growth media used with E.coli showing recipe and maximal growth

2.2.1.1 E.coli expression vectors used for protein expression

Recombinant proteins can be cloned onto different expression vectors in order to optimise expression or allow the addition of an affinity tag. One of the most common expression vectors are the pET vectors (Novagen). These vectors contain the lacI gene for production of the lac repressor, the lacO DNA for the lac repressor binding, binding T7 promoter for use with DE3 cell lines, and the ampicillin resistance gene. Other vectors in use include the pT7-7, pQE, pGEX and pMAL represented in table 2.2. pT7-7 expresses native proteins under control of the T7 promoter like that of the pET expression vectors. pQE, pGEX and pMAL express 6×HIS, Glutathione-s-transferase (GST) and maltose binding domain (MBD) tags respectively. Protein expression of pQE expression vectors is controlled by the T5 promoter, pGEX and pMAL expression vectors are controlled by the Ptac promoters. This allows the protein of interest different expression systems and different affinity tags which can facilitate increased expression and aid purification.

2.2.1.2 E.coli growth media used for protein expression

Expression of the protein of interest can be improved by the use of different growth media, these include, Luria-Bertani (LB), 2×tryptone and yeast extract (2×TY) and Overnight Express™ Autoinduction media (table 2.3). LB is the most common growth media used, providing a rich growth media to allow high concentration of cell densities (Anderson, E. H. 1946). 2×TY media is similar to LB media with double quantities of tryptone and yeast extract, which allows higher cell densities to be achieved. Autoinduction media does not require the introduction of IPTG for protein expression. This media also allows significantly increased cell densities, ~6fold that of LB (Grossman, T. H. *et al* 1998).

2.2.1.3 Standard methods of protein expression by E.coli

During protein expression in DE3 cell lines the culture has IPTG added commonly at an optical density at a wavelength of 600nm (OD_{600}) of 0.6AU. At this time, it can be advantageous to modify the temperature of the cell culture. IPTG is commonly used at a final concentration in the media of 1mM, optimal concentration for expression can be between 0.1 and 2mM. The maximum IPTG concentration for use is ~3mM as the number of RNA polymerase molecules in a cell is limited. The protein expression is commonly induced with IPTG when the cell suspension reaches an OD_{600} of 0.6AU, however can be between 0.4 and 1.0AU is suitable as cells at this OD are in the logarithmic growth phase and are most stable. Cells are grown at 37°C and recombinant protein can be expressed at this temperature, some protocols however use a lower temperature of protein expression. An expression temperature of 20-30°C results in slower growth and an increased yield of folded protein. This can be influenced by the production of cold shock chaperone proteins to improve the expression of appropriately folded proteins (Schein, C. H. 1989).

After protein expression by addition of IPTG, the cells can require differing lengths of time to reach optimal expression levels. During the initial analysis of protein expression, samples are removed at 1hour intervals and analysed by SDS-PAGE, in order to determine the level of expression. This can demonstrate the length of time required for adequate levels of protein expression.

2.2.2 Materials and methods for the expression of PCNA

HsPCNA pET-28a construct was transformed into BL21 (DE3) PLysS and plated out onto ampicillin, 100µg ml^{-1}, LB agar plates. A single colony was grown to an OD_{600} of 0.4AU in 50ml 2×TY media. 5ml of this culture was used to inoculate 1

litre of 2×TY, 100µg ml^{-1} ampicillin and 50µg ml^{-1} chloramphenicol, media pre-warmed to 37°C; this culture is grown to an OD_{600} of 0.6AU. 1ml of 1M IPTG (Sigma-Aldrich) was added to the culture and the cells were left to express for 3hours, at all times shaken to 180rpm. Cell lines were verified as expressing HIS-HsPCNA by western blotting (Towbin, H. *et al* 1979) using an anti-poly-HIS antibody conjugated to alkaline phosphatase. The presence of the alkaline phosphatase enzyme was determined by the NBT/BCIP reaction (Roche).

SpPCNA construct was transformed into BL21 (DE3) PLysS and plated out onto ampicillin, 100µg ml^{-1}, LB agar plates. A single colony was grown overnight in 50ml LB media. Overnight culture was used to inoculate 1 litre of LB, 100µg ml^{-1} ampicillin and 50µg ml^{-1} chloramphenicol, media to an OD_{600} of 0.1AU. The media is pre-warmed to 30°C; this culture is grown to an OD_{600} of 0.7AU. 1ml of 1M IPTG was added to the culture and the cells were left to express for 7hours, at all times shaken to 180rpm.

LmPCNA construct was transformed into BL21 (DE3) PLysS and plated out onto kanamycin, 50µg ml^{-1}, LB agar plates. A single colony was grown to an OD_{600} of 0.4AU in 50ml LB media. 5ml of this culture was used to inoculate 1 litre of LB, 50µg ml^{-1} kanamycin and 50µg ml^{-1} chloramphenicol, media pre-warmed to 37°C; this culture is grown to an OD_{600} of 0.6AU. 1ml of 1M IPTG (Sigma-Aldrich) was added to the culture and the cells were left to express overnight (12-14hours), at all times shaken to 180rpm.

Cells were harvested by centrifugation at 9000RPM (JLA9.1000), and pellets were stored as 1litre cultures in 50ml falcon tubes. These pellets were flash frozen by immersion in liquid nitrogen and stored at -20°C.

2.2.3 The expression of PCNA from three species in E.coli cell lines

HsPCNA was expressed in BL21 (DE3) PLysS cell lines, grown in 2×TY media, 100µg ml^{-1} ampicillin. Cells were grown at 37°C to an OD_{600} of 0.6AU, expressed with 1.0mM IPTG and for three hours to achieve optimal expression (figure 2.3).

SpPCNA was expressed in BL21 (DE3) PLysS cell lines, grown in LB, 100µg ml^{-1} ampicillin. Cells were grown at 30°C to an OD_{600} of 0.7AU, expressed with 1.0mM IPTG and for six hours to achieve optimal expression.

LmPCNA was expressed in BL21 (DE3) PLysS cell lines, grown in LB media, 100µg ml^{-1} kanamycin. Cells were grown at 37°C to an OD_{600} of 0.6AU. The temperature of the culture was then reduced from 37°C to 20°C, and expression was induced with 1.0mM IPTG and overnight to achieve optimal expression.

2.3 Protein purification from bacterial cell lines

Bacterial cells are lysed to release expressed proteins. The protein of interest is in a mixture of cell membranes, nuclei, other cellular organelles and cellular proteins. Larger impurities e.g. organelles can be removed by centrifugation, leaving a protein mixture. The protein of interest is then purified from other cellular proteins by various purification techniques, outlined below.

2.3.1 Principles of protein purification

Protein purification ideally starts with high expression levels to provide large quantities of protein. Each purification step results in a loss of the total protein yield. This loss is the result of protein denaturation and sample loss. At each stage of purification, the sample may be exchanged into a new buffer which may result in denaturation, proteins may be required to be concentrated which can result in protein

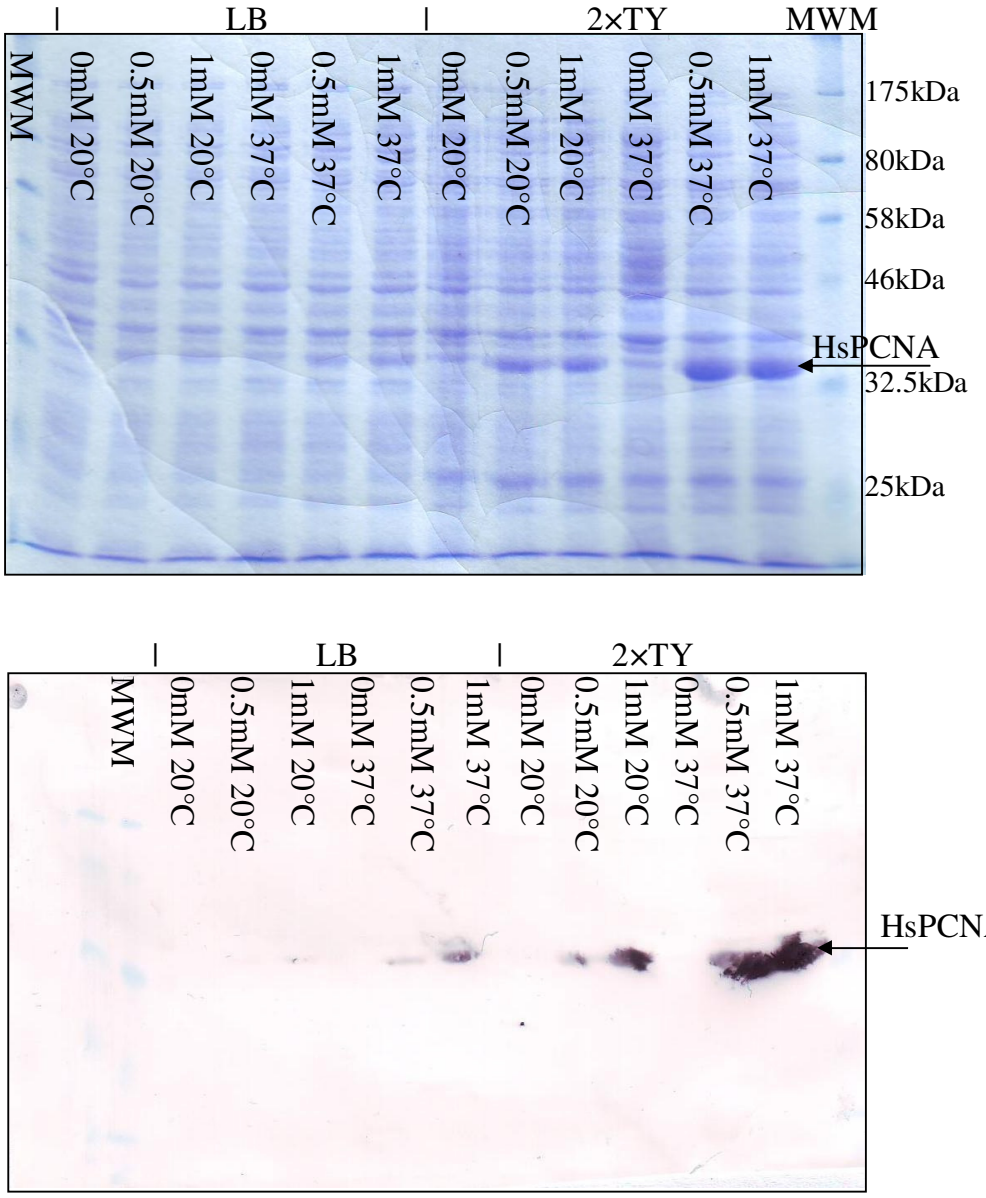


Figure 2-3. **The expression of HsPCNA**
Top) 12% gel showing the trial expression of HsPCNA-pET28a construct. These lanes show four trial expressions, in LB, 2×TY and at 20°C, 37°C. The lanes show the cell lines with 0mM, 0.5mM and 1mM IPTG. It is clearly shown that protein expressed at 37°C in 2×TY media with 0.5mM and 1mM IPTG show the greatest levels of expression. Visulalised by coomasie blue.
Bottom) Western blot of the gel (middle) with anti-poly-HIS antibody, demonstrating the increased expression is the HsPCNA-pET28a construct visualised as in 2.2.2.

associating to filter membranes, and during purification methods, sample may associate to chromatographic resins irreversibly.

The protein of interest can be isolated from cellular proteins by many techniques, usually a combination of affinity chromatography (Lee, W. C. and Lee, K. H. 2004), protein precipitation "salting out" (Coen, C. J. *et al* 1995), size exclusion chromatography (Kostanski, L. K. *et al* 2004), ion exchange chromatography (Peterson, E. A. and Sober, H. A. 1956) and hydrophobic interaction chromatography (Vailaya, A. and Horvath, C. 1996). A final quantity of protein required for useful crystallisation and other biophysical studies is a minimum of 1mg at >97% purity. Careful purification from bacterial expression lines using affinity chromatography can commonly produce protein yields of 5-20mg.

2.3.1.1 Methods of protein purification

The purification techniques each allow different methods for increased purity. These different techniques will result in separation of the recombinant protein from bacterial cellular proteins by association, precipitation or molecular sieving. Most techniques use direct association of the protein of interest to a resin e.g. affinity chromatography, ion exchange chromatography and hydrophobic interaction chromatography (HIC). An example of protein separation by precipitation is salting out, and an example for molecular sieving is size exclusion chromatography.

2.3.1.2 Affinity based purification techniques

Affinity chromatography is one of the most common methods of protein purification, this method uses a proteins natural binding affinity for a ligand or requires the addition of an affinity tag associated to the protein of interest e.g. the 6×HIS, GST or MBD tag. The affinity tag directly binds to a partner immobilised to

a solid resin and is held within the affinity column. Ni-NTA affinity columns bind specifically to the poly-HIS tag. The 6×HIS tag, used in this work, anneals to the resin via a Ni^{2+} ion immobilised to a Nitrilotriacetic acid (NTA) group (figure 2.4A). This Ni^{2+}-6×HIS interaction can be dissociated by exposure to imidazole, mimic of the histidine ring structure (Hochuli, E. *et al* 1987). GST affinity columns are an agarose resin with glutathione immobilised. Proteins conjugated to the GST tag will specifically bind the immobilised glutathione (figure 2.4B) and bacterial proteins will not associate via this functional group. The immobilised GST can then be released form the agarose resin by the addition of high concentrations of reduced glutathione (Smith, D. B. and Johnson, K. S. 1988). Recombinant proteins, conjugated to the maltose binding domain, are capable of binding maltose with high affinity. Resin with maltose immobilised can be used for affinity purification and subsequently eluted using maltose (Diguan, C. *et al* 1988).

2.3.1.3 Ion exchange purification techniques

Ion exchange chromatography is a method of isolation based upon the net charge of the protein. This method immobilises the protein to a charged resin by modification of the pH; this will then allow other proteins to be removed based upon differences in pI. Samples are dialysed in a pH buffer that will confer a charge on the protein near the pI of the protein. Once proteins are immobilised to the resin, the buffer pH can be changed on a gradient and as the pH passes the proteins pI, the protein can dissociate from the resin (Luo, R. G. and Hsu, J. T. 1997). This is an accurate method of protein purification and is most beneficial as a second step after bulk impurities are removed.

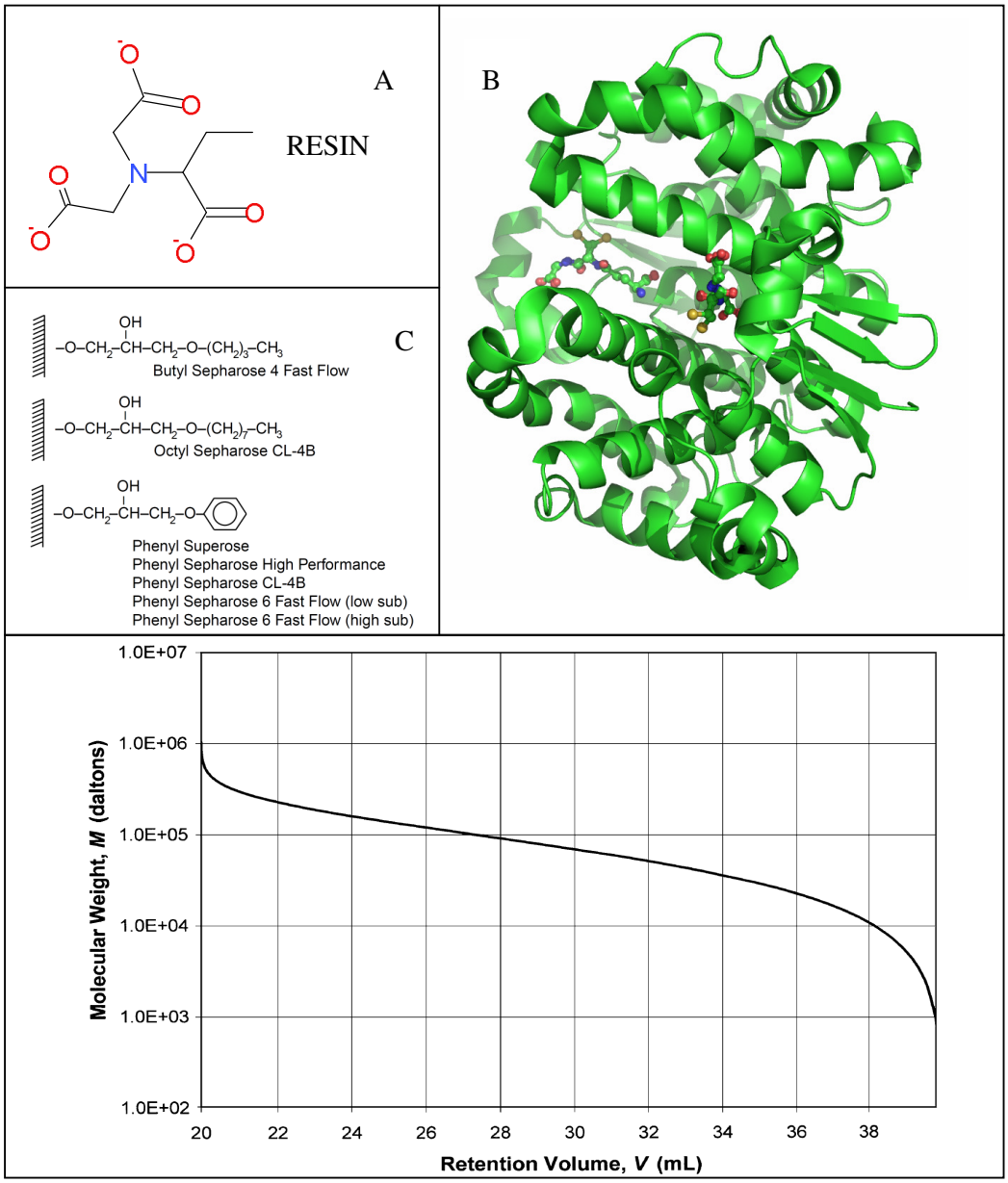


Figure 2-4. **Images demonstrating differing purification techniques, showing functional groups and elution characteristics**

A) The NTA functional group, demonstrating the position of the association to resin. The Ni^{2+} is associated to NTA via the negative charges on the carboxylic acid groups. Created using Symyx®Draw.

B) Crystal structure of the synthetic glutathione-S-transferase (GST) (1AGS) cartoon representation with glutathione represented in ball and stick form (Zeng, K. *et al* 1994).

C) Functional groups present on hydrophobic interaction columns. Shows linear groups (top, middle) and phenyl groups (bottom). These functional groups are attached to the resin via a glycidyl-ether linker, modified from Amersham catalog (15/1/1993).

D) The relationship of molecular weight to elution volume of a theoretical size exclusion column, with a void volume of 20ml and a porous accessible volume of 20ml. This demonstrates the linear portion of the size exclusion elution which can allow for the estimation of protein molecular weight, from (Kostanski, L. K. *et al* 2004).

2.3.1.4 Hydrophobic interaction chromatographic techniques

Hydrophobic interaction chromatography (HIC) isolates proteins based on the level of surface hydrophobicity. This method requires proteins to be dialysed into high salt concentrations to promote hydrophobic interactions; this can result in protein precipitation. The resin is commonly sepharose beads with functional groups coupled via an ether bond to a monosaccharide. There are a range of functional groups for HIC, comprising of long alkyl chains or an aryl ring (figure 2.4C). The selection between these groups is made based upon the structural features of the protein to be immobilised. Surface ring structures will associate to an aryl ring via π-electrons and long alkyl chains will associate to broad hydrophobic surfaces, like those found in many protein-protein binding interfaces. As the protein solution is passed over the resin in solutions with high ionic strength, ~4M NaCl, ~1M $(NH_4)_2SO_4$ (Ammonium Sulphate), proteins are likely to aggregate via hydrophobic surfaces. The hydrophobic groups allow the association of the protein to the resin and results in the sample to be held in the column. This sample can then be eluted by the use of buffers with lower ionic strengths, reducing the propensity of the sample to bind the resin and allowing the protein to pass through the column (Machold, C. *et al* 2002).

2.3.1.5 Salting out as a bulk purification technique

Protein "salting out" is a method of protein precipitation and uses ammonium sulphate, pH or thermal precipitation. These methods precipitate proteins in the cell lysate using specific concentrations of ammonium sulphate, changes in pH or temperature shifts. Should precipitation of the protein of interest occur at higher ammonium sulphate, pH or temperature, then all protein precipitated at that point can

be removed by centrifugation. This is a common first step method as it rarely has sufficient specificity to achieve high purity. It is also cheap and allows a simple method for rapid removal of bulk impurities (Grover, P. K. and Ryall, R. L. 2005).

2.3.1.6 Size exclusion chromatography purification techniques

Size exclusion chromatography (SEC) is a method of separating proteins based upon size (Porath, J. and Flodin, P. 1959). There are many types of SEC resin e.g. Superdex, Superose, Sephacryl, Sephadex, Toyopearl etc. all of which are inert in order to reduce non-specific interactions. The resins are made of a variety of materials and compositions. Superdex resins are a composite of agarose and dextran; superose is a highly cross-linked agarose; sephacryl is a composite of dextran and methylenebisacrylamide; sephadex is cross linked dextran and toyopearl resins are hydroxylated methacrylic polymers. These resins are produced as small spheres that act as porous beads and can have different ranges of pores sizes to allow proteins of differing sizes to pass through. Proteins larger than the resin pore sizes will be unable to enter resin (reducing the volume accessible) and will pass around the resin relatively quickly. Small proteins will be able to enter the resin (increasing the volume accessible) and take longer times for elution. Selection of appropriate resins is accomplished using the molecular weight of the proteins that are to be separated, the elution of proteins from the size exclusion column occurs at a linear portion of retention volumes. When appropriately selected, the protein of interest will elute in this section to achieve the greatest resolution (figure 2.4D).

Size exclusion chromatography also allows the approximation of a proteins molecular weight, based upon the elution volume (Dai, H. *et al* 1998). Molecular weights determined from this method are not based on mass but the hydrodynamic

radius (Stokes radius). The Stokes' equation relates the radius of a hard sphere (r_s) with the diffusion rate of the molecule in solutions of a specific viscosity at a specific temperature. The Stokes radius is given by equation 2.1

Equation 2.1) $$r_s = \frac{k_B T}{6\pi\eta D}$$

where r_s is the Stokes radius in meters, k_B is the Boltzmann constant (1.38×10^{-23} J K^{-1}), T is the temperature in kelvin, η is the viscosity of the solution in Pascal seconds and D is the diffusion constant in meters per second. This Stokes radius will be affected by the shape and solvation of the protein, so the Stokes radius is not necessarily an accurate representation of molecular weight. The elution volume of a protein can be affected by the shape and solvation but also the oligomeric state of the protein in solution. As demonstrated by equation 2.1, the value of the Stokes radius and elution volume can be affected by the temperature and viscosity of the buffer. Therefore all elution volumes determined in SEC are condition-dependant. The mass of globular proteins can be estimated by measuring elution rates of protein standards (Potzschke, H. *et al* 1996). Protein diffusion coefficients have been linked to molecular weight using equation 2.2

Equation 2.2) $$D = 8.34\left(\frac{T}{\eta MW^{1/3}}\right)$$

where MW is the protein molecular weight from (Young, M. E. *et al* 1980). This is still an estimate for a globular protein (figure 2.4D).

This method is not commonly used as a method of bulk purification as separation of proteins is only effective with an order of magnitude difference in size. This method can however determine the purity and homogeneity of the sample, and

determine if the protein is folded. These factors result in this method being employed more as an analytical technique.

2.3.2 Materials and methods of Purification of three PCNAs

Cells were resuspended in 30ml, ~10% (w/v), of Ni-NTA binding buffer (50mM NaH_2PO_4, 300mM NaCl, 10mM Imidazole, pH8.0) in which one mini EDTA-free protease inhibitor tablet (Roche) is dissolved. Cells are lysed by sonication with the Soniprep150 (MSE) at full power for 6x20seconds at 10second intervals. Cellular debris are removed by centrifugation at 23000rpm (Beckman 25.50) for 30minutes at 4°C, then passed through a 0.2µm filter. Every step of the lysis of the cell cultures is carried out on ice.

Cleared lysate was loaded on a Ni-NTA affinity column, pre-equilibrated with Ni-NTA binding buffer, and extensively washed to baseline with binding buffer, ~8 column volumes. This is run at a flow-rate of 2ml min^{-1} and a maximum pressure 0.3MPa. A 5 column volume step elution of 26mM Imidazole elutes bulk impurities. A gradient elution of 26-250mM Imidazole over 150mls elutes the protein of interest. Protein concentration was determined by the A_{280} assuming an extinction coefficient of 16305 M^{-1} cm^{-1}, 9190 M^{-1} cm^{-1} and 20775 M^{-1} cm^{-1} for HsPCNA, SpPCNA and LmPCNA respectively as determined by ProtPram (Gasteiger, E. *et al* 2003). PCNA was concentrated to no more than 8 mg ml^{-1} and passed over the Superdex 200 column 10x300 in HEPES biochemistry buffer (10mM HEPES, 150mM NaCl, pH 7.4). LmPCNA prepared for crystallography is run in LmPCNA crystallisation buffer (25mM Tris, 25mM NaCl, 2mM DTT, pH7.5). Final purified protein samples were stored at 4°C for no more than 1 week prior to use.

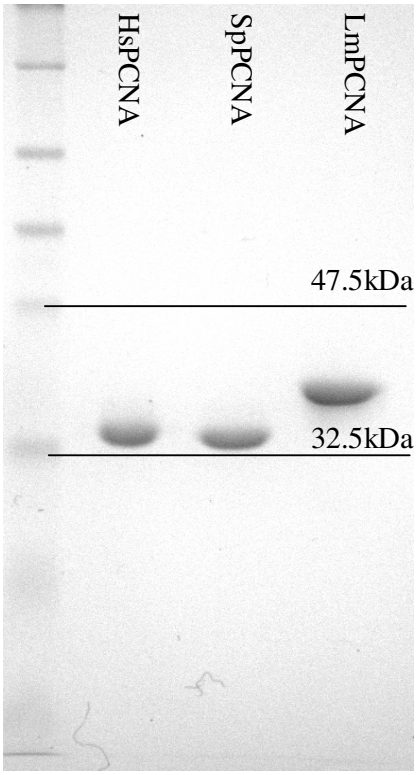
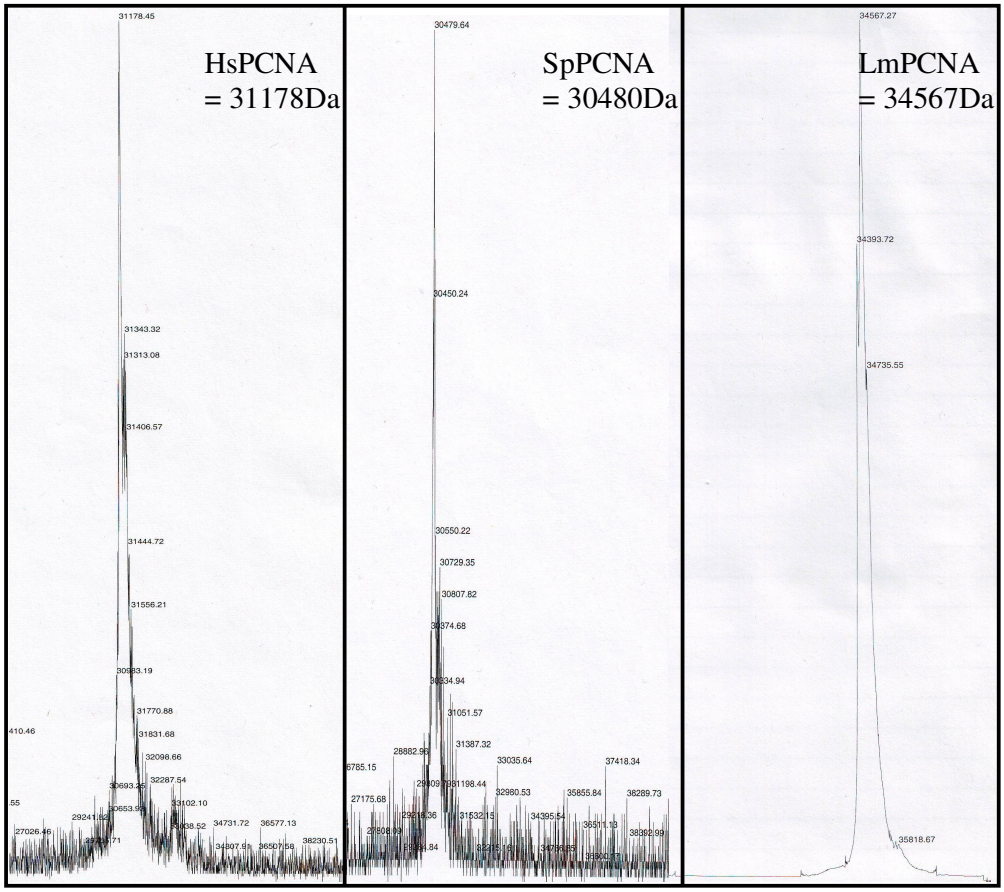


Figure 2-5. **Analysis of the purified PCNA showing a greater than 97% purity and bands corresponding to appropriate sizes of recombinant PCNA**

Top) 12% SDS-PAGE Coomassie stained gel of 5µg each of HsPCNA, SpPCNA and LmPCNA. This gel demonstrates the high level of purity achieved for each protein. The protein bands run to a slightly greater molecular weight than expected.

Bottom) The full length mass spec analysis of HsPCNA, SpPCNA and LmPCNA from left to right respectively. The determined values of molecular weight gave good correlation to the predicted masses; 31.3kDa, 30.5kDa and 34.4kDa for HsPCNA, SpPCNA and LmPCNA respectively.



2.4 Conclusions of the purification of three homologous PCNAs

The *H. sapiens* PCNA has been successfully cloned into pET-28a and PCNA from all three sources have allowed expression to a high level. This has allowed all three PCNAs to be purified to greater than 97% purity in large enough quantities to allow further structural and biophysical characterisation.

CHAPTER 3. The X-ray structural analysis of LmPCNA

The three dimensional structure of a protein can be solved by only a few techniques e.g. X-ray crystallography, nuclear magnetic resonance (NMR), cryo-electron microscopy (Cryo-EM) (Sali, A. *et al* 2003). Computational approaches can also be used and a variety of programs exist which can predict the three dimensional structure of a protein e.g. homology modelling by 3DJigsaw (Bates, P. A. *et al* 2001).

The different methods of structure solution provide differing levels of information and each have different associated errors. X-ray crystallography can provide the best resolution (~1Å), however may not be an accurate representation of the protein in solution as the sample is restrained within the crystal lattice (Smyth, M. and Martin, J. 2000). NMR commonly provides a lower resolution (~3Å) and gives the average position of protein atoms relative to other protein atoms in solution. The size of the macromolecule to be solved is limited using NMR and also isotopic tags may be required (Sali, A. *et al* 2003). Cryo-EM is a technique which has become, in recent years, more widespread for protein structure determination; this is however usually at a lower resolution (~8Å) (Sali, A. *et al* 2003). In this work only X-ray crystallography has been implemented and will be discussed in detail.

An overview of the steps involved solving protein structures using X-ray crystallography is given in figure 3.1. X-ray crystallography requires the protein of interest to be crystallised and good quality crystals are essential for high resolution diffraction (Wiencek, J. M. 1999). These crystals are commonly cryo-protected and frozen in order to increase crystal stability during data collection (Weik, M. *et al* 2001). Frozen crystals are placed in a continuous nitrogen stream and maintained at

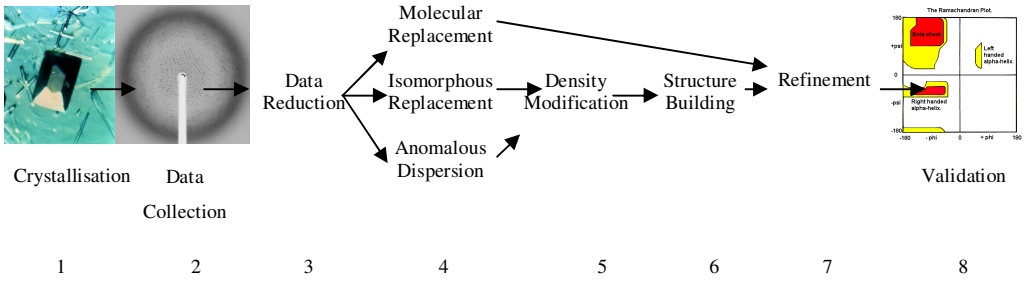


Figure 3-1. **An overview of the steps involved in the solution of protein structures**
Figure showing an overview of the general steps in X-ray crystallography. 1) the crystallisation of the protein, 2) the collection of diffraction patterns, 3) reduction of the diffraction data, 4) phasing of experimental data, 5) density modification using known high resolution electron density to improve phases, 6) the building of protein structure based on electron density, 7) refinement of the built protein model with electron density, 8) validation of the determined protein structure based on fits to the diffraction data and bond angles and lengths.

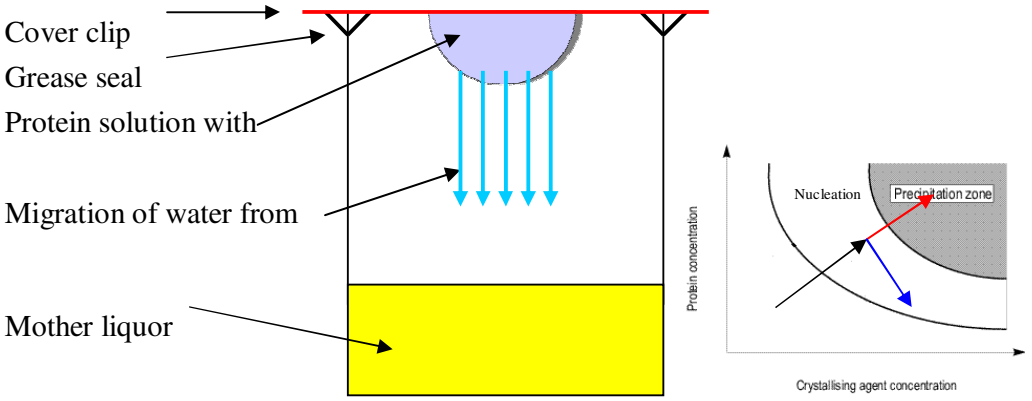


Figure 3-2. **The crystallisation of proteins by hanging drop vapour diffusion**
Figure showing the set up for crystal growth by the hanging drop method. This shows a protein solution with mother liquor in an air tight system. Grease seals at the cover slip prevent the system from completely drying, and water can migrate from the protein solution to the mother liquor, slowly increasing the protein and precipitant concentrations. The protein and precipitant concentrations are shown on the right where protein drops precipitate (red) and crystallise (blue)

100K in the beam path. In order to obtain a diffraction pattern, the X-rays used are of high energy (~1Å wavelength, commonly 0.7Å-1.7Å). This diffraction pattern is measured on a charge coupled device (CCD) detector. Only intensities and positions of Bragg reflections are captured and measured using the collection device and no "phase" information is available. The unit cell (the smallest unit by rotation and translation) dimensions of the crystal can be determined from the spacing of the Bragg reflections. Phasing can be accomplished by different experimental methods e.g. anomalous dispersion, isomorphous replacement or molecular replacement. Calculated electron density maps are used to fit the protein sequence. The final steps involve refinement and validation of the structure, as discussed below.

3.1 Protein crystallisation for X-ray crystallography

The crystallisation of a protein arranges the molecules into regular arrays. Protein is commonly crystallised using the slow dehydration of a high concentration protein sample (Wiencek, J. M. 1999). The protein is slowly dehydrated by the migration of water from the protein solution. The dehydration can be achieved by sitting or hanging drop methods (Mcpherson, A. 1985) in an air tight systems in which protein solutions are exposed to a dehydrating environment. The mother liquor is placed in a reservoir within this enclosed environment. The mother liquor is a solution containing required levels of precipitant, pH and additives. Water from the protein solution migrates to the mother liquor (figure 3.2A). Dehydration of the protein solution slowly increases its concentration. The increase in protein concentration increases the propensity for protein precipitation. Within precipitant e.g. PEG or ammonium sulphate, at a set pH, a protein will have a specific concentration for precipitation or crystallisation (figure 3.2B).

Forty-nine structures of PCNA from different species have been solved and deposited in the PDB. PCNA crystals have formed in a wide variety of conditions. Twenty crystals were formed in Poly-Ethylene-Glycol (PEG) of average molecular weights of 400Da (27.5-30%) through to 20000Da (7.5%). Nineteen PCNA crystals formed in ammonium sulphate (1.6-2.5M). Other precipitants have been used to form crystals, such as Methyl-2,4-Propane-di-ol (MPD) (25-30%), Na/K phosphate (1.4-1.6 M, pH 9.0). The crystals for all solved structures were formed between pH 4.5 and 9.0, commonly at pH 6.0. From these determined crystallisation conditions, it is inferred that LmPCNA can crystallise in a range of conditions.

3.1.1 Results of the crystallisation of LmPCNA

LmPCNA was crystallised by the hanging drop method. LmPCNA was concentrated to 10mg ml^{-1}. A precipitant solution of 10%w/v PEG 6000, 0.1M MES pH6.0 was used as a mother liquor. Hanging drops were prepared in a 2:2:0.5 ratio of LmPCNA, mother liquor and 0.1M tri-methyl-amine hydrochloride and crystals grew in 3-4 days; these crystals diffracted to 8Å resolution. Conditions were refined to 12%w/v PEG 400, 0.1M MES, 0.1M tri-methyl-amine, pH6.5 at 4°C, crystals diffracted to ~7Å. LmPCNA crystals were ~0.5×0.5×0.2 mm in size (figure 3.3A).

LmPCNA co-crystallised using protein at 10mg ml^{-1} with a final concentration of 10µM peptide 1 (*H. sapiens* p21 141-152) present in the protein solution. The LmPCNA peptide solution was incubated for 20minutes prior to mixing with mother liquor. Crystallisation drops were set up as above using the same mother liquor and additive in the same ratios. Crystals were of similar range of sizes and diffracted to a similar resolution.

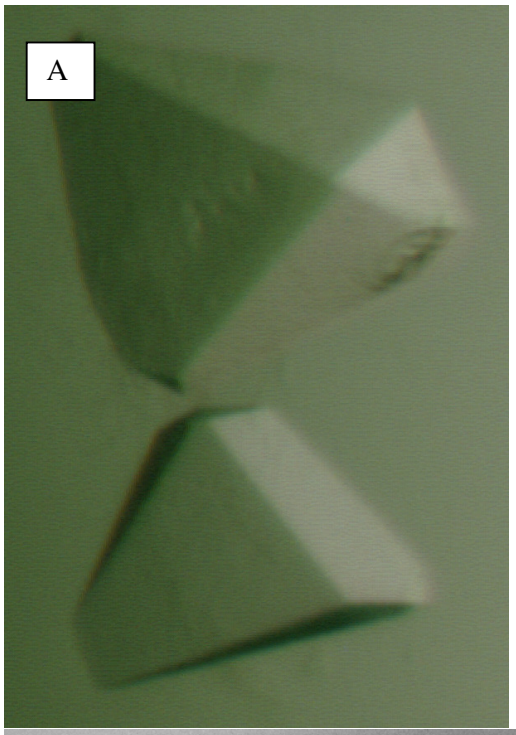
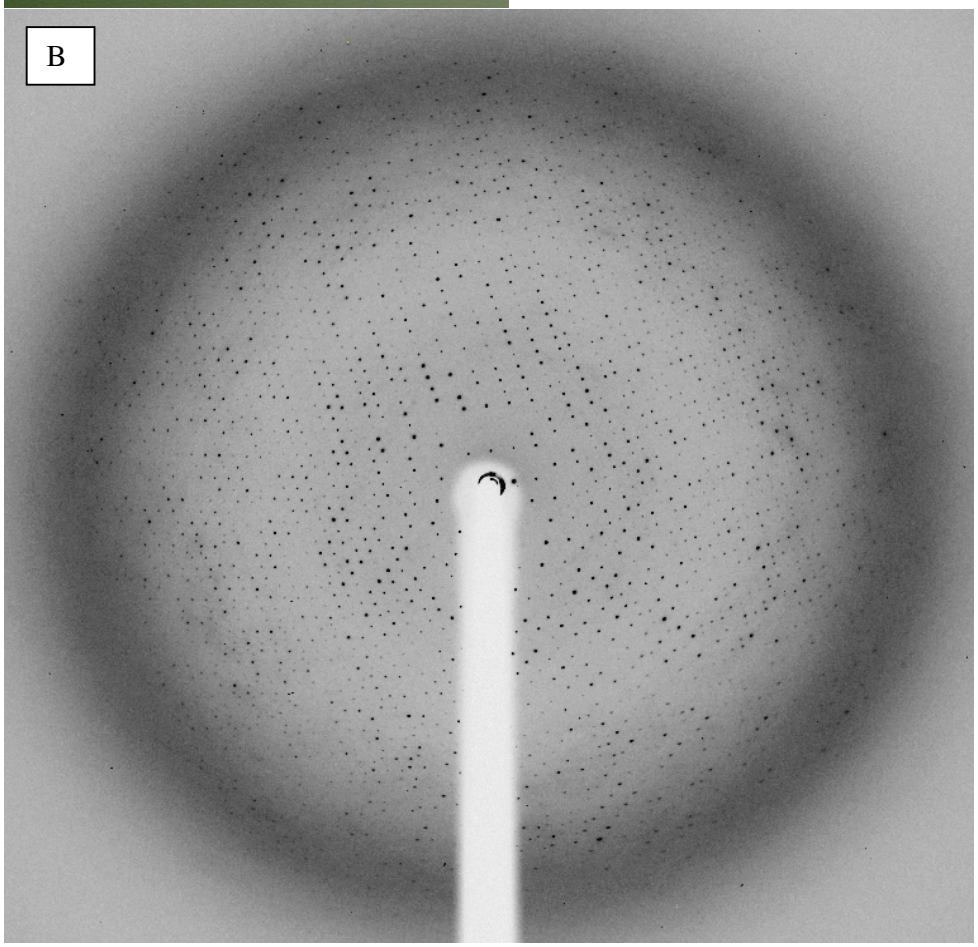


Figure 3-3. **The crystal and diffraction pattern of LmPCNA**

A) Two crystals of LmPCNA in a hanging drop. These crystals grew in 15%w/v PEG400, 0.1M MES, 0.1M tri-methyl-amine, pH6.5; and required 3-4 days to form. The top crystal is approximately 0.5×0.5×0.2 mm in size and not optically birefringent.

B) The diffraction pattern of the LmPCNA crystals presented in A after the crystal dehydration. The crystal diffracted to 3Å and shows clearly defined diffraction spots. This diffraction pattern shows no strong ice rings and the frequency of the diffraction spots infers a large space group.



3.1.2 Cryo-protection and dehydration of protein crystals in crystallography

Crystals are used in high intensity X-ray beams in order to obtain a diffraction pattern. Exposure to high energy X-rays can cause the release of free radicals resulting in damage to the crystal. Damage to a protein crystal reduces the observed diffraction resolution. Damage to protein crystals by free radicals can be reduced by maintaining the crystal below 155K and commonly at 100K during the data collection process (Weik, M. *et al* 2001). In order to collect data at 100K, the sample needs to be flash frozen in liquid nitrogen after being removed from the crystallisation drop.

Protein crystals contain a high percentage of solvent (~30-65%) (Matthews, B. W. 1968). Crystals with a high solvent content may form ice crystals when frozen. Ice formed in the protein crystal results in intense ice rings on the diffraction pattern which may obscure data from protein diffraction. The formation of ice within a protein crystal may reduce resolution by damaging the crystal lattice and increasing mosaicity.

The formation of ice crystals can be prevented by cryo-protection. Cryo-protection is the process of the removal of waters by either dehydration or replacement using a variety of different chemicals e.g. glycerol, PEG, MPD etc. The method of cryo-protection can damage crystals. A common method of dehydration is the slow removal of water from the whole drop and is accomplished by vapour diffusion. Replacement of water with other molecules is achieved by a series of small injections of cryo-protectant directly into the drop, slowly increasing the cryo-protectant concentration (Heras, B. and Martin, J. L. 2005).

Dehydration of protein crystals can be used as a means to improve resolution (Heras, Bego a *et al* 2003). Protein crystal dehydration can be achieved by different means e.g. vapour diffusion over dehydration solution and direct addition of dehydration solution; both of which are described above as a means of cryo-protection. In one example, direct contact to air has been used to almost completely dehydrate a crystal for 28 months and then was re-hydrated over crystallisation solution (Izasrd, Tina *et al* 1997). The Free Mounting System (FMS) is an instrument designed to control the dehydration of protein crystals whilst mounted on a goniometer system head. The FMS monitors the level of humidity at a standard cryo-loop. The level of humidity can be altered and changes the solvent content of the crystal. Crystals can then be frozen in-situ at the optimum relative humidity. This dehydrated crystal can diffract at a higher resolution (Chotiyarnwong, P. *et al* 2007).

3.1.3 Results of the cryo-protection of LmPCNA crystals for crystallography

Crystals of LmPCNA were crystallised in 12% PEG 400 and dehydrated over 30% PEG 400, 10%glycerol by vapour diffusion. LmPCNA crystals would not grow at 30% PEG 400. Crystals were not optically birefringent prior to dehydration. Dehydration treatment of crystals took ~12hours. Once dehydrated, these crystals could then rotate plane polarised light, indicative of the crystal becoming more ordered. The dehydration of LmPCNA crystals increased the resolution from 7Å to 3Å. The increase in resolution may be explained by the increased regularity due to decreased solvent content of the crystal. LmPCNA co-crystallised with *H. sapiens* p21 (141-152) responded in the same way to dehydration as the un-complexed protein and resolution improved from ~8Å to ~3Å.

3.1.4 The X-ray sources for X-ray crystallography

The wavelength of X-rays used in most protein crystal studies is between 0.7 and 2.0Å, commonly at ~1.0Å. X-ray beam can be generated by different sources: X-ray tubes, rotating anode sources and particle storage rings. A new generation of sealed X-ray tubes are commonly available and can be used for data collection. Rotating anode sources are the common "in-house" source and are ten times more powerful than most X-ray tubes. Modern rotating anode sources can acquire a complete data set in ~6hours (Delcueto, J. A. and Shevchik, N. J. 1978). The most powerful and versatile X-ray generators are the synchrotron radiation sources. Synchrotron energy allows complete X-ray data sets to be collected in minutes and usually at a greater resolution (Sorensen, T. L. M. *et al* 2006). Data sets can be obtained from small crystals using synchrotron radiation (Sorensen, T. L. M. *et al* 2006). The intensity at undulator beams in third generation synchrotron sources is sufficient for data collection from crystals of ~500μm^3 size (Riekel, C. *et al* 2005).

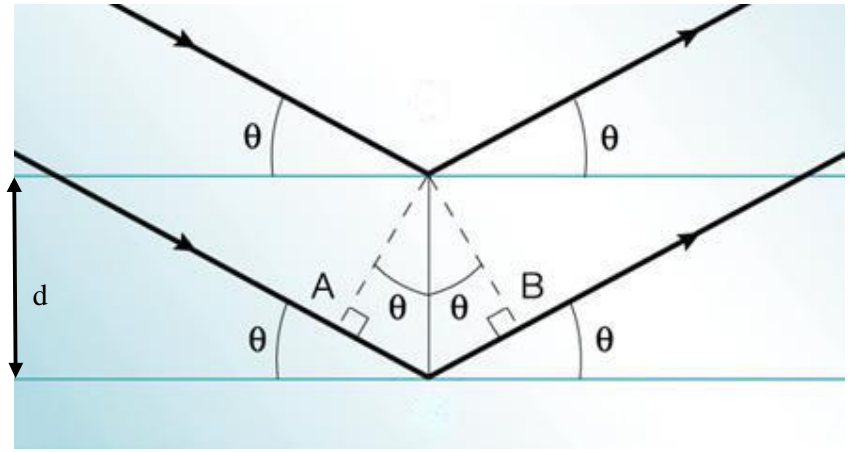
3.1.5 Results from the X-ray data collection of LmPCNA crystals

Crystals of LmPCNA were studied by X-ray diffraction at the Daresbury synchrotron radiation source station 10.1. LmPCNA crystals were studied using synchrotron X-ray radiation at a wavelength of 0.97Å. Data were recorded using a CCD and diffraction images were taken for 0.5° over 180°. The samples were maintained at 100K in a continuous cryo-stream and diffracted to a resolution of 3Å. No strongly diffracting ice rings were apparent on the diffraction pattern (figure 3.3B).

3.1.6 The diffraction of X-rays for X-ray crystallography

High energy photons are scattered by the electron clouds of atoms within the protein crystal. The crystal is comprised of unit cells (the smallest repeating unit that can make up a crystal by translation). Within the unit cells is the asymmetric unit, the smallest repeating unit that can make up a crystal by rotation and translation. Each spot on the diffraction pattern can be regarded as the result of a reflection from a set of parallel planes within the crystal. These planes slice through each unit cell within the crystal, and the measured Bragg reflection provides an average signal from the whole crystal. If the reflected beams from the planes are in phase, the signal undergoes constructive interference and creates a diffraction spot, a Bragg reflection (figure 3.4). The position of observed reflections is a function of the unit cell dimensions. The reciprocal relationship between d (the spacing between the Bragg planes) and sin θ (figure 3.4 Bragg's law) means that the larger the d-spacing, the smaller the Bragg angle (θ). Therefore, larger unit cells show a more crowded diffraction pattern; the Bragg reflections are closer together. The reflection intensities and phase of each spot contain information of the content of these unit cells.

To collect a complete data set, the diffraction pattern for the crystal at a range of angles is required. The diffraction images are scaled to regularise the intensity of different frames and merged to give the position of repeated reflections and overall intensities at each spot (Powell, H. R. 1999). The diffraction image at different angles through the crystal can give the space group, unit cell dimensions and approximate solvent content. The space group and unit cell dimensions can be determined as discussed above. The solvent content can be approximated by the Matthews equation determined in 1968 (equation 3.1) (Matthews, B. W. 1968).



Bragg's Law) $n\lambda = 2d \sin\theta$

Figure 3-4. **An example of a diffraction grating demonstrating Bragg's law**
Figure demonstrating the principals of Bragg's law for the formation of diffraction spots and the equation for Bragg's law. The diffraction grating, shown as two lines with a separation of distance d. The incident X-ray beams with angle of incidence and reflection, θ. The additional path length of the lower reflection can be determined by trigonometry if d, θ and λ are known (right angled triangles represented at positions A and B). If this path length is a whole number, n; then total constructive interference would occur.
Modified from (Datta, S. and Grant, D. J. W. 2004)

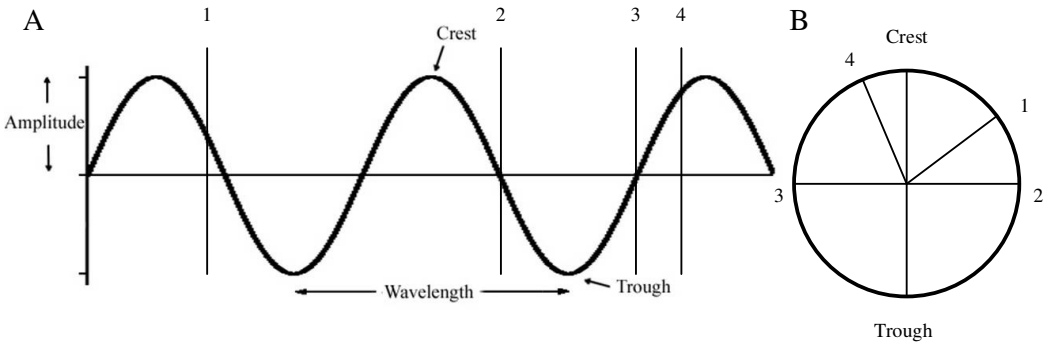


Figure 3-5. **A sinusoidal wave representation of light to demonstrate phase**
Figure showing a representation of light as a wave. A) This figure shows the amplitude, wavelength (λ), crest and trough of the wave form. The phase of light is a representation of the position of the photon on this wave at any point in time. This is commonly represented as a circle, B) presenting the crest at 0° and the trough at 180°. The numbers on the vertical line shown in the wave path correspond to numbers of the positions on the circle diagram. This position in the wave path is known as the phase.
Modified from http://www.physics.ubc.ca/

Eq 3.1) $$V_M = \frac{Volume_{uc}}{Mass_{uc}}$$

where V_M is the Matthews co-efficient, $Volume_{uc}$ is the volume of the unit cell and $Mass_{uc}$ is the mass of the unit cell.

The study by Matthews was based on the evaluation of the crystal structures of 116 proteins (1968) and 226 proteins (1976) (Matthews, B. W. 1976). Based on the size of the unit cell and the molecular weight of the monomer, a series of potential monomer numbers are related to solvent content. The solvent content is related to the number of monomers within the unit cell by the Matthews co-efficient (equation 3.1). The Matthews co-efficient is the crystal volume per unit of protein molecular weight (Matthews, B. W. 1968). The Matthews equation was reassessed in 2003 with a larger data set and was determined to be accurate (Kantardjieff, K. A. and Rupp, B. 2003). From this 2003 survey, it was possible to show a strong correlation between solvent content and the resolution of the X-ray data set. The higher the solvent content, the lower the resolution (the molecular weight of the protein per se does not necessarily reduce the resolution) (Kantardjieff, K. A. and Rupp, B. 2003).

The completeness is the ratio of the unique reflections measured as a number of those theoretically measurable for a given resolution. If all data is present for this resolution, then the data is 100% complete. The multiplicity is sometimes referred to as redundancy. The multiplicity is the number of times each unique reflection is independently measured. R_{merge} (shown in equation 3.2) demonstrates the similarity of the intensity of identical reflections measured independently. This gives a value of reliability for the diffraction data at that resolution. Higher resolution shells are commonly more error prone as the data are generally weaker. Incompleteness or

crystal damage may contribute to poor R_{merge} statistics. This evaluation method may not be a genuine evaluation as to the quality of the data as this statistical analysis is dependent on the multiplicity of the data set (Weiss, M. S. and Hilgenfeld, R. 1997).

Eq 3.2) $$R_{merge}(I) = \frac{\sum_{hkl}\sum_{i}\left|I_{hkl,i} - \langle I_{hkl}\rangle\right|}{\sum_{hkl}\sum_{i} I_{hkl,i}}$$

where $R_{merge}(I)$ is the R_{merge} for a specific reflection, $\langle I_{hkl}\rangle$ is the average intensity of symmetry related reflections from (Weiss, M. S. and Hilgenfeld, R. 1997).

The $\langle I\rangle / \langle \sigma I\rangle$ is a relationship between signal and noise of different reflections. $\langle I\rangle$ is the mean reflection intensity and $\langle \sigma I\rangle$ is the mean standard deviation of intensity. Low $\langle I\rangle / \langle \sigma I\rangle$ infers poor quality data as data spots are similar to background (noise) levels.

3.1.7 Results of the space group determination of LmPCNA

Diffraction data from both un-complexed and complexed LmPCNA crystals were processed using MOSFLM (Leslie, A. G. W. 1999) in the $P2_12_12_1$ space group. The frames were scaled using SCALA (Collaborative Computational Project, Number 1994).

The un-complexed LmPCNA crystal gave 492387 observed reflections of which 44892 are unique. The unit cell dimensions have been determined to be: a = 130Å, b = 149Å, c = 169Å; α = β = γ = 90.0°. Using the Matthews equation the un-complexed LmPCNA crystal was determined to have a Matthews co-efficient of 5.05 (Matthews, B. W. 1968). Assuming a protein molecular weight of 34.4kDa, the

Matthews co-efficient suggests there are six monomers in the asymmetric unit with a solvent content of ~70%.

The LmPCNA complexed with *H. sapiens* p21 (141-152) crystal gave 263630 observed reflections of which 69468 are unique. The unit cell dimensions have been determined to be: a = 132Å, b = 151Å, c = 173Å; $\alpha = \beta = \gamma = 90.0°$. Using the Matthews equation, the un-complexed LmPCNA asymmetric was determined to have a Matthews co-efficient of 5.18 (Matthews, B. W. 1968). Assuming a protein molecular weight of 35.9kDa, the Matthews co-efficient suggests there are six monomer in the asymmetric unit with a solvent content of ~69% (the calculated molecular weight included that of the protein and peptide combined). The LmPCNA crystals contain a high solvent content, which can result in poor resolution. The high solvent content and the increase in resolution suggest that the dehydration of the crystal was an important step to obtain higher resolution data.

The diffraction data gave a completeness of 100% for the un-complexed LmPCNA and 88.4% for the complexed LmPCNA structures. This reduction in completeness is linked to a decrease in diffraction strength of the complexed crystal. The high solvent content of the crystal may cause an increase in mosaicity and be linked to the poor resolution, which in the case of the complexed LmPCNA structure gave a reduced completeness when solved at 3Å.

The R_{merge} in this case is 9.9% and 23% for un-complexed and complexed LmPCNA respectively. The increase in R_{merge} was associated with a decrease in multiplicity of the diffraction data (7.4-3.8), suggesting that this increase in R_{merge} is an unbiased statistic and not the result of increasing numbers of reflections. The increase in R_{merge} of the complexed protein is most likely a result of much weaker

diffraction. The $\langle I \rangle / \langle \sigma I \rangle$ of the data sets were 17.4 and 6.0 for un-complexed and complexed LmPCNA respectively.

3.1.8 Solving the phase problem of X-ray crystallography

The diffraction data from an X-ray crystallography experiment is recorded as a flat image. The diffraction data is a measurement of signal intensity and spot position and gives no information regarding the phase of the diffracted ray (figure 3.5). The loss of this phase information is known as the phase problem.

The phase problem can only be solved experimentally by interference experiments. Interference experiments commonly used for protein crystallography are isomorphous replacement (IR) and anomalous dispersion (AD). IR is based on a change in spot intensity between un-complexed crystals and those soaked with heavy atoms. The change in intensity with the addition of heavy atoms can allow the construction of a Harker diagram which can determine the phases of the diffraction spots (Harker, D. 1956). Scattering of photons becomes anomalous if the electrons are tightly held in a lower electron shell. AD requires the absorption component (Cromer, D. T. and Liberman, D. 1970) and the dispersive component (Kronig, R. D. L. 1926) to be known. The scatter of anomalous dispersion and non-anomalous dispersion are affected differently by the addition of heavy atoms. Triple beam experiments are a method of partial phase determination and used to support IR and AD. This method is based on a single wave being reflected, which suffers interference from a second wave which has been subject of two reflections (Weckert, E. and Hummer, K. 1990). Phase determination by triple beam experiments are novel to IR and AD. Triple beam experiments can be very accurate, however they are time

consuming and return a small number of reflections phases'. This is more commonly used to support IR and AD (Soares, A. S. *et al* 2003).

Molecular replacement is a method of phase determination by purely computational means. The practical application of molecular replacement was outlined as early as 1962 (Rossmann, M. G. and Blow, D. M. 1962). A pdb file of similar three-dimensional structure is required.

The Patterson map contains information about intramolecular interactions and intermolecular interactions. The intermolecular interactions can be the result of the repeating unit cell, a repeated asymmetric unit, non-crystallographic symmetry with another identical protein or a non-identical protein (Grosse-Kunstleve, R. W. and Adams, P. D. 2001). The technique involves rotating the Patterson map of the model structure against the Patterson map from the X-ray data, which can be calculated through reflection intensities alone (i.e. no phase information is required). The Patterson map calculated from the model can be rotated round orthogonal (or Eulerian) angles to exhaustively cover all possible orientations (Rossmann, M. G. and Blow, D. M. 1962). At each orientation, the overlap of the Patterson peaks is calculated. The best solution will have maximum overlap of the Patterson peaks between the model and the measured data.

Once a good orientation is found, a translational search can be carried out (Rossmann, M. G. and Blow, D. M. 1962). The correct translational solution will give a consistent set of intermolecular Patterson vectors which should match peaks in the Patterson calculated from the measured X-ray data.

3.1.9 Results of molecular replacement for LmPCNA

The molecular replacement of LmPCNA was first performed using the structure of HsPCNA, 1AXC (Gulbis, J. M. *et al* 1996). The use of this HsPCNA structure resulted in a poor fit to the data and failed to return a valid structure. Other structures of PCNA were attempted in both monomeric and trimeric forms and all searches were unsuccessful. A homology model of LmPCNA, based on the 1AXC structure (Gulbis, J. M. *et al* 1996), was generated by the use of the program 3D-jigsaw and included the elongated loops (Bates, P. A. *et al* 2001). This homology model was used to create a trimeric ring structure (using the HsPCNA trimer as a template) and this trimer was used for molecular replacement. The program MOLREP (Vagin, A. and Teplyakov, A. 1997) was used to determine the position of the model by molecular replacement. Two trimeric rings were determined to be present in the asymmetric unit. The determined model underwent twenty rounds of refinement by refmac (Murshudov, G. N. *et al* 1997). The returned structure was examined to ensure a good fit to the determined electron density.

The structure of LmPCNA complexed with *H. sapiens* p21 peptide (141-152) was also solved by molecular replacement. The search model used for the complexed protein was that of the un-complexed LmPCNA structure with a *H. sapiens* p21 peptide (143-152) peptide modelled into the PIP box binding interface. The model was generated by the superposition of the 1AXC structure (Gulbis, J. M. *et al* 1996) over the un-complexed LmPCNA.

Among the top twelve scoring rotation function solutions from MOLREP theta and phi angles were related by 3 fold rotation axes with little deviation around chi (data not shown). All the highest scoring molecular replacement results had very

similar translational values with the twelve different rotation function solutions. The rotation function solutions may be a representation of three fold symmetry of the PCNA trimer.

The packing arrangement determined by MOLREP program showed two trimeric rings in the asymmetric unit packed face to face with the elongated loops extending out into the solvent (figure 3.6).

Trimer	θ	φ	χ	Frac. x	Frac. y	Frac. z	R_{factor}	Score
1	180.00	0.00	0.50	0.448	0.073	-0.034	0.539	0.602
2	136.33	-85.49	177.33	0.297	0.209	-0.222	0.476	0.708

Best fit of trimers in the complexed PCNA crystal

Trimer	θ	φ	χ	Frac. x	Frac. y	Frac. z	R_{factor}	Score
1	90.00	-179.93	180.00	-0.067	-0.083	0.045	0.525	0.486
2	98.34	-91.45	99.31	-0.187	-0.197	0.212	0.435	0.656

3.1.10 Structure refinement and model building from X-ray data

The refinement of crystal structures commonly have methods of verification to determine the reliability of the solved structure; the main values for validation are the R_{factor}, R_{free}, R_{merge}, R_{symm} and the Ramachandran plot.

The R_{factor} (equation 3.3) is a determination of the reliability of the structure and how well the determined structure fits the data used to generate it.

Eq. 3.3) $$R_{factor} = \frac{\sum_{hkl} \left\| F_{obs}(h,k,l) \right| - \left| F_{calc}(h,k,l) \right\|}{\sum_{hkl} \left| F_{obs}(h,k,l) \right|}$$

where F_{obs} is the intensity of the measured reflection and F_{calc} is the intensity at the same reflection calculated from the structure.

R_{free} is the reliability of the structure, similar to the value determined to the R_{factor}. Data is removed from the structure determination process. Data which is not incorporated into the solution can be used as a comparison to reduce bias during refinement. This value is a comparison of a diffraction pattern based on the generated protein model to that of the removed data. A lower value represents increased similarity therefore, greater reliability (Morris, A. L. *et al* 1992).

R_{merge} demonstrates the similarity of the intensity of identical reflections on different frames; this gives a value of reliability for the structure at that resolution. Higher resolution shells are commonly more error prone as the data may not be complete or damage may occur during collection. This is a result of the crystal not demonstrating a sufficiently strong diffraction (described above in section 3.1.6).

R_{symm} (equation 3.4) is the agreement in signal intensities for each reflection between different diffraction images which are determined to be symmetry-related. Symmetry-related reflections should present identical intensities and variances in the reflection intensities, and can infer an error in the measurement technique.

Eq. 3.4) $$R_{symm} = \frac{\sum_{hkl} (I_{hkl1} - I_{hkl2})|}{\sum_{hkl} I_{hkl2}}$$

where I_{hkl1} and I_{hkl2} are the intensities of two symmetry-related reflections

The Ramachandran plot determines the validity of the generated structure by the determination of residues demonstrating bond angles in the allowed regions. The structure can be determined to be true if significant numbers of residues are in the allowed regions as a result of the steric hinderance of the amino acid side chains. Prolines are restricted to smaller areas and glycines are significantly more flexible and are allowed to occupy almost any area of the Ramachandran plot.

3.1.11 Results of the model building and refinement of LmPCNA structures

The structure created by the molecular replacement of the LmPCNA diffraction was examined for areas of poorly fitting structure. Amino acids not fitting the electron density were deleted or the side chains were rotated to better fit the data using the program COOT (Emsley, P. and Cowtan, K. 2004). After each successive COOT modification, the structure underwent 20 rounds of refmac. The final result was analysed with Procheck (Laskowski, R. A. *et al* 1993), this returned a final R_{factor} of 24% and an R_{free} of 29% for un-complexed LmPCNA with no residues in the disallowed portion of the Ramachandran plot. This gave a final model of the asymmetric unit represented in a unit cell (figure 3.6), a monomeric structure (figure 3.7) and the data and refinement statistics (table 3.2). LmPCNA peptide complex has a final R_{factor} of 25% and an R_{free} of 32% with no residues in the disallowed portion of the Ramachandran plot (table 3.2). This structure shows the peptide PCNA binding is conserved and is in a similar form to HsPCNA. It should be noted that not all LmPCNA binding sites were occupied to the same level; three of the six binding interfaces were occupied in the asymmetric unit.

The output structure of un-complexed LmPCNA was used as the molecular replacement model for the LmPCNA peptide complex data. The determined structures were validated using the program Procheck (Laskowski, R. A. *et al* 1993). The structure solution of this protein was similar to that of the un-complexed LmPCNA (figure 3.6). The structure of the monomer electrostatics and refinement statistics are represented in figure 3.7 and table 3.2.

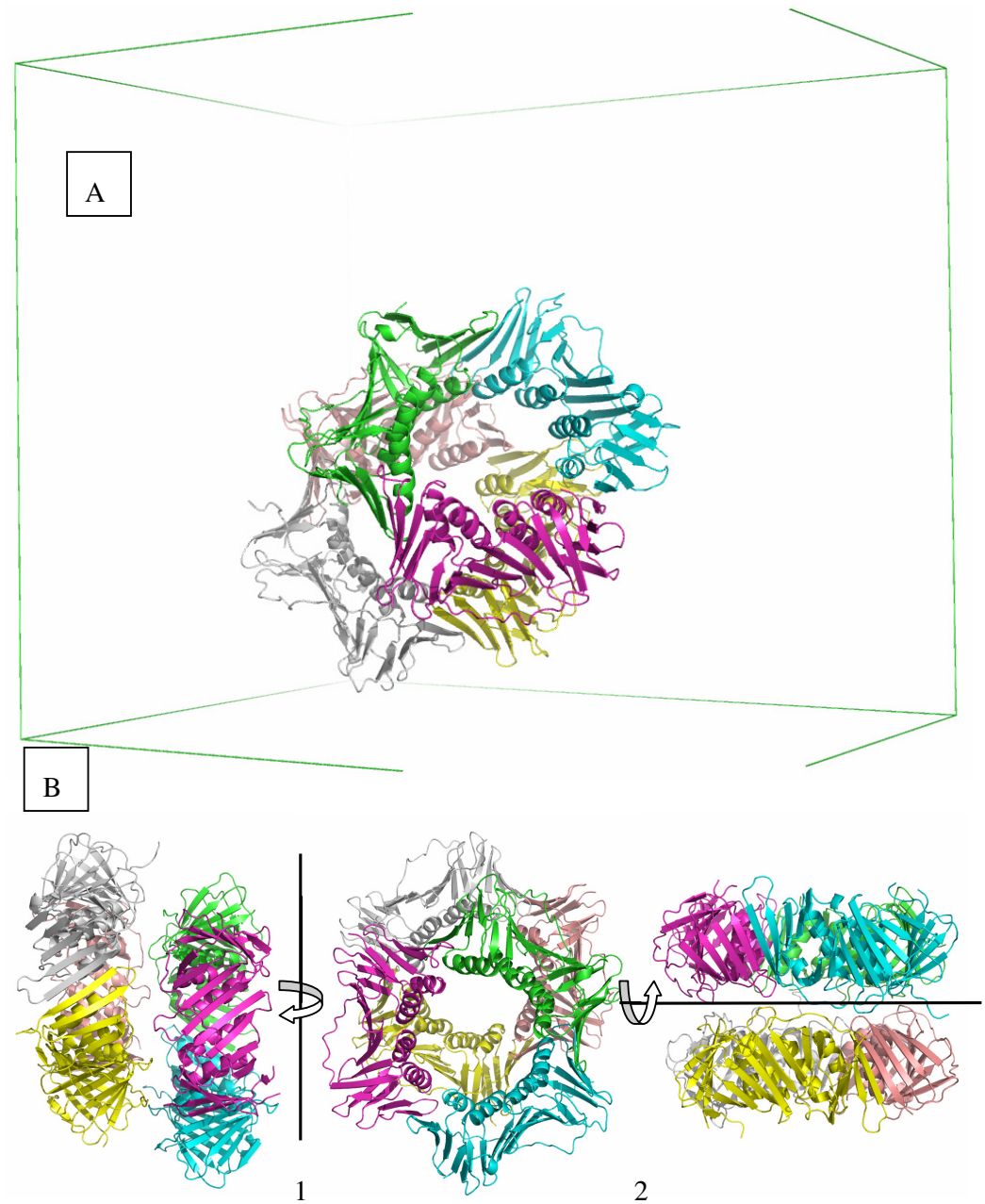


Figure 3-6. **The determined structure of LmPCNA**
A) Shows the structure of LmPCNA asymmetric unit presented within the unit cell of the crystal. This shows six monomers of LmPCNA per asymmetric unit however there was no data for the elongated loops which would branch out into solvent away from the two trimers (discussed in chapter 1). The lack of these loops is most likely due to the flexibility of these regions giving no consistent electron density. Both complexed and un-complexed structures shared these features.
B) Three views of the determined structure of LmPCNA asymmetric unit. These three views show the clear trimeric ring formation and the potential hexameric form. 1: The structure is rotated 90° along the vertical axis. 2: The structure is rotated 90° along the horizontal axis.

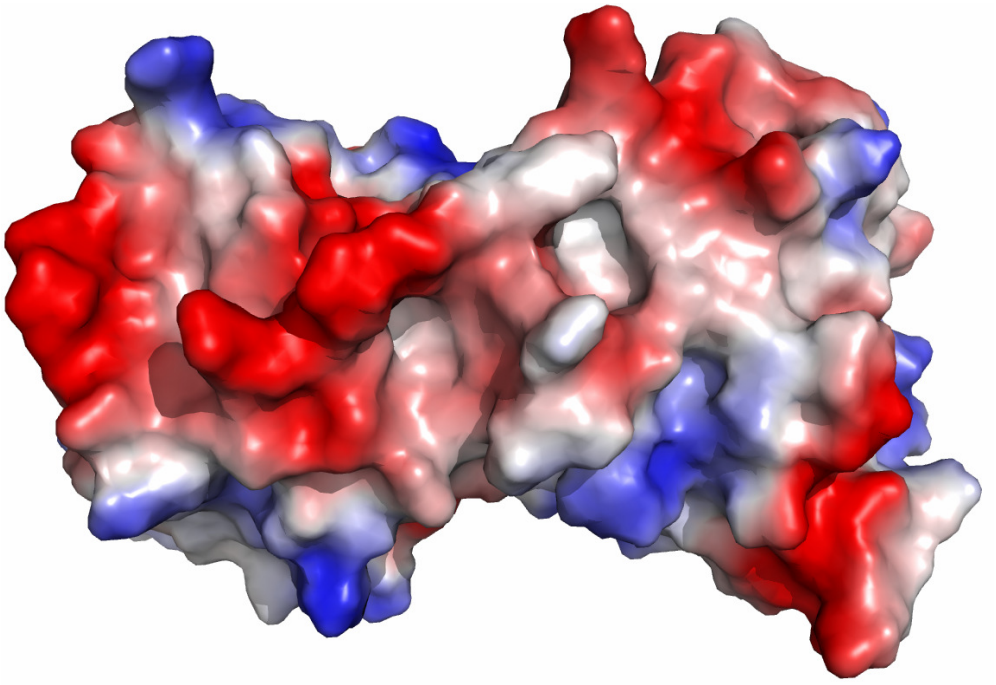


Figure 3-7. **The crystal structure of the un-complexed LmPCNA**
This structure shows LmPCNA as an electrostatic potential determined by PyMOL (DeLano, W. L. 2002). This is a single monomer of the solved structure.

Table 3-1. **The crystallographic statistics of un-complexed LmPCNA solution**

Data Statistics		
Space Group	$P2_12_12_1$	
Unit Cell parameters		
a	130Å	
b	150Å	
c	169Å	
$\alpha = \beta = \gamma$	90°	
Resolution range	30.4Å - 3.00Å	
Solvent content	75.4%	
Molecules per ASU	6	
Observed reflections	492387	(71019)
Unique reflections	66702	(9606)
I/σI	17.4	(2.0)
Completeness	100%	(100%)
Multiplicity	7.4	(7.4)
R_{merge}	9.9%	(84.3%)
Refinement Statistics		
R_{factor}	24%	
R_{free}	29%	
Correlation coefficient	0.75	
No. of protein atoms	11650	
Mean B-factor	84Å^2	
Residues in disallowed regions	0	

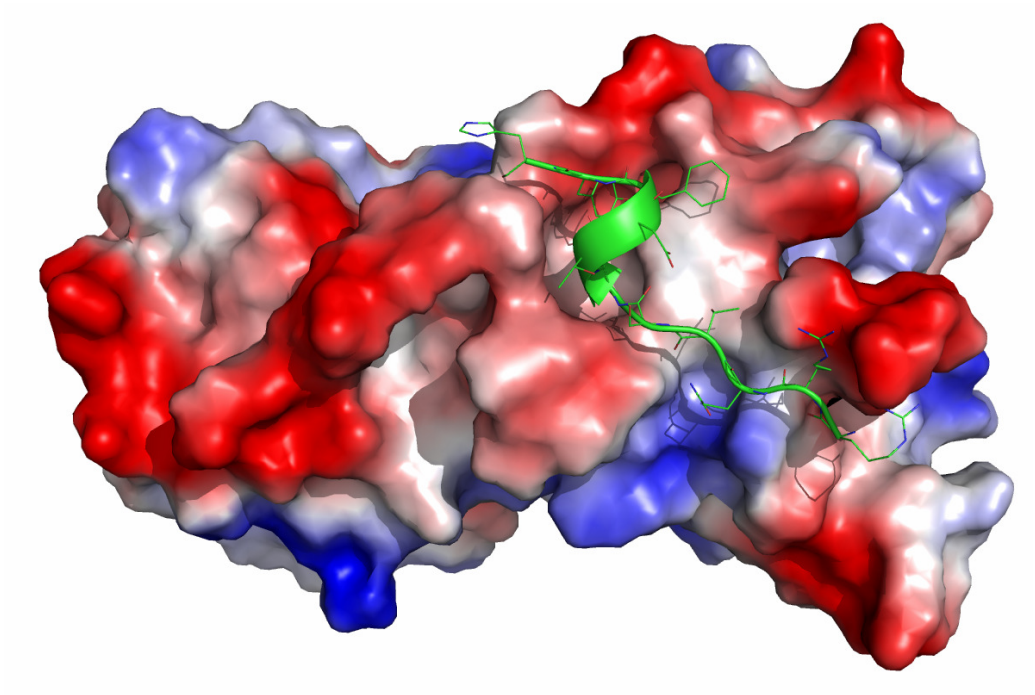


Figure 3-8. **The crystal structure of the LmPCNA peptide complex**
This structure shows LmPCNA (electrostatic potential) with the *H. sapiens* p21 peptide (143-152) bound (green). This is a single monomer of the solved structure.

Table 3-2. **The crystallographic statistics of LmPCNA peptide complex structure**

Data Statistics		
Space Group	$P2_12_12_1$	
Unit Cell parameters		
a	132Å	
b	151Å	
c	173Å	
$\alpha = \beta = \gamma$	90°	
Resolution range	47.8Å - 3.00Å	
Solvent content	76.1%	
Molecules per ASU	11	
Observed reflections	263630	(37467)
Unique reflections	69468	(10082)
I/σI	6.0	(0.4)
Completeness	99.3%	(88.4%)
Multiplicity	3.8	(3.7)
R_{merge}	23%	(81%)
Refinement Statistics		
R_{factor}	25%	
R_{free}	32%	
Correlation coefficient	0.76	
No. of protein atoms	11946	
Mean B-factor	88Å^2	
Residues in disallowed regions	0	

3.2 Conclusions from the X-ray crystallographic study of LmPCNA

The LmPCNA monomer shares a 38% sequence identity to HsPCNA and a sequence identity of 48% to SpPCNA; HsPCNA and SpPCNA share a 50% sequence identity. PCNA from these three species all have a trimeric structural architecture (figure 3.11).

The secondary structure of LmPCNA shows four main helices and sixteen β-strands. LmPCNA demonstrates a similar topology as HsPCNA and SpPCNA (described in section 1.2). The LmPCNA monomer exists as two globular domains (figure 3.10) connected by the LmPCNA IDCL (figure 3.9). An LmPCNA globular domain is made of two βαββββ folds (figure 3.10). There are four βαββββ folds per monomer and twelve in the trimer, identical to that of HsPCNA and SpPCNA.

The sixteen β-strands of each monomer associate to form three β-sheets (figure 3.10). These three β-sheets provide structural rigidity to LmPCNA (figure 3.10). The β-sheets also aid in the formation of the monomer-monomer (discussed in section 3.1.1) and PIP box binding interfaces (discussed in section 3.1.5). The four main helices associate and are supported by the β-sheets (figure 3.10). The secondary structural features of LmPCNA are similar to that of HsPCNA and SpPCNA.

The PCNA monomers from different species can be superposed (figure 3.11). The level of structural similarity can be derived when analysing the superposition of the backbone atoms. Of the three PCNAs, SpPCNA and HsPCNA (using 912 atoms) are structurally the most similar (0.654Å "Root Mean Square Deviation" R.M.S.D.) and LmPCNA and HsPCNA (using 880 atoms) are the least similar (1.01Å R.M.S.D.) (figure 3.11). This trend in the R.M.S.D. of the determined structure of PCNA agrees with the level of sequence identity shown above.

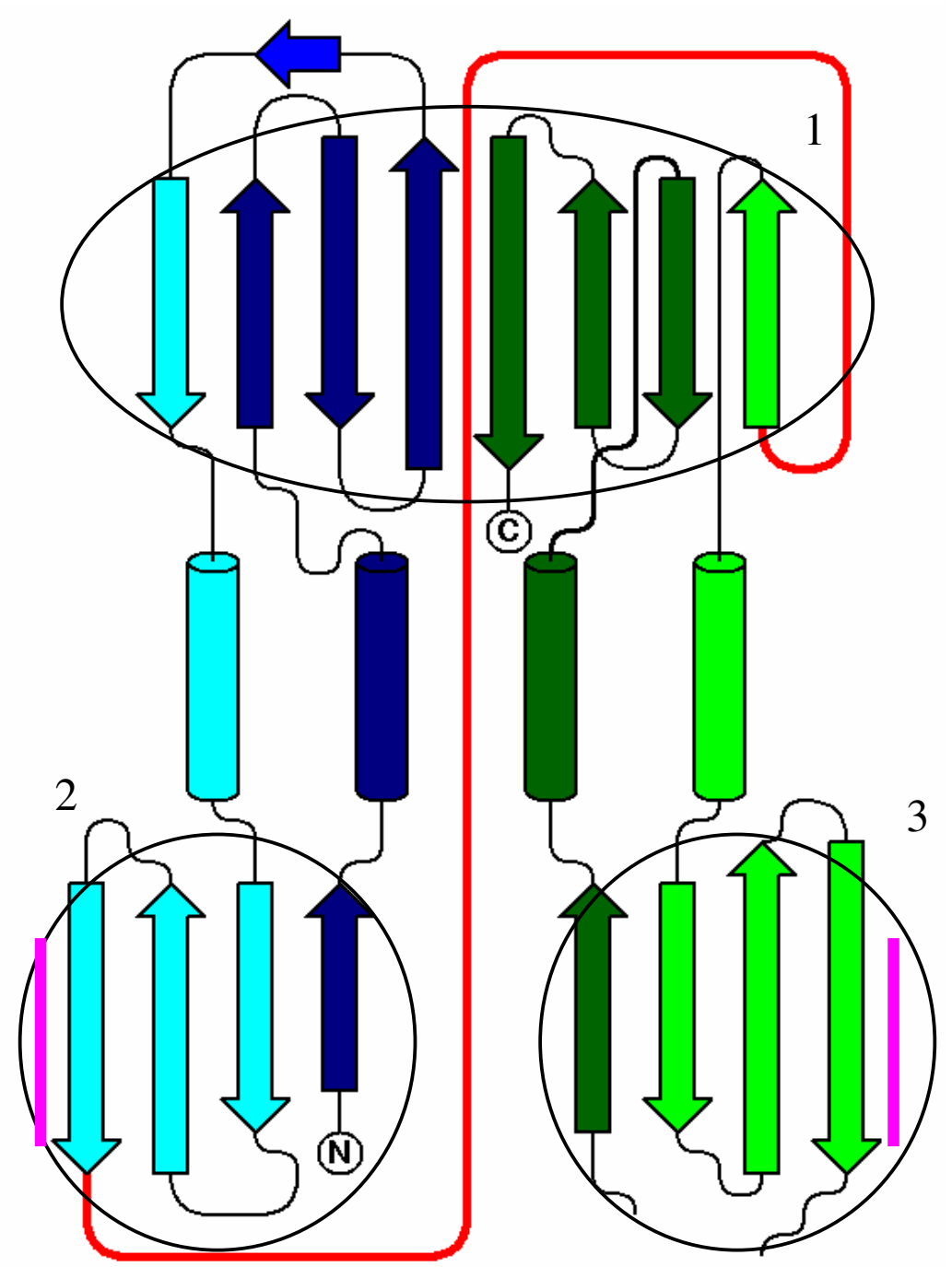


Figure 3-9. **The topology diagram of LmPCNA**
Figure showing the topology of the LmPCNA monomer. The four βαβββ are represented in dark blue, light blue, light green and dark green. The IDCL is shown in red and demonstrates the clear separation in the two domains.

1) the central β-sheet of the LmPCNA monomer which forms the PIP box binding interface.
2) 2&3) the β-sheets forming the monomer-monomer binding interfaces at the magenta line. Generated using TopDraw (Bond, C. S. 2003).

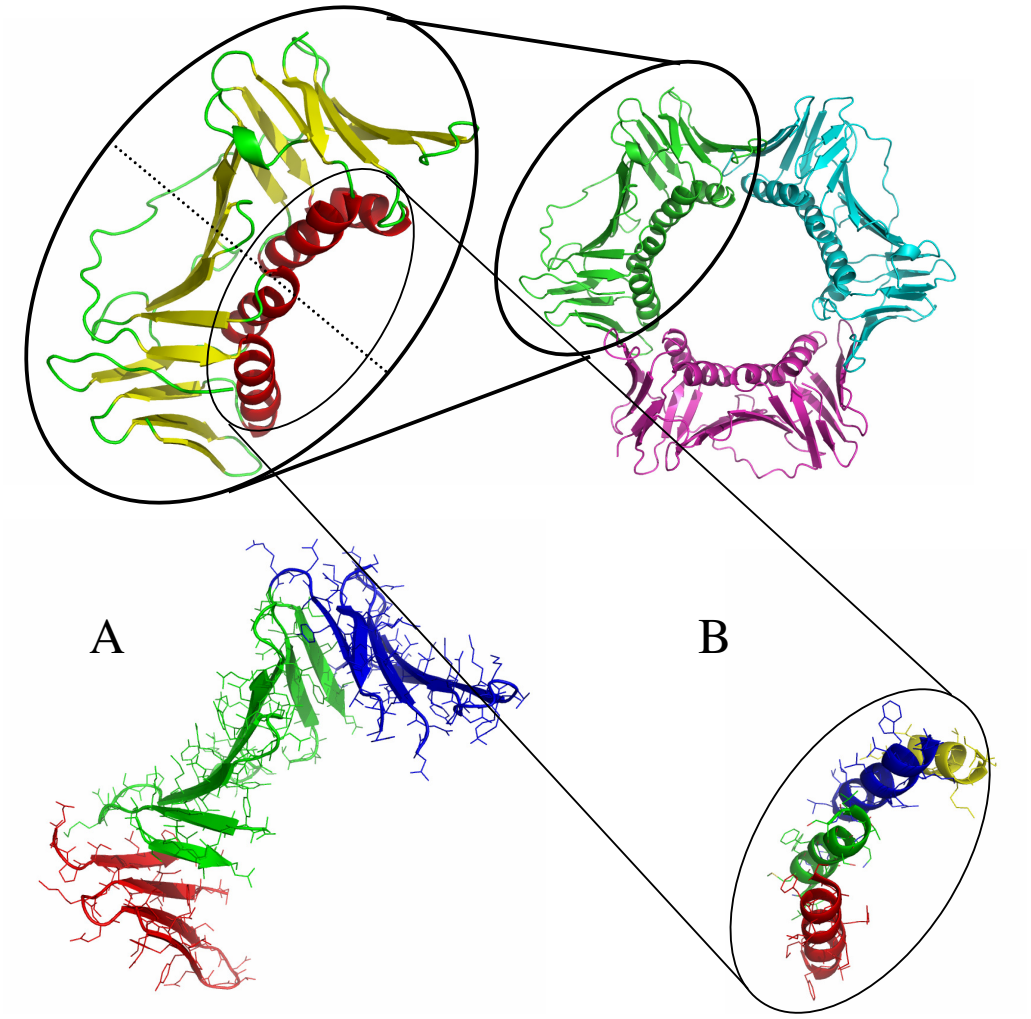
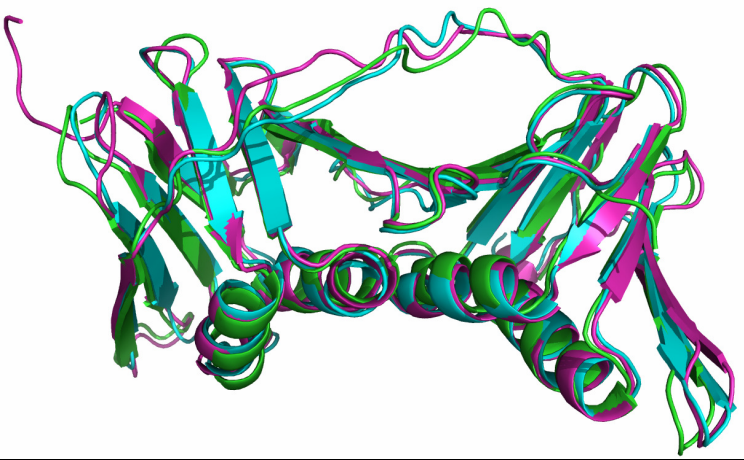
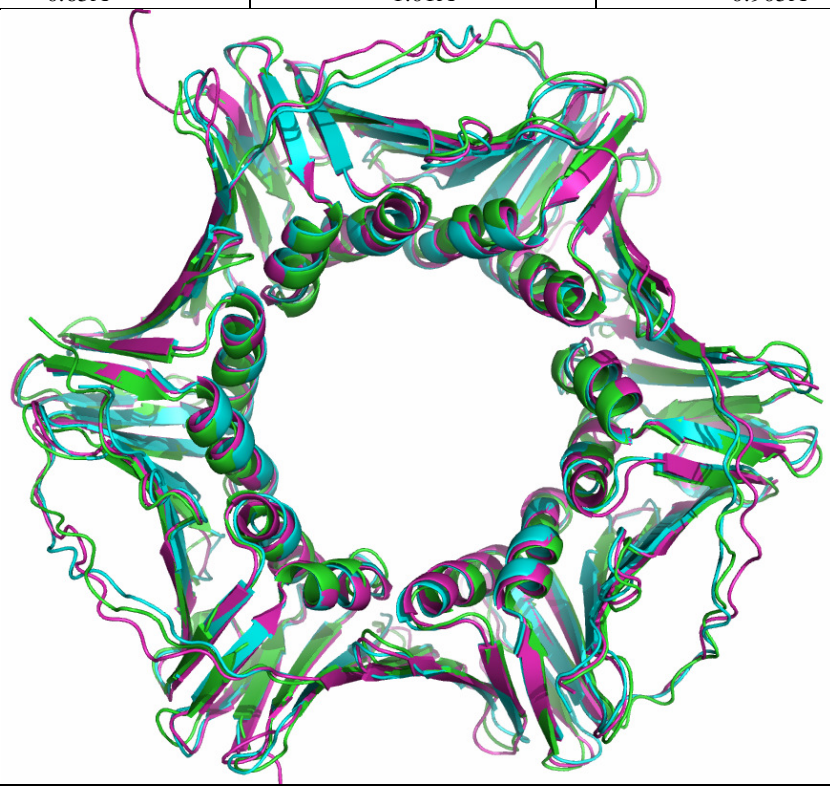


Figure 3-10. **The secondary structure of the LmPCNA monomer**
Figure showing the secondary structure of the LmPCNA monomer in an exploded form separating β-strands from α-helices. The dashed line of the monomer shows the separation of the two globular domains.
A) The arrangement of the sixteen β-strands to form three β-sheets (red, green and blue). These three β-sheets provide structural rigidity to the monomer. The green β-sheet aids in the formation of the PIP-box binding interface and the red and blue β-sheets aid in the formation of the monomer-monomer binding interface.
B) The arrangement of 4 α-helices, red, green, blue and yellow into a helix bundle. Hydrophobic residues are positioned facing the β-sheets and positively charged residues are exposed to solvent (described in greater detail in figure 3.17).



The RMSD of the alignment of PCNA monomers from three different species		
HsPCNA - SpPCNA	HsPCNA - LmPCNA	SpPCNA - LmPCNA
0.65Å	1.01Å	0.905Å



The RMSD of the alignment of PCNA trimers from three different species		
HsPCNA - SpPCNA	HsPCNA - LmPCNA	SpPCNA - LmPCNA
0.73Å	1.11Å	1.05Å

Figure 3-11. **The structural alignment of HsPCNA, SpPCNA and LmPCNA**

Figure showing the alignment of PCNA from three different sources. Top) The alignment of PCNA monomers. Middle) The alignment of PCNA trimers showing HsPCNA (cyan), SpPCNA (magenta) and LmPCNA (green). Bottom) The R.M.S.D. of the alignment of PCNA monomers and trimers (Å). This shows a greater R.M.S.D. of trimers compared to monomers, suggesting a shift in the structural architecture.

3.2.1 The monomer-monomer binding interface of LmPCNA

The LmPCNA-LmPCNA binding interface is comprised of twenty-three residues (figure 3.12A). Ten residues from one monomer interface are associated to thirteen residues from the neighbouring monomer (figure 3.12A). These twenty-three residues of the binding interface were defined as any residue within 4Å of the neighbouring chain, calculated by PyMOL (DeLano, W. L. 2002).

The twenty-three residues for the LmPCNA-LmPCNA interaction have only a 26% sequence identity with HsPCNA and a 38% sequence identity with SpPCNA (figure 3.12A). This binding interface has a lower level of sequence identity then the average between PCNA from these species. Conserved sequences on one strand hydrogen bond with conserved residues on the opposing strand. Sequences with little conservation hydrogen bond with poorly conserved residues on the opposing strand (figure 3.12A) (figure 3.13). Retention of sequence identity at the PCNA-PCNA interface does not seem to be essential as the hydrogen bonds form primarily between backbone atoms.

Hydrogen bonds form across an anti-parallel β-sheet between residues K108-N115 and Q177-Q183 (figure 3.13C). HsPCNA forms this interface with β-strands from E109-K117 and L175-S183 and SpPCNA forms this interface with β-strands from I111-K117 and D174-T181 (figures 3.13AB). Away from the 18 residues making up these β-strands, there are a further 5 residues at the PCNA monomer-monomer interface. These residues are on the neighbouring helix and provide additional hydrogen bonding (E143···R110) and create hydrophobic surfaces (V153 and I81) (figure 3.13). The residues which form this binding interface vary between PCNA from these three sources. Changes in the sequence of the opposing

```
        107       81
H.sap   EKVSDYEMK C
S.pom   DRISDYDVK C
L.maj   TRKCEYQLK I
         : .:*::*

        174       143 146 150 153
H.sap   LGNGNIKLS E   R   D   H
S.pom   IGNGSTTYK E   R   D   T
L.maj   VGQGYTFLQ E   K   D   V
        :*:*    . *   :   *
```

Figure 3-12. **A) Residues at the PCNA monomer-monomer binding interface**
Figure showing the alignment of residues at the PCNA monomer-monomer binding interface. These residues are shown to be involved in the binding interface in HsPCNA from (Gulbis, J. M. *et al* 1996). These residues form a β-sheet resulting in hydrogen bonds to form between E107 and S183 through to K117 to L175. This hydrogen bonding network facilitates the association of two monomers. This figure shows that the K117/L175 side of this β-sheet has the greatest level of conservation in both strands. Changes in residue composition at the E107 side of the β-strand may be compensated for by modifications to the S183 side of the associating strand.

```
        27 29 40 44   67   97 118           131 208 232 234 250 252
H.sap   C  D  M  HVS  AMG  D  LMDLDVEQLGIP  Q   A   D   P   Y   APKI

        27 29 40 44   67   97 118           131 208 232 234 249 251
S.pom   N  D  M  HVA  ALG  E  LMDIDQEHLGIP  I   S   D   P   Y   APKI

        27 29 40 44   67   97 118           131 242 266 268 284 286
L.maj   N  D  M  HVA  ILG  D  LLEIEAESMGIP  M   A   D   P   Y   APKV
           *  *  **:   :*  :  *:::: * :***      :   *   *   *   ***:
```

Figure 3-12. **B) Residues at the PCNA-PIP box binding interface**
Figure showing the alignment of residues at the PCNA-PIP box binding interface. These residues are conserved showing a sequence identity of 43% and residues with similar shape and charge of 80%. The high sequence similarity of this binding interface suggests that LmPCNA has a similar binding motif as HsPCNA. Residues in red represent hydrogen bonding partners. This suggests that LmPCNA will have similar cellular activities. The chain of 12 residues starting at 118 is the IDCL.

```
H.sap   181 KLSQT------------SNVDKEE----------------------EAVTI 197
S.pom   181 TLKQH------------TDLSDQD----------------------QSIEI 197
L.maj   181 FLQAAGVSDRSAKSEVKSEVKAEARDEDEHEPISRRYNKAEGGNGAIGVEV 231
            EEE             HHHHHHHHHHH       HHH           EEE
```

Figure 3-12. **C) Residues at the LmPCNA elongated loop domain**
Figure showing the novel loop domain of LmPCNA. This loop has 41 residues in LmPCNA opposed to 7 residues in HsPCNA. This additional loop domain is predicted to have a helical region in LmPCNA but not in either HsPCNA of SpPCNA. This could suggest this loop may act as a scaffolding domain.

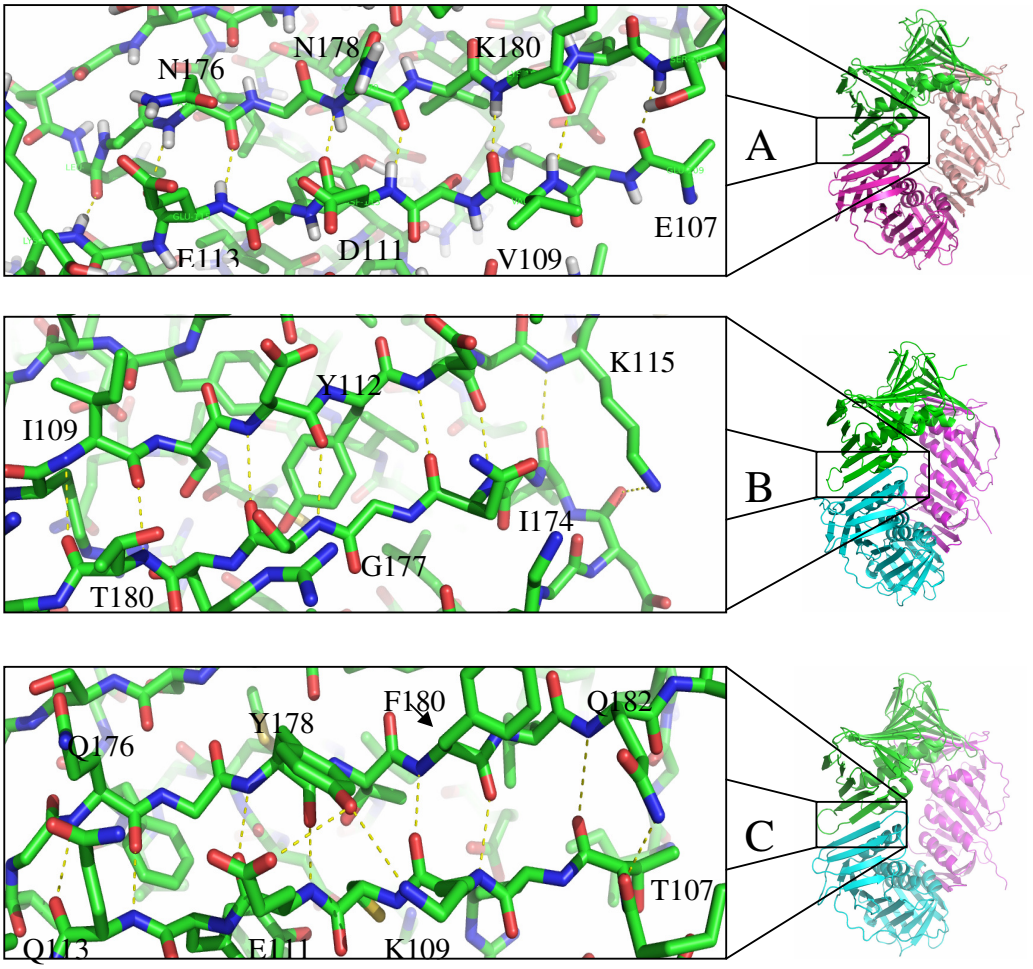


Figure 3-13. **The monomer-monomer binding interface of HsPCNA, SpPCNA and LmPCNA**

Figure showing the monomer-monomer binding interface of PCNA from *H. sapiens*, *S. pombe* and L.major. The interaction of these monomers is essential for ring formation. The number of hydrogen bonds has been linked to protein stability (section 1.2). A) The hydrogen bonding network of HsPCNA showing the potential formation of 8 hydrogen bonds. The hydrogen bonding of HsPCNA monomers occurs at β-strands E109-K117 and L175-S183.

B) The hydrogen bonding network of SpPCNA showing the potential formation of 8 hydrogen bonds. The hydrogen bonding of SpPCNA monomers occurs at β-strands I111-K117 and D174-T181.

C) The hydrogen bonding network of HsPCNA showing the potential formation of 10 hydrogen bonds. The hydrogen bonding of HsPCNA monomers occurs at β-strands K108-N115 and Q177-Q183.

strand or position of the hydrogen bonding network may compensate for a loss of sequence conservation (figure 3.12A).

The hydrogen bonding between the monomers at this β-sheet is maintained in PCNA from all sources. The three PCNAs investigated here all demonstrate a trimeric structural architecture and similar binding properties at the monomer-monomer interface (figure 3.13). The un-complexed structure of LmPCNA shows that twelve hydrogen bonds can be formed at the LmPCNA monomer-monomer interface (figure 3.13C). The crystal structure 1AXC (Gulbis, J. M. *et al* 1996) and the crystal structure of SpPCNA (unpublished data) show that PCNA from these species form twelve and eight hydrogen bonds at the monomer-monomer interface respectively (figures 3.13AB).

The LmPCNA monomer-monomer interaction buries 1370 $Å^2$ of surface area, as determined by PISA (table 3.3) (figure 3.14) (Krissinel, E. and Henrick, K. 2007a). The monomer-monomer interaction of HsPCNA and SpPCNA buries 1380 and 1780 $Å^2$ of surface area respectively (table 3.3) (figure 3.15). HsPCNA and SpPCNA have been shown to involve three and seven water molecules at the monomer-monomer interface. The position of one of these water molecules is conserved between HsPCNA and SpPCNA (figure 3.15). The electron density of the hexamer showed a potential water molecule at this position between chains E & D and A & C. The density for this potential water was poor and only occurred at a sigma level of 0.36, suggesting this may or may not be genuine. The SpPCNAs greater buried surface area and larger numbers of interacting water molecules may compensate for the reduced number of hydrogen bond interactions at the monomer-monomer interface. Based on the crystal structures of PCNA from these three

Formula	Monomers within structure	Is this a stable interaction	Surface area (Å^2)	Buried area (Å^2)	ΔG of formation (kcal.mol^{-1})	ΔG of dissociation (kcal.mol^{-1})
The following complexes appear to be genuine						
A_3	ACE	Yes	32340	4150	-14.0	4.9
AB	AF	Yes		1200	-9.0	5.0
C_3	BBB	Yes	34440	5350	-7.8	1.6
CB	BC	Yes		1240	-10.0	6.2
D_3	MNO	Yes	34990	4110	-15.0	2.4
DB	OH	Yes		1450	-12.3	9.5
The following complexes fall into a grey area and may or may not be genuine						
D_6	EGIMNO	No	68290	10020	-33.4	-2.0

Table 3-3. **The energetic properties of different binding interfaces as determined by PISA (Krissinel, E. and Henrick, K. 2007a)**
This table give the values of surface area and buried surface area (Å^2) and the calculated ΔG of formation and dissociation (kcal mol^{-1}) of different protein-protein interactions as determined by (Krissinel, E. and Henrick, K. 2007a). HsPCNA is represented by A and is the structure 1AXC as determined by (Gulbis, J. M. *et al* 1996). SpPCNA is represented by C and is the solved structure of the complexed SpPCNA (unpublished data). LmPCNA is represented by D and is the solved structure with peptide bound. Values were also derived for the PCNA-p21 (R143-H152) interaction with the three different PCNAs. The species are grouped by the complex validity of interaction.
This determines an interaction to be genuine based on the buried surface area in relation to the size of the immobilised protein. The ΔG of formation should be negative and the ΔG of dissociation should be positive.
PISA concluded that the trimeric ring structures are most likely a genuine interaction. The PCNA p21 (R143-H152) are a genuine interaction. The LmPCNA dimerisation of trimers falls into a grey area and may or may not be genuine.

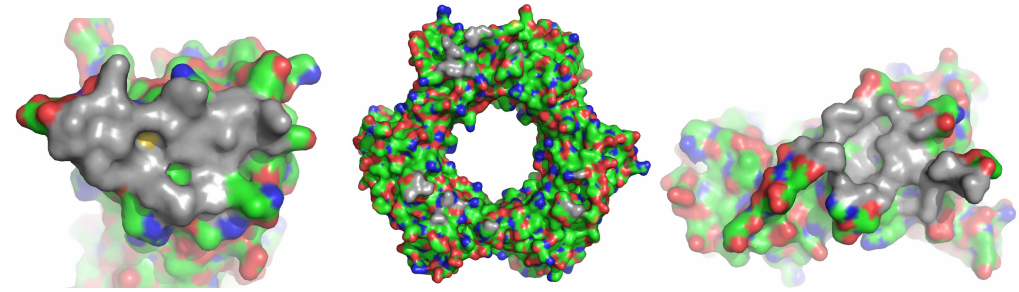


Figure 3-14. **The buried surface areas of LmPCNA binding interfaces**
Figure showing the buried surface areas of LmPCNA when complexed at the monomer-monomer interface (left), trimer-trimer interface (middle) and the PCNA-p21(143-152) interface (right). These images show carbons (green) oxygens (red) nitrogens (blue), sulphurs (yellow) and residues within 4Å of the interacting molecules (grey).

Interactions at HsPCNA interface	Interactions at SpPCNA interface	Interactions at LmPCNA interface
E143(SC)-K110(SC)	R146-S81	E143-R110
R146-K79	I175-K117	K(SC)146-I81
R146-A82	N177-D115	Q177-Q115 (×3) (1×SCs)
L175-K117	S179-D113 (×2)	Y(SC)179-K(SC)111
N177-E115 (×3) (1×SCs)	T(SC)180-S(SC)112	Y179-E113 (×3) (1×SCs)
N179-D113 (×2)	T181-I111 (×2)	F181-K111 (×2)
K181-V111 (×2)	K183-D109	Q183-T109
S183-E109	D(SC)186-K(SC)107	
Interactions involving a single water		
D150(SC)-K110(SC)	T(SC)181-I111	
R146(SC)-N84(SC)	I175-K117	
L175-K117	N177(SC)-K117(SC)	
	T153(SC)-K77	
	T153-K77(SC)	
	R149(SC)-R80	
	D150(SC)-R80	

Table 3-4. **The hydrogen bonds formed at the monomer-monomer interface of PCNA**

Table showing the hydrogen bond interactions at the monomer-monomer interface of PCNA. HsPCNA forms twelve hydrogen bonds at the interface and three hydrogen bonds through water. SpPCNA forms ten hydrogen bonds at the interface and seven hydrogen bonds through water. LmPCNA forms twelve hydrogen bonds at the interface however there was no density for water molecules in the crystal structure.

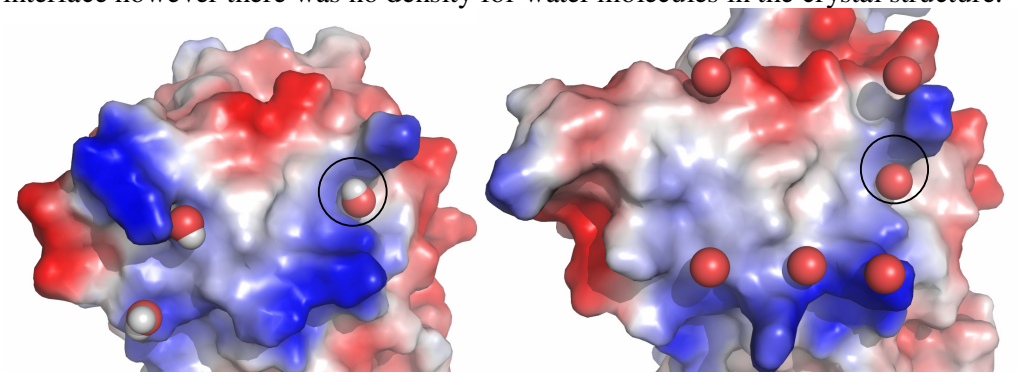


Figure 3-15. **The monomer-monomer binding interfaces of HsPCNA and SpPCNA showing water positions**

Figure showing the monomer-monomer binding interface of HsPCNA and SpPCNA with interacting waters shown as spheres. The position of waters determined by crystal studies have been solved and are shown to interact with the binding interface. The interacting waters associate near the edges of binding interface away from the hydrophobic core. The position of only one water molecule is conserved between these two crystal structures. SpPCNA forms fewer hydrogen bonds at the monomer monomer interface, this may be compensated for by the larger hydrophobic buried surface and greater numbers of interaction waters.

species, these monomer-monomer interactions forming the trimeric ring have been determined to be genuine by PISA (table 3.3). Values determined by PISA gave a ΔG of formation of the monomer-monomer interaction. LmPCNA and HsPCNA have similar ΔG of formation, 19.5 $kJmol^{-1}$ and 20.9 $kJmol^{-1}$ (table 3.3). SpPCNA gave a ΔG of formation at the monomer-monomer binding interface of 10.9 ΔG $kJmol^{-1}$ (table 3.3). This apparent reduction in the ΔG of formation from HsPCNA and LmPCNA to SpPCNA is a result of the reduced level of hydrogen bonding in SpPCNA. Trimer stability has been studied in solution and is discussed in chapter 6.

The R.M.S.D. of the monomeric PCNAs have been determined (section 3.1) by superposition of the backbone atoms alone. This can also be performed with the trimer backbone atoms of aligned residues. The superposition of the whole trimer can give information about the conservation of the quaternary structure. An identical R.M.S.D. of monomer (determined in section 3.1) and trimer superposition would suggest that there is no deviation in quaternary structure. Superposes of the trimers were performed using only the backbone atoms and the R.M.S.D. were compared to the R.M.S.D. of the monomer superposition. The trimer superposition of the three PCNAs gave a greater R.M.S.D. implying that there are indeed small differences in the quaternary structure. The greatest change in R.M.S.D. from monomeric super position to trimeric superposition was between SpPCNA and LmPCNA. This showed the monomeric superposition to have an R.M.S.D. of 0.91Å (using 931 atoms) and a trimeric R.M.S.D. of 1.05Å (using 2786 atoms) resulting in a change of +0.14Å (figure 3.11). The lowest difference between monomeric and trimeric superposition of PCNA was between HsPCNA and SpPCNA (using backbone 2732 atoms) resulting in a change of +0.071Å (figure 3.11). These values follow the trends in

sequence identity (section 3.2) between PCNA from these three species. The R.M.S.D. of the trimeric rings is only a slight increase from the monomer deviations, which suggests well conserved ring architecture. The maintenance of appropriate ring formation may be essential for optimal PCNA activity.

The structure of LmPCNA shows a potential hexameric state (figure 3.13A). The hexamer interface of LmPCNA shows only four hydrogen bond contacts (figure 3.13B). The potential hexameric state of LmPCNA was analysed by PISA (Krissinel, E. and Henrick, K. 2007a) in order to determine if this is a genuine interaction. The formation of the LmPCNA hexamer from two trimers results in a buried surface area of 1760Å^2, and based on this analysis, it could not be conclusively determined if this interaction is stable (table 3.3). The potential hexameric state has been studied in solution (discussed later in chapter 5) to determine if this hexameric state of LmPCNA is a crystallographic artefact

3.2.2 The central pore of the PCNA ring

The central pore of the LmPCNA trimer (figure 3.16A) is a conserved feature in both HsPCNA (Kelman, Z. and ODonnell, M. 1995) and SpPCNA (unpublished data). HsPCNA has previously been determined to be ~35Å and of sufficient size to surround B-form DNA (Kelman, Z. 1997). The distance between the Cα of E17 and R149 of HsPCNA, SpPCNA and LmPCNA has been measured. This defined pore diameter gave LmPCNA a pore of ~37Å (figure 3.17A) compared to HsPCNA and SpPCNA of ~36Å. The residues which make up the inner lining of PCNA pore are positively charged to facilitate the association to DNA. The pattern of positively charged residues is conserved in PCNA from *H. sapiens*, *S. pombe* and *L. major* (figure 3.17B).

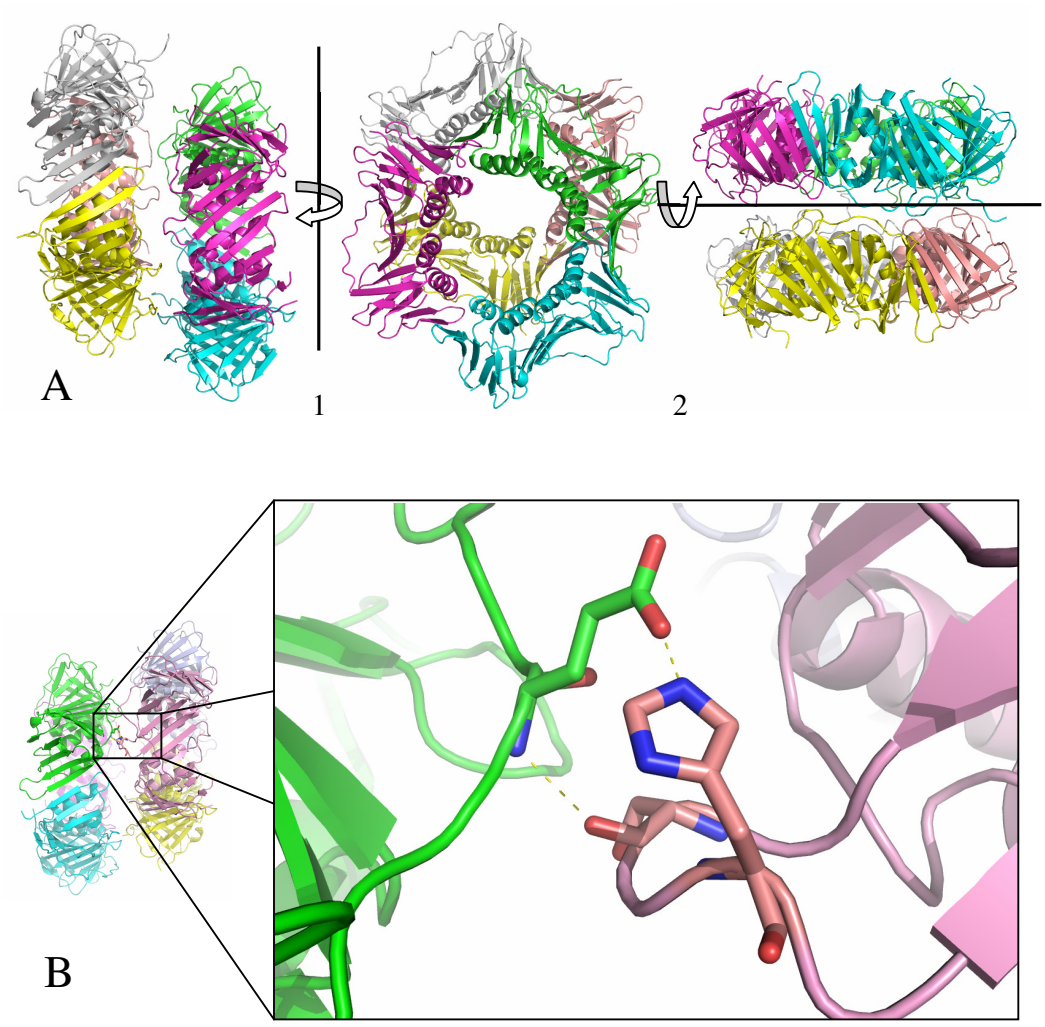
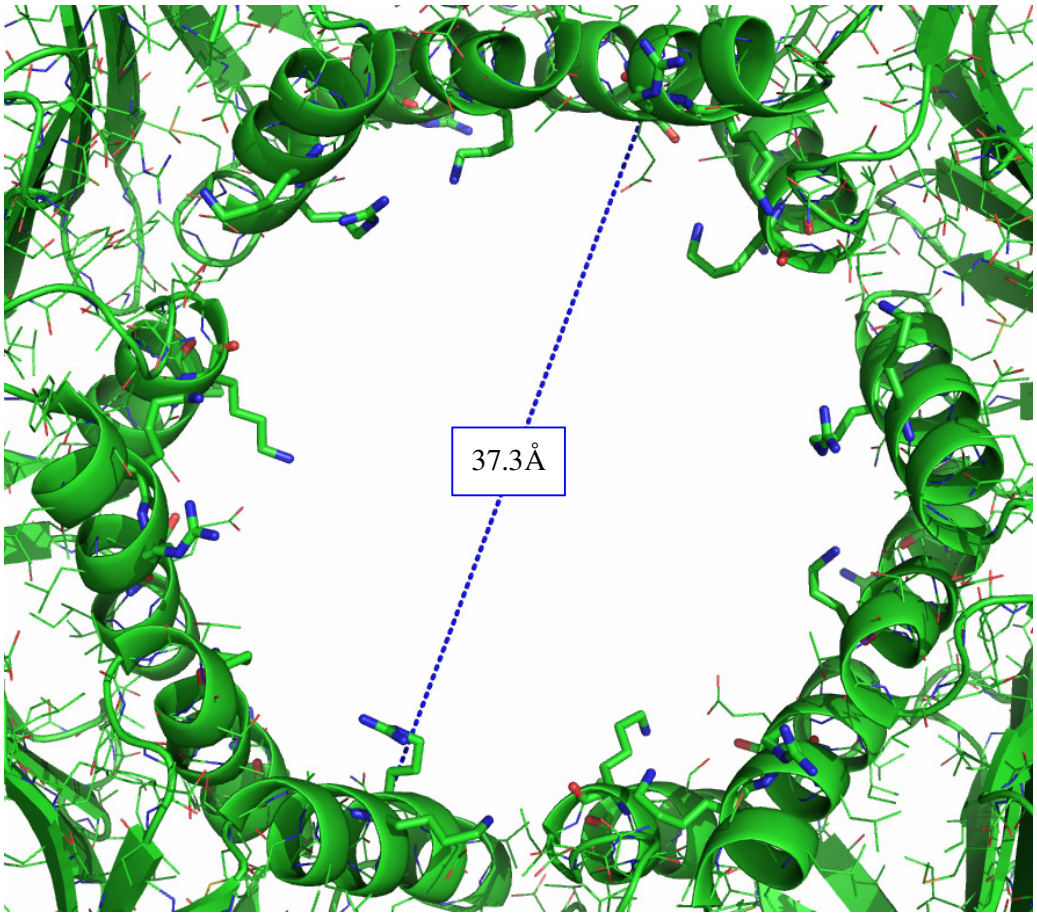


Figure 3-16. **Structure of LmPCNA**
Figure showing the structure of LmPCNA in the asymmetric unit. There are 6 LmPCNA monomers in the asymmetric unit and these form an apparent hexameric structure. A) Two homo-trimers of LmPCNA showing the potential dimer of trimeric rings as presented in the asymmetric unit. 1: The structure is rotated 90° along the vertical axis. 2: The structure is rotated 90° along the horizontal axis.
B) The potential hexameric form of LmPCNA contains 4 hydrogen bonds at the binding interface, this structure shows a close up of two of these hydrogen bonds.




```
                13 20 77 80 146 149 210 217
H.sap           KK K  K  K  C   R   R   K

                13 20 77 80 146 149 210 217
S.pom           KK K  K  R  R   R   K   K

                13 20 77 80 146 149 244 251
L.maj           KR N  K  K  K   R   R   K
                *: :  *  :      *   :   *
```

Figure 3-17. **The pore of LmPCNA**
Figure showing the inner lining of the LmPCNA pore with a conserved architecture to that of HsPCNA. Top) LmPCNA pore with diameter of ~37Å is of sufficient size to surround DNA. The inner lining of the LmPCNA trimeric ring comprised of 12 α-helices. Each helix has at least two positively charged residues for the association to DNA. There are 27 positively charged resides represented in stick form. Bottom) the conservation of the 27 positively charged resides which make up the surface of the pore. This demonstrates a high level of sequence identity and similarity within these residues. This high level of conservation suggests that these residues are essential for the action of PCNA within the cell.

There are approximately two positively charged residues per helix lining the central pore which show a high level of conservation. The positively charged residues in the PCNA monomers have a sequence identity of 44% with (89% similarity) (figure 3.17B).

3.2.3 The elongated loop region (186-223) of LmPCNA

An elongated loop region in protozoan PCNA exists in LmPCNA (section 1.2, figure 1.2). This loop domain is not present in eukaryotic PCNA (figure 3.12C). Eukaryotic PCNA has a loop of 7 residues at this region. The loop domain in LmPCNA is 41 amino acids in length (figure 3.12C). The increased size of the loop may give LmPCNA another mode of interaction with DNA or other partner proteins.

The elongated loop domain of LmPCNA has a flexible base and is protruding into solvent. (figure 3.18). The flexibility of this loop is the result of a glycine rich sequence at the N-terminal side of the loop domain and hydrophilic residues at the C-terminal side (figure 3.12C). This loop branches from the second globular domain between the third and fourth βαββββ folds (figure 3.9).

The corresponding loop in HsPCNA is not resolved in the structure of HsPCNA 1AXC (Gulbis, J. M. *et al* 1996). The structure of SpPCNA loop is seven residues in length and is solved (unpublished data). In SpPCNA, this region has a high B-factor (47Å^2) indicative of a high degree of flexibility. The presence of the 41 residue loop domain in LmPCNA results in the trimeric rings being separated, packing LmPCNA in the observed form. This reduces the proteins density and increases the solvent content of the crystal to ~76%. Crystals with high solvent content are linked to poor resolution crystal data (Souza, D. H. F. *et al* 2000). Electron density of this loop domain was observed in the crystal structure of both

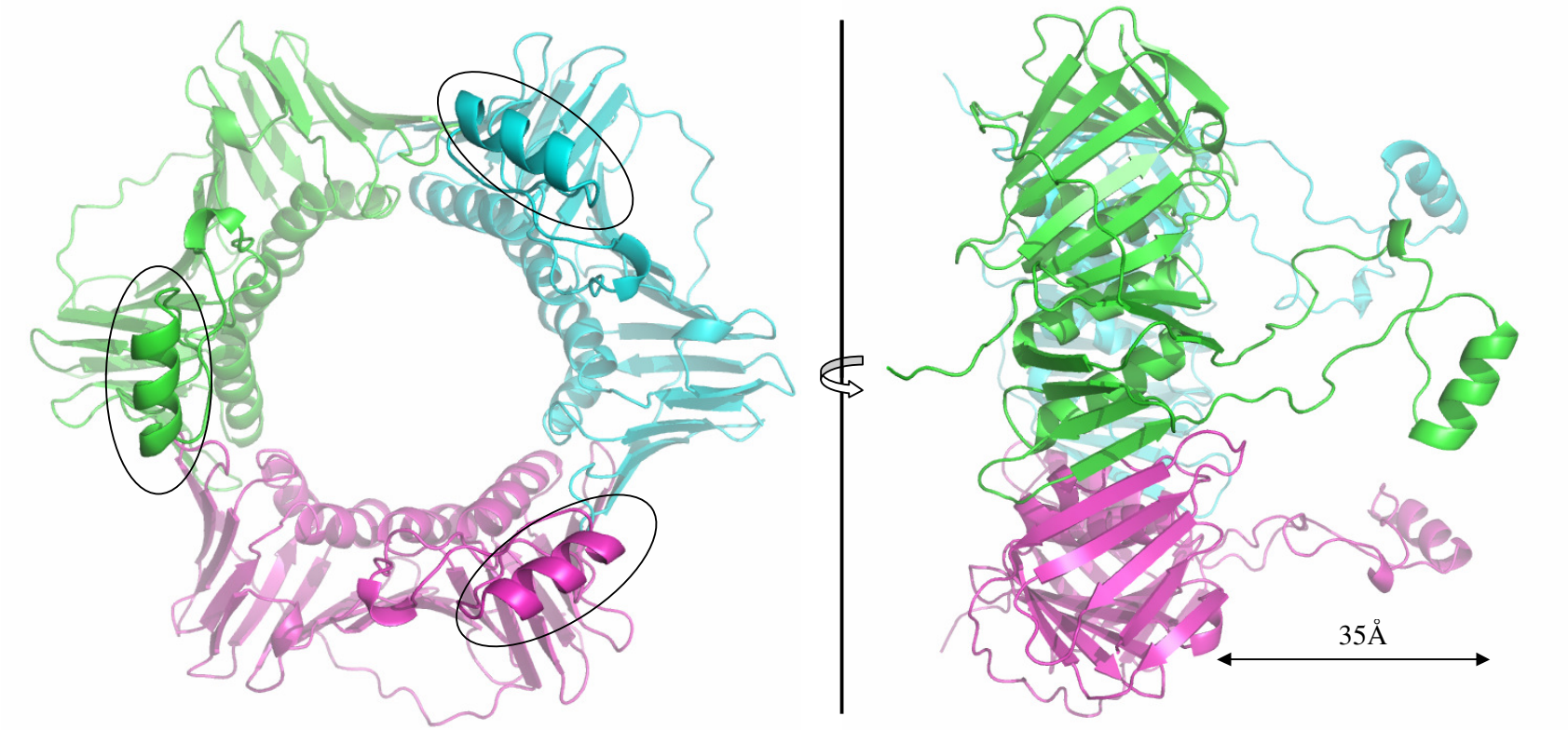


Figure 3-18. **The predicted position of the elongated loop region of LmPCNA**
Figure showing the model of LmPCNA generated using 3D-JIGSAW (Bates, P. A. *et al* 2001). This model was superposed onto each PCNA monomer of the structure 1AXC (Gulbis, J. M. *et al* 1996) generating a trimeric structure. Left) The ring form of the LmPCNA trimeric model showing the position of the loop helix (figure 3.19). This helix is in alignment with the helical bundle region of the LmPCNA trimer suggesting a potential interaction with DNA. Right) The trimeric model of LmPCNA with a 90° right handed rotation. This shows that the 41 residue loop region allows the helix to be held 35Å from the ring.

complexed and un-complexed LmPCNA. Approximately three residues of the LmPCNA loop domain could be modelled in either LmPCNA structure. It was not possible to build the remaining residues of this loop domain resulting in the loop region not being resolved in either X-ray crystal structures (figure 3.16A). This loop is not resolved because (as described above) the base of the loop is likely flexible and disordered in the solvent.

The elongated loop domain of LmPCNA was modelled using 3DJigsaw (Bates, P. A. *et al* 2001). and was predicted to contain an α-helix (figure 3.19A). This helix is not present in HsPCNA or SpPCNA (figure 3.12C) and the function of the loop in *L. major* cell lines has not been investigated. The predicted helical part of the loop is positively charged down one face and negatively charged down the other (figure 3.19B). This differential charge could allow association to DNA or other binding partners. The predicted loop is positioned near the helices lining the pore, protruding ~35Å from the N-terminal side of LmPCNA (figure 3.18). The theoretical helix may be near the path of the associated DNA molecule (figure 3.18). Association of the loop to the DNA molecule may create another binding interface in LmPCNA.

3.2.4 The comparison of un-complexed and complexed structures of LmPCNA

The crystal structure of LmPCNA has been determined with and without *H. sapiens* p21 (141-152) bound. The complexed protein-peptide crystal was formed by co-crystallisation using the same conditions as the un-complexed crystal. The un-complexed and complexed crystals formed were of similar size and shape. The crystals belong to the same space group and have very similar unit cell dimensions, a

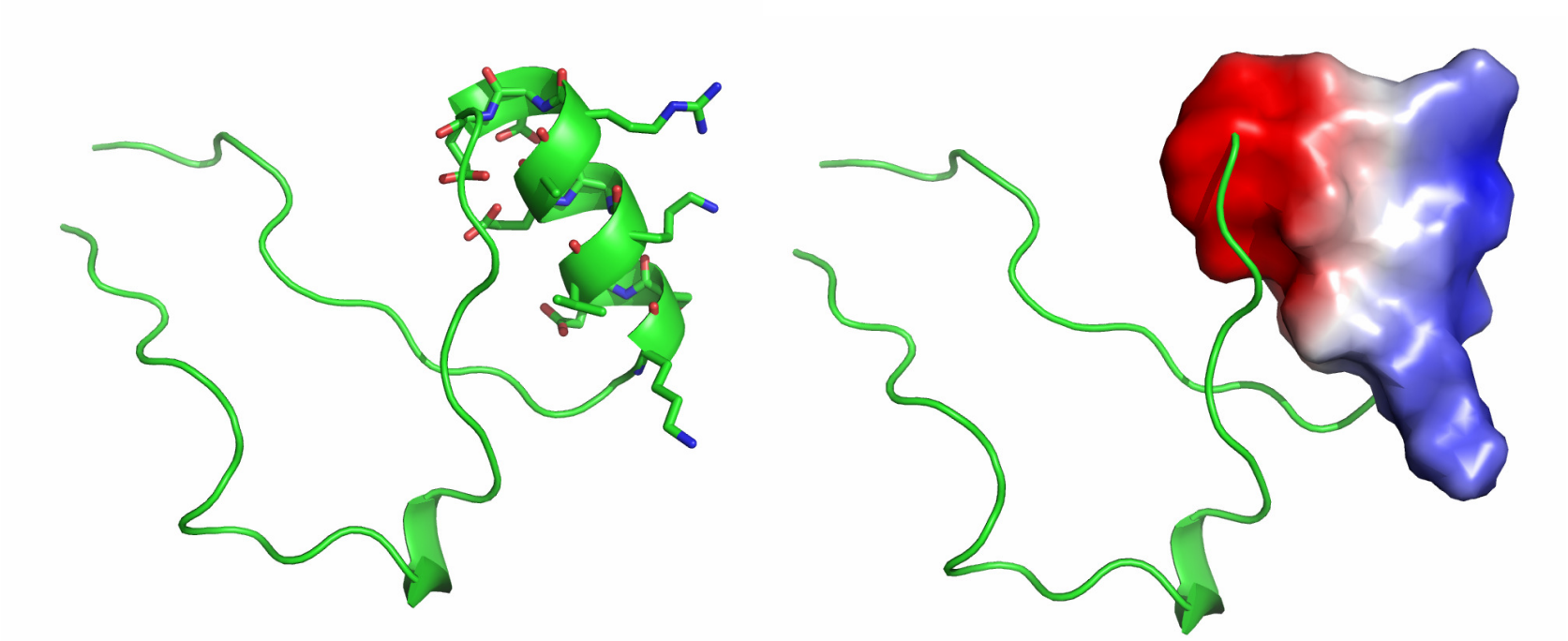


Figure 3-19. **The predicted loop domain of LmPCNA**
Figure showing the novel loop domain of LmPCNA. This loop has 41 residues in LmPCNA opposed to 7 residues in HsPCNA. The loop which may be flexible and shows a large quantity of random coiled region. The loop is determined to contain a helix (left), contains three positively charged residues and four negatively charged residues. These residues are positioned on opposing sides of the predicted helix giving a differential charge. The surface electrostatics of this differentially charged helix (right) show a clear positive charge on one side which may facilitate DNA association.

result of the crystal packing. A neighbouring asymmetric unit sits near the PIP box binding interface. As a result of peptide association to the PIP box binding interface, the LmPCNA hexamers will be pushed apart. The separation of LmPCNA hexamers resulted in a slight increase in the solvent content of the crystals, 75.4% for the un-complexed and 76.1% for the complexed. The resolution of the un-complexed and complexed protein crystals could be dramatically improved when the crystals were dehydrated (as discussed in section 3.1.3) from 7Å to 3Å.

The structure solutions of un-complexed and complexed LmPCNA were very similar showing the trimeric and potential hexameric structures. Monomers, trimers and hexamers were superposed using backbone atoms only. The superposition of the un-complexed and complexed LmPCNA monomers gave a 0.615Å R.M.S.D. superposition of the trimers gave a 0.902Å R.M.S.D. and superposition of the hexamers gave a 6.64Å R.M.S.D. (figure 3.20). The large increase in the R.M.S.D. of the hexamer alignment infers that this potential hexamer of LmPCNA may be able to associate in different orientations.

Both un-complexed and complexed structures have high B-factors and poor resolution as a result of the high solvent content. The un-complexed structure of LmPCNA gave an average B-factor of 84.3Å^2 for all 5884 backbone atoms. LmPCNA from the complexed structure gave an average B-factor of 81.3Å^2 for all 5992 backbone atoms. The average B-factors for different regions of the protein were calculated. Flexible regions of the protein like the N-terminus, C-terminus, elongated loop domain branch points and IDCL showed on average the highest B-factors (figure 3.21, table 3.5). These B-factors show reasonable similarity between the two LmPCNA structures. The complexed structure uniformly showed higher B- factors in

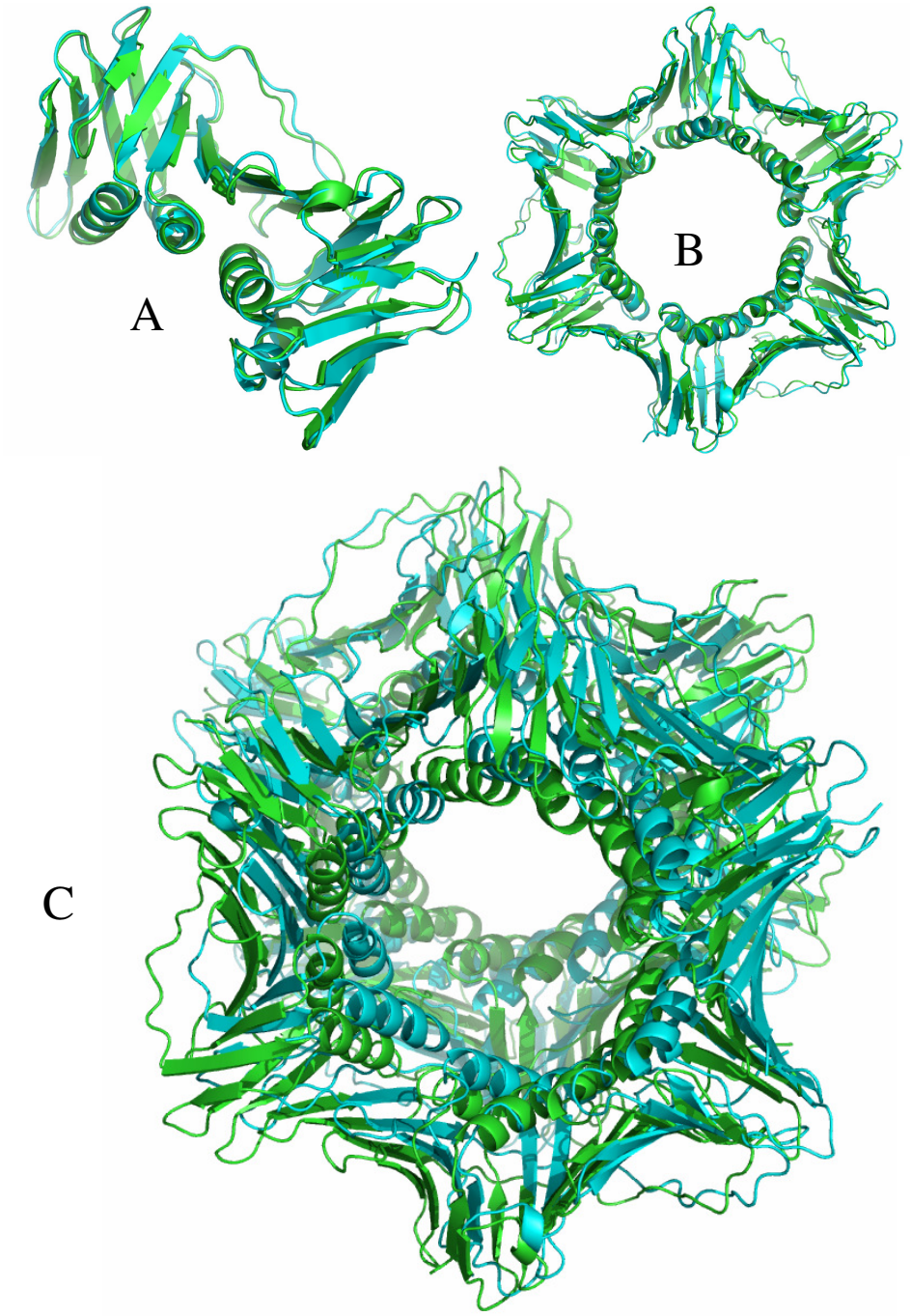


Figure 3-20. **The superposition of un-complexed and complexed LmPCNA monomers, trimer and hexamers**

Figure showing the superposition of LmPCNA monomers, trimers and hexamers of un-complexed (green) and complexed (blue) LmPCNA. A) The superposition of monomeric LmPCNA with an RMSD of 0.615Å. B) The superposition of trimeric LmPCNA with an RMSD of 0.902Å. C) The superposition of hexameric LmPCNA with an RMSD of 6.64Å. This shows little conservation of the hexameric form suggesting LmPCNA can exist in different hexameric orientations.

Protein region	B-factor of un-complexed structure (Å^2)	B-factor of complexed structure (Å^2)
Average of whole protein	84.3	87.6
N-terminus	98.3	100
α helicies	78.7	87.8
β sheets	96.7	98.0
IDCL	111	106
N-terminus of loop	98.3	94.2
C-terminus of loop	93.4	92.3
C-terminus	104	102

Table 3-5. **Table showing the average B-factors of regions of LmPCNA in un-complexed and complexed structures**
Table showing the B-factors of LmPCNA; in both un-complexed and complexed crystal structures. The B-factors of LmPCNA have been shown with regards to different regions of the protein with the greatest flexibility compared to the average secondary structure. It is shown that some regions demonstrate an increased B-factor in the complexed crystal structure e.g. N-terminus and secondary structures. The IDCL and C-terminus do show a reduced B-factor as a result of peptide binding. This is due to the involvement of these areas in the PIP-box binding site of LmPCNA which are stabilised by this interaction.

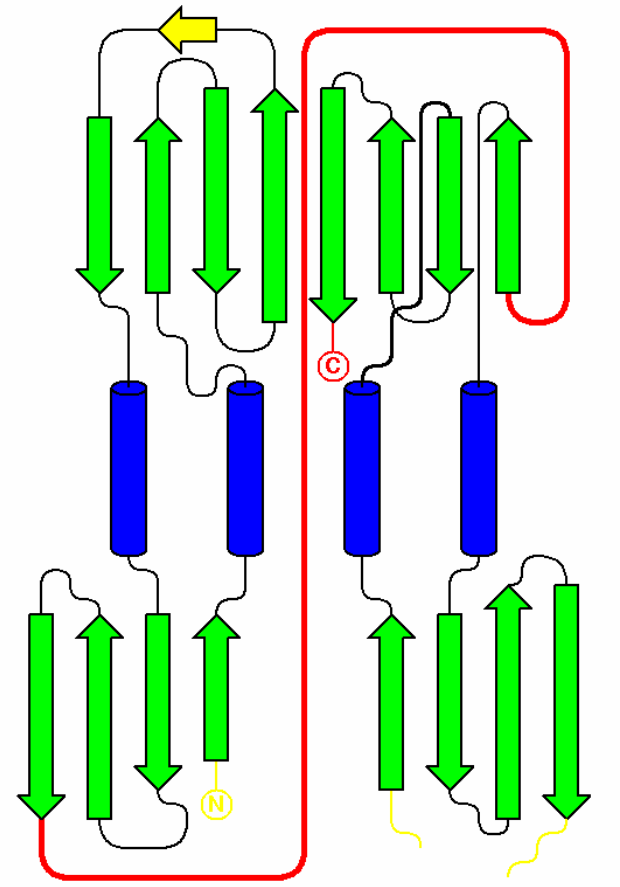


Figure 3-21. **Topology diagram showing areas with colour related to B-factor**

The topology diagram of the LmPCNA monomer with different areas (noted in table 3.5) coloured by B-factor. The colours are (from high) red, yellow, green and blue (to low). The IDCL and C-terminus show the greatest B-factor and it is these areas which make up the PIP box binding interface. The N-terminus and elongated loop braches have high B-factors and are solvent exposed. The helical regions are the most highly supported structures and show the lowest B-factor.

these flexible regions than that of the un-complexed structure. This increase in B-factor is likely the result of poorer resolution in the complexed structure. The IDCL and C-terminus had a slightly reduced B-factor in the complexed structure. This is the result of the peptide binding the PIP box binding interface which is formed by the IDCL and C-terminus. This binding event stabilised the IDCL and C-terminus giving the reduced B-factor in the complexed structure.

3.2.5 Structural analysis of the PIP-box binding interface

The PCNA PIP box binding interface (described in section 1.2.1) is comprised of 30 residues. These residues are required for association to the PIP-box sequence of partner proteins e.g. DNA polymerase. The interaction with partners involves both hydrophobic interactions and hydrogen bond interactions. These thirty PIP box binding residues of LmPCNA have a sequence identity of 58% (84% similarity) (figure 3.12B) to the equivalent residues in HsPCNA. The sequence identity between LmPCNA and HsPCNA over 261 residues is of 38% suggesting higher conservation at the PIP box binding interface.

LmPCNA can bind the PIP-box sequence, demonstrated in biochemical studies chapters 6, 7 and 8. In this work, we have studied crystallographically and biochemically the LmPCNA complex formation with *H. sapiens* p21 (141-152). The conservation of this binding activity is due to the conservation of residues at the PIP box binding interface (figure 3.12B).

The crystal structure of HsPCNA complexed with H-sapiens p21 (143-160) 1AXC (Gulbis, J. M. *et al* 1996) showed out of the thirty PIP-box binding residues twelve were involved in forming hydrogen bonding contacts. The twelve residues involved in the hydrogen bonding of HsPCNA to the PIP-box are highly conserved

(figure 3.12B). Residues involved in hydrogen bonding at the PIP box binding interface have a sequence identity of 50% (75% similarity) with equivalent residues in HsPCNA. The level of sequence identity in the residues forming hydrogen bonds is no higher than the average of the PIP box binding interface (figure 3.12B).

The structures of LmPCNA have been solved with and without peptide bound. Figure 3.22 shows the residues which make up the binding interface in stick form overlaid on a cartoon representation. The PIP-box binding interface is formed from the IDCL, central β-sheet, C-terminus and two loop regions (highlighted in figure 3.22). The PIP-box binding interface is shown with peptide bound. The binding interaction of *H. sapiens* p21 (141-152) shows a helix at the hydrophobic pocket (figure 3.22). This helix is a conserved feature in SpPCNA and HsPCNA (Gulbis, J. M. *et al* 1996).

There are small structural re-arrangements of the LmPCNA PIP-box binding interface upon association of *H. sapiens* p21 (141-152). The bulk of these small rearrangements are the movement of amino acid side chains away from the hydrophobic pocket e.g. H44, M126, P129, D266 (figure 3.23). The C-terminus has previously been determined to be important for PCNA binding activity (Fukuda, K. *et al* 1995). The solution of LmPCNA complexed to a PIP box peptide showed a significant structural rearrangement in the C-terminus as shown in figure 3.23, the Cα of D266 shifts over 8Å.

Crystal structures of HsPCNA, SpPCNA and LmPCNA bound to *H. sapiens* p21 have been solved. HsPCNA has been solved with p21 (143-160) (Gulbis, J. M. *et al* 1996), SpPCNA has been solved with p21 (142-159) (unpublished data) and LmPCNA has been solved with p21 (142-152) (this work). The main PIP box

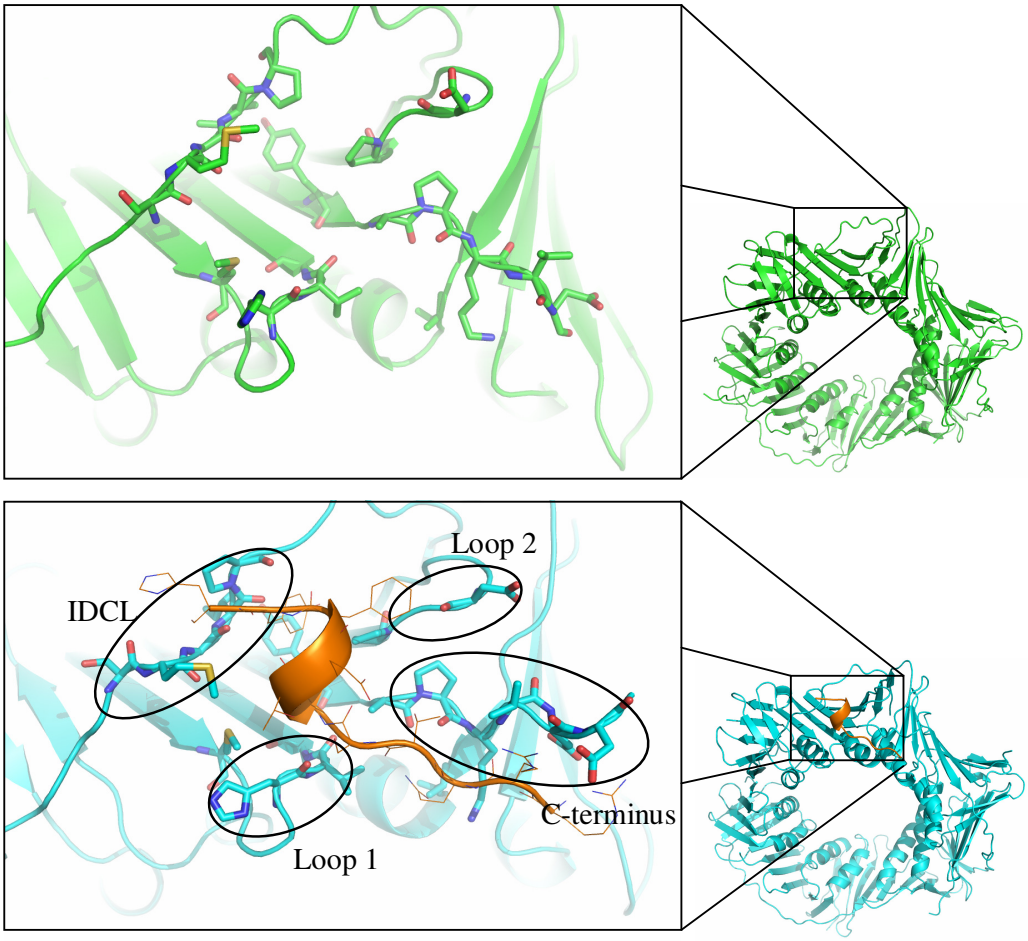


Figure 3-22. **The PIP box binding interface of LmPCNA in structures determined both complexed and un-complexed forms**
Figure showing the PIP box binding interface of LmPCNA from two crystal structures. The green structure (top) is the PIP box binding interface of the un-complexed form of LmPCNA. Residues involved in the interaction with p21 (143-152) are shown in stick form. The blue structure (bottom) is the crystal structure complexed with p21 (143-152). Residues involved in the PIP box binding interface are shown in stick form and the p21 (143-152) peptide is shown in orange.

HsPCNA (pep-PCNA)	SpPCNA (pep-PCNA)	LmPCNA (pep-PCNA)
R143-I252	R143(CO)-I254(N)	R143(εN)-D291
Q144-A252	R143(N)-I254(CO)	R143(CO)-V289
T145(εO)-P253	Q144-A251	E144-K288
T145(CO)-P253	Q144-P252	T145(CO)-P287
M147-H44	T145-P252	T145(εO)-P287
Y151-N131	M147-H44	M147-H44
H152-G127	H152-G127	Y151-Y133

Table 3-6. **The hydrogen bonding network at the PCNA p21 (143-152) interface**
Table showing the residues involved in the hydrogen bonding to PCNA in the three species of interest. The first residue is the peptide position and the second is on the PCNA main chain.

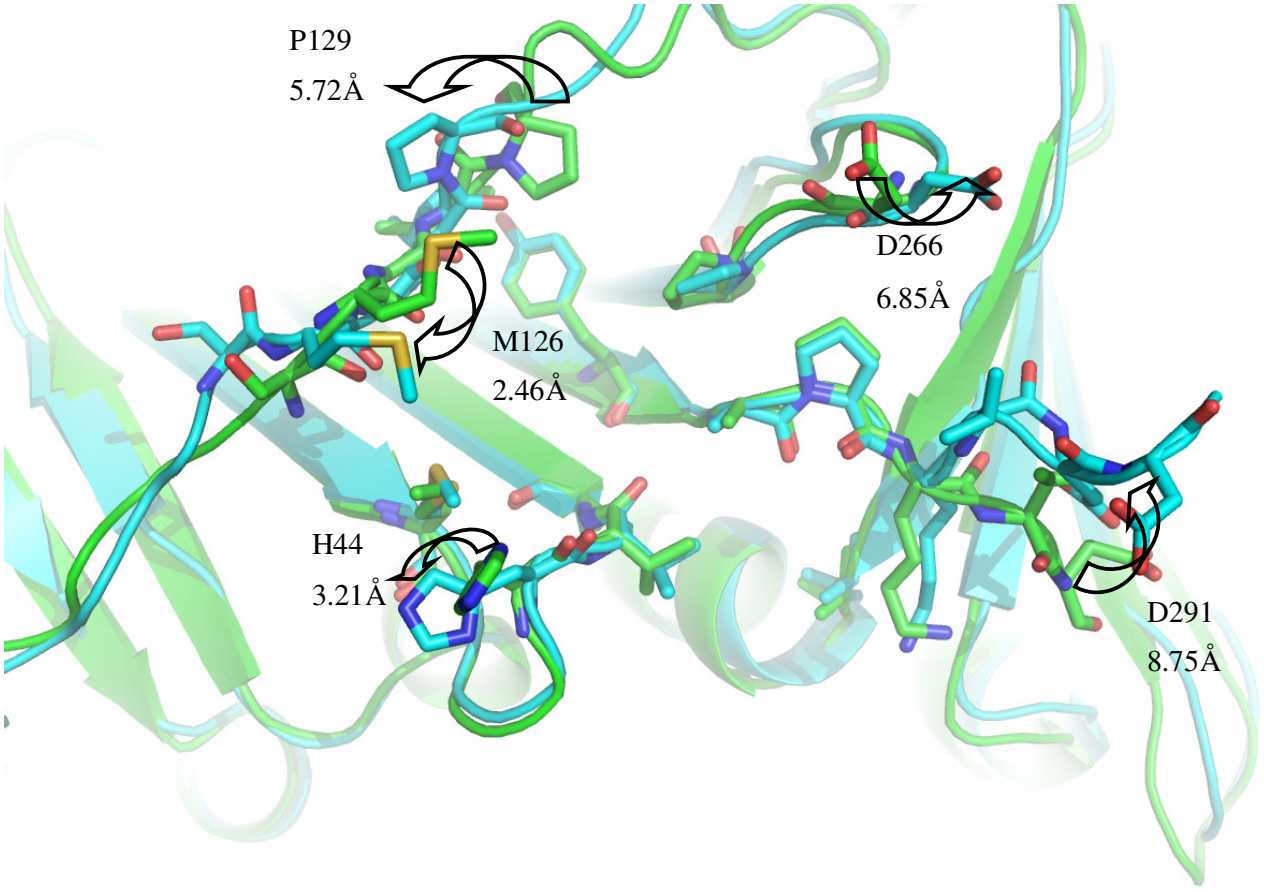


Figure 3-23. **The movement of amino acids in LmPCNA upon peptide binding**

Figure showing the changes in position of residues at the PIP box binding interface upon peptide binding. Comparison of the un-complexed crystal structure (green) and complexed crystal structure (blue) showing residues involved in the PIP box binding interface in stick form. The greatest deviations are observed at residues H44, M126, P129 and D266, the shift in the side chains of these residues are shown in Å. The C-terminus demonstrated a significant shift in position. The movement of the C-terminus was measured from Cα of D291.

sequence (Q-x-x-h-x-x-a-a) and PCNA is present in all (HsPCNA, SpPCNA and LmPCNA) co-crystal structures. The number of hydrogen bonds were analysed from structural information. Between the PCNA peptide (143-152) interaction seven hydrogen bonds were formed (figure 3.24), this was consistent in PCNA from all species (table 3.6).

The hydrophobic pocket of the PIP box binding interface of PCNA is conserved in HsPCNA, SpPCNA and LmPCNA. A comparison of the thirty residues making up the PIP-box binding interfaces of the three PCNAs was made by superposition of the backbone and side chain atoms. Superposition of 97 atoms from HsPCNA and SpPCNA had the smallest R.M.S.D. (0.61Å). The R.M.S.D. of the PIP box binding interface of LmPCNA and SpPCNA is the greatest (2.00Å). This large increase in R.M.S.D. may be a result of the flexibility of the IDCL at poor resolutions.

The structures of PCNA p21(143-152) complex were analysed by PISA. This gave predicted values of buried surface area and ΔG of formation (table 3.3). PISA analysis showed a buried surface area upon PCNA p21 (143-152) interaction of 1200, 1240 and 1450 $Å^2$ for HsPCNA, SpPCNA and LmPCNA respectively (table 3.3). The predicted ΔG of formation upon PCNA p21 (143-152) interaction of -37.7, -41.8 and -51.5 kJ mol^{-1} for HsPCNA, SpPCNA and LmPCNA respectively (table 3.3). The values returned from the PISA analysis show that the computed buried surface area and ΔG of formation follow similar trends. The values infer that the LmPCNA-peptide interaction may be the tightest with HsPCNA and SpPCNA sharing similar affinities. The values determined by PISA are predicted from static structures and these associations have been analysed in solution (chapters 6, 7 and 8).

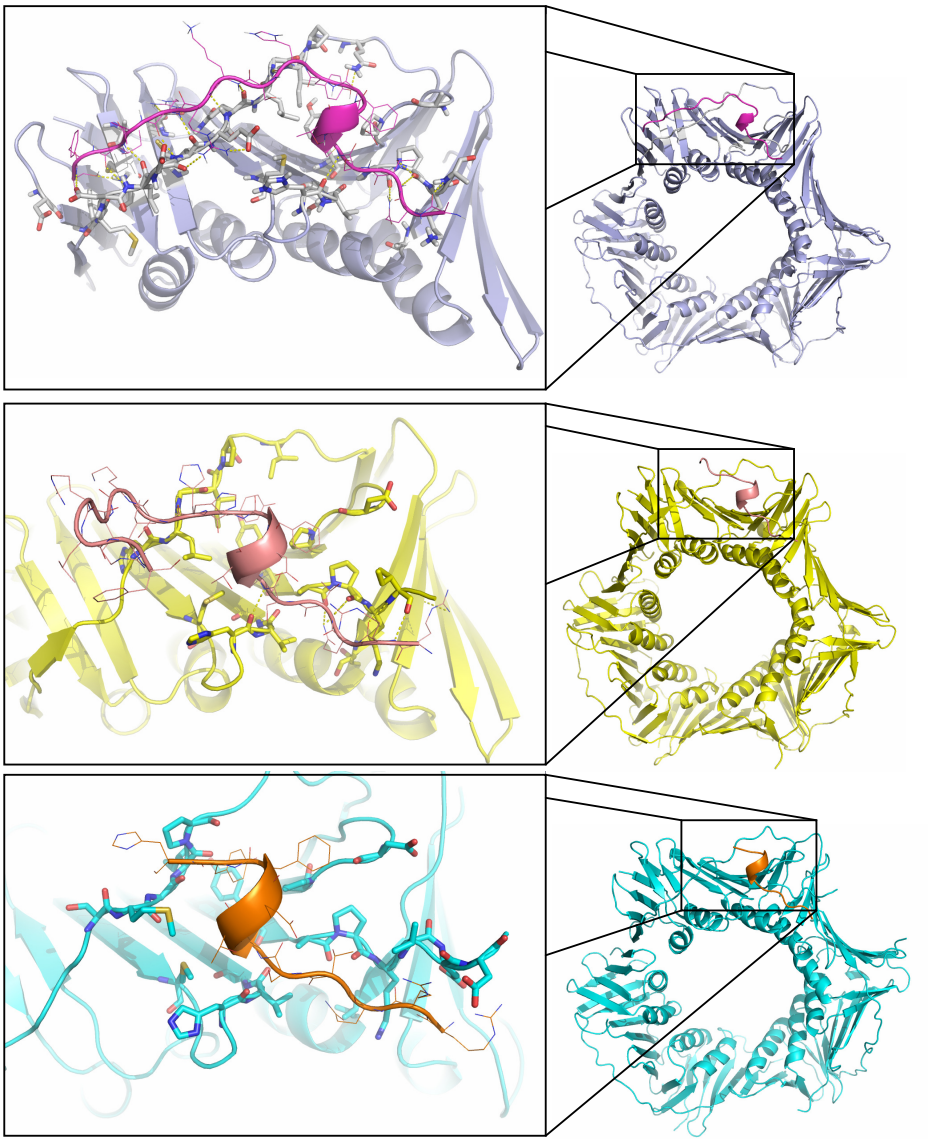


Figure 3-24. **The PIP box binding interfaces of HsPCNA, SpPCNA and LmPCNA.**
Figure showing the binding interfaces of PCNA from *H. sapiens* (top) (Gulbis, J. M. *et al* 1996), *S. pombe* (centre) (unpublished data) and *L. major* (bottom). These three crystallographic structures show the position of the p21 peptide of varying lengths bound to HsPCNA (magenta) SpPCNA (pink) and LmPCNA (orange). The binding interactions from all these crystal structures share the p21 (143-152) sequence containing the PIP box. This area of the peptide forms the helical region of the peptide and is a conserved feature in the interaction of all three PCNA structures. The interaction of the p21 (143-152) peptide forms 7 hydrogen bonds in all determined structures.

CHAPTER 4. The oligomeric state of PCNA

In this chapter the functional oligomeric state of PCNA is discussed. The oligomeric state can be difficult to determine as all subunits are identical and crystallographic data may not be a genuine representation of the protein in-vivo. The oligomeric state observed by crystallographic analysis may be imposed due to the high concentration of protein within a crystal.

Solution studies of the size of proteins can give information on a more genuine interaction. Solution methods include size exclusion chromatography, dynamic light scattering and chemical cross-linking. These methods have been implemented to give information about the potential oligomeric state of LmPCNA. These methods each rely on different facets of protein oligomeric states. Factors relating to size, mass, protein interface surface and direct chemical contacts can present information of the oligomeric state (Harding, S. E. 1995). The crystal structure of LmPCNA presented a potential hexameric state. This could be a crystallographic artefact and not the result of a biologically relevant interaction; alternatively this could be a mechanism of auto-regulation or providing a different platform for PCNA-protein interactions.

4.1 The potential oligomeric states of PCNA

PCNA is presented as a ring structure in all homologues studied as discussed in section 1.2. This ring structure has differing oligomeric states in different species, commonly a trimeric ring in eukaryotes (Gulbis, J. M. *et al* 1996; Shamoo, Y. and Steitz, T. A. 1999) or a dimer in prokaryotes (Kong, X. P. *et al* 1992). Drosophila PCNA has been argued to exist in a hexameric structure as a back-to-back double trimer (Henderson, D. S. *et al* 2000). This was later demonstrated in HsPCNA by

activity studies and chemical cross-linking studies (Naryzhny, S. N. *et al* 2005) (discussed below in section 5.1.1). This could give information regarding the actions of HsPCNA in-vivo, facilitating the coordination of multiple PCNA binding partners.

4.1.1 The potential oligomeric states of HsPCNA

Human PCNA is a trimeric ring structure as first shown in cryo-electron microscopy studies (Gogol, E. P. *et al* 1992). There is an ambiguity in the literature about the functional oligomeric state of PCNA. A previously published crystal structure of HsPCNA (Kontopidis, G. *et al* 2005) demonstrates an apparent hexameric structure (figure 4.1A); this has been determined to be a weak, but genuine interaction (Naryzhny, S. N. *et al* 2005). The hexamerisation of HsPCNA allows the binding interfaces (the front) to be exposed (figure 4.1A). It is believed that this can allow HsPCNA to function in both directions and facilitate greater co-ordination of DNA damage repair mechanisms. HsPCNA has been previously determined to bind a single co-factor per trimer, this was shown by competition with DNA Ligase I (Pascal, J. M. *et al* 2004). DNA Ligase I, a protein with a molecular weight of ~100kDa, totally encircles DNA occluding all other binding interfaces of PCNA. Native HsPCNA molecules have been demonstrated to co-purify with two PCNA binding partners in (Naryzhny, S. N. *et al* 2005), supporting the theory of a back-to-back hexameric structure. The existence of the hexameric structure of HsPCNA would allow greater speculation about function of PCNA within the cell.

4.1.2 The potential oligomeric states of LmPCNA

The oligomeric state of LmPCNA is unknown and had not been studied before this work. The X-ray crystal studies of LmPCNA clearly demonstrate a

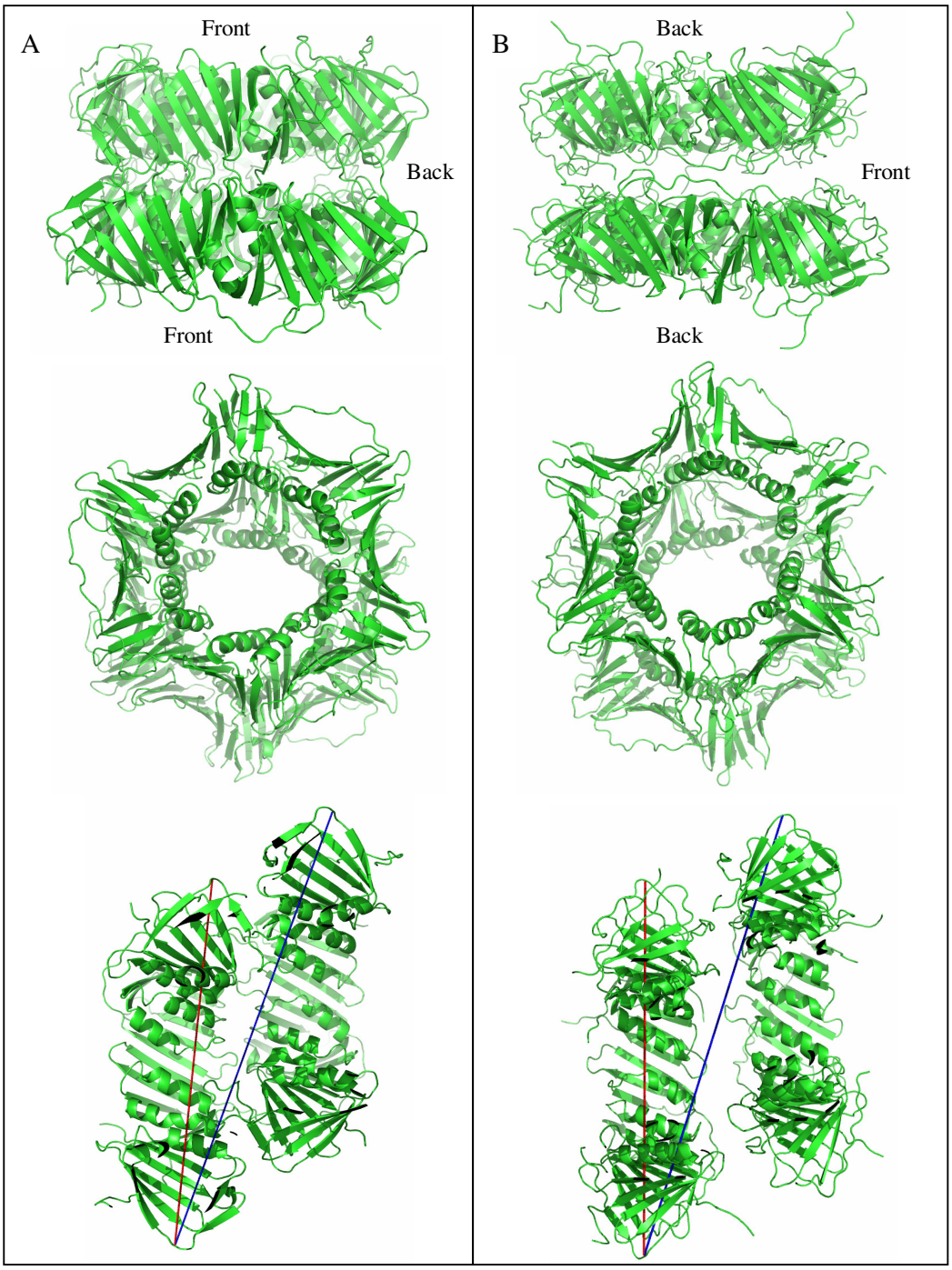


Figure 4-1. **The potential hexameric structures of HsPCNA and LmPCNA**

A) The alignment of the HsPCNA trimeric rings, top; the view along the axis of the DNA, middle; the side view shown in slab mode with the measurement of the diameter of the trimer and potential hexamer as 90.0Å and 111Å respectively, bottom.
B) The alignment of the LmPCNA trimeric rings, top; the view along the axis of the DNA middle; the side view shown in slab mode with the measurement of the diameter of the trimer and potential hexamer as 89.6Å and 110Å respectively, bottom.

trimeric ring structure, but also present a doublet of trimers in a hexameric state (figure 4.1B) similar to that present in HsPCNA (Kontopidis, G. *et al* 2005). Within the crystal, this could be an energetically favourable form of packing or a pre-existing hexamer. The question is therefore whether this is a genuine trimer-trimer binding interface or a crystallographic artefact. It should be noted that the trimer-trimer interaction in HsPCNA has a back to back configuration (figure 4.1A top), presenting the binding interfaces to the solvent, whereas the LmPCNA trimer-trimer interaction is face to face (figure 4.1B top) partially occluding the binding interface.

4.2 The use of PISA to analyse oligomeric state

The crystal structure once determined can be used to derive the validity of the trimer-trimer interaction. Bioinformatic analysis of X-ray crystal structure protein assemblies calculates buried surface area, charged states of the buried surfaces and the size of the final complex as opposed to the individual subunits. Balanced against these factors is the observation that there is an increased instability of larger protein complexes due to increasing entropic cost (Jones, S. and Thornton, J. M. 1996). The protein contacts of a test set of protein complexes have been analysed by quantifying molecular interactions of a number of protein interfaces and these have been scored to provide viability scores (Krissinel, E. and Henrick, K. 2007b). The molecular contacts are estimated from buried and free surface areas of the structure. The entropic terms of the structure are a sum of the translational entropy, rotational entropy, vibrational entropy and side-chain entropy at the binding interface (Krissinel, E. and Henrick, K. 2007b). The equations used to estimate these values from crystal structure data are given in (Krissinel, E. and Henrick, K. 2007b). In broad terms, the greater the increase in complex symmetry, the greater the entropic

cost and this is used to derive the Gibbs free energy of dissociation (Krissinel, E. and Henrick, K. 2007b). The validation of oligomeric state from crystal structures can be accomplished by the use of a computer program: Protein Interfaces, Surfaces and Assemblies (PISA) (Krissinel, E. and Henrick, K. 2007b). This program takes into account the chemical and physical properties of interactions to determine the viability of the protein interaction relative to the complex size. This program will return values of the surface area, buried area, Gibbs free energy change (ΔG) of complex formation, a ΔG of dissociation and an approximation of the stability of the complex. Complexes determined to be stable would show a negative ΔG for complex formation and a positive ΔG of complex dissociation. This program was shown to be 80-90% accurate for a test set of 218 previously solved proteins structures with experimentally determined oligomeric states (Krissinel, E. and Henrick, K. 2007b).

4.2.1 Size exclusion chromatography to analyse oligomeric state

As stated in chapter 2, size exclusion is, in general, a means of separating proteins by size. This separation is based on the Stokes radius of the protein and can be affected by not only the molecular weight of the protein, but other factors. As such, if the assumed molecular weight, based on size exclusion data, appears to have a greater molecular weight than the calculated, this could be the result of shape, buffer viscosity or oligomeric state (Dai, H. *et al* 1998).

A pure protein sample should demonstrate a smooth equidistant peak. This peak may broaden or split as a result of different oligomeric states in the protein sample. If the protein concentration is sufficiently high, all protein may appear in the higher order oligomeric state and result in a shift in the whole peak.

4.2.2 Dynamic light scattering as a means to analyse oligomeric state

Dynamic light scattering is a method of determining the Stokes radius of a macromolecule in solution. It relies on similar variables as that of size exclusion chromatography but is measured by a significantly different technique. This method measures the light scattered by a solution. Light is scattered by macromolecules, and the reflected light from these molecules may interfer with other reflections. Where two or more reflections combine, there can be constructive or destructive interference. The total intensity of scattered light will fluctuate as a result of the random movements of the particles in the solution. The light scatter is measured again at various time intervals. This intensity of light will differ from the original scatter as the time course proceeds. The rate in change of the scatter is the correlogram, which in dynamic light scattering can be used to give data on the sizes of the species present in the cuvette (Jachimska, B. *et al* 2008). Rapid changes in scatter are indicative of a solution containing species with a high diffusion constant and slow changes in scatter are indicative of a solution containing species with a low diffusion constant. The correlation function can also give data regarding the quantity of each different molecular weight species in the solution using the dynamic Zimm plot (Zimm, B. H. 1948).

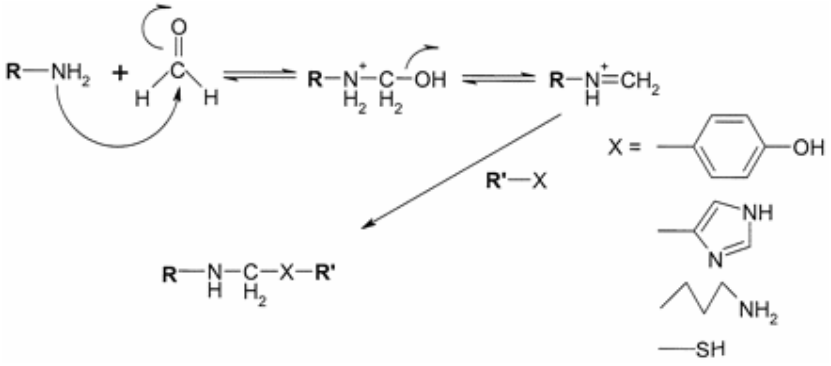
This method has similar strengths and weaknesses to those of size exclusion chromatography. These methods both determine the Stokes radius of the protein sample, and are dependent upon protein shape and buffer viscosity. Dynamic light scattering can also present errors in the predicted molecular weight of a protein as the samples Stokes radius is computationally calculated from the correlogram (Harding, S. E. 1995). Due to the increased scatter resulting from a larger particles, a sample

with equal concentrations of two species and order of magnitude difference in molecular weight will result in the larger species presenting a signal 10^6 times stronger. As a result, the presence of a large impurity, e.g. dust, will mask signal returned form the protein of interest.

4.2.3 Chemical cross-linking to analyse protein oligomeric state

Proteins form higher order oligomeric states by close protein interactions. The close association of two polypeptide chains can be stabilised by side group interactions as in section 5.1. As a result, protein samples can be covalently coupled by chemical cross linking. This will occur when protein functional groups are in close proximity. The chemical cross links formed by this method are not broken by the reducing environment of an SDS-PAGE gel (Melcher, K. 2004). The molecular weight shown by SDS-PAGE analysis will be a multiple of the monomer size.

This method is one of the most widely accepted methods of oligomeric state determination and can use a variety of chemical reagents that facilitate the cross-linking. One of the most common reagents is formaldehyde, which reacts with primary amines to form a Schiff base. The Schiff base can react with a variety of protein side-chains e.g. tyrosine, histidine, lysine and cysteine (Fujita, N. and Wade, P. A. 2004).



This results in an irreversible covalent bond suitable for separation using SDS-PAGE gels. Formaldehyde is also a short linker molecule resulting in the only interactions determined to be in close proximity (~2Å) (Orlando, V. 2000).

A common error of chemical cross-linking is the potential of poly-peptide monomers to associate due to the addition of the cross-linking reagent. Should protein solution be incubated for long periods of time with formaldehyde present, then it is possible for protein to form progressively higher order structures when analysed by SDS-PAGE gels. These higher order structures may not be a true representation of the native protein solution. The imposition of a higher order structure can be determined by use of a size exclusion column. Retention volumes of the native and non-cross-linked protein solution should remain similar if the protein oligomeric state is genuine.

4.3 Materials and methods of oligomeric state determination of LmPCNA

In the study of the oligomeric state of LmPCNA, the protein was purified as in chapter 2, and prior to any experiment, the sample was size excluded into the appropriate buffer using the Superdex S200 10/300. Should protein concentration be required, this was achieved using a Vivaspin 20 ultra filtration column (GE Healthcare) with a 10kDa molecular weight cut off. These columns were centrifuged at 3800rpm at 4°C in a bench top centrifuge. Protein concentrations were determined using the NanoDrop™ 2000 (Thermo Scientific) set at 280nm, assuming an extinction co-efficient of 20775 $M^{-1}cm^{-1}$. Analysis of protein interactions using PISA were performed as in (Krissinel, E. and Henrick, K. 2007a), accessed 8/8/07.

LmPCNA size exclusion analysis: The protein size exclusion analysis was performed as in section 2.3.2. The column was thoroughly equilibrated with buffer and run at 0.5ml min^{-1}, fractions were collected every minute.

LmPCNA dynamic light scattering analysis: Dynamic light scattering was performed on a series of LmPCNA protein concentrations, 0.025, 0.05, 0.1, 0.2, 0.5, 1.0 and 2.0mg ml^{-1}. This was performed at Malvern instruments by Andrew Jones on the 19/11/2007. This data was obtained using the Malvern ® Zetasizer® Nano ZS™, at 25°C in HEPES biochemistry buffer. The data was analysed using the Malvern Zetasizer Nano ZS version 5.02 software. All samples were filtered through a 0.2µm filter prior to transport and were stored at Malvern at -20°C until use, and data was returned as a printed report by mail.

Results obtained from Malvern were confirmed by the Wyatt DynaPro™ plate reader. LmPCNA in HEPES biochemistry buffer was concentrated to 2.0mg ml^{-1} and filtered through a 0.2µm filter. Light scattering was determined for this sample in a black 384 well plate (Greiner Bio-one) at 25°C with a laser wavelength of 828nm. The light scattering was measured and repeated every 0.5µs for approximately 7seconds, the scatter intensity returned for these measurements were then compared. This was then analysed to give a correlation co-efficient by the DYNAMICS V6™ version 6.9.1.15 assuming a sample refractive index and viscosity of water.

LmPCNA chemical cross-linking analysis: Protein samples for chemical cross-linking were concentrated to 1mg ml^{-1} in chemical cross-linking buffer (50mM HEPES, 0.1mM EDTA, pH 8.2). Protein solution had a 0.012M concentration of formaldehyde (Fisher) added, pipette mixed and was incubated for 2hours with

samples removed at 0s, 10s, 30s, 1min, 2min, 3min, 5min, 10min, 15min, 30min, 60min and 120min. These samples were quenched with SDS-PAGE loading buffer and run on a gel. This experiment determined the time required for the formation of higher order structures.

A 400µl sample of LmPCNA (1mg ml^{-1}) in chemical cross-linking buffer was split into two equal samples; sample one was chemically cross-linked by the addition of formaldehyde to a final concentration of 0.012M and incubated at 25°C for 10mins. This sample was then quenched by the addition of a 15 fold molar excess of primary amines from Tris to a final concentration of 0.18M, the sample was then incubated for a further 5mins at 25°C to couple the remaining formaldehyde molecules. This sample was then run over the Superdex S200 10/300 in pre-equilibrated LmPCNA crystallisation buffer. After this sample was eluted, the column was re-equilibrated in LmPCNA crystallisation buffer and the second sample was passed over the column. The eluted fractions were collected and analysed by 12% SDS-PAGE.

4.4 Oligomeric state determination of LmPCNA

The oligomeric state of LmPCNA has been determined by the PISA, size exclusion chromatography and chemical cross-linking. The results of these methods are described below. These experiments have demonstrated the potential oligomeric state of LmPCNA in-vitro.

4.4.1 Oligomeric state determination of LmPCNA by PISA

PISA analyses whole .pdb files and uses the solved asymmetric unit and known unit cell parameters to analyse all possible interfaces. The .pdb file generated

in chapter 4 was used as the input for the PISA analysis. This analysis returned values for the surface area and buried area (square Å), the Gibbs free energy change (ΔG) of complex formation and dissociation in kcal mol^{-1} and an approximation of the complex stability based on these values. Within the asymmetric unit, 6 monomers were present and formed two trimers. These two trimers, which are known to be a genuine interaction, were both found in the crystal structure and determined to be genuine (table 4.1). The two trimers of LmPCNA presented different values for the surface area, buried area, ΔG of complex formation and ΔG of complex dissociation. The differences were only slight for the surface area and buried area (figure 4.2), with an average calculated surface area of 35025 ±35 sq. Å (>0.1%) and an average calculated buried area of 4130 ±20 sq. Å (>0.5%). The calculated ΔG of formation presented an average of -15.2 ±1.25 kcal mol^{-1} (~8.22%) and the ΔG of dissociation presented an average of 2.1 ±1.4 kcal mol^{-1} (~66.7%). The program determined that both trimers were genuine complexes (table 4.1).

The hexamer (figure 4.1B) demonstrated a significantly greater buried surface area (10020sq.Å) (figure 4.2, table 4.1) and a reduced ΔG of formation (-33.4 kcal mol^{-1}) (table 4.1). This complex also showed a significantly greater surface area (68290sq.Å), and a reduced ΔG of dissociation (-2.0 kcal mol^{-1}) (table 4.1). The negative ΔG of dissociation was only slight and resulted in the hexameric state of the protein to fall within a grey area. As a result, the computational method determined that the LmPCNA hexamerisation may or may not be a genuine interaction (table 4.1).

The PISA analysis of the hexamer of HsPCNA has been determined to be genuine with a buried surface area (15980sq.Å) (table 4.1) and reduced ΔG of

Formula	Monomers	Stable	Surface area (sq.Å)	Buried area (sq.Å)	ΔG of formation	ΔG of Dissociation
The following complexes appear to be genuine						
A_3	ACG	Yes	35060	4150	-19.0	7.3
A_3	BDE	Yes	34990	4110	-15.0	2.4
The following complex falls in a grey region and may or may not be genuine						
A_6	ABCDEG	No	68290	10020 (Δ1760)	-33.4	-2.0
The following complexes are genuine in HsPCNA structures						
A_3	A_3	Yes	34900	4750	-16.0	6.3
A_6	A_3B_3	Yes	61670	15980 (Δ6480)	-18.8	1.3

Table 4-1. **The energetics and binding areas of potential LmPCNA complex formations**
Table showing a brief account of the results obtained from an analysis of the protein structure, as determined by PISA. This table shows the complex formula, the monomers of the crystal structure involved in the interaction, a determination of complex stability, the calculated surface area in sq.Å, the calculated buried area in sq.Å, the ΔG of formation and the ΔG of dissociation in kcal mol^{-1}. The first two complexes were determined to be genuine, as the trimeric rings present in the asymmetric unit. The third complex is a determination of the hexameric structure. It could not be determined if this complex was genuine or not. The lower panel shows the values determined for the trimeric and hexameric structures.

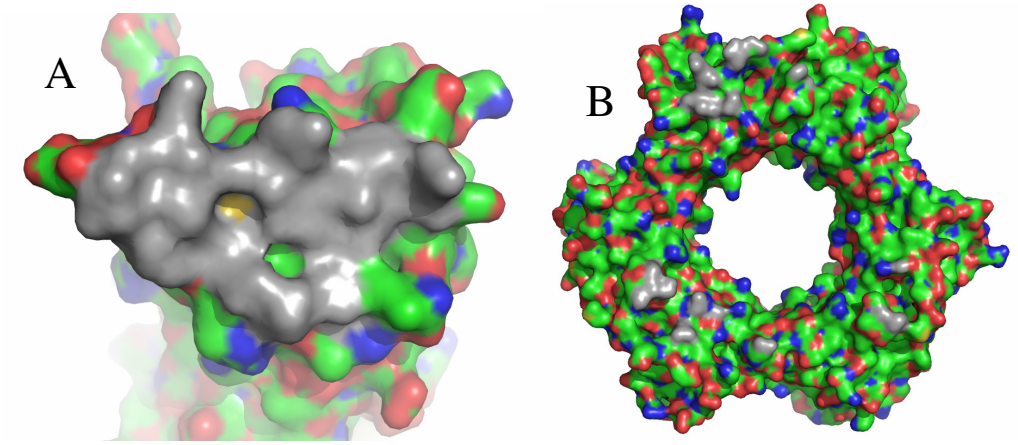


Figure 4-2. **The buried surfaces areas at the monomer-monomer and trimer-trimer interfaces**
Buried surfaces areas at the monomer-monomer interface (A) and the trimer-trimer interface (B). Both figures show carbon (green), nitrogen (blue), oxygen (red), sulphur (yellow) and the buried atoms (grey). These surfaces show the areas which make up the buried surfaces areas of LmPCNA in the crystal structure and give the values shown in table 4.1.

formation (-18.8 kcal mol^{-1}) (table 4.1). This complex also showed a significantly greater surface area (61670sq.Å), and a reduced ΔG of dissociation (1.3 kcal mol^{-1}) (table 4.1). The HsPCNA hexamer is determined to be genuine when no determination could be made regarding the LmPCNA hexamer. The HsPCNA interaction is deemed to be genuine due to the greater buried surface area at the trimer-trimer interface (6480sq.Å, an increase in 4720sq.Å) and the positive ΔG of dissociation (1.3 kcal mol^{-1}) (table 4.1).

4.4.2 Oligomeric state determination of LmPCNA by size exclusion

Size exclusion analysis of LmPCNA was performed during purification, section 2.3.2. This demonstrated that in HEPES buffer, the protein appeared to have a Stokes radius of a globular protein ~94.3kDa, the approximate molecular weight of the trimer. When analysed in Tris buffer, LmPCNA appeared to have a Stokes radius of a globular protein ~200kDa, the approximate molecular weight of the putative hexamer (figure 4.1B). This could indicate that the hexamer formation is condition-dependent, requiring different concentrations of ionic strengths or protein. This analysis, however, agrees with that of the computational method and determines that the LmPCNA hexamerisation may or may not be a genuine interaction.

4.4.3 Oligomeric state determination of LmPCNA by dynamic light scattering

LmPCNA was analysed by dynamic light scattering using a variety of similar equipment from different companies. This method gave the determined hydrodynamic diameter, double the Stokes radius, and the approximate molecular weight for a globular protein of this size. The two systems used returned results showing a good agreement. Protein samples were analysed at a concentration series

(from 0.025 to 2.0 mg ml^{-1}) in HEPES buffer. Protein concentration gave no significant change in hydrodynamic diameter and returned an average hydrodynamic diameter of 106Å over all concentrations (figure 4.3). This is the approximate diameter of a 165kDa globular protein, the approximate mass of a pentamer which can be approximated to the hexameric structure. The determined diameter of LmPCNA, from the crystal structure for the trimer is ~89.6Å, and for the hexamer is ~110Å (figure 4.1B). Thus, the determined diameter from dynamic light scattering is in close agreement with the hexameric structure. This data suggests that the hexameric structure is genuine; however this method is most sensitive if there is an order of magnitude shift in molecular weight which is not the case in this experiment.

4.4.4 Oligomeric state determination of LmPCNA by chemical cross-linking

The chemical cross-linking of LmPCNA was performed using formaldehyde solution as the cross-linking reagent. This method requires a 10 fold molar excess of formaldehyde to free amines in order to obtain cross-linking, so the number of free amines needs to be calculated from the sequence. This was determined to be 54. This reaction generated higher order structures at 25°C with 10minute incubation time after the addition of formaldehyde, as determined by SDS-PAGE gels. The formation of higher order structures is coupled with the reduction of monomeric protein bands at 35kDa. Less than 5 seconds after the addition of formaldehyde, two gel bands at a molecular weight lower than monomeric LmPCNA were formed. These gel bands are most likely due to internal cross-linking within the monomer and resulting in LmPCNA to exist in a more compact state, causing more rapid migration through the gel. No gel band appeared at the approximate size of the dimer; this could be the

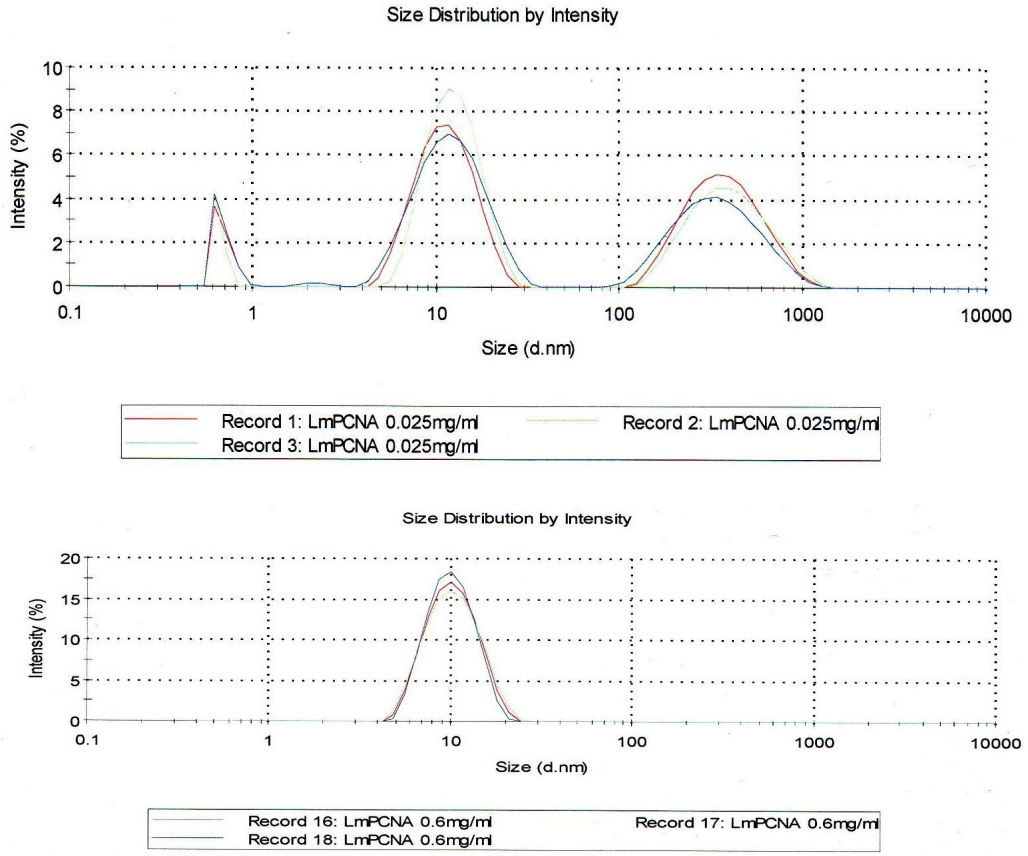


Figure 4-3. **The hydrodynamic diameter of species present in different concentration solutions of LmPCNA as determined by dynamic light scattering**

A chart showing the signal intensity generated from the correlogram resulting from the dynamic light scattering of a solution of LmPCNA. This solution was at concentrations of 0.025mg ml^{-1} and 0.6mg ml^{-1} and repeated three times each. Each repeat returns consistent values for the calculated hydrodynamic diameter of ~10nm (100Å). The low concentration of LmPCNA shows an increase in the proportion of large molecular weight impurities as the signal from the protein becomes less significant. A concentration series shows this trend to be repeated and gives an average hydrodynamic diameter of 10.6nm (106Å), approximately that of the LmPCNA hexamer. This has been correlated to a mass of a globular protein of 165kDa, similar to that of a hexameric LmPCNA.
Dynamic light scattering data provided by Melvern Instruments LTD.

result of the protein existing strongly in a trimeric form and the time required for a inter-monomer cross-link to form would result in a larger proportion of molecules cross-linked to more than one other poly-peptides.

A sample was split into two and one was passed over the Superdex 200 size exclusion column and a peak was observed to elute at a volume of ~11.5ml. This is the native state of the protein in Tris buffer (figure 4.4B). The second portion of this same sample was chemically cross-linked as above. The cross-linked sample was passed over the Superdex 200 column and presented two elution peaks (figure 4.4C). The peaks eluted at 11.1ml and 13.1ml, and when run on a SDS-PAGE gel, the peak eluted at 11.1ml corresponded to bands of ~200kDa and ~100kDa, which consist of hexameric and trimeric LmPCNA respectively. This data is in agreement with (Naryzhny, S. N. *et al* 2005). The peak eluted at 13.1ml, when run on a SDS-PAGE gel, corresponded to bands of ~35kDa and ~32kDa (figure 4.4A), these bands consist of the monomer and cross-linked monomer LmPCNA. Change in the elution volume from 11.5ml to 11.1ml could be the result of the LmPCNA monomers being cross-linked and presenting increased rigidity leading to an increased Stokes radius. The chemical cross-linking analysis of LmPCNA demonstrates the clear formation of hexameric state protein eluting at the same volume as the native, confirming that this is a genuine interaction in the environment of this experiment.

4.5 Conclusions from the oligomeric state determination of LmPCNA

These experiments clearly demonstrate that the hexamerisation of LmPCNA does occur to some extent in solution. Energetics determined by the PISA analysis of the crystal structure contacts showed that the hexamerisation falls in a grey area and may or may not be a genuine interaction. The requirement of the formation of a

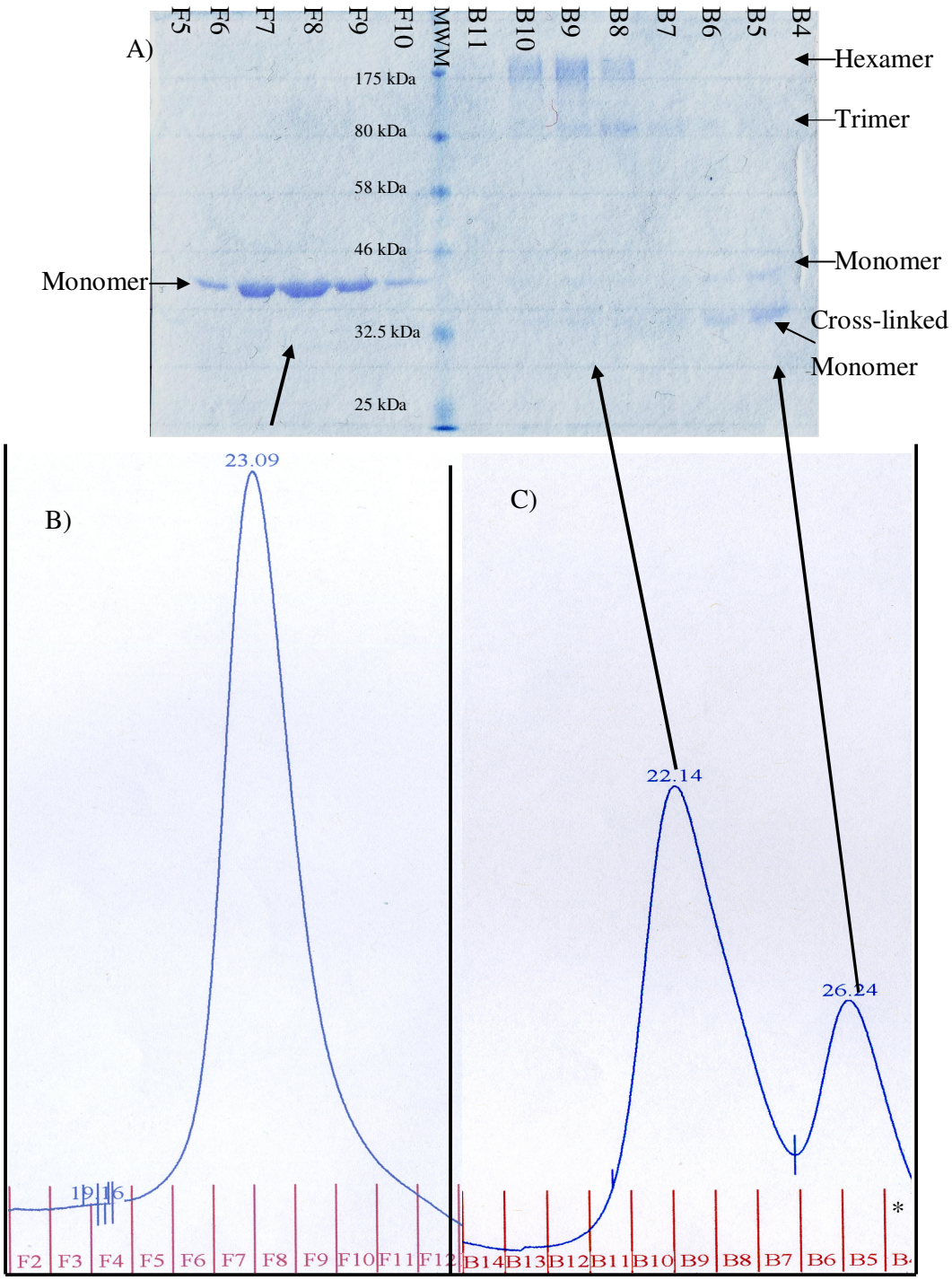


Figure 4-4. **Size exclusion and SDS-PAGE of native and cross-linked LmPCNA** A) 10% SDS-PAGE analysis of the size exclusion fractions (F5-F10 and B11-B4 indicated by a *). The first size exclusion peak demonstrates hexamers and trimers, the second demonstrates monomers and cross-linked monomers. As visualised by coomasie stain. B-C) The size exclusion of the native (B) and chemically cross-linked (C) LmPCNA at a 0.5ml min^{-1} flow rate, with the elution time shown above the elution peak. B) shows a single uniform peak C) eluting at two separate peaks.

genuine interface requires a large buried surface area and the nature of the surface gives a negative ΔG of formation and a positive ΔG of dissociation. The hexamer showed a slight negative ΔG of dissociation (-2.0 kcal mol^{-1}) however this technique has been determined to have a ± error of ~5.0 kcal mol^{-1} (Krissinel, E. and Henrick, K. 2007b). The trimer is shown to be a likely oligomeric state, however the hexameric state could not be determined to be genuine or false from this method.

The size exclusion shows LmPCNA eluting at two volumes in a buffer dependent fashion. In high salt buffer (150mM NaCl), LmPCNA elutes to an approximate volume of a 94.3kDa globular protein, the approximate molecular weight of the trimer. In low salt buffer (25mM NaCl) LmPCNA elutes at a volume of a 200kDa globular protein, the approximate molecular weight of the hexamer. When the PCNA peak was eluted from the size exclusion column there was no second peak so no approximate change in state was observed. A link between buffer ionic strength and PCNA hexamerisation has previously been observed in (Naryzhny, S. N. *et al* 2005).

Dynamic light scattering was performed in high salt buffer and returned a hydrodynamic radius of 106Å, the approximate diameter of a 165kDa globular protein, in agreement with a potential hexamer. These methods are most capable of distinguishing the mass of proteins that differ by an order of magnitude. It is possible that these effects upon the size exclusion and dynamic light scattering are the effect of the PCNA trimer existing as a ring structure with a central hole, as this can affect the diffusion rates. Loop regions present on LmPCNA can also affect the diffusion rate of this protein which could account for difference in the Stokes radius. These loops could increase the Stokes radius and result in an apparently higher oligomeric

state. As a result these methods may present potential errors which must be taken into account when determining the oligomeric state for a protein.

Chemical cross-linking has determined that the protein can form a hexameric state. This is a genuine interaction, present in the native sample, as determined by comparison of cross-linked and native size exclusion data. LmPCNA forms a hexameric structure when at 1mg ml^{-1}, in cross-linking buffer (low salt buffer). The link between buffer ionic strength and hexamerisation has been previously observed in (Naryzhny, S. N. *et al* 2005).

The potential oligomeric states of PCNA were determined with a range of PCNA concentrations in DLS and size exclusion analysis. Neither of these techniques showed any change with changing concentration which may be the result of low sensitivity of these methods. When LmPCNA was analysed by Dynamic Scanning Fluorimetry (DSF) (chapter 5), two increasing concentrations of LmPCNA gave a slight increase in the melting temperature of the protein (~3°C). This slight increase occurred around the 4μM concentration of LmPCNA, which infers a potential K_d of interaction of 4μM. Changes in the melting temperatures of HsPCNA and SpPCNA were not observed with changing concentration.

In-vivo it is possible that LmPCNA does not form a hexameric structure and remains in a trimeric state. This could be the result of increased ionic strength within the system, or the association of other proteins. The crystal structure presented demonstrates that this protein may not be in a functional state. The interaction interfaces shown in (Kontopidis, G. *et al* 2005; Naryzhny, S. N. *et al* 2005) determined that the "front" of the PCNA molecule are exposed in HsPCNA hexamer; in LmPCNA however these "front" sides are within the binding interface (figure 4.1

A and B). A potential reason for the hexamerisation in LmPCNA could be that a method of auto regulation prevents the effect of PCNA up-regulation or a novel PCNA-protein mode of interaction. This novel interaction could exist as a result of the flexible loop regions, which remains unresolved by X-ray crystallography.

CHAPTER 5. The study of PCNA by melting curve

The melting curve of a protein can provide information about structure and binding properties. Protein melting curves can be analysed by different methods, both direct and indirect. These methods involve the measurement of shift in secondary or tertiary structure with temperature. Direct methods involve the measurement of secondary structure in real-time for example with the use of circular dichroism (Chen, Y. H. *et al* 1974). Indirect methods involve a measurement of the shift in fluorescence quenching with an increase in temperature for example differential scanning fluorimetry (DSF) (Vedadi, M. *et al* 2006). The melting curve of protein complexes can also be used to determine ligand binding affinities (Ericsson, U. B. *et al* 2006; Lo, M. C. *et al* 2004).

5.1 Introduction to Circular Dichroism

Circular dichroism can be used to quantify the secondary structure of a protein in solution. This is achieved by determining the ellipticity of plane polarised light at different wavelengths. The shape of the resulting trace (e.g figure 5.2) contains information about the relative amounts of helix sheet and coil structures.

5.1.1 Principles of Circular Dichroism

Plane polarised light that is produced in the Circular dichroism photo-spectrometer can be regarded as a mixture of Left and Right (clockwise and anticlockwise) circularly polarised light. Practically, incident light is switched between the clockwise and anticlockwise components and measured separately, the beams are then amalgamated in-silico to give the ellipticity of the resultant beam (Kelly, S. M. *et al* 2005). Should light be absorbed differentially between clockwise

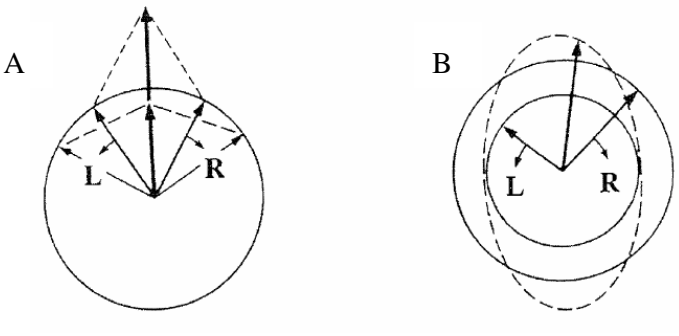


Figure 5-1. **Example of the formation of ellipticity**

A) Shows the formation of plane polarised light due to the constructive interference from both the left and right hand form of equal magnitude.
B) Shows the formation of ellipticity due to the disproportionate absorption of the left hand form opposed to the right, resulting in the reconstructed beam having a lower magnitude with a degree of ellipticity. From (Kelly, S. M. *et al* 2005)

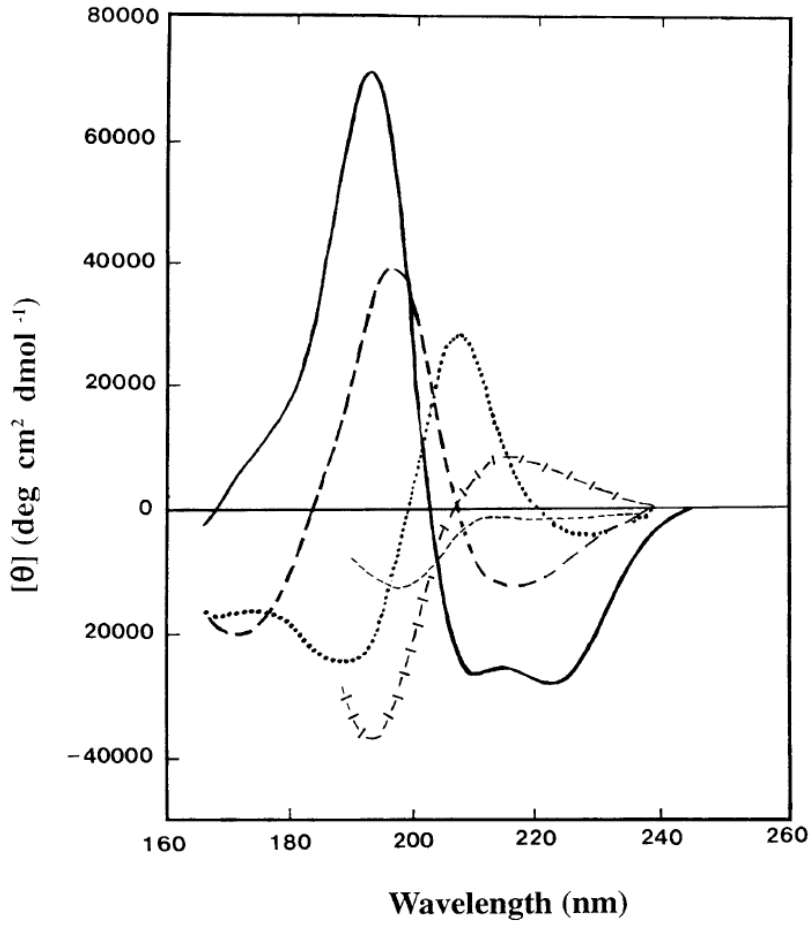


Figure 5-2. **Graph showing the components of the CD spectra**
Graph showing the ellipticity on the y-axis against the wavelength on the x-axis. Solid line, α-helix; long dashed line, anti-parallel β-sheet; dotted line, type1 β-turn; crossed dashed line, extended 3_{10}-helix or polyproline II helix; short dashed line, irregular structure. From (Kelly, S. M. *et al* 2005)

and anticlockwise components, then the re-constructed light will be as elliptically polarised (figure 5.1). The two beams of light are absorbed by the protein solution. If this absorption occurs due to protein precipitant or dust particles, then the signal will show no ellipticity as absorption will occur in both clockwise and anti-clockwise components equally.

The CD spectrum is obtained by scanning the wavelength of the incident light commonly from 300nm to 190nm. The absorption maxima of a protein solution by peptide bonds occurs at ~190nm (Goldfarb, A. R. and Saidel, L. J. 1951). The different amino acid side chains of a protein each have different absorption maxima (260-320nm) and disulphide bonds give a broad absorption maxima at ~260nm (Kelly, S. M. *et al* 2005). Absorption as a result of side chains and disulphide bonds are only a small component of the CD spectra. Each major type of secondary structure has a characteristic CD spectrum (figure 5.2). The absorption curves show the ellipticity against the wavelength of the measured light, of which all or some of these components will be present at differing levels in a CD spectra.

5.1.2 Deconvolution of the CD data

The CD spectra can be de-convoluted using various computer models hosted by online servers. Older programs that were in use did not require the use of a training set (a set of known CD spectra). These programs however are generally more unreliable. Two of the most reliable algorithms will be discussed here, SELCON and CDSSTR. The program SELCON (Sreerama, N. and Woody, R. W. 1993) is a self consistent model. This algorithm uses a reference set of CD spectra with known secondary structure. SELCON3 selects the reference set closest to the CD spectrum of interest and attributes the same characteristics initially. This is

termed the initial guess and these values are used for further calculations. The quantity of secondary structure is associated to the CD spectrum with the use of the singular value decomposition algorithm and the locally linearised model (Vanstokkum, I. H. M. *et al* 1990). For a model to be accepted using this method, the sum of all secondary structure components must be between 0.95 and 1.05 of the total protein; the quantity of each fraction must be greater than -0.025 and the RMSD of the calculated against the experimental spectra must be less than 0.25 $\Delta\varepsilon$ (change in ellipticity) (Dalmas, B. and Bannister, W. H. 1995). The fractions of each secondary structure component are then used as the model and the algorithms are run again. This program continues until there is a consistent result returned which has a sum of fractions between 0.90 and 1.10 (Sreerama, N. *et al* 1999).

CDSSTR (Johnson, W. C. 1999) is a program which uses training sets of CD spectra and fits the eight most relevant curves, thus preventing less relevant curves from influencing the results. These eight data sets are recombined to form the final data and matched to the query data set. The data will be accepted only if the sum of characteristics are between 0.952 and 1.05, no characteristic is <-0.03 and the RMSD to the query set is less than 0.25 $\Delta\varepsilon$ (Dalmas, B. and Bannister, W. H. 1995). This will present a varied number of combinations, composed of different training data. When the combined data are returned, CDSSTR modifies the quantity of α-helix as this value usually presents the greatest level of error. The average quantity of α-helix returned for this analysis allows the selection of the best training set combinations as described in (Johnson, W. C. 1999). Finally the accepted combination was used as the model, and as in (Sreerama, N. and Woody, R. W. 1993) and the SELCON program, the analysis was performed until the results became self-consistent.

5.1.3 Common errors associated to CD analysis of protein structure

CD spectra can be an unreliable method for secondary structure determination as there are numerous components to the shape of the spectra, as explained above. Inaccuracies in the CD spectrum arise by most systems which use a Xe arc light source, and this does not allow low wavelength measurements (<180nm) (Kelly, S. M. *et al* 2005). The lack of information in the CD spectrum results in only five components of the spectra to be deconvoluted. If the data is truncated to 190nm, only 3-4 components can be determined (Johnson, W. C. 1999). When each of the five individual spectra are recombined, the $\Delta\varepsilon$ RMSD of the fit is at its lowest (Hennessey, J. P. and Johnson, W. C. 1981). There are however four main secondary structure forms which which have the greatest contribution upon the CD spectra, and as such, all remaining features can still be compensated for with the final variable (Johnson, W. C. 1999).

5.1.4 Melting Curve Measurement Using Circular Dichroism

The measurement of the melting point of a protein by CD is possible due to the reduced level of secondary structure in a melted protein. The reduction of the ellipticity of a CD signal to ~0 at a specific wavelength is a result of all secondary structure being removed. A melting curve can be plotted by monitoring the ellipticity of the protein solution at a specific wavelength, commonly at 222nm or a wavelength with a large $\Delta\varepsilon$ signal from 0. With continual measurement of ellipticity at this wavelength, the protein solution is heated by the use of a cuvette thermal jacket and a Peltier system. A melting temperature can be determined from a plot of the temperature against the shift in ellipticity.

5.2 Materials and methods of protein analysis by CD

CD analysis was performed using a Jasco J-810 spectropolarimeter thoroughly flushed with nitrogen to increase signal stability and device lifetime. The temperature of the sample was controlled using a Peltier system Jasco PTC-4235. CD equipment was controlled using the Spectra manager version 1.53 software (Jasco).

PCNA was buffer-exchanged into CD buffer (10mM NaF, 10mM NaH_2PO_4, pH7.4) using a G-25 desalting column (Pharmacia). HsPCNA was analysed at 2.04 mg ml^{-1}, SpPCNA was analysed at 1.86 mg ml^{-1} and LmPCNA was analysed at 1.25 mg ml^{-1}. PCNA was placed into a 1mm path-length cuvette and allowed to equilibrate to 20°C for 30minutes. CD spectra were obtained from 300 to 175nm wavelength and the data was analysed using the DICHROWEB server (Whitmore, L. and Wallace, B. A. 2008) and the analysis program SELCON3 (Sreerama, N. and Woody, R. W. 1993) using a reference set 3 optimised for 185-240nm. For melting temperature determination, the samples were heated to 95°C at 0.5°C intervals. The CD signal at 222nm wavelength was recorded upon temperature stabilisation.

5.3 Results of CD measurements

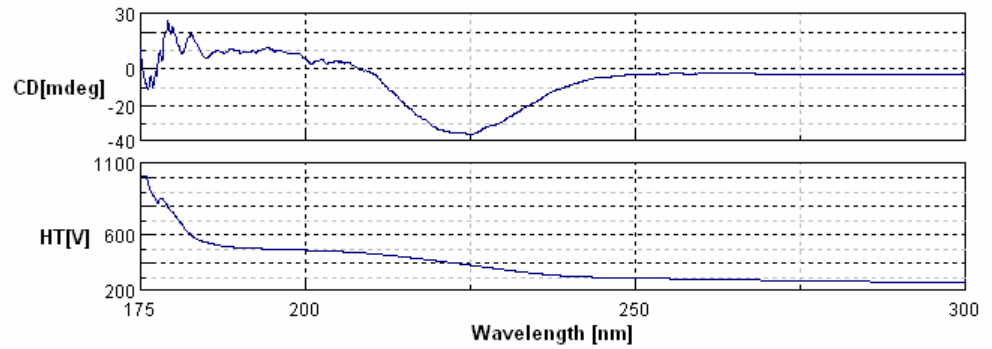
5.3.1 CD analysis of HsPCNA

SELCON analysis of the CD measurement of HsPCNA (figure 5.3A) returned values of 19.3% α-helix, 43.6% β-strand, 15.8% turns and 17.8% unordered. The secondary structure content determined by x-ray crystallography was 18% α-helix and 46% β-strand (Gulbis, J. M. *et al* 1996), showing a good correlation with the CD data obtained. The melting curve of HsPCNA (figure 5.3B) showed an increase in ellipticity of the sample with an increase in temperature. The ellipticity of the HsPCNA sample started increasing at 321K. Once the sample was heated to

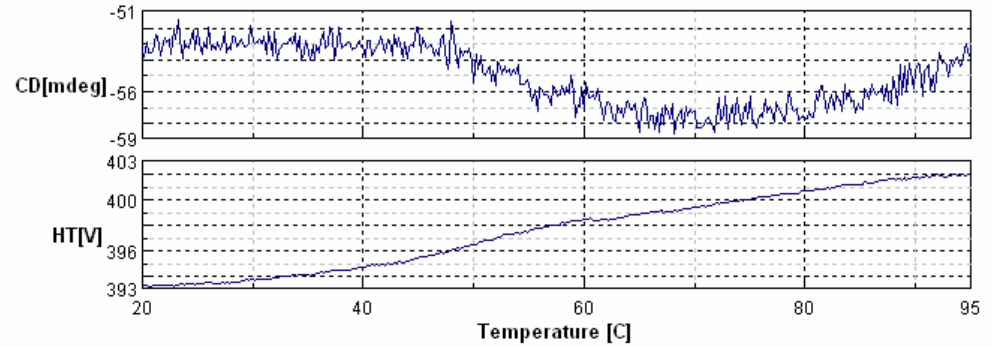
368K (95°C), the CD spectra was taken once again and this returned secondary structure values of 6.5% α-helix, 34.2% β-strand, 9.1% turns and 44.2% unordered showing a significant increase in the unordered portion, however maintaining a large portion of β-strand. The decrease in the ellipticity of the sample with increasing temperature is unexpected. A control experiment with L.mexicana Pyruvate kinase shows a more usual behaviour of a protein melting and the ellipticity at the melting temperature increasing to zero (figure 5.4). The further decrease in ellipticity obtained in HsPCNA may be related to the retention of β-strand structures.

5.3.2 CD analysis of SpPCNA

SELCON analysis of the CD measurement of SpPCNA (figure 5.5A) returned values of 5.4% α-helix, 42.5% β-strand, 24.9% turns and 25.4% unordered. The secondary structure content of S.cerevisiae PCNA (a close homologue of SpPCNA) determined by X-ray crystallography was 17% α-helix and 44% β-strand (Vijayakumar, S. *et al* 2007), showing a reasonable correlation with the CD data obtained. The melting curve of SpPCNA (figure 5.5B) showed an increase in ellipticity of the sample with an increase in temperature. The ellipticity of SpPCNA started to shift at 322K, demonstrating a similar stability to HsPCNA. Once the sample was heated to 368K, the CD spectrum was taken once again and this returned secondary structure values of 2.7% α-helix, 37.5% β-strand, 18.1% turns and 39.7% unordered, again showing a significant increase in the unordered portion, however maintaining a large portion of β-strand. The decrease in the ellipticity of the sample with increasing temperature is in contradiction with expected trends (figure 5.4), but is however is in agreement with the retention of β-strand. The maintenance of β-strand is consistent with data obtained for HsPCNA.



A)



B)

Figure 5-3. **The CD analysis of HsPCNA**

A) Top panel, the CD spectra of HsPCNA from 300 to 175nm wavelength. This panel shows increased noise in the signal at low wavelengths. Bottom panel, shows the laser voltage with a shift in wavelength. As the voltage increase above 600V, this is associated with high noise levels; data taken above this level is unreliable. B) Top panel, the CD in milli-degrees of HsPCNA at 222nm wavelength with increasing temperature, demonstrating an increase in ellipticity with increasing temperature. This shift occurs at 321K. Ellipticity of the sample decreased with higher temperatures due to the increased voltage demonstrated in the bottom panel.

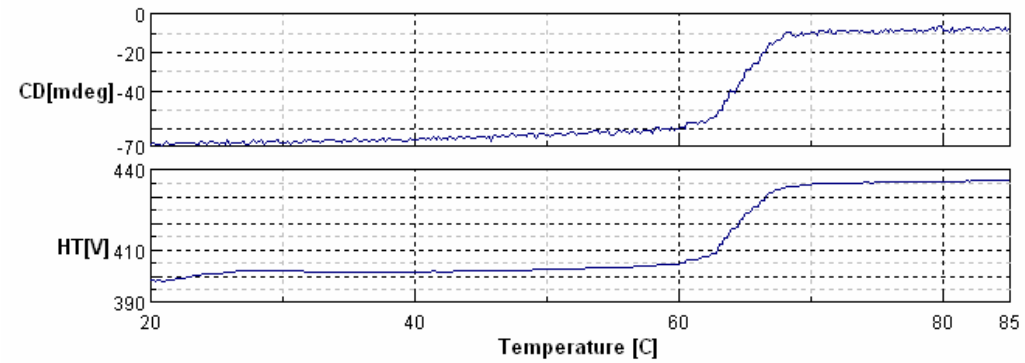
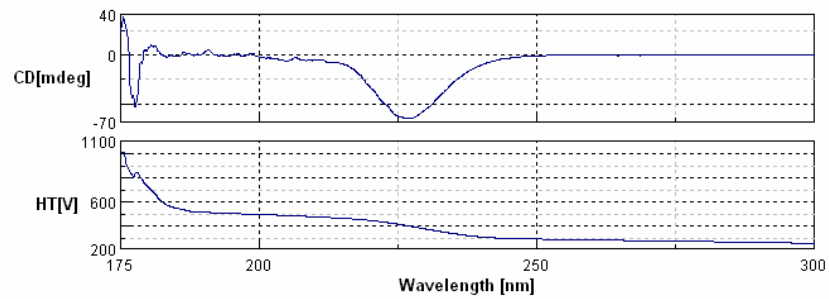
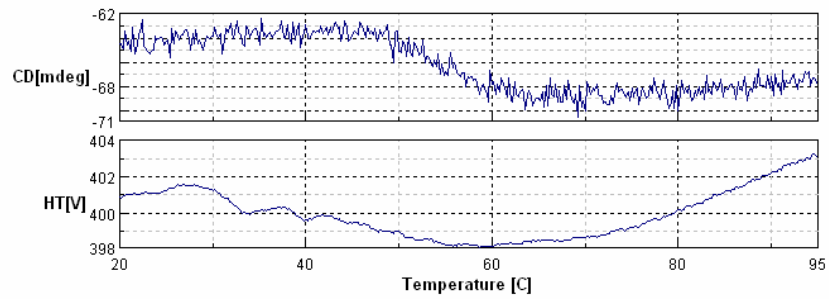


Figure 5-4. **The CD melting curve analysis of Pyruvate Kinase (PYK)**

Top panel, the CD in milli-degrees of PYK at 222nm wavelength with increasing temperature. This demonstrates a significant decrease in ellipticity with increasing temperature approaching 0mdeg, the expected result of CD analysis of protein melting. A significant increase in voltage was required, associated to protein melting, due to large precipitant formation in the sample cuvette.



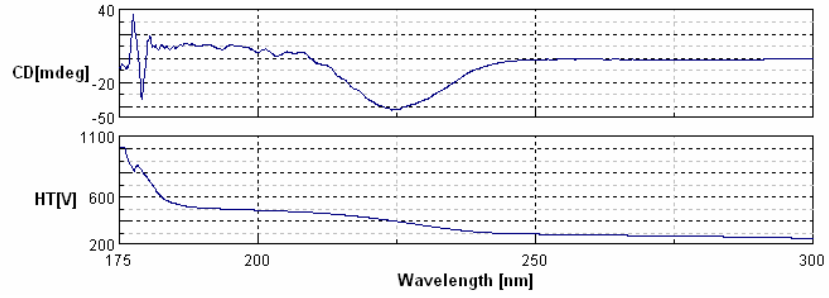
A)



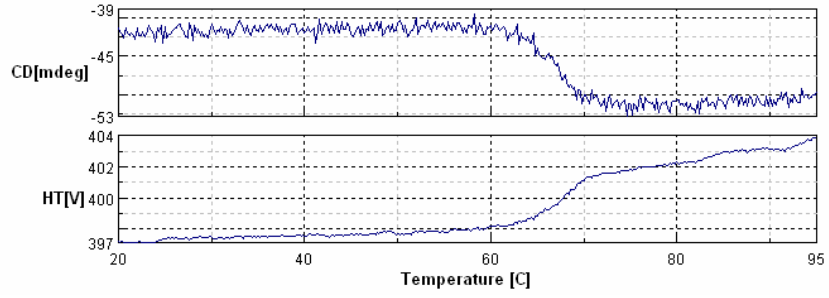
B)

Figure 5-5. **The CD analysis of SpPCNA**

A) Top panel, the CD spectra of SpPCNA from 300 to 175nm wavelength. Bottom panel, shows that the laser voltage increase is responsible for increased noise at lower wavelengths. B) Top panel, the CD of SpPCNA at 222nm wavelength with increasing temperature, demonstrating an increase in ellipticity with increasing temperature at 322K.



A)



B)

Figure 5-6. **The CD analysis of LmPCNA**

A) Top panel, the CD spectra of LmPCNA from 300 to 175nm wavelength. Bottom panel, shows that the laser voltage increase is responsible for increased noise at lower wavelengths. B) Top panel, the CD of LmPCNA at 222nm wavelength with increasing temperature, demonstrating an increase in ellipticity with increasing temperature at 335K.

5.3.3 CD analysis of LmPCNA

SELCON analysis of the CD measurement of LmPCNA (figure 5.6A) returned values of 5.8% α-helix, 37.4% β-strand, 21% turns and 34% unordered. The secondary structure content determined by X-ray crystallography was 16% α-helix and 33% β-strand (unpublished data), showing a reasonable correlation with the CD data obtained. The melting curve of LmPCNA (figure 5.6B) showed an increased ellipticity of the sample with an increase in temperature; the ellipticity of LmPCNA started to shift at 335K. This demonstrates a significant increase in protein stability from HsPCNA and SpPCNA. Once the sample was heated to 368K, the CD spectra was taken once again, and this returned secondary structure values of 6% α-helix, 35.9% β-strand, 19.8% turns and 37.9% unordered. This showed a small increase in the level of unordered protein and maintains similar levels of secondary structure. LmPCNA did not show a large decrease in α-helix like that observed in HsPCNA and SpPCNA. This may be due to the small amounts determined to be present in the pre-heated sample. Like results obtained for HsPCNA and SpPCNA, there was the retention of β-strand structures in LmPCNA

5.4 Conclusions from the CD analysis of PCNA

The calculated secondary structure derived from the CD analyses shows reasonable agreement with known secondary structure content of PCNA as determined by X-ray analysis. In the cases of SpPCNA and LmPCNA, which contain lower α-helix content, the value of α-helix present is underestimated. This underestimation is a common error in CD analysis (Johnson, W. C. 1999).

The melting curve analysis showed an unexpected trend of increasing ellipticity with increased temperature for all three proteins. This unexpected shift was

not a result of the experimental technique, as demonstrated in figure 5.4, but seems to be a genuine characteristic of PCNA. A similar increase in ellipticity with increasing temperature has been previously observed (Miller, P. *et al* 1976), and was determined to be an increase in amyloid-like aggregates in the sample (Sadqi, M. *et al* 2002; von Bergen, M. *et al* 2000). Heated PCNA samples were analysed by a full CD spectrum at 95°C in order to determine the effect of heat treatment upon the abnormal secondary structure. This demonstrated a significant drop in the quantity of α-helix and an increase in the unordered protein, the level of β-strand however remained similar to that of the native folded structure. The maintenance of the relative quantity of β-strand supports the theory that denatured PCNA forms amyloid-like aggregates. Samples were cooled in order to determine if PCNA could spontaneously re-fold to a native form: no shift was observed and the quantity of secondary structure remained unchanged from analysis performed at higher temperatures.

5.5 Introduction to DSF protein melting studies

DSF is a rapid and cheap method of determining small molecules which bind and stabilise proteins. Fluorescence of a dye increases with an increase in the quantity of unfolded protein and this can be monitored as a function of temperature (Lo, M. C. *et al* 2004). The increased fluorescence is a result of hydrophobic residues becoming exposed as the protein melts (Niesen, F. H. *et al* 2007).

A fitting procedure outlined in (Lo, M. C. *et al* 2004) can give the transition midpoint. The change in this midpoint when in the presence of a binding partner is related to the affinity of the interacting molecule. The stabilising ligand can be a low molecular weight small molecule, peptide or nucleic acid (Boeckler, F. M. *et al* 2008). DSF can be commonly performed using equipment developed for Real-Time PCR (RT-PCR). This method of analysis allows results to be obtained in a single day.

The equipment used in this method was originally designed for the monitoring of DNA or RNA production in Real Time PCR (RT-PCR) experiments (Wong, M. L. and Medrano, J. F. 2005). This system can rapidly and accurately increase the temperature of a 96 well plate with continual measurement of the fluorescence of the well. This ability allows this equipment to be used to monitor the denaturation of fluorescence based protein in melting curve assays (Niesen, F. H. *et al* 2007).

5.5.1 Principles of protein melting studies by DSF

Protein denaturation can be monitored by indirect methods using the shift of fluorescence of a dye. The dye used is SYPRO® Orange (Invitrogen), a dye with an excitation maximum at 473nm wavelength and emission maximum of 570nm

wavelength. This has also been successfully used with an excitation maximum of 490nm and detection at 575nm (Ericsson, U. B. *et al* 2006). The structure of this dye is proprietary information, however, it is hydrophobic and in association with water, fluorescence is quenched. Once a protein is melted, then it is possible to observe an increase in the fluorescence of the dye (Ericsson, U. B. *et al* 2006), as the dye associates with the hydrophobic side groups released from the core of the protein (figure 5.7).

Protein solutions require sufficient concentration of protein and dye to obtain a signal; this can be determined by trial and error. Some protein concentrations previously used have been of 1 μM (Lo, M. C. *et al* 2004) and 10 μM (Boeckler, F. M. *et al* 2008; Vedadi, M. *et al* 2006). Once optimum reagent conditions are achieved, then it is possible to monitor the unfolding of the protein as the well temperature increases. The iQ5-RT-PCR can raise the temperature of the wells by increments of 0.1°C, allowing accurate thermal shift data to be measured. In practice, this experiment will increase the temperature by increments of 0.5°C with a one minute lag time at each step to allow each well to stabilise to the increased temperature. After stabilisation, the samples fluorescence is measured. This method has been previously adopted to observe interactions with a range of melting temperatures e.g. 20°C - 90°C at 0.2°C increments (Ericsson, U. B. *et al* 2006) and 25°C - 75°C at 1°C min^{-1} (Vedadi, M. *et al* 2006).

This experiment can be used to determine the K_d of interaction of the melted protein with a binding partner from its shift in melting temperature. This shift can be used to approximate the K_d of interaction only at the melting temperature of the protein-ligand complex and if the ligand is in excess. The value of K_d at the melting

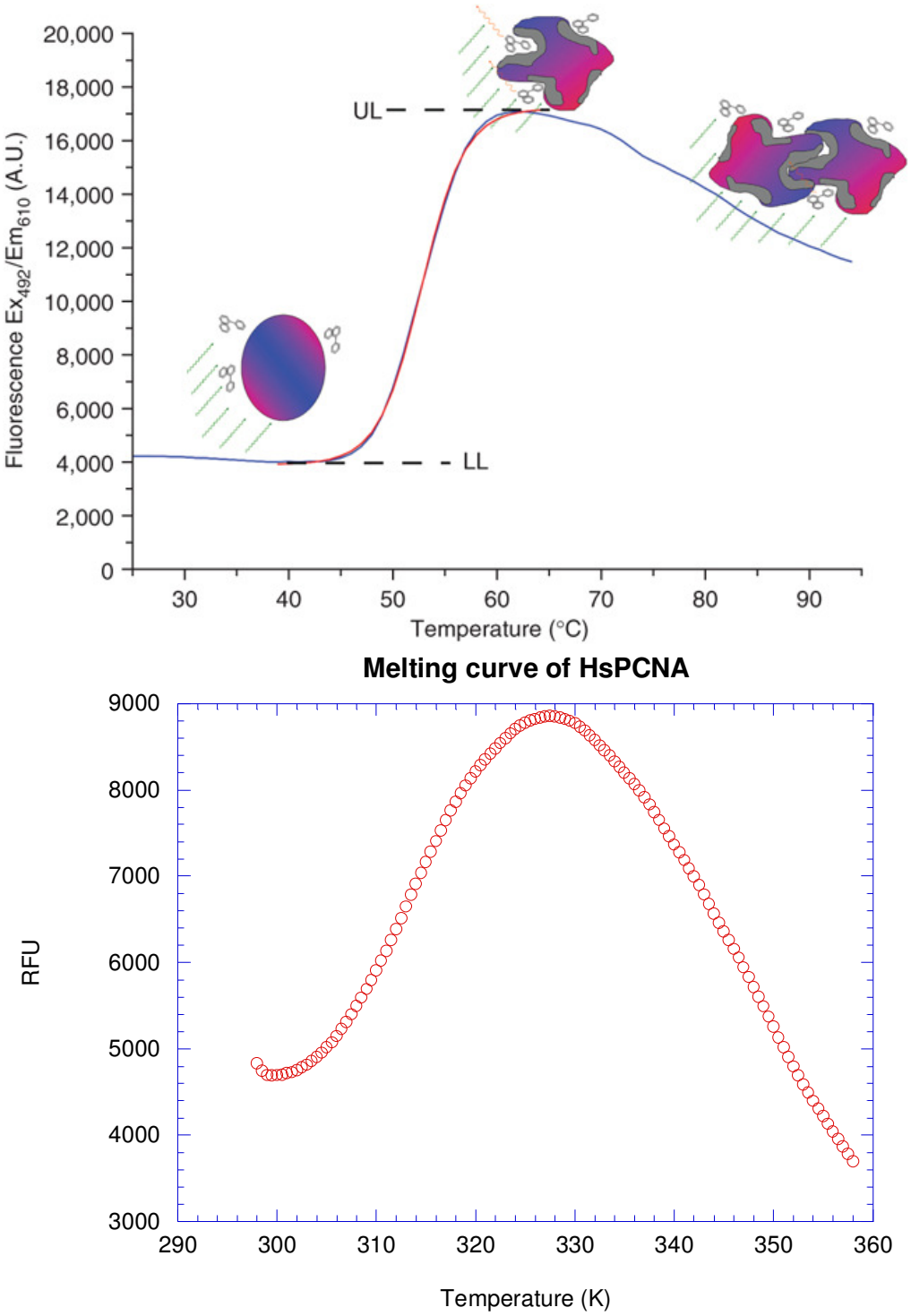


Figure 5-7. **The fluorescence of protein solution in DSF with temperature**
Top) the melting of a protein showing little association of dye to folded protein (LL) the association to unfolded protein (UL) and then a reduction due to self association of unfolded proteins from (Niesen, F. H. *et al* 2007)
Bottom) The curve determined for HsPCNA (showing a curve as above).

temperature lacks biological relevance. However, if the enthalpy of interaction at 298K can be approximated, then a value for K_d at 298K can be determined. This method has been used to determine affinities (K_d) for interactions of 140μM at 37°C (Boeckler, F. M. *et al* 2008), 70 nM to 46 μM at 25°C (Lo, M. C. *et al* 2004).

5.5.2 Melting curve measurement obtained by DSF

Data returned from the melting curve experiment is commonly plotted on a graph of temperature versus relative fluorescence units (RFU). This shows an increase in the RFU with temperature until all protein is unfolded (figure 5.7). This method refers to the melting temperature of the protein as the T_{m50}, temperature at which 50% of the protein within the solution has been unfolded. The values of the enthalpy of unfolding (ΔH_u), specific heat capacity of unfolding (ΔC_{pu}), T_{m50}, RFU after protein denaturation ($F_{(post)}$) and RFU before denaturation ($F_{(pre)}$) can be determined from the data curve. This is achieved by fitting to equation 5.1.

Equation 5.1) $$F_{(T)} = F_{(post)} + \frac{F_{(pre)} - F_{(post)}}{1 + \exp\left\{\frac{-\Delta H_u}{R}\left(\frac{1}{T} - \frac{1}{T_{m50}}\right) + \frac{\Delta C_{pu}}{R}\left[\ln\left(\frac{T}{T_{m50}}\right) + \frac{T_{m50}}{T} - 1\right]\right\}}$$

where T is the temperature at any point, $F_{(T)}$ is the RFU at T and R is the universal gas constant (Lo, M. C. *et al* 2004).

5.5.3 Melting shift as a means to derive K_d

In order to derive an approximate value of K_d of the protein ligand interaction, a further experiment is performed. A protein melting curve is created with excess ligand present to obtain a shift in the melting temperature. To derive the affinity of the protein-ligand complex, the protein-ligand complex T_{m50} (T_{mL50}), ΔH_u, ΔC_{pu} and the T_{m50} of the protein melting curve are required (equation 5.2).

Equation 5.2) $$K_{a(T_{mL50})} = \frac{\exp\left\{\frac{-\Delta H_u}{R}\left(\frac{1}{T_{mL50}} - \frac{1}{T_{m50}}\right) + \frac{\Delta C_{pu}}{R}\left[\ln\left(\frac{T_{mL50}}{T_{m50}}\right) + \frac{T_{m50}}{T_{mL50}} - 1\right]\right\}}{L_{(T_{mL50})}}$$

where $Ka_{(TmL50)}$ is the equilibrium association constant at (T_{mL50}) and $L_{(TmL50)}$ is the free ligand concentration at T_{mL50}, $L_{(TmL50)}$ can be assumed to be the ligand concentration only when ligand concentration is in excess, replicated from (Lo, M. C. *et al* 2004).

The K_a returned from equation 5.2 can be modified to a value at 298K only with an approximate enthalpy of interaction. This is achieved using equation 5.3.

Equation 5.3) $$K_a^{\theta} = K_{a(T_{mL50})} \exp\left\{\frac{-\Delta H^{\theta}}{R}\left(\frac{1}{T_0} - \frac{1}{T_{mL50}}\right) + \frac{\Delta C_p^{\theta}}{R}\left[\ln\left(\frac{T_0}{T_{mL50}}\right) + 1 - \frac{T_0}{T_{mL50}}\right]\right\}$$

where K_a^{θ} is the equilibrium association constant at the reference temperature, T_0 is the reference temperature (298K), ΔH^{θ} is the enthalpy of interaction at the reference temperature and ΔC_p^{θ} is the specific heat capacity of interaction at the reference temperature.

5.6 Materials and methods

Fluorescence based thermal denaturation was studied using the iQ5 Multicolor Real-Time PCR Detection System, iCycler (Bio-Rad) set to excite at 485nm and read at 575nm. SYPRO® Orange was used as the fluorophore, excitation wavelength 473nm, emission wavelength 570nm. Data were recorded and viewed by the iQ5 optical system software version 2.0 and processed using Microsoft Excel and Kaleidagraph version 4.03.

5.6.1 Protocols of determining fluorescence based melting curves

Initial melting curves were obtained with PCNA at concentrations of 0.04 to 0.12mg ml^{-1} (1-4μM). SYPRO® Orange (stored at 1000 fold concentration) is diluted to 0.5 fold concentration in the final reaction volume with buffer 10mM HEPES, 150mM NaCl, 1%DMSO (Fisher), pH7.4. The solution is made up to a final volume of 50μl with buffer. These experiments were used to determine the optimum protein concentrations and buffer conditions.

Protein melting in the presence of peptide was carried out at a protein concentration of 0.2mg ml^{-1} (Hs 6.39μM, Sp 6.56μM, Lm 5.81μM) in 1%DMSO, SYPRO® Orange (at a 10 fold concentration), 10mM HEPES, 150mM NaCl, pH7.4. In reactions with peptide complex, peptide 0 was present at 50, 5, 0.5μM and Peptides 1-5 were present at 100, 50, 5μM. Reaction solutions were pipette-mixed and left to stand at room temperature for 2 minutes prior to use.

The fluorescence of the reaction solution is monitored with increasing temperature from 25°C to 85°C. The thermal denaturation experiment was performed by equilibration to 25°C and then sample plate will "dwell" at this temperature for 30 seconds prior to the fluorescence signal being measured. The temperature was subsequently increased by 0.5°C and once stabilised the plate would dwell for 30 seconds and the fluorescence re-measured. This was repeated until the sample reached 85°C.

5.7 Melting curves of PCNA observed using the iQ5-RTPCR

Melting curves were generated for HsPCNA and SpPCNA. Preliminary results were used to determine optimal protein concentrations. HEPES buffer was found to give sufficient signal for analysis. LmPCNA was determined to have a T_{m50}

of 342K. The melting temperature of LmPCNA showed a consistent concentration dependency, i.e. 239K at 1μM and 342K at 5μM. As discussed in chapter 4, this may be representative of a change in the oligomeric state at this concentration. The melting temperature of LmPCNA is significantly higher than HsPCNA (~315K) and SpPCNA (~316K) (figure 5.8). As stated above, the affinity is determined at the protein melting temperature. For this reason, LmPCNA would require high peptide concentrations to obtain a shift in the melting temperature. Similarly, peptide 6 (a Phe150Ala point mutant) was not tested as the K_d of this interaction would result in too high concentration of peptide to be required.

5.7.1 ΔT_{m50} of HsPCNA and SpPCNA peptide complexes

The associations of six different peptides –

Peptide0 – KKRQTSMTDFYHSKRRLIFS

Peptide1 – KRRQTSMTDFYH

Peptide2 – Bio-KRRQTSMTDFYH

Peptide3 – Bio-PEG-KRRQTSMTDFYH

Peptide4 – Bio-PEG-PEG-KRRQTSMTDFYH

Peptide5 – Bio-PEG-PEG-Ahx-KRRQTSMTDFYH

with HsPCNA and SpPCNA were analysed for a change in their melting temperature. When HsPCNA and SpPCNA were melted, a smooth curve was observed demonstrating an increase in RFU with increasing temperature, as expected (figure 5.9).

The presence of peptides resulted in an increase of melting temperature (figure 5.9). These curves were repeated four times with three different peptide concentrations to obtain averages. The observed change in melting temperature was clear for both

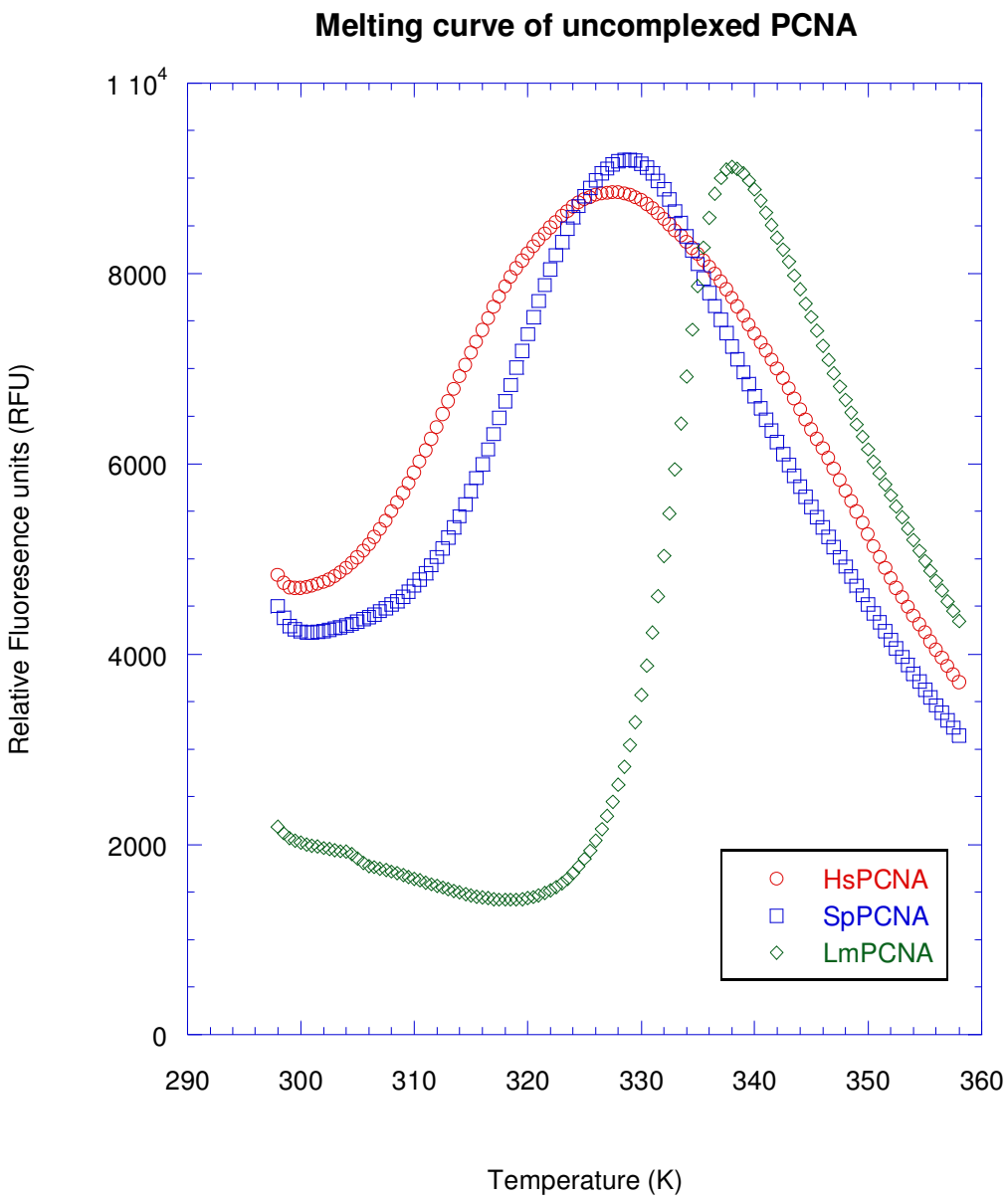
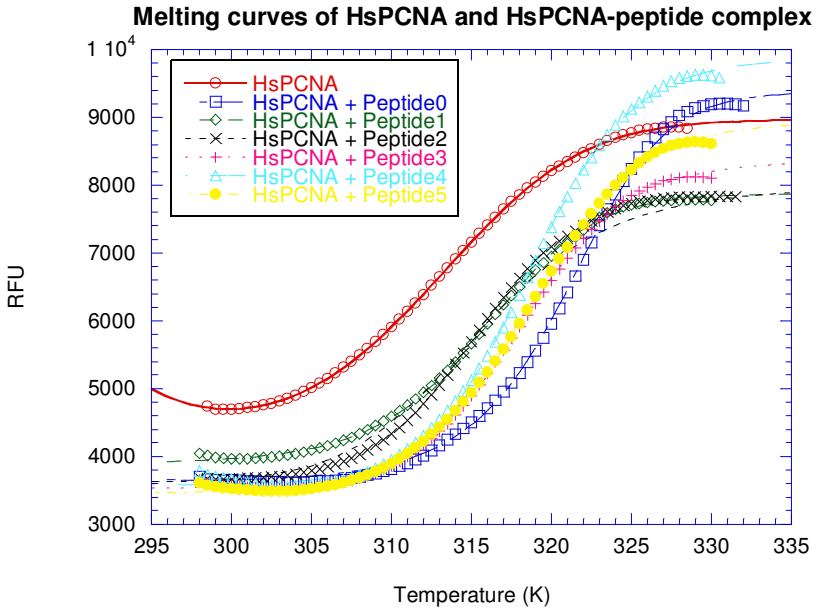
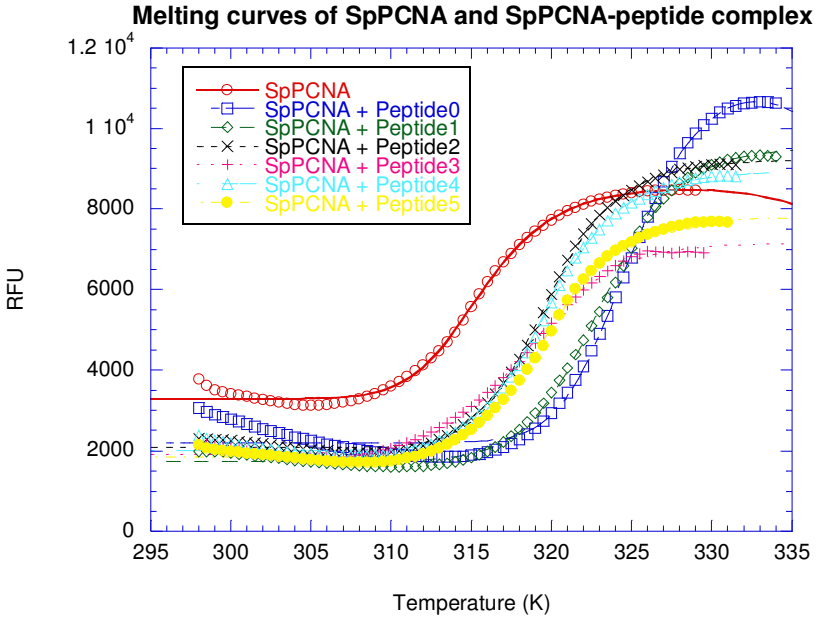


Figure 5-8. **The melting curves of three uncomplexed PCNAs**

Figure showing the melting curves of PCNA from three different species, HsPCNA (red circles) SpPCNA (blue squares) and LmPCNA (green diamonds). Fitting these curves to equation 5.1 (fit not shown) gives the melting temperature of the uncomplexed protein. This figure demonstrates the similar melting temperatures of HsPCNA (~315K) and SpPCNA (~316K) and the significantly more stable LmPCNA (~342K).



A)



B)

Figure 5-9. **The melting curve of PCNA, unbound and complexed**
A) Graph showing the RFU of HsPCNA solution with increasing temperature. Individual melting curves of HsPCNA alone (red) and peptide complexes all show a smooth increase in RFU with temperature. Plotted curves are accurately fit to the data points. Complex curves demonstrate a shift in the melting temperature with the HsPCNA-peptide0 complex demonstrating the greatest shift.
B) Graph showing the RFU of SpPCNA solution with increasing temperature. This graph shows a similar trend to that of HsPCNA with peptides 2-5 having similar melting temperatures, peptide0 demonstrating the greatest shift in the melting temperature.

HsPCNA (314.5K to 322.5K, a change of 8K) (figure 5.10, table 5.1) and SpPCNA (315.7K to 326K, a change of 10.3K) (figure 5.10, table 5.2) when complexed with peptide 0. It is clearly demonstrated that the melting temperature of protein complexed with peptide 0 was the greatest (figure 5.10). HsPCNA had a ΔT_{m50} of 8°C when complexed with 50μM peptide 0 and 3°C when complexed with 100μM peptide 1 (figure 5.10, table 5.4). These changes in the ΔT_{m50}, of PCNA infer a stronger interaction for peptide 0 than peptide 1.

Melting curves were repeated at each peptide concentration and the standard deviations (shown in figure 5.10) were formed from these replicates. The standard deviations allowed a T-test to be performed and data which was determined to not be statistically significant was not used further. These curves could be fit to equation 5.1 to give the required values for K_d determination.

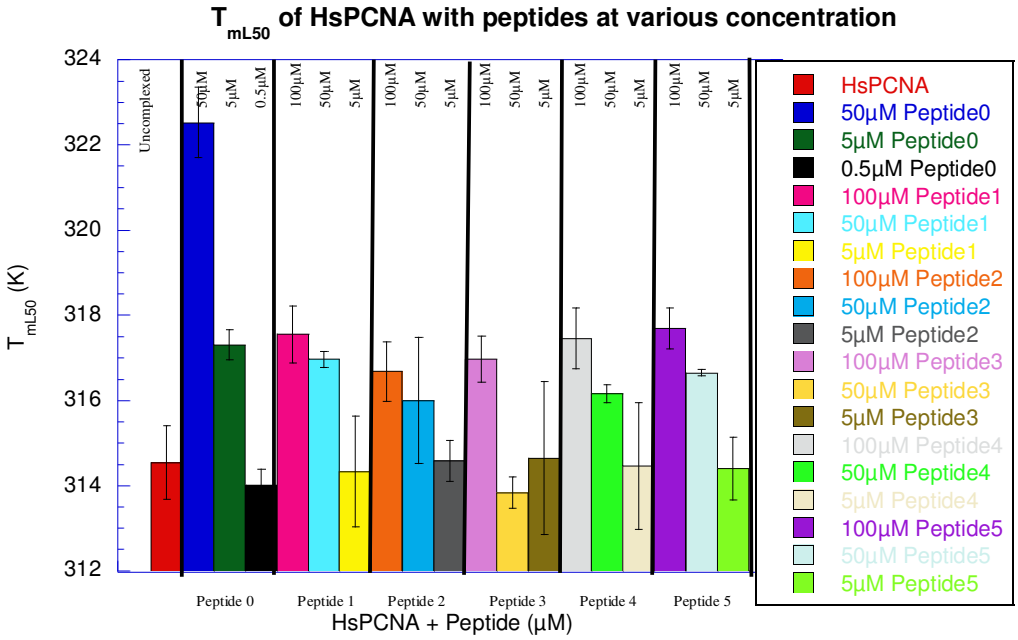
5.7.2 Evaluation of the K_d between HsPCNA and SpPCNA for peptides

The T-test allowed data which was determined to not be statistically significant to be excluded form the calculation of K_d. The K_d at the melting temperature of each experiment could then be determined for each significant test using equation 5.2 (Table 5.3A and Table 5.4A).

Equation 5.2) $$K_{a(T_{mL50})} = \frac{\exp\left\{\frac{-\Delta H_u}{R}\left(\frac{1}{T_{mL50}} - \frac{1}{T_{m50}}\right) + \frac{\Delta C_{pu}}{R}\left[\ln\left(\frac{T_{mL50}}{T_{m50}}\right) + \frac{T_{m50}}{T_{mL50}} - 1\right]\right\}}{L_{(T_{mL50})}}$$

For equation 5.2, the ΔH of unfolding, the ΔC_p of unfolding and the T_{m50} are required as obtained from the data fit for uncomplexed protein. The T_{mL50} is obtained from the melting curve data fit for complexed protein. As the peptide was present in excess, the concentration of peptide added was assumed to be the concentration of free

A)



B)

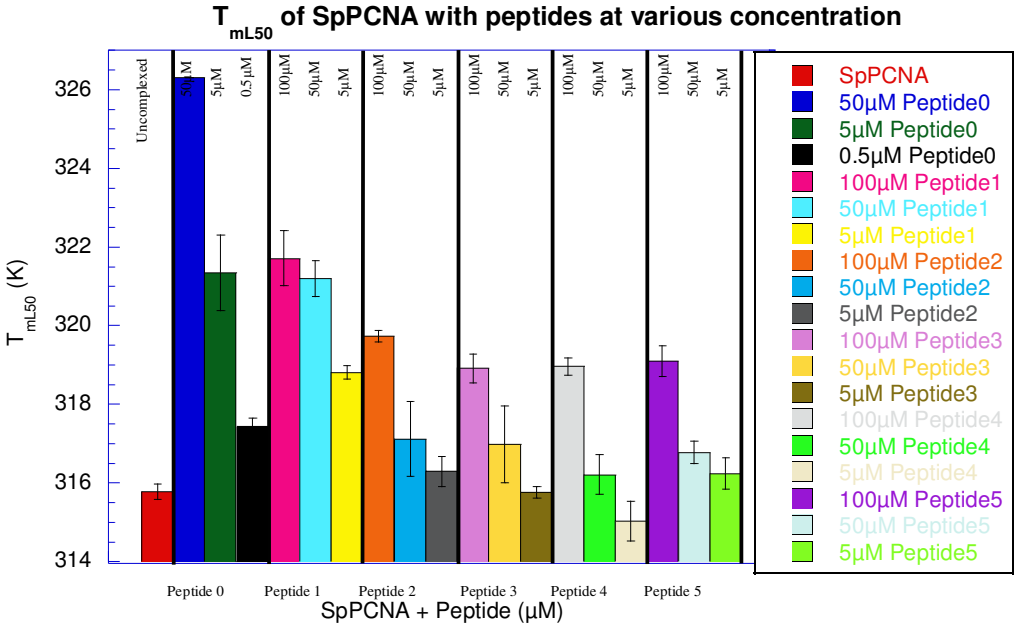


Figure 5-10. **The melting temperature of PCNA with a concentration series of peptides**

A) The melting temperatures of HsPCNA showing the T_{m50} on the y-axis. HsPCNA alone is shown in red on the left and three concentrations of peptide, as explained by the figure legend. Error bars are generated using replicate runs. B) The melting temperature of SpPCNA showing the T_{m50} on the y-axis. The order of peptide complexes is identical to that of HsPCNA, with error bars generated as above.

Peptide	Peptide concentration within series		
	High	Medium	Low
0	322.5 (+8.0)	317.3 (+2.8)	314.0 (-0.5)
1	317.6 (+3.1)	317.0 (+2.5)	314.3 (-0.2)
2	316.7 (+2.2)	316.0 (+1.5)	314.6 (+0.1)
3	317.0 (+2.5)	313.8 (-0.7)	314.7 (+0.2)
4	317.5 (+3.0)	316.2 (+1.7)	314.5 (no change)
5	317.7 (+3.2)	316.7 (+2.2)	314.4 (-0.1)

Table 5-1. **The change in melting temperature of HsPCNA with peptides**

The peptide concentration series for peptide 0 is 50μM (high) 5μM (medium) and 0.5μM (low) and peptide 1-5 is 100μM (high) 50μM (medium) and 5μM (low). The melting temperature of the protein is shown in Kelvin and the change in melting temperature from the uncomplexed protein is shown in brackets in Kelvin. Values in red were not statistically significant as determined by t-test.

Peptide	Peptide concentration within series		
	High	Medium	Low
0	326.3 (+10.5)	321.3 (+5.5)	317.4 (+1.6)
1	321.7 (+5.9)	321.2 (+5.4)	318.8 (+3.0)
2	319.7 (+3.9)	317.1 (+1.3)	316.3 (+0.5)
3	318.9 (+3.1)	317.0 (+1.2)	315.8 (no change)
4	319.0 (+3.2)	316.2 (+0.4)	315.0 (-0.8)
5	319.1 (+3.3)	316.8 (+1.0)	316.2 (+0.4)

Table 5-2. **The change in melting temperature of SpPCNA with peptides**

The peptide concentration series for peptide 0 is 50μM (high) 5μM (medium) and 0.5μM (low) and peptide 1-5 is 100μM (high) 50μM (medium) and 5μM (low). The melting temperature of the protein is shown in Kelvin and the change in melting temperature from the uncomplexed protein is shown in brackets in Kelvin. Values in red were not statistically significant as determined by t-test.

A)

Peptide Conc.	Peptide 0	Peptide 1	Peptide 2	Peptide 3	Peptide 4	Peptide 5
High	20.8	75.3	84.8	81.6	76.3	73.9
Medium	3.89	38.9	57.7		45.6	42.6
Average	12.3	57.1	71.3	81.6	61.0	58.3

B)

Peptide Conc.	Peptide 0	Peptide 1	Peptide 2	Peptide 3	Peptide 4	Peptide 5
High	0.136	17.5	17.4	16.2	12.1	11.8
Medium	0.079	9.48	12.6		8.22	7.53
Average	0.107	13.5	15.0	16.2	10.2	9.68

Table 5-3. **The K_d (M) of HsPCNA for peptides**
A) The determined K_d of affinity in μM of each peptide for HsPCNA at the T_{mL50} as determined by equation 5.2. Only affinities of statistically significant data are shown. The average is the value of each of the peptide concentrations returned
B) The determined K_d of affinity in μM of each peptide for HsPCNA at 298K as determined by equation 5.3. Melting curves determined using lower peptide concentrations gave higher affinities, average binding affinities demonstrate peptide 0 to have the highest affinity, with peptide 1-5 having similar affinities.

A)

Peptide Conc.	Peptide 0	Peptide 1	Peptide 2	Peptide 3	Peptide 4	Peptide 5
High	2.33	15.0	27.0	34.8	34.3	32.8
Medium	0.834	8.70	31.4	32.8		35.2
Low	0.282	1.80				
Average	1.15	8.51	29.2	33.8	34.3	34.0

B)

Peptide Conc.	Peptide 0	Peptide 1	Peptide 2	Peptide 3	Peptide 4	Peptide 5
High	0.013	3.12	3.62	4.75	3.91	4.30
Medium	0.013	1.88	5.44	5.45		5.85
Low	0.009	0.462				
Average	0.012	1.82	4.53	5.10	3.91	5.08

Table 5-4. **The K_d (M) of SpPCNA for peptides**
A) The determined K_d of affinity in μM of each peptide for SpPCNA at the T_{mL50} as determined by equation 5.2. Only affinities of statistically significant data are shown. The average is the value of each of the peptide concentrations returned
B) The determined K_d of affinity in μM of each peptide for SpPCNA at 298K as determined by equation 5.3. Melting curves determined using lower peptide concentrations gave higher affinities, average binding affinities demonstrate a binding affinity of peptide 0>>peptide 1>peptides 2-5.

peptide ($L_{(TmL50)}$) at T_{m50}.

The determined value for K_a at the melting temperature can be used with equation 5.3 to give the K_a at 298K.

Equation 5.3) $$K_a^{\theta} = K_{a(T_{mL50})} \exp\left\{\frac{-\Delta H^{\theta}}{R}\left(\frac{1}{T_0}-\frac{1}{T_{mL50}}\right)+\frac{\Delta C_p^{\theta}}{R}\left[\ln\left(\frac{T_0}{T_{mL50}}\right)+1-\frac{T_0}{T_{mL50}}\right]\right\}$$

$$K_a^{\theta} = 13283 \exp\left\{\frac{49220}{R}\left(\frac{1}{T_0}-\frac{1}{318}\right)+\frac{-1000}{R}\left[\ln\left(\frac{T_0}{318}\right)+1-\frac{T_0}{318}\right]\right\}$$

The $Ka_{(TmL50)}$ is the equilibrium affinity constant between HsPCNA and peptide 1 at the melting temperature. The values are shown for the HsPCNA peptide 1 interaction where ΔH^{θ} is the enthalpy of interaction at 298K. This is assumed to be that determined in chapter 8. R is the universal gas constant (8.314472 J K^{-1} mol^{-1}), T_0 is the reference temperature (298K) and T_{mL50} is the melting temperature of HsPCNA peptide 1 complex (318K). The value for the ΔC_p^{θ} is the specific heat capacity of HsPCNA peptide interaction (~ -1000 J K^{-1} mol^{-1}), approximated from the non-linear van't Hoff analysis in chapter 8. The value used for ΔC_p^{θ} has been previously argued to only give a slight effect to the determined K_a of interaction (Lo, M. C. *et al* 2004). Equation 5.3 was used to derive the K_d at 298K represented in tables 5.3B and 5.4B.

The final values determined for the K_d of interaction showed similar affinities for both HsPCNA and SpPCNA. The K_d of interaction at 298K between HsPCNA and the six peptides has been shown to have a range of 107nM to 16.3μM (table 5.3B). The K_d of interaction at 298K between SpPCNA and the six peptides has been shown to have a range of 11.7nM to 5.10μM (table 5.4B). Peptide 0 demonstrated a high affinity of interaction for HsPCNA (107nM) and SpPCNA (11.7nM) (table 5.3B and table 5.4B). Peptide 0 has a lower K_d of interaction than peptides 1-5 for

HsPCNA (9.68μM to 16.3μM) and SpPCNA (1.82μM to 5.10μM). Of the interactions of peptides 1-5, peptide 1 had the lowest K_d.

5.8 Conclusions from the melting curve analysis

5.8.1 Agreement between melting curves derived from CD and fluorescenc

The melting curves derived from CD analysis showed melting temperatures of 321K, 322K and 335K for HsPCNA, SpPCNA and LmPCNA respectively. The melting temperature derived from fluorescence based analysis gave melting temperatures of 315K, 316K and 342K for HsPCNA, SpPCNA and LmPCNA respectively. The differences between the two measurement techniques e.g. different buffers (10mM NaH_2PO_4, 10mM NaF pH7.4 for CD and 10mM HEPES, 150mM NaCl and 1% DMSO pH 7.4 for DSF), the presence of dye and the change in protein concentration (1.25 – 2.04 mg ml^{-1} for CD and 0.2 mg ml^{-1} for DSF) may be responsible for the slight shift in melting temperature. The trends in melting temperature are consistent, with HsPCNA and SpPCNA having similar melting temperatures and LmPCNA being significantly more stable. This increase in the stability of the LmPCNA could be an evolutionary trait as *L. major* exists in insect vectors in tropical areas.

5.8.2 Peptide binding affinities derived from DSF studies

Data returned from the DSF experiments had statistical significance verified from the shift in melting temperature (figure 5.10). The affinities determined by this method were achieved with protein present in HEPES buffer (10mM HEPES, 150mM NaCl, 1%DMSO, pH7.4). The HEPES buffer is similar to that used for ITC (chapter 7) and SPR (chapter 8) experiments and should reduce errors caused by

differences in buffer effects. The use of the ΔH^{θ} value from SPR experiments (chapter 8) in equation 5.3 will result in a slight bias toward the values determined from section 8. This however will only have a slight effect. The values for ΔC_p^{θ} used in this method were determined in SPR experiments (chapter 7) for peptides 1-5 to be in the range of -600 to -1300 J K^{-1} mol^{-1}. This value was approximated to -1000 J K^{-1} mol^{-1} and this will not result in any significant modification in the final result as explained in (Lo, M. C. *et al* 2004).

The affinity for HsPCNA with peptide 0 has been previously determined by ITC at 298K as having a K_d of 87.7nM (Zheleva, D. I. *et al* 2000), this is in close agreement to the results obtained here as an average of 107nM (table 5.3 and table 5.4). The close agreement in value for peptide 0 supports the validity of all the affinities determined using this method.

The values determined in this chapter will be compared with those determined by ITC (chapter 6) and SPR (chapter 7) in chapter 8.

CHAPTER 6. The study of PCNA peptide interaction by ITC

Isothermal Titration Calorimetry (ITC) can determine the thermodynamic properties of protein-ligand binding interactions. ITC can determine the association equilibrium constant of an interaction at different temperatures; with this data it is possible to ascertain the Gibbs free energy of binding using equation 6.1 (Okhrimenko, O. and Jelesarov, M. 2008).

6.1) $$\Delta G^{\theta} = -RT \ln K_a$$

where ΔG is the change in Gibbs free energy and the K_a association equilibrium constant of the interaction. The interactions determined by ITC use the enthalpy of the interaction, with accurately known concentrations of reactants (Jelesarov, I. and Bosshard, H. R. 1999). The enthalpy of interaction is a relationship of the bonds (intra and inter-molecular) formed to bonds broken. Should the energy required to break the number of bonds in the system (desolvation and protein rearrangements) be greater than the energy released by bond formation, then the interaction would be endothermic (Privalov, Peter L. and Potekhin, Sergey A. 1986). The result of an endothermic interaction is to make this an enthalpically unfavourable binding event. Strength of binding interactions are also mediated by entropy, this is a reference to the level of disorder within the system. The greater the disorder, the more favourable the interaction; this is related to the Gibbs free energy and enthalpy shown in equation 6.2 (Okhrimenko, O. and Jelesarov, M. 2008).

6.2) $$\Delta G^{\theta} = \Delta H^{\theta} - T\Delta S^{\theta}$$

where T is the absolute temperature in Kelvin and ΔS is the change in entropy of the system.

The interactions of PCNA with peptides have been documented. Many of these interactions have been shown between HsPCNA and the PCNA interacting peptide (PIP) box (Warbrick, E. 2000). The interactions of peptides to PCNA have been previously shown to be enthalpically favourable and entropically unfavourable (Bruning, J. B. and Shamoo, Y. 2004).

6.1 PCNA interactions previously quantified by ITC

Previous examples of PCNA-peptide binding studies have been performed by ITC. These have shown interactions between peptides and PCNAs of different origins, eukaryotic and prokaryotic. The interactions between HsPCNA and peptides are mediated by the PIP box, a common sequence in HsPCNA interacting proteins (Warbrick, E. 2000). Binding interactions of HsPCNA with p21 (141-160) have previously demonstrated a stoichiometry of 1 and a K_d of 87.7nM (Zheleva, D. I. *et al* 2000). These values are in agreement with a more comprehensive analysis covering other protein interactors. It has been shown that p21 (139-160), p66 (451-466) (451-472) and FEN1 (332-353) have K_ds of 82.6nM, 15.6μM, 1.54μM and 59.9μM respectively. These interactions were all shown to have stoichiometries of 1 and were enthalpically favourable and entropically unfavourable (Bruning, J. B. and Shamoo, Y. 2004). HsPCNA has also been shown to bind with DNA polymerase η with a K_d of 81μM (Bomar, M. G. *et al* 2007). Due to the level of conservation of PCNA, homologues can also demonstrate similar binding affinities to these peptides. E.coli PCNA showed similar thermodynamic values for peptide interactions to that of HsPCNA (Georgescu, R. E. *et al* 2008). This recent work also determined a small molecule inhibitor for PCNA (DNA polymerase subunit β) and demonstrates a potential site for inhibition.

6.2 The principles of ITC experiments

ITC is a method of very accurately measuring biophysical interactions. ITC measures energy released or absorbed as heat upon binding. It is the method of choice for the measurement of protein ligand binding events (Okhrimenko, O. and Jelesarov, M. 2008). Ligand is contained within a syringe and injected into a solution of binding partner and the energetic exchange is measured as a function of temperature (MicroCal, LLC 2002). With the subsequent injection of more ligand, there is a reduced level of energetic exchange. As the ligand concentration approaches saturation, the energetic exchange approaches that of the dilution effect of the ligand. This shift allows the determination of stoichiometry, dissociation equilibrium constant (K_d), enthalpy of interaction (ΔH) and the entropy of interaction (ΔS) (Jelesarov, I. and Bosshard, H. R. 1999).

6.2.1 Determination of thermodynamic parameters from raw ITC data

Thermodynamic parameters of stoichiometry, K_a and ΔH can be determined from an ITC experiment. Resolving these values requires raw data to have a baseline determined, integrated and fitted to a curve. This is done reliably using the Origin software, however, an understanding of the principles of the experiment and method of determining these values is ideal (Jelesarov, I. and Bosshard, H. R. 1999).

Precise volumes of ligand are added at specific, regular time points whilst stirred by the injection syringe (figure 6.1). Upon each injection of ligand into binding partner, heat energy is either released or absorbed relative to the reference cell. This thermal energy is directly related to the enthalpy of the interaction measured. The thermal jacket of the calorimeter provides or removes energy in order

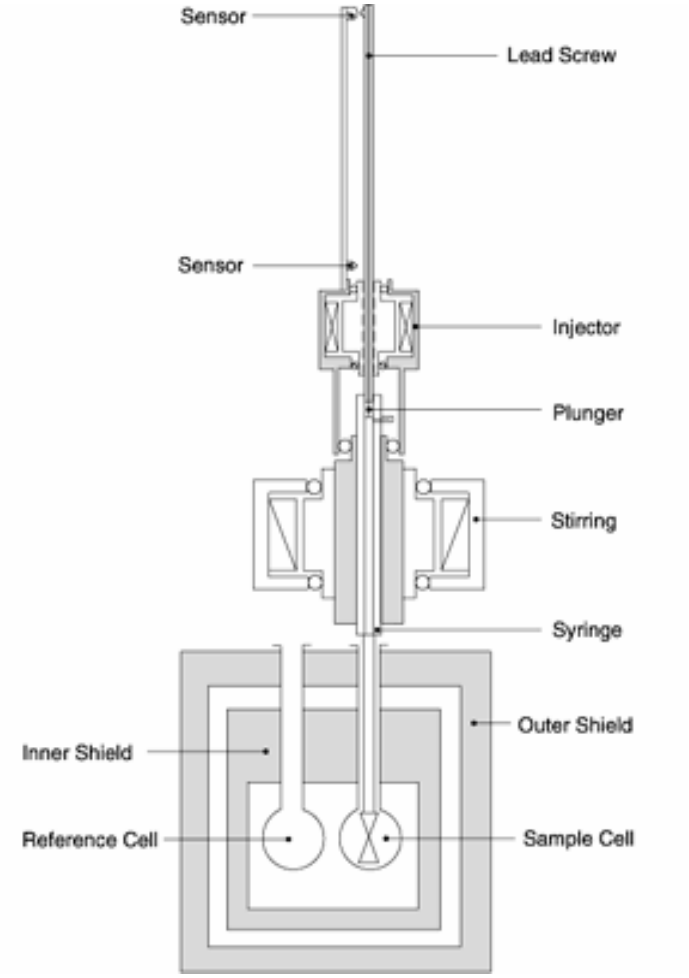


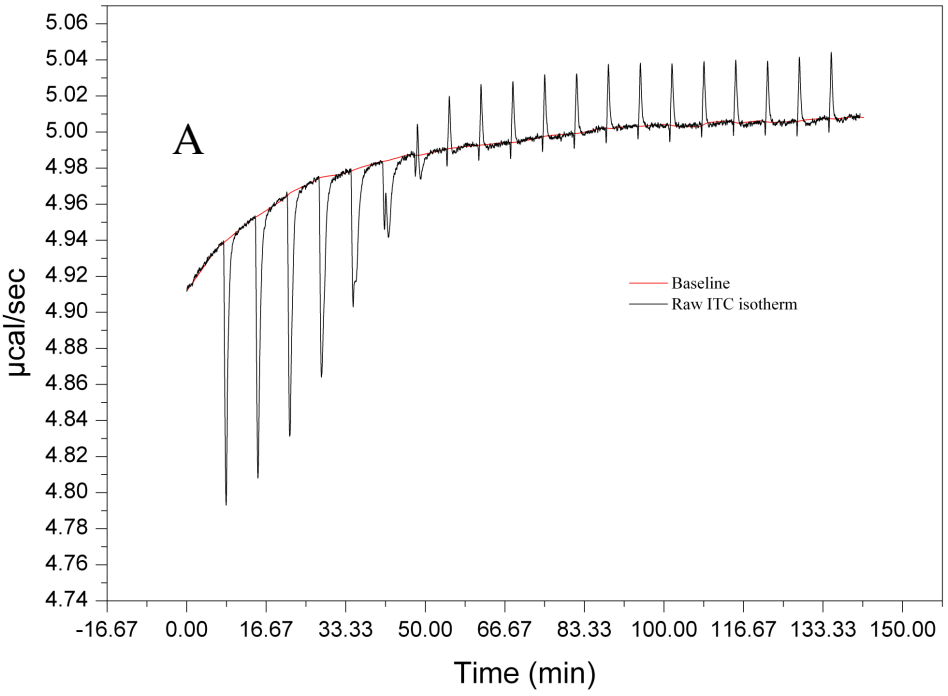
Figure 6-1. **Schematic representation of ITC equipment and sample positions from "VP-ITC MicroCalorimeter users manual"**

Schematic diagram showing the reaction cell, reference cell and injection syringe of the VP-ITC MicroCalorimeter, *MicroCal LLC*. The thermal jacket surrounding the cells is accurately temperature controlled and the required energy difference between the reference and reaction cells are recorded. The thermal jacket is capable of maintaining temperatures of 275-353K. The injection syringe is used for the injection and stirring of the reaction cell. This syringe holds 780µl of ligand and can inject at a minimum of 1.5µl, however this volume can be inaccurate so it is more common for 10µl injections. The injection syringe maximum stirring speed is 580rpm, but at this speed, the potential for applying mechanical stress upon the protein in solution can be great, typically the sample is stirred at 310rpm. Replicated from (MicroCal, LLC 2002).

to maintain a constant temperature; this required power is measured in µcal s^{-1}. The peaks are recorded as a differential power signal, showing power (µcal s^{-1}) against experimental time (minutes) (Freire, E. *et al* 1990; Okhrimenko, O. and Jelesarov, M. 2008). An ITC experiment can span hours and, with fluctuations in ambient temperature, there is a constant drift in required power to maintain jacket temperature. When the final ITC differential power signal is generated, a baseline signal is created comprised of data points prior to each injection. This potentially curved baseline is determined to be the level of no differential energy and is plotted to 0µcal s^{-1} (figure 6.2A).

A single peak from the differential power signal shows the heat change upon addition of a known volume of known concentration ligand to the reaction cell. The total energy exchange in this peak is the sum of contributions from the energy of dilution, energy of non-specific binding and the energy of association. This energy is represented by the area of the peak and can be integrated to give a total energy per ligand concentration (Jelesarov, I. and Bosshard, H. R. 1999). As ligand concentration increases, all available binding partners in the reaction cell become occupied. This results in reduced level of energy change in subsequent injections. Once all binding partners are occupied, the measured energy change is a dilution effect of concentrated ligand into the reaction cell. The energy of dilution can be accounted for by performing a second control experiment with ligand injected into a buffer sample. The integrated energy of dilution can then be deducted from the integrated energy of binding data to determine the energy due to interaction (equation 6.3) (Jelesarov, I. and Bosshard, H. R. 1999).

6.3) $$\Delta E_{ass} = \Delta E_{tot} - \Delta E_{dil} - \Delta E_{ns}$$



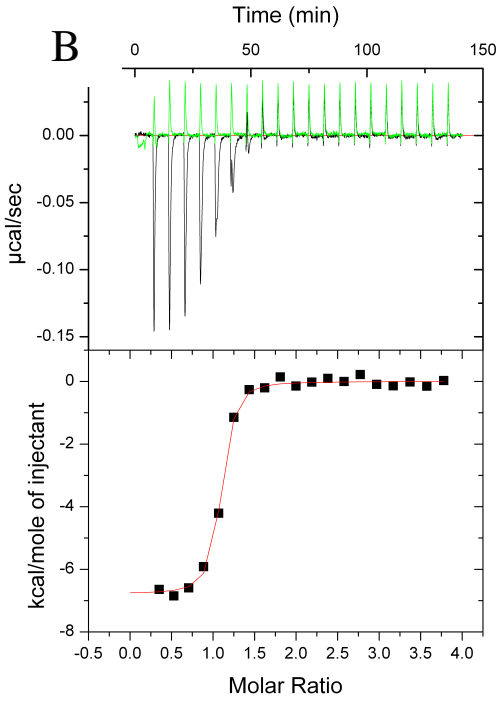


Figure 6-2. **Isotherms of SpPCNA binding to peptide0**

Figures showing isotherms of SpPCNA binding to peptide0. (A) Raw ITC data (black) demonstrating an exothermic interaction upon injection of ligand coupled with an endothermic dilution effect. This isotherm also demonstrates a baseline drift (red) over the course of the experiment.
(B) (top) change in applied energy from previously determined baseline of the peptide PCNA interaction (black) and peptide buffer energetics (green). (bottom) integrated binding energy of each injection with dilution energy deducted and fit as in section 6.1.1.

where ΔE_{ass} is the energy of association, ΔE_{tot} is the total change in energy, ΔE_{dil} is the energy of dilution and ΔE_{ns} is the energy related to non-specific interactions. The integrated data is fit to a non-linear least squares fit (figure 6.2B) (Siligardi, G. *et al* 2004). Assuming a single uniform set of protein binding sites, the energy of injection can be related to thermodynamic parameters using equation 6.4 (Jelesarov, I. and Bosshard, H. R. 1999).

6.4) $$\Delta E_{ass} = N[M]_{tot}.V_{cell}.\Delta H_{tot} \times ¡$$

where N is the stochiometery, $[M]_{tot}$ is the total concentration of binding partner in the reaction cell, V_{cell} is the volume of the reaction cell, ΔH_{tot} is the total enthalpy of the interaction and ¡ is the root of the quadratic equation. The root of the quadratic equation is as equation 6.5 (Jelesarov, I. and Bosshard, H. R. 1999).

6.5)

$$\left(\frac{\Delta[L]_{bound}}{[M]_{tot}}.[L]_{tot}\right)^2 - \frac{\Delta[L]_{bound}}{[M]_{tot}}.[L]_{tot}.\left(1+\frac{1}{N.Ka.[M]_{tot}}+\frac{[L]_{tot}}{N[M]_{tot}}\right)+N[L]_{tot}.[M]_{tot}=0$$

where $[L]_{bound}$ is the concentration of ligand associated to binding partner, $[L]_{tot}$ is the total concentration of ligand in the reaction cell after injection and K_a is the equilibrium association constant of the ligand-binding partner interaction.

Integrated ITC heat data with dilution energies subtracted can therefore be accurately fitted to give values of N, ΔH and K_a using equations 6.2 and 6.3. Using standard equations the values of K_a and ΔH can also yield the entropy (ΔS) of interaction. Finally, change in specific heat capacity (ΔC_p) due to the interaction can be determined as the change in enthalpy of interaction divided by the change in temperature. So, should ITC experiments be carried out on a temperature series, a ΔC_p of the solution due to the interaction can be solved.

6.2.2 Sources of experimental error in ITC methodologies

In previously published data, there have been limitations on the ITC method. These limitations have been observed in interactions with little change in enthalpy (Knipscheer, P. *et al* 2007), weak interactions (Li, H. *et al* 2008; Rainaldi, M. *et al* 2007; Siligardi, G. *et al* 2004), dilution effects (Huang, Shir Ly *et al* 2005) and due to poorly determined concentrations (Rainaldi, M. *et al* 2007).

The heat produced or absorbed in an ITC experiment is directly related to the enthalpy of an interaction by equation 6.2. In examples of binding interactions that present a small ΔH, it is difficult for the integrated data to be fit, as system noise would present a proportionally larger effect. Previous experiments with protein-peptide interactions have shown useable data for one peptide and poor data for a second due to low ΔH (Knipscheer, P. *et al* 2007). In this example, this was a result of short peptide sequence resulting in reduced favourable contacts and reduced ΔH. Other examples have demonstrated maintenance of peptide length, however changes in sequence which resulted in no significant alteration in ΔG. These alterations in sequence resulted in significant changes in ΔH (Lim, N. S. *et al* 2006). This is termed an enthalpy-entropy compensation effect (Rainaldi, M. *et al* 2007; Yamniuk, A. P. and Vogel, H. J. 2004) and presents little alterations in K_a. However, should ΔH become small, (Lim, N. S. *et al* 2006) then no measurement can be made by ITC.

6.2.3 The error associated to fitting weak binding interactions

Weak binding interactions result in poor curve fitting to integrated data. The amount of curvature between two baselines, initial injections and injections demonstrating only dilution effects, is a result of the K_a, and the weaker this interaction is the broader the curve will be. In weak interactions, a large ligand

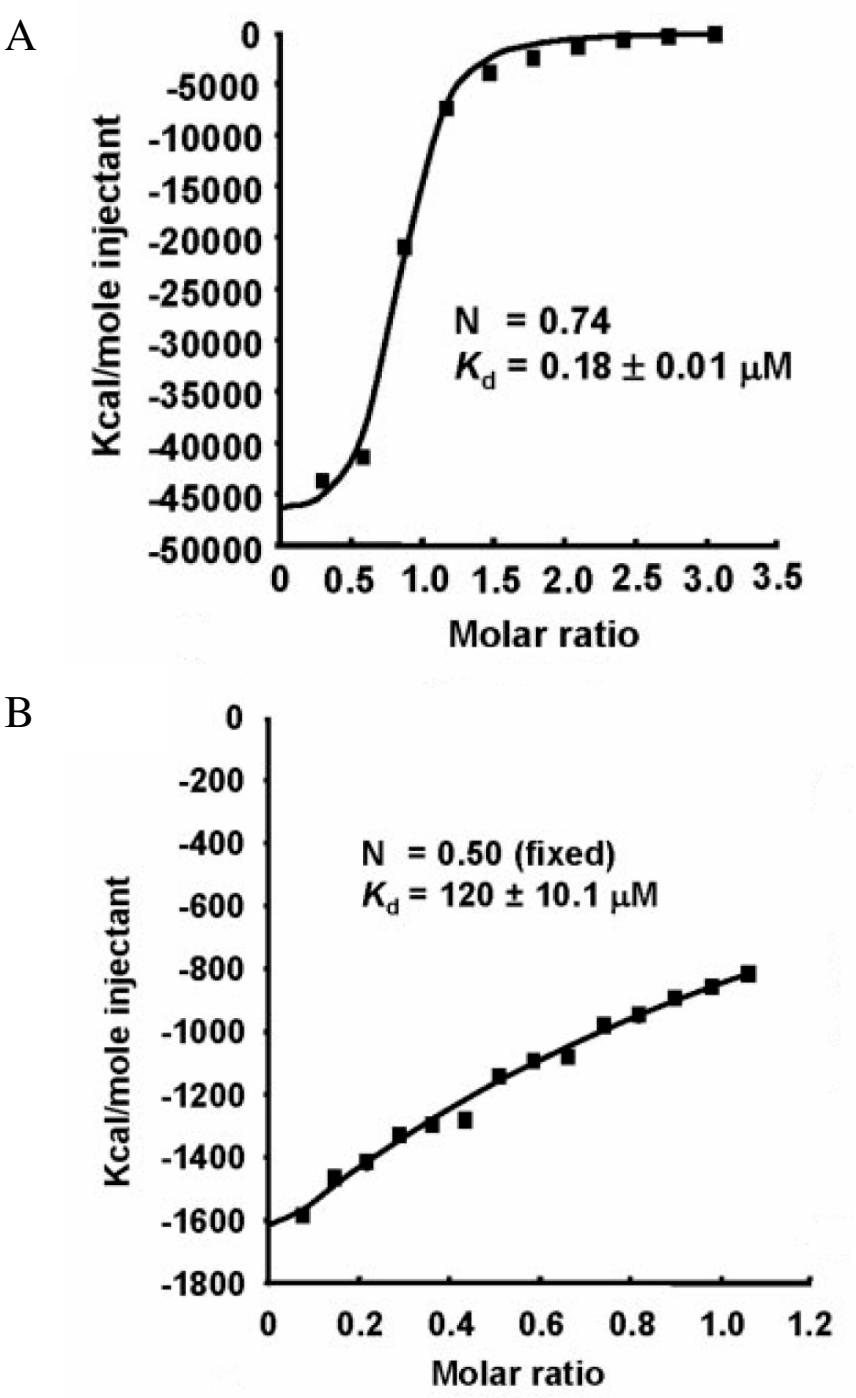


Figure 6-3. **Representations of strong and weak ITC data**

Interactions of Hsp90 with (A) Sti1, (B) Sba1 showing a strong interaction between Hsp90 and Sti1 due to the presence of baseline at both the high and low molar ratios. There is a weaker interaction between Hsp90 and Sba1 which does not present a baseline, resulting in poor determination of thermodynamic values and stoichiometry.

concentration is required in the reaction cell to drive the reaction to completeness. This can be problematic should ligand concentration be limited due to solubility or availability. Protein-peptide interactions have previously been weakened beyond the ability of ITC to determine an accurate K_a with reduction in peptide size (Li, H. *et al* 2008). Protein peptide interactions have also been shown to weaken with differences in ionic strength whilst remaining stable with increased temperature (Rainaldi, M. *et al* 2007). Protein-protein interactions have also been shown to be affected by point mutations, where there is no appreciable difference in structure, which resulted in binding interactions of such weak K_a that the data was unusable (Siligardi, G. *et al* 2004).

6.2.4 The error in ITC experiments associated to dilution effects

Dilution effects of high ligand concentration being injected into the reaction cell can pose a significant alteration upon the energetics of interaction, as shown in equation 6.3. The effects of dilution energy can change the determined energy of association. Dilution energies are accumulated by slight differences in temperature, ionic strength, detergents and mechanical stress (Huang, Shir Ly *et al* 2005). These differences can result in large dilution energies, however, they are simple to take into account by performing ligand-buffer controls. It has been shown that dilution enthalpies are increased with high temperatures and ionic strength (Huang, Shir Ly *et al* 2005; Tellinghuisen, J. 2007) but are however most prominent with buffer mismatches.

6.2.5 The error related to the sample concentrations

Should one or both the reactants, ligand or binding partner, have unknown or imprecise concentrations, then data returned can be highly unreliable (Jelesarov, I. and Bosshard, H. R. 1999). The dependence this method has upon accurate concentrations is shown in equation 6.5. Concentrations when using dissolved peptides can be difficult to determine. There are three main methods of determination of peptide concentration, 1) measuring the mass of lyophilised peptide dissolved, 2) measuring the A_{280} of dissolved peptide solution and 3) quantitative amino acid analysis. The first method is simple and common (Rainaldi, M. *et al* 2007) however can be inaccurate due to bound waters in the solid mass or salts. The amount of bound water can vary depending upon peptide sequence from 5-40%; this difference in mass can vastly change the peptide concentration. The second method of determining peptide concentration, measuring the A_{280}, is more accurate provided the peptide contains tyrosine or tryptophan. However, if the peptide is long and hydrophobic, it may show propensity to fold. Finally, quantitative amino acid analysis is not available within many laboratories and can also be expensive, requiring peptide that can not be recovered.

6.2.6 Weak data as a result of low enthalpy of interaction

ITC isotherms require a significant enthalpic shift with micro-molar affinity during binding in order to result in a useable isotherm. Should this isotherm not show strong curvature, then it can be difficult to determine thermodynamic properties. In some examples, point mutants have reduced the binding affinity too much to determine any binding properties using ITC (Knipscheer, P. *et al* 2007; Li, H. *et al* 2008). Protein binding partners can present two binding sites with different affinities.

In these examples, the isotherms have contributions from both binding sites which can result in poorer determination of each site (Rainaldi, M. *et al* 2007). Examples demonstrating biphasic interactions can also show these poor determinations of thermodynamic properties with increasing temperature. Increasing temperature should present a fairly uniform effect on enthalpy, however, some published data have values which fluctuate high and low with subsequent increases in temperature (Rainaldi, M. *et al* 2007). These fluctuations imply that the results may not be completely accurate. Some examples with protein-peptide interactions have compared similar peptides and shown significant shifts in thermodynamic values with little shift in affinity (Lim, N. S. *et al* 2006). Many of these peptides, of the same length, have very low shifts in enthalpy upon binding. These interactions show similar Gibbs free energy and significant fluctuations in entropy, to the extent that immobilisation of peptides is entropically favourable. It is highly unlikely that the immobilisation of a large flexible peptide will be an entropically favourable event. When enthalpic return of the interaction raised to greater levels, the entropy of the interaction became unfavourable as would be expected (Lim, N. S. *et al* 2006).

6.2.7 Thermodynamic parameters of protein interactions determined by ITC

A wealth of information regarding protein interactions has been determined by ITC. This information follows general thermodynamic trends, dependent upon the interaction in question, protein-protein (Knipscheer, P. *et al* 2007; Nishimiya, Y. *et al* 2000; Shiroishi, M. *et al* 2001; Yokota, A. *et al* 2003; Zhou, Y. L. *et al* 2005), protein-peptide (Brokx, R. D. *et al* 2001; Jin, R. S. *et al* 2006; Rainaldi, M. *et al* 2007; Yamniuk, A. P. and Vogel, H. J. 2004) or protein-small molecule (Burnett, J.

C. *et al* 2007; Engel, M. *et al* 2006; Gulbis, J. M. *et al* 1996; Yang, F. *et al* 2007). In some cases, there are deviations from the standard (Brokx, R. D. *et al* 2001; Jin, R. S. *et al* 2006) interactions and appear to get stronger with increased temperature (Rainaldi, M. *et al* 2007; Yamniuk, A. P. and Vogel, H. J. 2004; Zhou, Y. L. *et al* 2005). In some cases, weaknesses have also been observed in ITC data which resulted in fluctuations in thermodynamic properties returned (Knipscheer, P. *et al* 2007; Lim, N. S. *et al* 2006; Rainaldi, M. *et al* 2007).

Protein-protein interactions which demonstrate a negative Gibbs free energy have commonly shown to be enthalpically favourable and entropically unfavourable (Knipscheer, P. *et al* 2007; Nishimiya, Y. *et al* 2000; Shiroishi, M. *et al* 2001; Yokota, A. *et al* 2003; Zhou, Y. L. *et al* 2005). Protein-peptide interactions will generally show the same trends in enthalpic return and entropic cost (Rainaldi, M. *et al* 2007; Yamniuk, A. P. and Vogel, H. J. 2004). Protein-small molecule interactions however, present a small enthalpic return and also be entropically favourable (Burnett, J. C. *et al* 2007; Engel, M. *et al* 2006; Yang, F. *et al* 2007). These general trends are common, however, this is not always the case. Protein-peptide interactions can show fluctuations in thermodynamic properties. Peptide length and hydrophobicity can result in thermodynamic parameters deviating from the common (Brokx, R. D. *et al* 2001; Jin, R. S. *et al* 2006).

6.2.8 Enthalpy Entropy compensation effects

Published ITC interactions have demonstrated an increase in K_a with an increase in temperature, an effect which seems counter-intuitive. This is reported to be due to an enthalpy entropy compensation effect. The interaction shows an increase in enthalpy with increasing temperature, which would result in a weaker interaction.

There is however a greater increase in the TΔS of the interaction; these two factors combined result in a slight increase in affinity (Rainaldi, M. *et al* 2007; Yamniuk, A. P. and Vogel, H. J. 2004; Zhou, Y. L. *et al* 2005).

6.2.9 Use of ITC to determine changes in specific heat capacity of solutions

ITC can reveal the change in specific heat capacity (ΔC_p) of a solution upon protein binding interactions (Jelesarov, I. and Bosshard, H. R. 1999; Liang, Y. 2008). This is possible due to fluctuations of ΔH with temperature, as a direct result of the ΔC_p of an interaction (Zhou, Y. L. *et al* 2005). Changes in the quantity of immobilised waters, the level of solvation, are responsible for the ΔC_p of an interaction. The quantity of immobilised waters is a relationship with the number of exposed hydrophobic and hydrophilic residues, so the levels of these residues have a relationship with ΔC_p (Liang, Y. *et al* 2001; Loladze, V. V. *et al* 2001; Rainaldi, M. *et al* 2007; Spolar, R. S. and Record, M. T. 1994; Yokota, A. *et al* 2003). Protein-protein interactions have demonstrated a general $-\Delta C_p$ of interaction, suggesting that interactions are driven by the burial of hydrophobic residues (Brokx, R. D. *et al* 2001; Jin, R. S. *et al* 2006; Nishimiya, Y. *et al* 2000; Yamniuk, A. P. and Vogel, H. J. 2004).

The value of ΔC_p for a protein interaction can be determined by ITC by obtaining a value of ΔH at different temperatures. The enthalpy is plotted against the temperature and the gradient of this plot is the ΔC_p of the interaction ($\Delta H/\Delta T$). In solution, proteins have waters bound to their surfaces, a solvation shell. The quantity of immobilised waters is related to the ΔC_p of the solution. During a binding event, the total quantity of bound waters could change as peptide side groups become solvent accessible, this is the ΔC_p of an interaction. The quantity of bound waters are

related to the number of solvent exposed hydrophobic and hydrophilic residues, these can be used to approximate ΔC_p. Values for ΔC_p have been determined for full protein folding (Liang, Y. *et al* 2001) which presented a ΔC_p of -156 kJ mol^{-1} K^{-1}. This large negative value is a result of a large number of hydrophobic residues becoming buried in the centre of a protein structure. Conversely, a large $+\Delta C_p$ of protein unfolding has also been determined using differential scanning calorimetry (DSC) (Loladze, V. V. *et al* 2001). Ubiquitin has been analysed with point mutations, some of which increased the number of buried hydrophilics, others which increased the number of buried hydrophobics. If hydrophilic residues were buried in the centre of the protein, there was a decrease in the ΔC_p of unfolding. If hydrophobic residues were buried there was an increase in the ΔC_p of unfolding (Loladze, V. V. *et al* 2001). ΔC_p has been previously determined using structural information of binding related to the change in exposed hydrophobics and hydrophilics (Spolar, R. S. and Record, M. T. 1994). In other examples, this calculation has been shown to be grossly inaccurate, and it is argued that this inaccuracy is a result of waters immobilised in the binding interface (Yokota, A. *et al* 2003).

Many protein-protein interactions have been determined, the bulk of which have shown to have a negative ΔC_p (Brokx, R. D. *et al* 2001; Jin, R. S. *et al* 2006; Rainaldi, M. *et al* 2007; Yokota, A. *et al* 2003). These interactions demonstrate a burial of hydrophobic residues at the binding interface. There are protein interactions demonstrating a positive ΔC_p where there are significant numbers of charged residues in the binding interface (Nishimiya, Y. *et al* 2000). Many reasonable values for ΔC_p, have been determined using ITC for protein interactions, and commonly

these values can be rationalised by the solved complex structures, as are shown in (Morton, C. J. and Ladbury, J. E. 1996; Nishimiya, Y. *et al* 2000; Sleigh, S. H. *et al* 1999; Yokota, A. *et al* 2003).

6.2.10 Role of bound waters in protein thermodynamic interactions

Waters within a protein binding site can affect different aspects of the interaction. A protein binding site can facilitate binding to peptides with different side groups from the ideal as a result of waters compensating for the mutation. The immobilisation of waters can result in alterations in the enthalpic, entropic and specific heat capacity shifts of interactions (Morton, C. J. and Ladbury, J. E. 1996; Nishimiya, Y. *et al* 2000; Shiroishi, M. *et al* 2001; Sleigh, S. H. *et al* 1999; Yokota, A. *et al* 2003).

Mutations can present decreased binding affinity in protein-peptide interactions. Interactions which should be reduced to only very slight affinities can still demonstrate binding. These persisting interactions are due to the recovery of peptide binding by water. When bound to mutant peptide, water can occupy now redundant atom sites and bridge interactions between protein and peptide. Water bridges result in binding recovery and can present only slight reductions in affinity even with significant alterations in residues e.g. Gln→His (Sleigh, S. H. *et al* 1999). It is documented in published data that waters can be common in many protein-protein interactions, and be pivotal for tight binding affinity (Yokota, A. *et al* 2003). In antibody-antigen complexes, waters are commonly immobilised at the binding interface. When point mutations are introduced to the antibody, significant reductions in affinity were not observed and it is argued that this was due to the shift

in water position, presenting a compensation for the mutation (Yokota, A. *et al* 2003).

6.2.11 Changes in the thermodynamic values of interaction as a result of bound water

In some interactions which use water in the binding interface, the effect of this water has been demonstrated by point mutations to hydrophobic residues. Trans-water interactions are shown to be enthalpically favourable, however do present an entropic cost due to the immobilisation of another molecule (Nishimiya, Y. *et al* 2000; Yokota, A. *et al* 2003). In a similar example, a close homologue in which this water is not shown, has been tested. This experiment mutated tyrosine to hydrophobics, and resulted in an increased enthalpic return with increased entropic cost (Nishimiya, Y. *et al* 2000). These examples presented data that immobilisation of waters in protein binding interactions may return an enthalpic yield but are however entropically expensive.

ΔC_p is directly related to the freedom of waters in a system (Liang, Y. 2008; Loladze, V. V. *et al* 2001; Morton, C. J. and Ladbury, J. E. 1996). This freedom is a result of solvent exposed hydrophobics, and has previously been calculated from crystal structures alone (Spolar, R. S. and Record, M. T. 1994). This theoretical derivation of ΔC_p has been shown to be inaccurate in systems when determined experimentally by ITC. It is argued that this is a result of water molecules being immobilised on the binding interface and result in a decrease of ΔC_p (Brokx, R. D. *et al* 2001; Liang, Y. 2008; Loladze, V. V. *et al* 2001; Morton, C. J. and Ladbury, J. E. 1996).

6.3 Auto-ITC materials and methods

ITC experiments were carried out on the Auto-ITC HT Microcalorimeter (MicroCal) and data was de-convoluted using the Origin7 (SR4) v7.0552 and the returned values were graphed using the program KaleidaGraph v4.03. The reference cell was given two hours to equilibrate thoroughly with sample prior to the start of the experiment. Samples were stored for runs in 96well 2ml round bottom plates (Greiner) at 15°C. Protein concentration was determined by A_{280}, and peptide concentrations were determined by the dissolved mass and from ±5% A_{280}, assuming peptide extinction coefficient (ε=1490$M^{-1}cm^{-1}$) generated using ProtPram (Gasteiger, E. *et al* 2003). Controls were performed with peptide only to give dilution energy. The energy required for dilution was deducted as a constant from sample curves. 1mg of lyophilised peptide was dissolved in 50µl of DMSO, 10µl of peptide solution was diluted in HEPES biochemistry buffer (10mM HEPES, 150mM NaCl, pH7.4) to obtain a 1% DMSO final concentration. Higher peptide concentrations required the increase to 2% DMSO solutions.

6.3.1 PCNA peptide interaction quantification by ITC

The associations of the six different peptides –

Peptide0 – KKRQTSMTDFYHSKRRLIFS

Peptide1 – KRRQTSMTDFYH

Peptide2 – Bio-KRRQTSMTDFYH

Peptide3 – Bio-PEG-KRRQTSMTDFYH

Peptide4 – Bio-PEG-PEG-KRRQTSMTDFYH

Peptide5 – Bio-PEG-PEG-Ahx-KRRQTSMTDFYH

with Hs, Sp and LmPCNAs were analysed by ITC to determine the K_d of interaction free in solution. This was performed at two temperatures, 283 and 298K, in order to determine consistency for BIAcore data (section8). These experiments were performed by injecting peptide as ligand into PCNA solution contained within the reaction cell, (see section 7.5). These binding affinities could present insight into the effect of spacers and differences between the PCNA homologues.

6.3.2 PCNA peptide concentration determination

Evaluation of the ITC data was performed using the accurate concentrations of protein, determined by A_{280}, and peptide, determined by mass of dissolved peptide. This resulted in stoichiometry and entropy which appeared to be incorrect. The stoichiometry of this protein to the peptide is known from ITC and crystal structures to be 1:1 (Bowman, G. D. *et al* 2004; Bruning, J. B. and Shamoo, Y. 2004) (1AXC,1SXJ) and the entropy is most likely to be negative, demonstrating a decrease in the level of disorder in the solution. When the peptide concentration is determined by A_{280} stoichiometry and entropy appear to approach expected values (Tables 1.5, appendix). The concentration determined by A_{280} was always lower than that determined by dissolved mass of the peptide. This could be due to salts in the lyophilised peptide, or water coordinated around the hydrophilic portions of the peptide. The peptide concentration as determined by the A_{280} is most likely to be accurate, therefore data shown is calculated using peptide concentrations based on peptide A_{280}.

6.3.3 ITC study of HsPCNA

Both reference cell and sample cells contained identical solutions of 5µM HsPCNA. HsPCNA and peptide solutions were buffered in ITC buffer (10mM HEPES, 150mM NaCl, 1%DMSO, pH7.4). 30 times 8µl injections of peptides0-1 were injected into the sample cell at 600 second increments. 17 times 17µl injections of peptides2-5 were injected into the sample cell at 400 second increments. Average peptide concentrations were: peptide 0 29.4µM; peptide 1 71.0µM; peptide 2 64.0µM; peptide 3 59.7µM; peptide 4 51.1µM; peptide 5 54.4µM.

6.3.4 ITC study of SpPCNA

SpPCNA in ITC buffer was used in both reaction and reference cell at 5µM concentration. 20 times 12µl injections of peptide 0 were injected into the sample cell at 400 second increments. 17 times 17 µl injections of peptide 1 were injected into the sample cell at 400 second increments. 15 times 18µl injections of peptides 2-5 were injected into the sample cell at 400 second increments. Average peptide concentrations were: peptide 0 27.7µM; peptide 1 78.8µM; peptide 2 72.4µM; peptide 3 56.7µM; peptide 4 47.9µM; peptide 5 59.1µM.

6.3.5 ITC study of LmPCNA

LmPCNA in 2% DMSO ITC buffer (10mM HEPES, 150mM NaCl, 2% DMSO, pH7.4) was used in both reaction and reference cells. 15 times 18µl injections of each peptide were injected into the sample cell at 400 second increments. Average peptide concentrations were: peptide 0 29.4µM; peptide 1 156µM; peptide 2 169µM.

6.4 Results of PCNA peptide binding affinities determined by ITC

6.4.1 Results of HsPCNA-Peptide interactions determined by ITC

The interaction of HsPCNA with different peptides as measured by ITC is shown in figure 6.4 with K_d of interaction presented in figure 6.5. The interactions were performed at 283K and 298K (figure 6.5) however only those at 283K are presented in figure 6.4.

The data shown in figure 6.4 demonstrates good signal levels (-20 kJ $mol_{(injectant)}^{-1}$ for peptide 0). The isotherms for peptides (1-5) gave little curvature with increasing molar ratio which can reduce the accuracy of the determined values. The lack of curvature is a result of the low peptide concentrations used in order to maintain DMSO concentrations of 1% (v/v). Due to the lack of curvature in the isotherm, the determined K_d of interaction may be error prone. The determined value of K_d is largely reliable as it is determined based on the mid-point of the curve and molar concentrations. The interaction with peptide 0 demonstrates the lowest K_d of interaction (5.15nM) (figure 6.5). Peptides 1-5 have similar affinities (5.57 – 0.97 μM) showing little effect of the biotin tag across different linker lengths.

The enthalpy and entropy of interactions are dependent on an accurate determination of K_d. With uncertainty in the K_d, the enthalpy and entropy of interaction may contain significantly greater levels of error. When peptide concentrations were determined by the A_{280}, the determined enthalpy of interaction is shown in each peptide to be negative (favourable). The enthalpy of interactions between HsPCNA and peptides 1-5 determined from these isotherms had a range of 31.8 kJ mol^{-1} (between -19.9 kJ mol^{-1} and -51.7 kJ mol^{-1}). The entropy of interaction for peptides 1-5 deviated by greater levels as this value is determined from the K_d

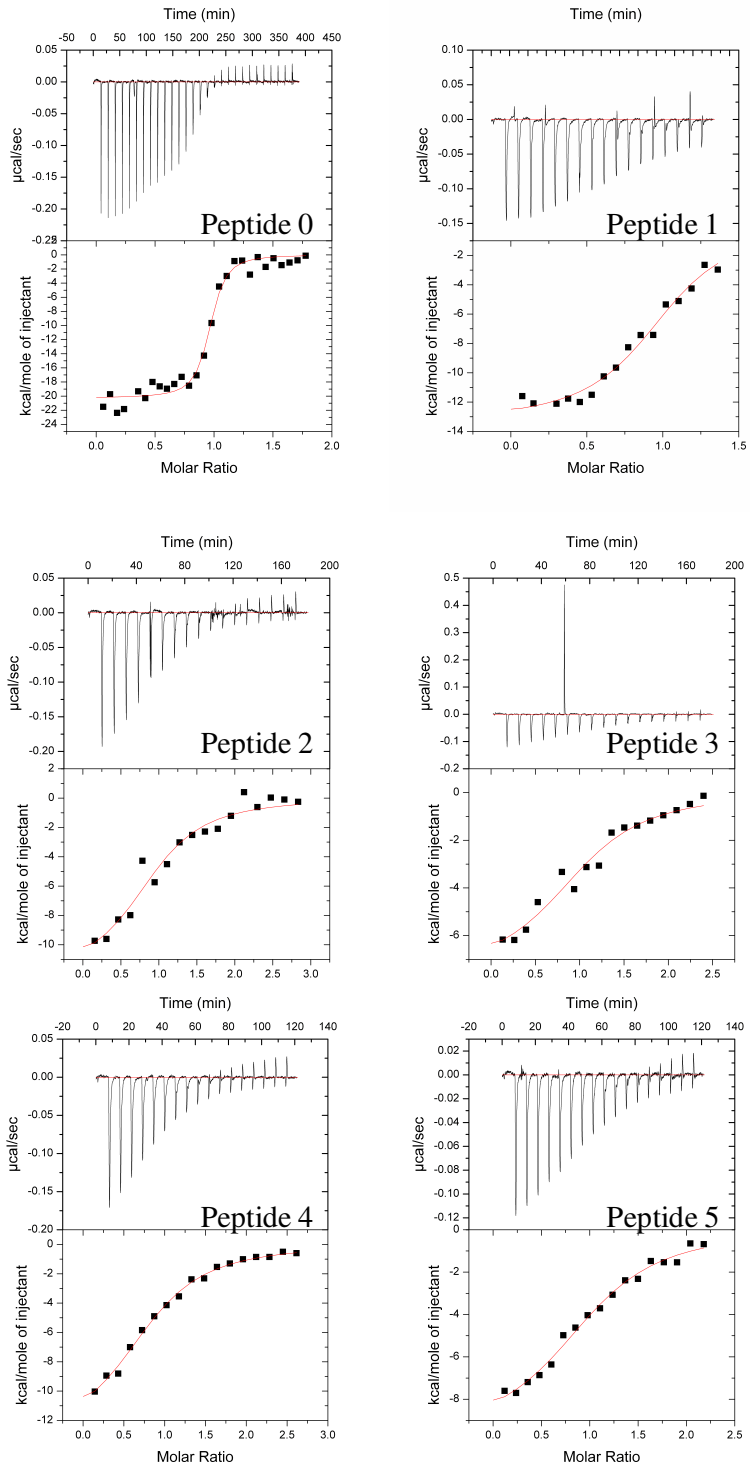
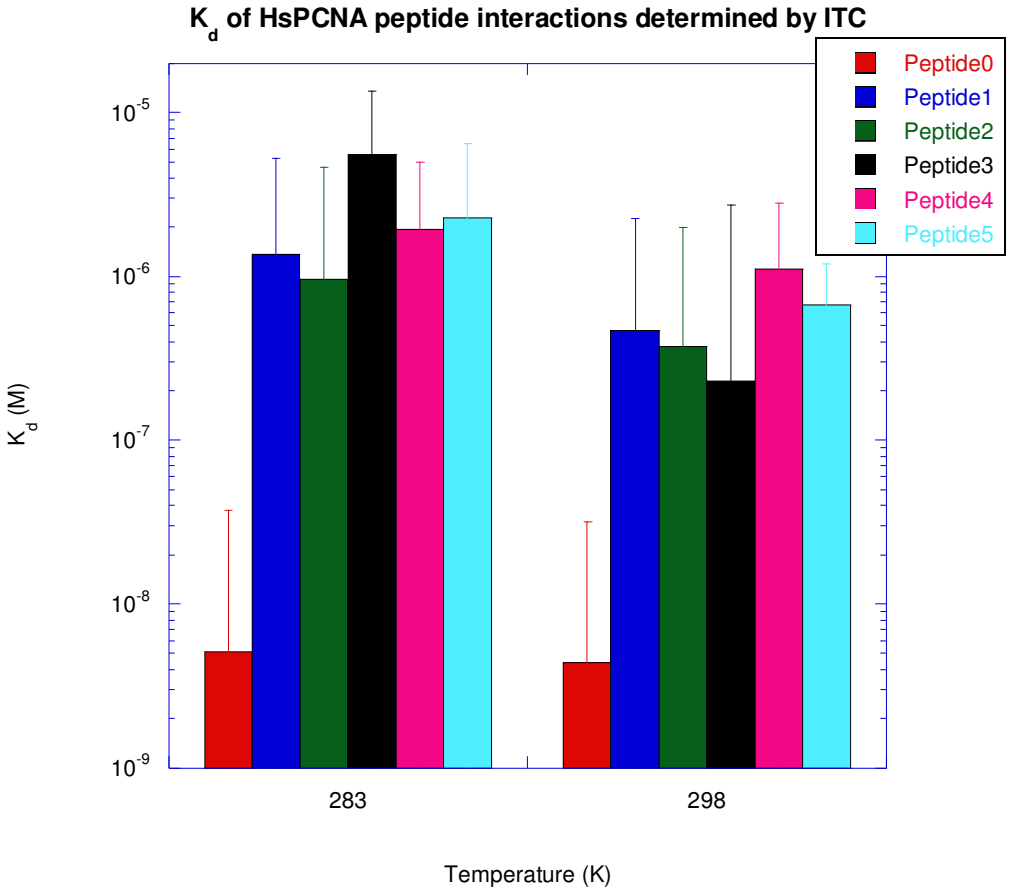


Figure 6-4. **The ITC isotherms of the HsPCNA peptide interactions at 283K**

Top panel) The ITC isotherm showing the energy input of the baseline sample (in µcal s^{-1}) against time (in minutes).

Bottom panel) The integrated data of the ITC injections fit as above plotted, as kcal $mol_{(injectant)}^{-1}$ against the molar ratio of protein and peptide.

These six different images show the ITC interactions of HsPCNA for peptide 0-5. The HsPCNA peptide 0 interaction shows clearly that there is a cooling upon peptide dilution. Peptide 3 shows a large data spike which gave an aberrant integrated data point removed from the fitted data.



Peptide	K_d at 283K (nM)	K_d at 298K (nM)
0	5.15	4.40
1	1370	470
2	965	375
3	5570	230
4	1940	1120
5	2300	670

Figure 6-5. **K_d of Peptides for HsPCNA at differing temperatures**
The bar chart (top) shows the K_d of peptide injected into HsPCNA solution with error bars demonstrating the shift in K_d due to the use of concentration determined by dissolved mass of peptide. This is present for both temperatures tested. The table (bottom) shows the figures for the K_d (in nM) using peptide concentration determined by the A_{280}. This data demonstrates that peptide0 is the strongest interaction, with peptides1-5 having similar K_ds. The error in this experimental method results in no clear shift as a result of temperature.

and enthalpy of interaction. The entropies of interaction between HsPCNA and peptides 2-5 are in the range of 113 J $Kmol^{-1}$ (46.9 to -69.5 J $Kmol^{-1}$).

Changes in the temperature of interaction resulted in a shift in the measured K_d of interaction. The affinity of interaction increased giving a slight (but consistent across all peptides) decrease in the K_d (e.g. peptide 0 drops from 5.15 nM to 4.40 nM). The change in K_d of interaction with temperature does vary greatly from a 1.2 (peptide 0) to a 24 fold (peptide3) increase in affinity. Peptides 1-5 gave an average of ~4-fold increase in affinity of interaction e.g. peptide 1 drops from 1.37 μM to 0.47 μM and peptide 5 drops from 2.30 μM to 0.67 μM.

6.4.2 Results of SpPCNA-Peptide interactions determined by ITC

The interaction of SpPCNA with different peptides as measured by ITC is shown in figure 6.6 with K_d of interaction presented in figure 6.7. The interactions were performed at 283K and 298K (figure 6.7) however only those at 283K are presented in figure 6.6.

The data shown in figure 6.6 demonstrates reasonable signal levels (-9 kJ $mol_{(injectant)}^{-1}$ for peptide 1). The isotherms for peptides 0, 1 and 2 gave reasonable curvature and peptides 3, 4 and 5 gave little curvature with increasing molar ratio. As stated above, isotherms with little curvature can reduce the accuracy of the determined values. The interaction with peptide 0 demonstrates the lowest K_d of interaction (18.5 nM) (figure 6.7). Peptide 1 demonstrated a greater affinity (78 nM) than peptides 2-5 (400 – 925 nM) (figure 6.7). The presence of the biotin tag slightly reduces affinity, however there is little effect of the biotin tag across different linker lengths.

The enthalpy of interaction is shown in each peptide to be negative (favourable). The enthalpy of interactions between SpPCNA and peptides 1-5,

determined from these data, had a range of 26.3 kJ mol^{-1} (between -30.9 kJ mol^{-1} and -57.2 kJ mol^{-1}). The entropy of interaction for peptide 1-5 deviated by greater levels as this value is determined from the K_d and enthalpy of interaction. The entropies between SpPCNA and peptides 2-5 are in the range of 118 J $Kmol^{-1}$ (31.5 to -87.0 J $Kmol^{-1}$). Variations in the thermodynamic values determined are consistent between HsPCNA and SpPCNA.

Changes in the temperature of interaction resulted in a shift in the measured K_d of interaction. The affinity of interaction increased giving a slight (but consistent across all peptides) decrease in the K_d (e.g. peptide 0 drops from 18.5 nM to 9.5 nM). The interactions at higher temperatures were however too weak for K_d values to be determined. The change in K_d of interaction with temperature does not vary greatly from a 1.25 (peptide 0) to a 2.26 fold (peptide 3) increase in affinity. Peptides 1-3 gave an average of ~1.7 fold increase in affinity of interaction e.g. peptide 1 drops from 77 nM to 62 nM and peptide 3 drops from 930 nM to 41 nM. These trends are consistent with that of HsPCNA; however with significantly less variation in numbers derived from SpPCNA.

6.4.3 Results of LmPCNA-Peptide interactions determined by ITC

The interaction of LmPCNA with different peptides as measured by ITC is shown in figure 6.8 with K_d of interaction presented in figure 6.9. The interactions were performed at 283K and 298K (figure 6.9) however only those at 283K are presented in figure 6.8.

The data shown in figure 6.8 demonstrates poor signal levels (-6 kJ $mol_{(injectant)}^{-1}$ for peptide 1). The isotherms for peptides 0 and 1 gave reasonable curvature and peptide 2 gave little curvature with increasing molar ratio. ITC

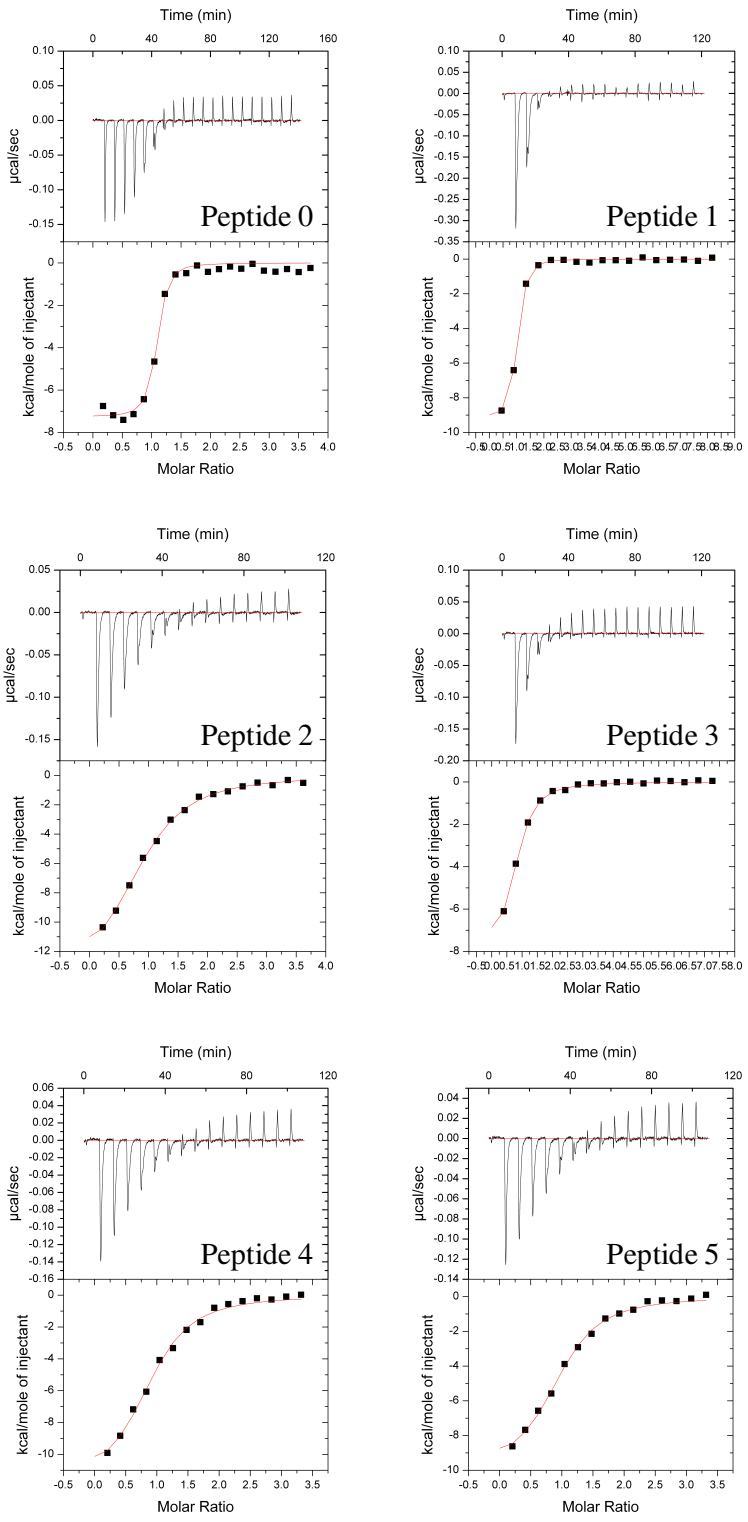
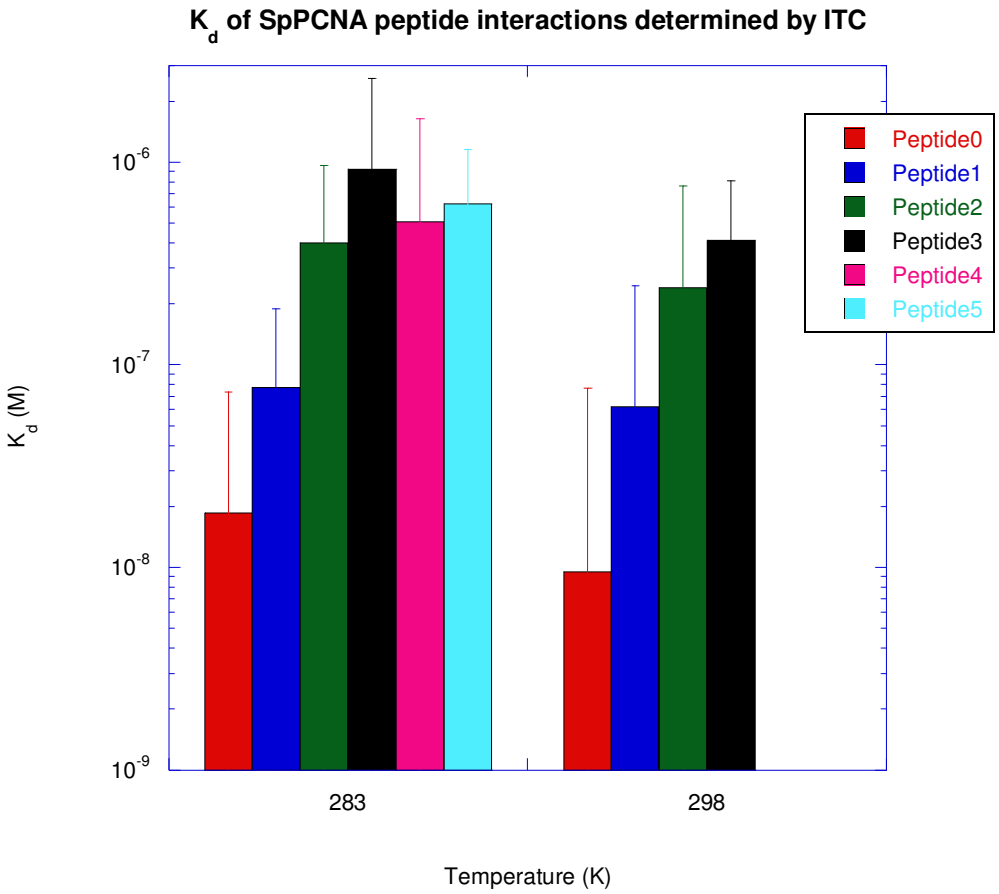


Figure 6-6. **The ITC isotherms of the SpPCNA peptide interactions at 283K**

Top panel) The ITC isotherm showing the energy input of the baseline sample (in µcal s^{-1}) against time (in minutes). Bottom panel) The integrated data of the ITC injections fit as above plotted as kcal $mol_{(injectant)}^{-1}$ against the molar ratio of protein and peptide. These six different images show the ITC interactions of SpPCNA for peptides 0-5. The SpPCNA peptide 0 interaction shows clearly that there is a cooling upon peptide dilution. Each of these interactions fit smoothly to a curve.



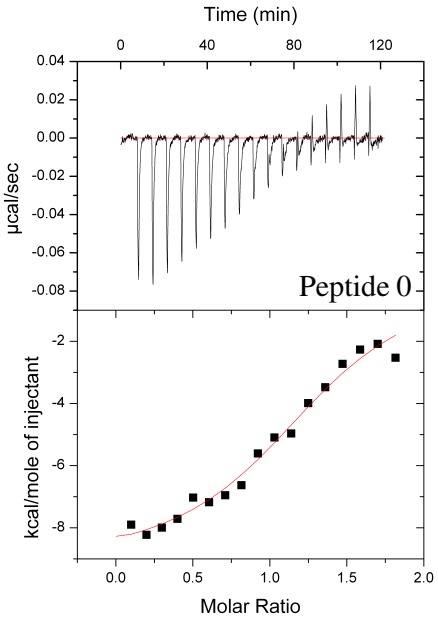
Peptide	K_d at 283K (nM)	K_d at 298K (nM)
0	18.5	9.50
1	77.5	62.0
2	400	240
3	925	410
4	505	
5	625	

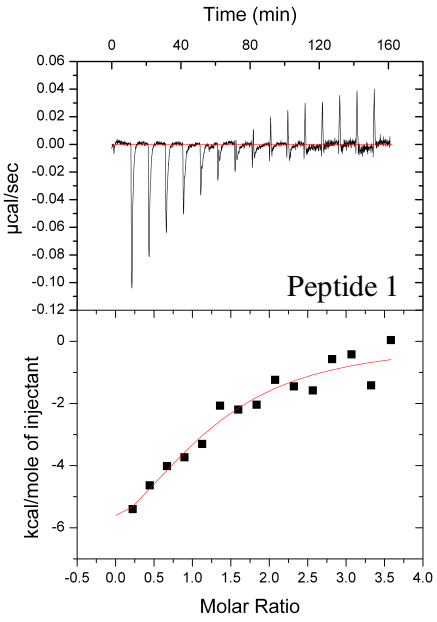
Figure 6-7. **K_d of Peptides for SpPCNA at differing temperatures**
The bar chart (top) shows the K_d of peptide injected into SpPCNA solution with error bars demonstrating the shift in K_d due to the use of concentration determined by dissolved mass of peptide. This is present for both temperatures tested. The table (bottom) shows the figures for the K_d (in nM) using peptide concentration determined by the A_{280}. This data demonstrates that peptide0 is the strongest interaction, followed by peptide1, with peptides2-5 having similar K_ds. The error in this experimental method results in no clear shift as a result of temperature, however interactions with peptides 4 and 5 were too weak for accurate values to be determined at 298K.

interactions between LmPCNA and peptides 3, 4 and 5 gave no useable data. The interaction with peptide 0 demonstrates the lowest K_d of interaction 408 nM (figure 6.9). Peptides 1 and 2 had similar affinities (2.46 – 2.50 μM) (figure 6.9) indicating little difference due to the presence of the biotin tag.

The enthalpy of interaction is shown in each peptide to be negative (favourable). The enthalpy of interactions between LmPCNA and peptides 1-2 determined from these isotherms had a range of 13.7 kJ mol^{-1} (between -17.9 kJ mol^{-1} and -31.6 kJ mol^{-1}). The entropy of interaction for peptides 1-5 deviated by greater levels as this value is determined from the K_d and enthalpy of interaction. The entropies between LmPCNA and peptides 1-2 are in the range of 44.3 J $Kmol^{-1}$ (40.7 to -3.60 J $Kmol^{-1}$). Variations in the thermodynamic values determined are consistent between HsPCNA, SpPCNA and LmPCNA.

Changes in the temperature of interaction resulted in a shift in the measured K_d of interaction. The affinity of interaction decreased giving a slight (but consistent across all peptides) increase in the K_d (e.g. peptide 0 increases from 408 nM to 426 nM). The interactions at higher temperatures were however too weak for K_d values to be determined. The change in K_d of interaction with temperature does not vary greatly from a 1.04 (peptide 0) to a 6.52 fold (peptide 2) decrease in affinity. Peptides 1-3 gave an average of ~3.1 fold decrease in affinity of interaction with a 15°C increase in temperature e.g. peptide 0 increases from 408 nM to 426 nM and peptide 2 increases from 2.5 μM to 16.3 μM. This decreasing affinity with increasing temperature is expected however not consistent with values determined for HsPCNA and SpPCNA.





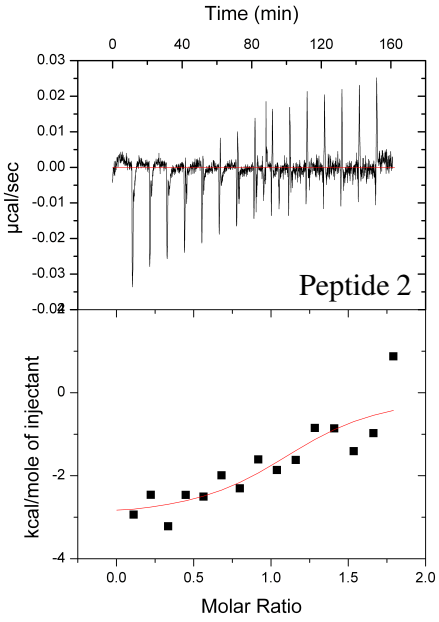
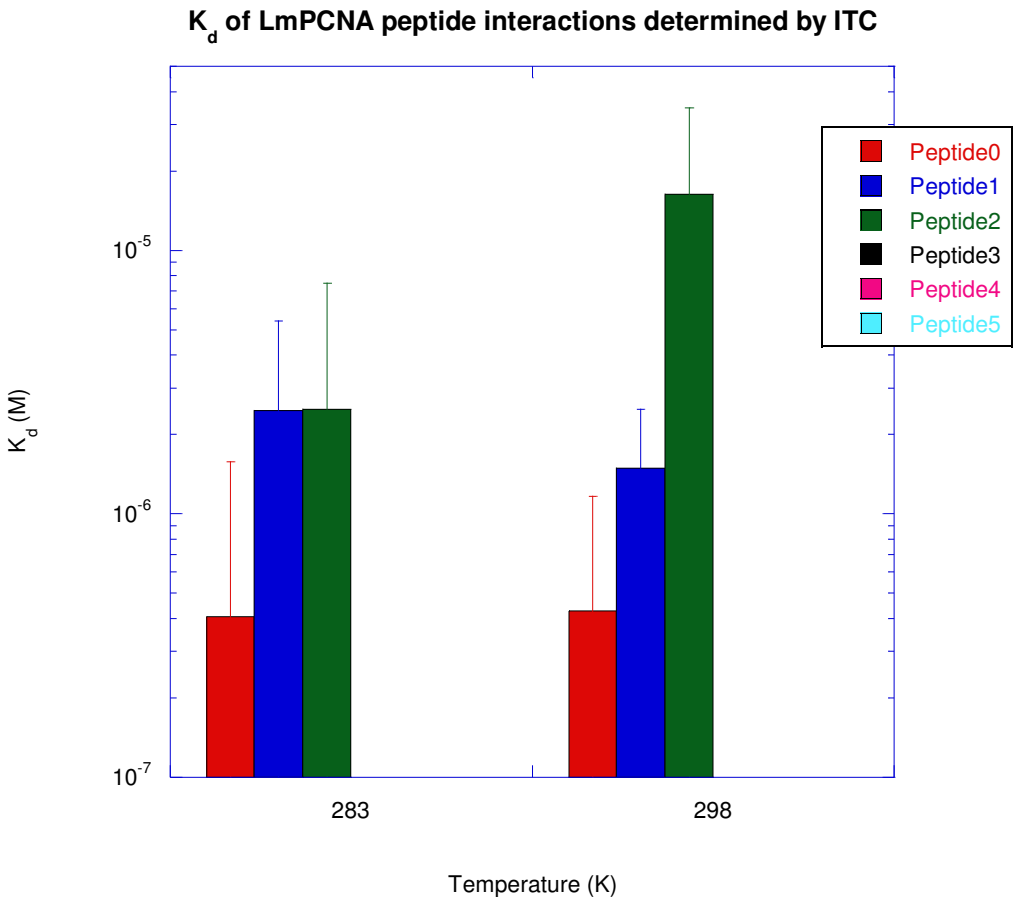


Figure 6-8. **The ITC isotherms of the LmPCNA peptide interactions at 283K**

Top panel) The ITC isotherm showing the energy input of the baseline sample (in µcal s^{-1}) against time (in minutes). Bottom panel) The integrated data of the ITC injections fit as above plotted as kcal $mol_{(injectant)}^{-1}$ against the molar ratio of protein and peptide.

These three different images show the ITC interactions of LmPCNA for peptides 0-2. The LmPCNA peptide 0 interaction shows clearly that there is a cooling upon peptide dilution. This peptide 0 interaction has the greatest affinity and peptide 2 gives only slight curvature. The isotherm for peptide 2 also shows a high noise to signal ratio reducing confidence in the fit numbers.



Peptide	K_d at 283K (μM)	K_d at 298K (μM)
0	0.408	0.426
1	2.46	1.49
2	2.50	16.3
3		
4		
5		

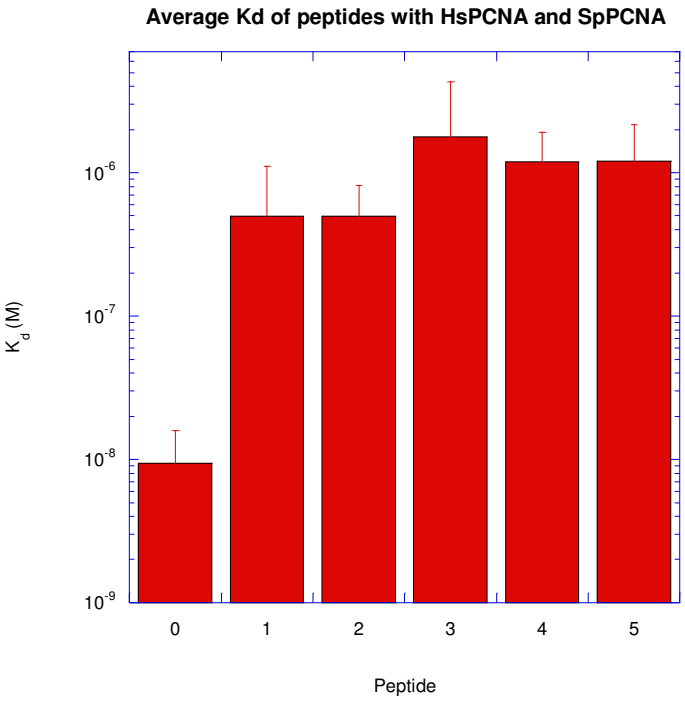
Figure 6-9. **K_d of Peptides for LmPCNA at differing temperatures**
The bar chart (top) shows the K_d of peptide injected into LmPCNA solution with error bars demonstrating the shift in K_d due to the use of concentration determined by dissolved mass of peptide. This is present for both temperatures tested. The table (bottom) shows the figures for the K_d (in μM) using peptide concentration determined by the A_{280}. This data demonstrates that peptide 0 has the strongest interaction, in agreement with interactions of Hs and SpPCNA. Weaker data returned for these interactions resulted in no values determined for peptides 3, 4 and 5 at either temperature.

6.5 Conclusions from the ITC analysis of PCNA-peptide interactions

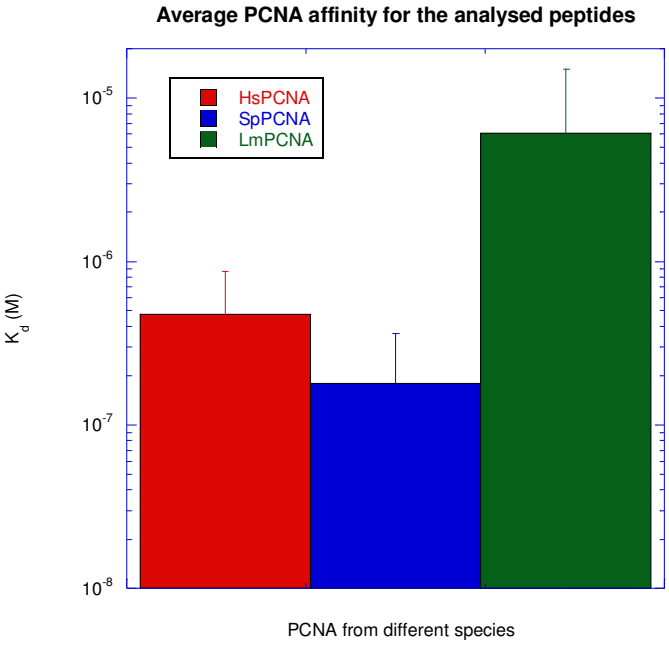
All three PCNAs from the different species demonstrated binding action for all six peptides. Peptide concentrations based upon the mass of dissolved peptide may be inaccurate therefore reliable thermodynamic data are difficult to obtain from this method. The available concentrations of peptide was also low, resulting in low signal levels. This poor signal level resulted in only affinity data to be determined (and Gibbs free energy). There also was a large dilution effect (figure 6.4 peptide 0). This is due to slight fluctuations in DMSO concentration between ligand solution and cell solution. LmPCNA and the Phe150Ala mutant peptide gave very weak interactions and so higher peptide concentrations were used but these still provided no reliable data. Higher peptide concentrations were unattainable without significant increases in DMSO concentration.

Robust numbers like those of the K_d and subsequently ΔG general trends can be observed. These numbers suggest that the trend in K_d of peptides was peptide 0<<peptides 1-5 (figure6.10) and that SpPCNA≈HsPCNA<<LmPCNA (figure 6.10B). The low K_d of peptide0 is most likely the result of the larger binding interface and multiple hydrogen bond formations. Conversely, the short peptides could demonstrate higher K_ds due to less binding surface. The similarity of affinity between peptides 1-5 is due to the identical sequence of peptide, and shows that the presence of linkers does not present any significant steric hinderance of the peptide PCNA interface. The presence of the biotin tag may result in a slight shift in affinity however this shift is consistent with all tagged peptides. Thermodynamic values for the interaction are more error prone. The determined enthalpy and entropy of interaction from these experiments can only infer that the interaction is enthalpically

Figure 6-10. **The average affinities of the PCNA peptide interactions as determined by ITC**



A) The average affinity of HsPCNA and SpPCNA for Peptides 0-5. This data clearly shows the reduction in affinity of the shorter peptide when compared to the longer peptide0. Each of the tagged peptides show very similar affinities for PCNA compared to that of the untagged peptide1, this demonstrated that there is little effect from the presence of the biotin linker (peptide2) and only slight shift with associated linkers (peptides 3-5). The standard deviation of the interactions is the deviation of affinity across all peptide binding data, including different target proteins (HsPCNA and SpPCNA) and different temperatures (283K and 298K).



B) The average affinity of the different PCNAs analysed for peptides with accurate data determination. This graph shows a slight shift between HsPCNA and SpPCNA however there is a significant reduction in affinity with LmPCNA. The averages and standard deviations were created using only data returned for all PCNAs, e.g. peptides 0, 1 and 2 at 283K and 298K.

favourable and entropically unfavourable.

As the temperature increased, the K_d of interaction appeared to drop slightly in HsPCNA and SpPCNA. This shift may be due to the large potential error in the experimental method from the imprecise concentrations of peptide. However, this could also be the result of enthalpy-entropy compensation (Rainaldi, M. *et al* 2007; Yamniuk, A. P. and Vogel, H. J. 2004; Zhou, Y. L. *et al* 2005). The increase in affinity with increase in temperature could also be due to potential difference in solubility of peptide, resulting in a higher free peptide concentration being injected at higher temperatures.

The values determined in this chapter will be compared with those determined by SPR and thermal denaturation assays in chapter 8.

CHAPTER 7. Study of PCNA-peptide interaction using BIAcore T-100

The method by which ligand is immobilised to the sensor chip surface may affect the affinity with binding partners. In the series of experiments described in this chapter we investigate PCNA proteins from three different species and their interactions with a series of six related peptides (figure 7.1). The peptides (figure 7.1) have been selected to examine differences in linker length. The tagged peptides (peptides 2-5 figure 7.1) all have an N-terminal biotin tag. Peptide 3 has a ~10Å PEG spacer between the biotin and the N-terminus, peptide 4 has a ~20Å PEG spacer between the biotin and the N-terminus, whereas peptide 5 has a ~20Å PEG and ~10Å hydrophobic spacer between the N-terminus and biotin (figure 7.1). Two peptides (peptide 1 and peptide 6) have been selected to examine the effect of a p21 F150A point mutant upon affinity.

HIS-tagged proteins can be immobilised to the sensor surface specifically by the N-terminus using NTA sensor chips (Wear, M. A. *et al* 2005). NTA surfaces can suffer from baseline drift as a result of ligand leaching from the sensor surface, which is made worse at higher temperatures (Nieba, L. *et al* 1997). In this work, we have adapted a protocol to covalently couple the PCNA to the sensor surface to circumvent this problem. The SPR experiments have been carried out with PCNA coupled to a NTA sensor surface and with peptide bound to a SA surface. SA surfaces have Streptavidin covalently coupled to the dextran matrix and can bind ligand using the tight ($\sim 1\times10^{-15}$M) Avidin-Biotin interaction (Green, N. M. *et al* 1975), one of the strongest non-covalent biological interactions. Differences in the affinity and thermodynamic values of interaction can be determined to be the result of the method of immobilisation. Experiments described here set out to examine

Peptide1 – KRRQTSMTDFYH

Peptide2 – Bio-KRRQTSMTDFYH

Peptide3 – Bio-PEG-KRRQTSMTDFYH

Peptide4 – Bio-PEG-PEG-KRRQTSMTDFYH

Peptide5 – Bio-PEG-PEG-Ahx-KRRQTSMTDFYH

Peptide6 – KRRQTSMTDAYH

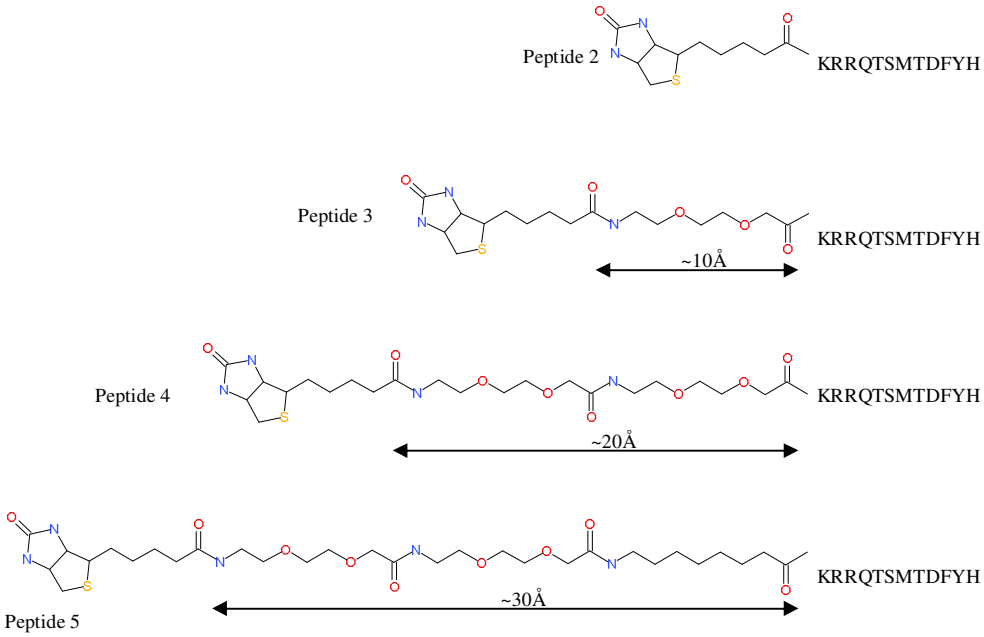


Figure 7-1. **The peptides investigated by SPR in this work**

Figure showing the peptides used in SPR experiments carried out in this study.
Top) Peptides 1-6 showing shorthand forms of all 6 peptides. Peptides 1-6 all have the same peptide sequence. Peptide 1 and 6 are untagged and peptide 6 has a F150A point mutation. Peptides 2-5 with biotin tags and linkers made of PEG and a hydrophobic spacer (Ahx).
Bottom) Peptides 2-5, showing the biotin tags and the chemical sequence of the linkers. The additional spacers have their approximate length shown below in Ångström.

whether peptides with short linkers have reduced affinity with immobilised peptide.

Affinity constants from these experiments can be derived at different temperatures 283-313K (10-40°C). K_ds derived over a temperature range can allow the determination of the enthalpy and entropy (Boysen, R. I. *et al* 1999) and the specific heat capacity of interaction (Murphy, K. P. *et al* 1990; Privalov, P. L. and Gill, S. J. 1988; Yoo, Seung Hyun and Lewis, Marc S. 2002). This series of experiments will also allow determination of the effect on binding properties of the interaction upon the mutation of Phe150Ala as well as the effect of biotin tags.

7.1 Materials and Methods of SPR experiments

All SPR experiments were carried out using the BIAcore T-100 (GE healthcare). Buffers used were consistent in all SPR experiments, BIAcore running buffer, 50µM EDTA, 0.05% P20 and 1% DMSO. 10mls of 100% DMSO (Fisher Scientific), 5mls of 10% v/v P20 (BIAcore) and 100µl of 0.5M EDTA were added to 100mls of HBS-N Buffer 10× (BIAcore), this was diluted to 1litre with ddH2O filtered through a 0.2µm filter. Prior to use, the BIAcore T-100 had the cleaning desorb run performed three times and the system was flushed thoroughly with buffer. Protein for these experiments was prepared as in chapter 2, eluted from the size exclusion column in HEPES biochemistry buffer. Data were evaluated using the BIAcore T-100 evaluation software Version 2.0.1 (GE healthcare) for the determination of the K_d of interaction. The thermodynamic analysis of interaction was performed using KaleidaGraph 4.0 (Synergy Software).

7.1.1 SPR Study of Triple Protein Sensor Chips

The study of immobilised PCNA surfaces was performed using NTA sensor surfaces (BIAcore). Sensor surfaces were washed for 30 seconds at 30µl min^{-1} with

50mM NaOH a minimum of three times prior to use, and equilibrated with BIAcore running buffer. Ni^{2+} is immobilised to the NTA surface using a 60 second injection at 10µl min^{-1} of Ni^{2+} charging solution; 1× HBS-N, 0.05% P20, 1%DMSO and 500µM $NiSO_4$. The sensor surface was activated for primary amine coupling using a 240 second injection at 5µl min^{-1} of 0.1M EDC, 25mM NHS solution. Protein at 100nM was injected at 30µl min^{-1} over the activated surface for 100 seconds. The surface was stripped with a 180 second injection of EDTA stripping buffer; 1× HBS-N, 0.05% P20 and 50µM EDTA. The level of immobilised protein was determined and if too low, Ni^{2+} was reloaded and more protein immobilised. Flow cell 2 had ~850RU of HsPCNA, flow cell 3 had ~800RU of SpPCNA and flow cell 4 had ~950RU of LmPCNA immobilised. The active surfaces were then quenched with a 300 second injection at 50µl min^{-1} of 1M Ethanolamine. Flow cell 1 was activated and quenched with no protein immobilised, and functioned as a reference cell. All surfaces were extensively washed with buffer untill all surfaces reached a steady baseline.

The peptides 1-5 were serially diluted from 50µM resulting in 25, 12.5, 6, 3, 1.5, 0.8, 0.4 and 0µM concentrations, peptide 6 was serially diluted from 100µM resulting in 50, 25, 12.5, 6, 3, 1.5, 0.8 and 0µM all in BIAcore running buffer. Peptides were passed over the surfaces at 50µl min^{-1} for 30 seconds with a repetition of the 3µM concentration. These experimental runs were performed three times at 283K, 288K, 293K, 298K, 303K, 308K, 313K (10-40°C at 5°C increments). Reference corrected data were used to determine the affinity of interaction for each peptide using the BIAcore T-100 Evaluation Software version 2.0.1 (GE Healthcare). The thermodynamic analysis was performed using KaleidaGraph 4.0.

7.1.2 Validation of the immobilised PCNA sensor surface activities

An NTA surface with 4 flow cells had ~850RU HsPCNA covalently immobilised on flow cell 2, ~800RU SpPCNA immobilised on flow cell 3, 950RU LmPCNA immobilised on flow cell 4 and flow cell 1 used as a reference cell. Protein were associated to the dextran matrix using the N-terminal 6xHIS tag Ni^{2+} affinity and then covalently immobilised using primary amine coupling in order to maximise surface activity. The theoretically obtained RU_{max} can be calculated with the known RU of PCNA immobilised (equation 7.1).

7.1) $$RU_{immobilised} = \left(\frac{MW_{ligand}}{MW_{analyte}} \right).N.RU_{\max}$$

The steady state fits of these interactions give values of the RU_{max}. When this is compared to theoretical RU_{max}, this ratio can give the surface activity. Hs, Sp and LmPCNA triple surfaces generated were ~57%, 44% and 99% active respectively. Three repeat runs were carried out using two separately created protein surfaces all of which had similar activities.

7.1.3 SPR Study of Immobilised Peptide Sensor Chips

Immobilised peptides were studied using an SA sensor surface which was cleaned and equilibrated as in section 7.1.1. Peptide was immobilised to the Streptavidin surface with a 15 second injection at 50µl min^{-1} of 50nM concentration of peptide. Injections were performed to reach a surface response of ~40RU on flow cells 2 and 4, with flow cells 1 and 3 used as reference cells. SA chip one had peptides 2 and 3 immobilised and SA chip two had peptides 4 and 5 immobilised. Once peptides were immobilised, the surfaces were equilibrated thoroughly with BIAcore running buffer to allow all surfaces to achieve a steady baseline.

Serial dilutions were made of HsPCNA and LmPCNA giving 50, 25, 12, 6, 3, 1.6, 0.8, 0.4 and 0µM concentrations and a dilution of SpPCNA giving 25, 12, 6, 3, 1.6, 0.8, 0.4, 0.2 and 0µM. These concentrations of PCNA were passed over the surfaces in 30 second injections at 30µl min^{-1} with repeat injections of 3µM PCNA. Experimental runs were performed three times at 283K, 288K, 293K, 298K, 303K, 308K, 313K (10-40°C at 5°C increments). Reference corrected data were used to determine the affinity of interaction for each immobilised peptide using the BIAcore T-100 Evaluation Software version 2.0.1 (GE Healthcare). The thermodynamic analysis was performed using KaleidaGraph 4.0.

7.1.4 Validation of the immobilised peptide sensor surface activities

The K_d of the three PCNAs (Hs, Sp and Lm) with surface of immobilised peptides were ascertained at seven temperatures. One SA surface had ~39.0RU of peptide 2 immobilised in flow cell 2 and ~38.4RU of peptide 3 immobilised in flow cell 4, with flow cells 1 and 3 used as reference cells respectively. A second SA surface had ~38.8RU of peptide 4 immobilised in flow cell 2 and ~37.6RU of peptide 5 immobilised in flow cell 4, with flow cells 1 and 3 used as reference cells respectively. Immobilisation of biotinylated peptides 2, 3, 4 and 5 to Streptavidin surfaces were ~0, 32, 44 and 41% active respectively (as determined using equation 7.1).

7.2 Results of SPR experiments

7.2.1 Determination of K_d between PCNA and free peptides at 25°C

Each chip with bound Hs, Sp and LmPCNA was used to examine the interaction with peptides 1-6 at different temperatures. Sensorgrams of the

interaction between PCNA and peptide 1 at 25°C are shown in figure 7.2. The immobilised PCNA peptide interactions showed no apparent mass transport problems (Karlsson, R. 1999). Analyte-ligand interactions which associate too rapidly can give a signal dependent on the rate of analyte delivery to the sensor surface. Mass transport problems are represented by the k_t of interaction in $RU.Ms^{-1}$ (figures greater than 1×10^7 $RU.Ms^{-1}$ for protein-protein interactions infer no mass transport problems). Immobilised PCNA gave k_t of 3.96×10^{15} $RU.Ms^{-1}$, 4.38×10^{13} $RU.Ms^{-1}$ and 7.61×10^8 $RU.Ms^{-1}$ for Hs, Sp and LmPCNA respectively.

The interactions presented in figure 7.2 have very fast on and off rates (k_{on} and k_{off}) (Oshannessy, D. J. *et al* 1993). At 25°C, the k_{on} of peptide 1 for Hs, Sp and LmPCNA are ~4.69×10^5 $M^{-1}s^{-1}$, 1.82×10^5 $M^{-1}s^{-1}$, 1.84×10^4 $M^{-1}s^{-1}$ respectively. The k_{off} of peptide 1 interaction with Hs, Sp and LmPCNA is ~0.8 s^{-1}, 0.9 s^{-1} and 1.4 s^{-1} respectively. Due to the rapid nature of the PCNA-peptide interactions, the data were analysed by the steady state of interaction (Adamczyk, M. *et al* 2000).

The steady-state analysis of the sensorgrams (figure 7.2) was used to determine the affinity of interaction. Affinity between HsPCNA, SpPCNA and LmPCNA for peptides 1-6 (at different temperatures) are shown in figures 7.4, 7.5 and 7.6 respectively. In contradiction with ITC data (chapter 6) but in agreement with literature data, the HsPCNA-peptide interactions were the strongest and the SpPCNA-peptide interactions were approximately 1.5 fold weaker; LmPCNA-peptide interactions were approximately 6 fold weaker than SpPCNA. Peptide 1 was determined to bind with the greatest affinity and peptides 2-5 demonstrated an approximately 5-fold weaker interaction, peptide 6 was shown to have an approximately 30-fold weaker interaction than peptides 2-5. K_d changed by ~1.7 fold

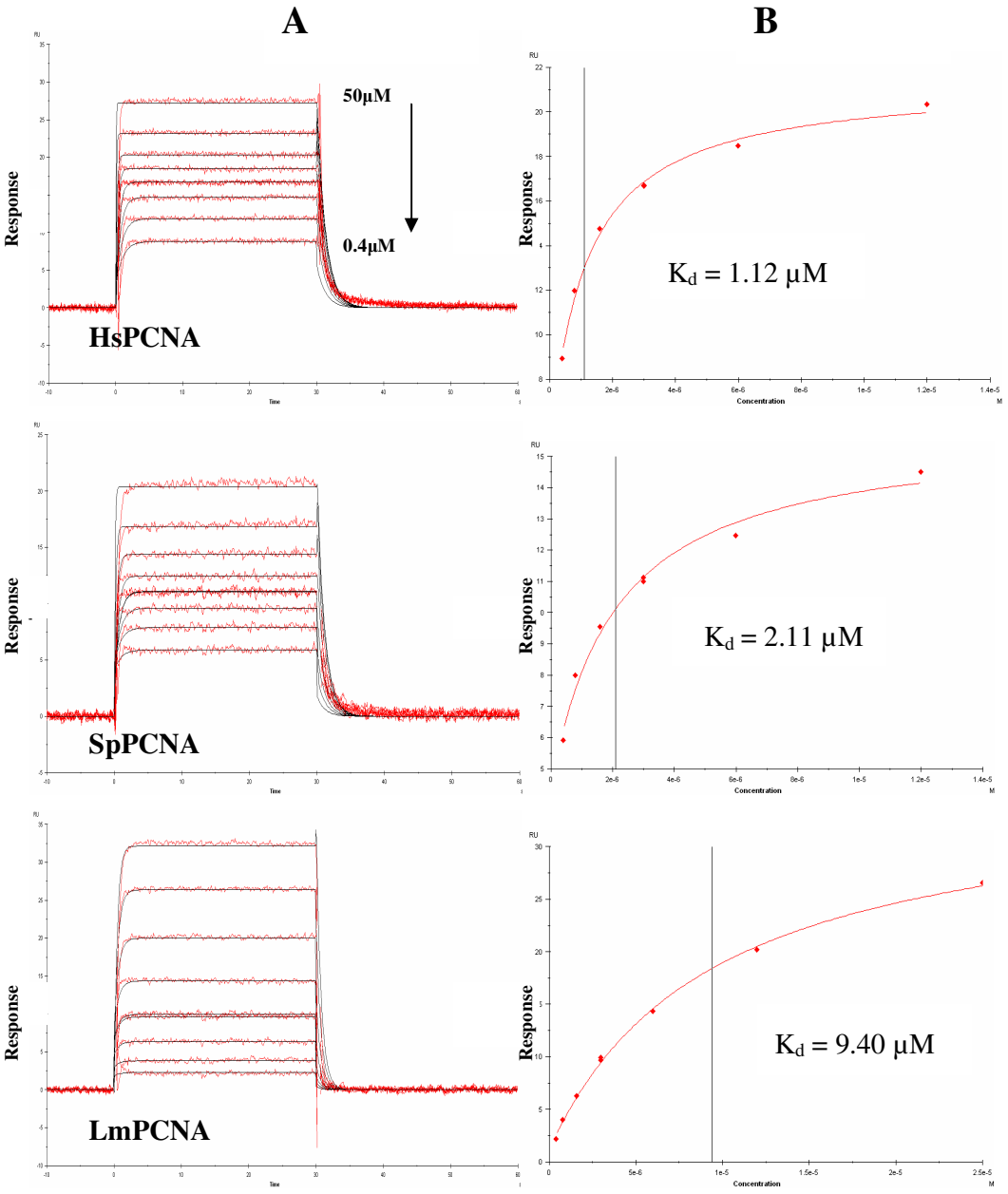
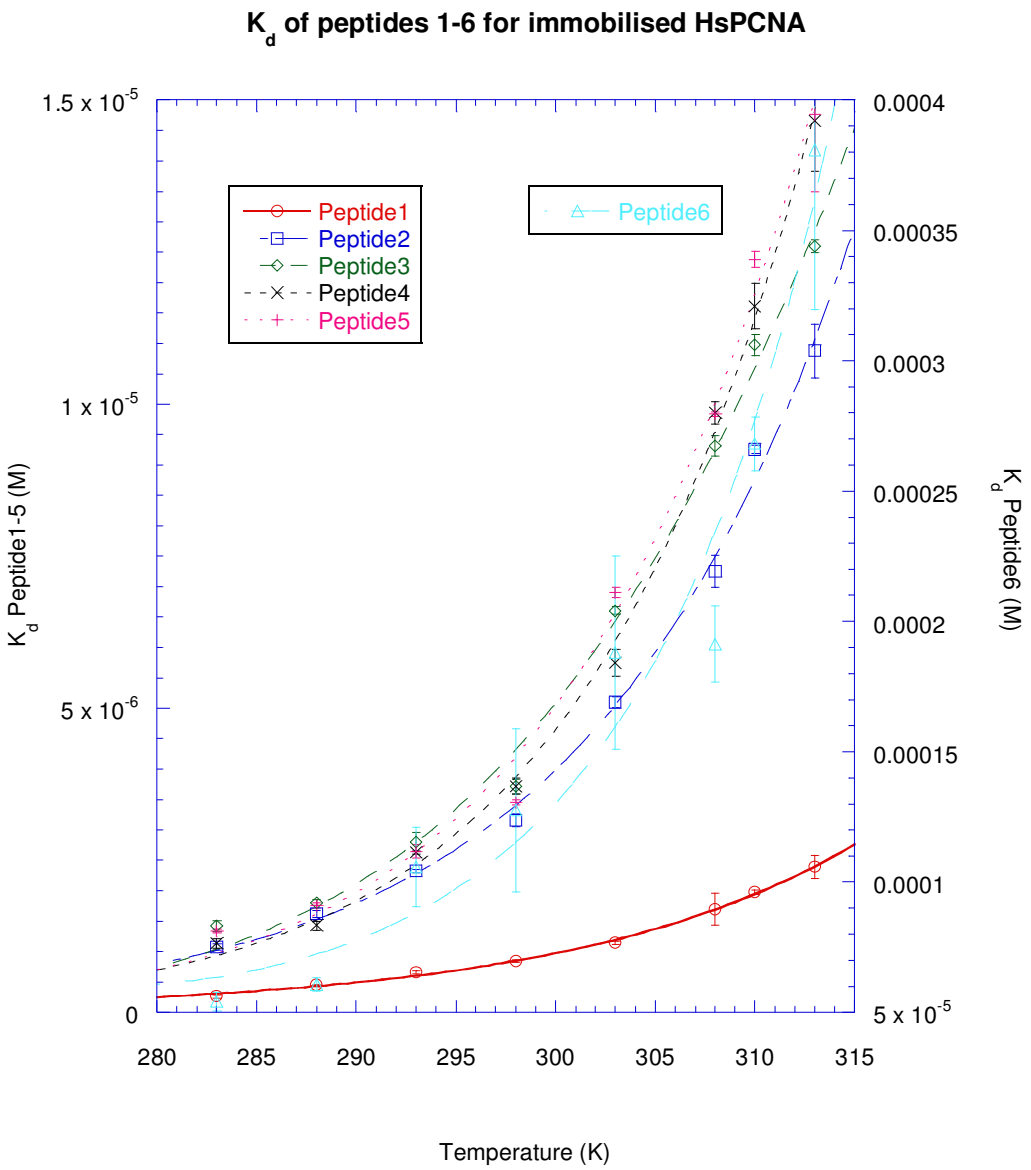


Figure 7-2. **BIAcore data for immobilised PCNA with peptide 1**

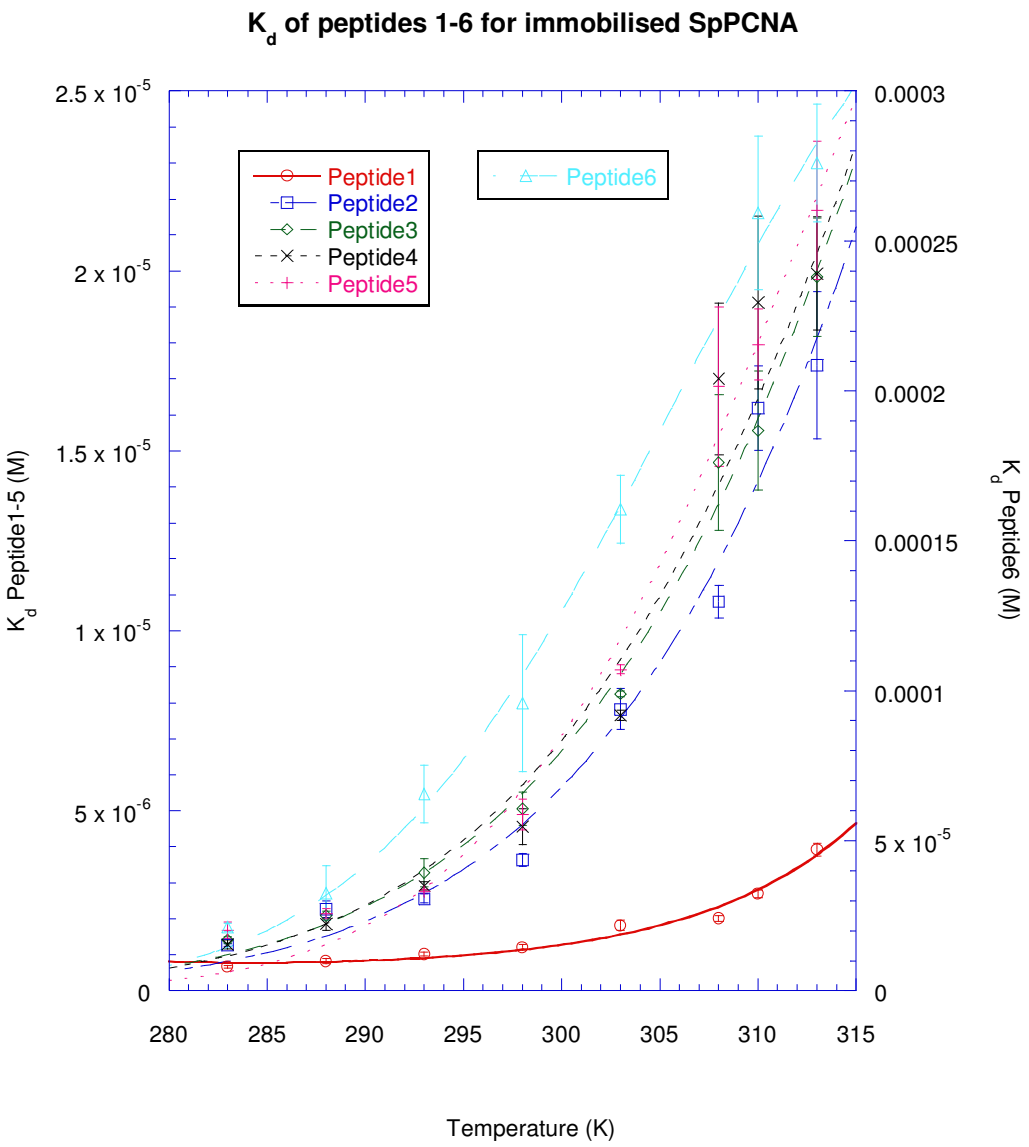
The interactions of peptide 1 for surfaces of BIAcore sensor chips with immobilised protein at 298K. **A)** Reference corrected results of BIAcore interaction plotted as RU against time in seconds. Serial dilutions of peptide 1 (50 to 0.4 µM) were passed over HsPCNA (top), SpPCNA (middle) and LmPCNA (bottom) immobilised surfaces. Peptide at 3 µM was passed over the surface twice in order to ascertain if there was any baseline drift or if the surface was losing activity during the course of the experiment. This shows the PCNA-peptide interaction is very rapid and gave no baseline drift. **B)** Steady state analysis of data plotting RU against peptide concentration was used to give a K_d. This is shown as a black line with the K_d value (in µM) inset.



Peptide	10 °C	15 °C	20 °C	25 °C	30 °C	35 °C	40 °C
1	0.263	0.450	0.656	0.843	1.14	1.70	2.39
2	1.08	1.62	2.33	3.16	5.10	7.25	10.9
3	1.42	1.80	2.80	3.71	6.61	9.32	12.6
4	1.13	1.43	2.63	3.72	5.74	9.86	14.7
5	1.34	1.76	2.65	3.46	6.91	9.84	14.8
6	54.1	60.7	106	128	188	191	381

Figure 7-3. **K_d of peptides to immobilised HsPCNA with shift in temperature.**

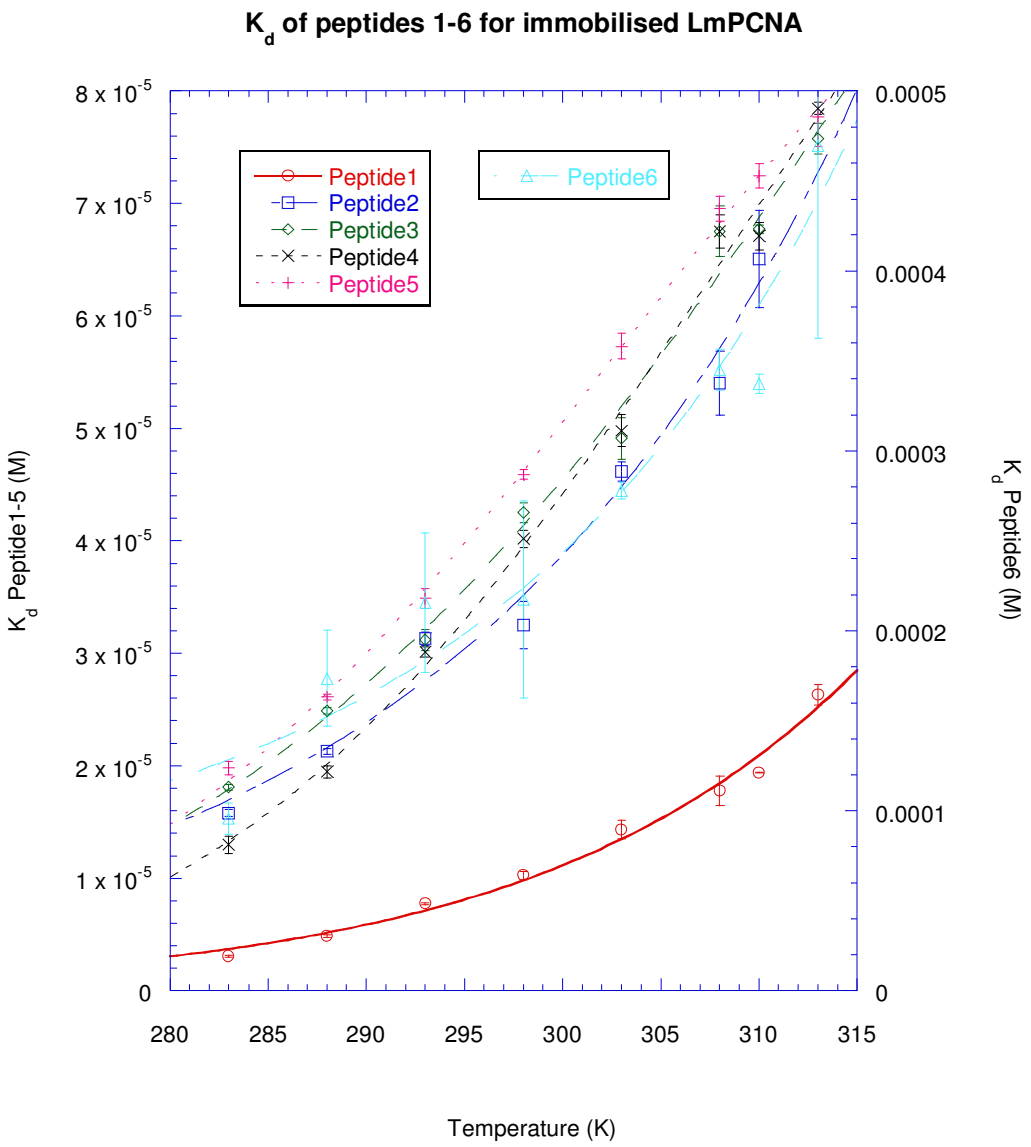
Figure showing the increase in K_d of peptides for immobilised HsPCNA with increase in temperature. Peptides 1-5 are shown on the left y-axis while Peptide 6 is shown on the right y-axis. The table (below) shows the K_d values (in μM) for each peptide at each temperature, each of these values are the average of three data sets.



Peptide	10 °C	15 °C	20 °C	25 °C	30 °C	35 °C	40 °C
1	0.667	0.821	1.02	1.20	1.81	2.00	3.92
2	1.27	2.27	2.56	3.64	7.83	10.8	17.4
3	1.42	2.08	3.28	5.07	8.25	14.7	19.8
4	1.29	1.85	2.91	4.55	7.64	17.0	19.9
5	1.68	2.14	2.76	4.90	8.93	16.8	21.7
6	21.1	32.6	65.6	95.9	161	266	276

Figure 7-4. **K_d of peptides to immobilised SpPCNA with shift in temperature.**

Figure showing the increase in K_d of peptides for immobilised SpPCNA with increase in temperature. Peptides 1-5 are shown on the left y-axis while Peptide 6 is shown on the right y-axis. The table (below) shows the K_d values (in μM) for each peptide at each temperature, each of these values are the average of three data sets.



Peptide	10 °C	15 °C	20 °C	25 °C	30 °C	35 °C	40 °C
1	3.06	4.84	7.75	10.3	14.3	17.8	26.3
2	15.8	21.3	31.3	32.5	46.2	54.0	56.4
3	18.1	24.9	31.2	42.5	49.1	67.5	75.8
4	13.0	19.5	30.1	40.2	49.8	67.5	78.5
5	19.8	26.1	34.9	45.9	57.3	69.6	77.7
6	95.6	174	216	217	278	345	470

Figure 7-5. **K_d of peptides to immobilised LmPCNA with shift in temperature.**

Figure showing the increase in K_d of peptides for immobilised LmPCNA with increase in temperature. Peptides 1-5 are shown on the left y-axis while Peptide 6 is shown on the right y-axis. The table (below) shows the K_d values (in μM) for each peptide at each temperature, each of these values are the average of three data sets.

with a temperature increase of 5K.

7.2.2 Van't Hoff analysis of PCNA surfaces with free peptide

Van't Hoff analysis of this data gives values of ΔH, ΔS and ΔC_p. The level of effect of ΔC_p can be small, so data fitting over a short temperature range is prone to error when performing non-linear van't Hoff analysis. Due to error of the non-linear van't Hoff plot, it is important to determine the shift in ΔH and ΔS by comparing both linear and non-linear fitting procedures. The linear (appendix 1.9) and non-linear van't Hoff plots gave similar values of enthalpy and entropy for the Hs, Sp and LmPCNA (figures 7.6, 7.7 and 7.8). The values derived from the non-linear van't Hoff plots for the three PCNAs with six peptides are shown in table 7.1.

7.2.2.1 Differences in the Gibbs free energy of interaction between peptides

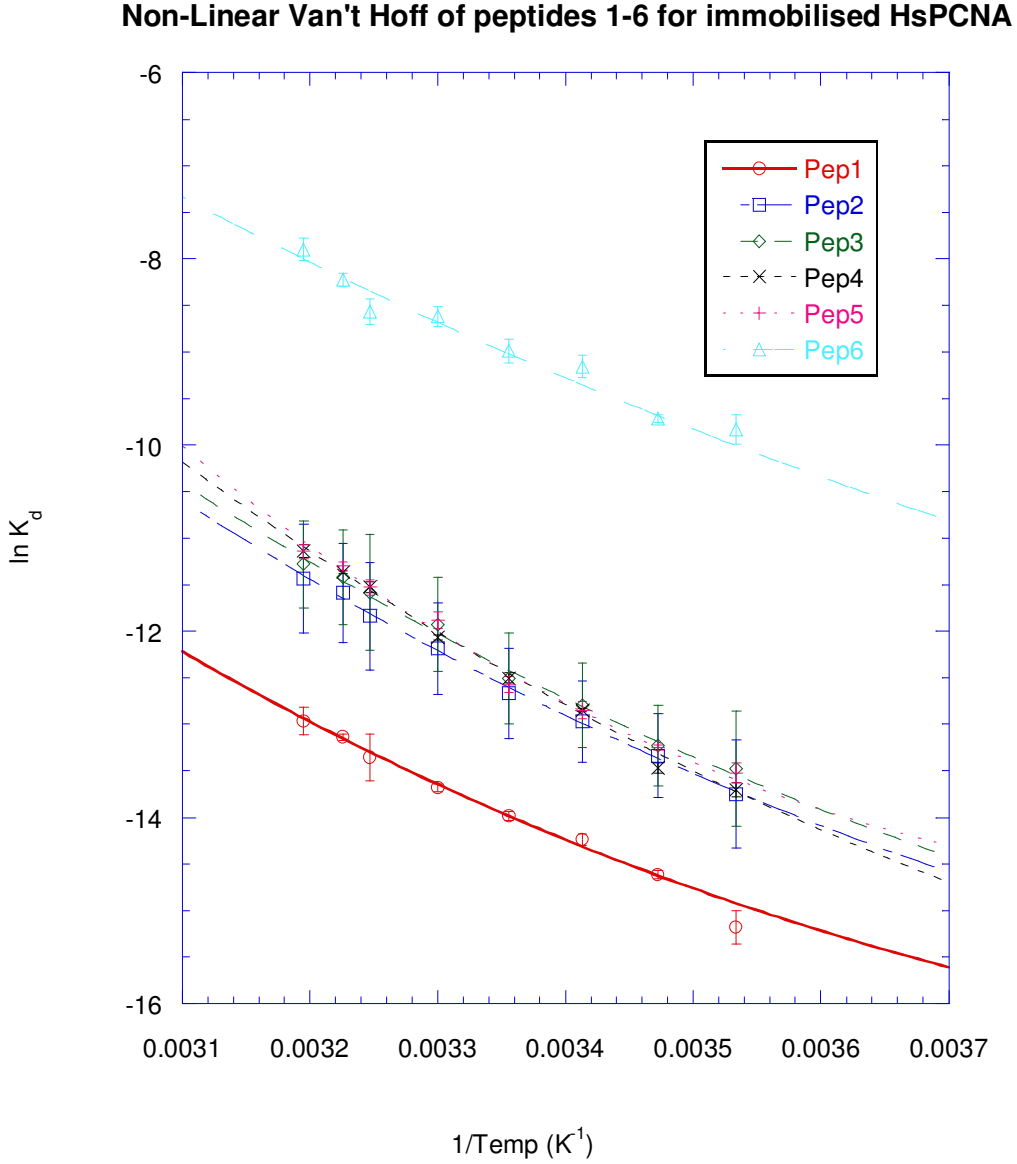
The Gibbs free energy of interaction for all peptides is negative, indicating that the interaction is favourable. The Gibbs free energy of interaction demonstrates that there is very little effect on binding as a result of linker length (-31.1 to -30.7 kJ mol^{-1}) (table 7.1). This small change in the Gibbs free energy of interaction with increasing linker size is a result of an enthalpy entropy compensation effect. The F150A point mutant gave a significantly increased Gibbs free energy (-31.6 to -22.3 kJ mol^{-1}) (table 7.1). The F150A point mutation resulted in an increase in the entropic cost of interaction and as a result, is a greatly weaker interaction. These trends in the Gibbs free energy of interaction are well conserved in both SpPCNA and LmPCNA peptide interactions (table 7.1). SpPCNA has very similar affinity to HsPCNA, however LmPCNA has a significantly weaker interaction e.g. HsPCNA and LmPCNA to peptide 1 is -31.6 and -28.4 kJ mol^{-1} respectively (table 7.1).

7.2.2.2 Changes in enthalpy and entropy with increasing linker length

All peptides analysed are enthalpically favourable and peptides with increasing linker length give a decrease in enthalpy of interaction for HsPCNA (-57.5 to -65.0 kJ mol^{-1}) (table 7.1). A decrease in enthalpy is consistent with an increase in the number of intermolecular bonds being formed between PCNA and the peptide. Peptide binding is entropically unfavourable in all cases, and the entropy of interaction decreases with increasing linker length (-88.0 to -115 J $Kmol^{-1}$) (table 7.1). This decreasing entropy is indicative of a longer more flexible peptide having greater disorder. These trends in enthalpy and entropy with increasing linker size are consistent with the values determined for binding to SpPCNA and LmPCNA (table 7.1).

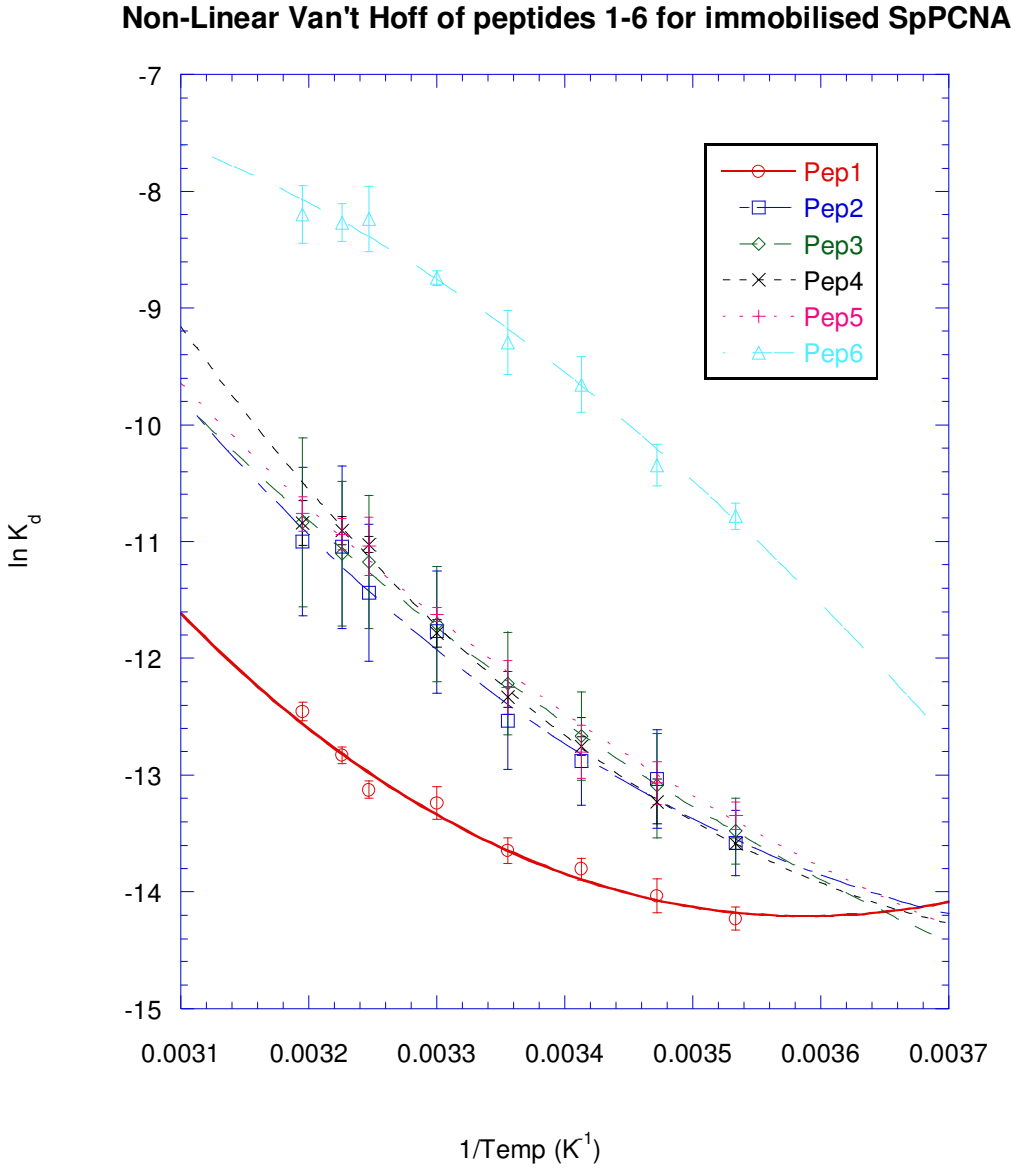
7.2.2.3 The enthalpy and entropy of p21 (141-152) and F150A mutant for PCNA

There is little change in the enthalpy of interaction between peptide 1 and peptide 6 (F150A) for HsPCNA (-49.2 to -49.6 kJ mol^{-1}) (table 7.1). The similar values of enthalpy between the two peptide interactions suggest that the phenylalanine ring provides little intermolecular contacts. There is a significant shift in the entropy of interaction between peptides 1 and 6 (-58.9 to -91.4 J $Kmol^{-1}$) (table 7.1). The change in entropy infers that the F150A mutant is either more disordered in solution or more ordered when bound to HsPCNA when compared to peptide 1. It is possible that the F150A mutation reduces the propensity of helix formation at the PIP box and result in greater entropic cost. SpPCNA has a decrease in enthalpy of interaction (-41.1 to -66.6 kJ mol^{-1}) and a significantly decreased entropy of interaction (-24.6 to -147 J $Kmol^{-1}$) (table 7.1). LmPCNA showed increase in both



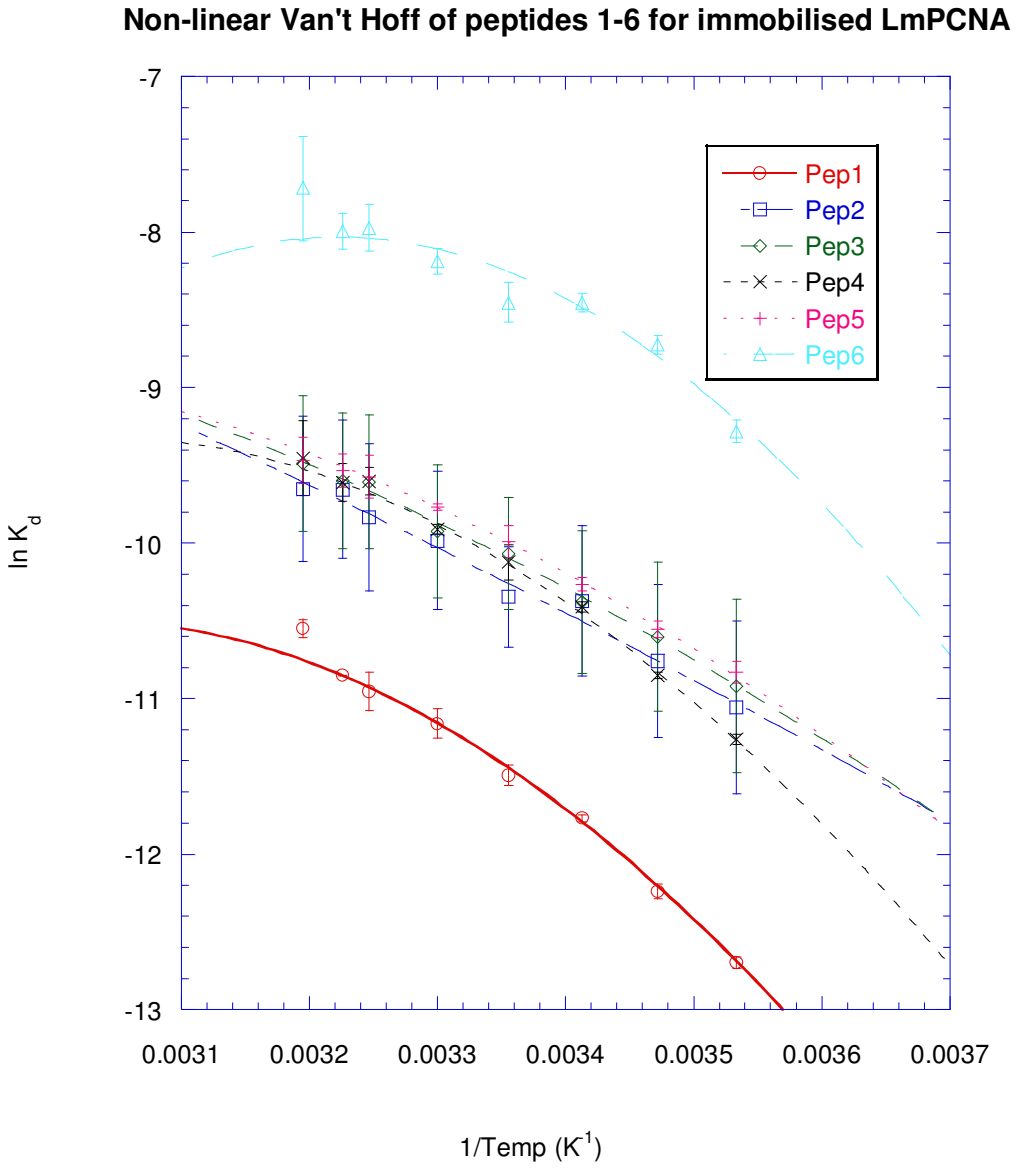
Peptide	Spacer Length (Å)	ΔH^{θ} (kJ mol^{-1})	ΔS^{θ} (J Kmol^{-1})	$\Delta C_p{}^{\theta}$ (J K^{-1} mol^{-1})	ΔG^{θ} (kJ mol^{-1})
1	-	-49.2	-58.9	-698	-31.6
2	10	-57.5	-88.0	-668	-31.3
3	20	-57.7	-90.2	-648	-30.8
4	30	-65.0	-115	-752	-30.7
5	40	-63.7	-110	-1310	-30.9
6	-	-49.6	-91.4	-436	-22.3

Figure 7-6. **Non-Linear van't Hoff of HsPCNA for peptides 1-6**
The linear Van't Hoff plot (top) with lnK_d on the y axis versus 1/T on the x axis plotted with Kaleidagraph. Table (bottom) shows thermodynamic properties and change in specific heat capacity of solution determined for each peptide to three significant figures.



Peptide	Spacer Length (Å)	ΔH^{θ} (kJ mol^{-1})	ΔS^{θ} (J Kmol^{-1})	ΔC_p^{θ} (J K^{-1} mol^{-1})	ΔG^{θ} (kJ mol^{-1})
1	-	-41.1	-24.6	-2110	-33.8
2	10	-66.5	-120	-1600	-30.7
3	20	-67.4	-125	-944	-30.2
4	30	-77.4	-158	-2010	-30.3
5	40	-67.8	-127	-1070	-30.0
6	-	-66.6	-147	1290	-22.8

Figure 7-7. **Non-Linear van't Hoff of SpPCNA for peptides 1-6**
The linear Van't Hoff plot (top) with lnK_d on the y axis versus 1/T on the x axis plotted with Kaleidagraph. Table (bottom) shows thermodynamic properties and change in specific heat capacity of solution determined for each peptide to three significant figures.



Peptide	Spacer Length (Å)	ΔH^{θ} (kJ mol^{-1})	ΔS^{θ} (J Kmol^{-1})	ΔC_p^{θ} (J K^{-1} mol^{-1})	ΔG^{θ} (kJ mol^{-1})
1	-	-46.6	-61.2	1480	-28.4
2	10	-34.9	-31.7	138	-25.5
3	20	-35.1	-33.7	419	-25.1
4	30	-42.0	-56.8	1390	-25.1
5	40	-35.2	-34.9	658	-24.8
6	-	-27.2	-22.7	2240	-20.4

Figure 7-8. **Non-Linear van't Hoff of LmPCNA for peptides 1-6**
The linear Van't Hoff plot (top) with lnK_d on the y axis versus 1/T on the x axis plotted with Kaleidagraph. Table (bottom) shows thermodynamic properties and change in specific heat capacity of solution determined for each peptide to three significant figures.

Peptide		1	2	3	4	5	6
HsPCNA	ΔH (kJ mol^{-1})	-49.2 (±1.72)	-57.5 (±13.8)	-57.7 (±13.3)	-65.0 (±1.12)	-63.7 (±1.45)	-49.6 (±2.20)
	ΔS (J $Kmol^{-1}$)	-58.9 (±5.77)	-88.0 (±46.3)	-90.2 (±44.5)	-115 (±3.78)	-110 (±4.94)	-91.4 (±7.49)
	ΔC_p (J K^{-1} mol^{-1})	-698 (±477)	-668 (±3360)	-648 (±3280)	-752 (±309)	-1310 (±558)	-436 (±808)
	ΔG (kJ mol^{-1})	-31.6	-31.3	-30.8	-30.7	-30.9	-22.3
SpPCNA	ΔH (kJ mol^{-1})	-41.1 (±2.38)	-66.5 (±13.0)	-67.4 (±13.3)	-77.4 (±2.62)	-67.8 (±3.87)	-66.6 (±4.75)
	ΔS (J $Kmol^{-1}$)	-24.6 (±7.90)	-120 (±43.5)	-125 (±44.5)	-158 (±8.64)	-127 (±12.7)	-147 (±15.6)
	ΔC_p (J K^{-1} mol^{-1})	-2110 (±614)	-1600 (±2680)	-944 (±2780)	-2010 (±577)	-1070 (±807)	1290 (±1000)
	ΔG (kJ mol^{-1})	-33.8	-30.7	-30.2	-30.3	-30.0	-22.8
LmPCNA	ΔH (kJ mol^{-1})	-46.6 (±0.66)	-34.9 (±12.7)	-35.1 (±12.4)	-42.0 (±1.94)	-35.2 (±2.13)	-27.2 (±3.09)
	ΔS (J $Kmol^{-1}$)	-61.2 (±2.18)	-31.7 (±42.6)	-33.7 (±41.3)	-56.8 (±6.48)	-34.9 (±7.03)	-22.7 (±10.4)
	ΔC_p (J K^{-1} mol^{-1})	1480 (±245)	138 (±2850)	419 (±2870)	1390 (±380)	658 (±501)	2240 (±645)
	ΔG (kJ mol^{-1})	-28.4	-25.5	-25.1	-25.1	-24.8	-20.4

Table 7-1. **Thermodynamic values derived from the non-Linear van't Hoff plots of immobilised PCNA interactions with free peptide**

Table showing the thermodynamic values of the PCNA peptide interactions determined by the non-linear van't Hoff analysis. The enthalpy, entropy, specific heat capacity and Gibbs free energy of interactions are grouped by PCNA species. In brackets the ± fitting error in non-linear van't Hoff plot for values presented above.

enthalpy and entropy of these interactions (table 7.1).

7.2.2.4 Differences in the ΔC_p of interaction between peptides

The change in the specific heat capacity for every peptide for HsPCNA is negative, inferring a release of water (burial of surface hydrophobics). The specific heat capacity of interaction with HsPCNA has a slight decrease with linker length (-668 to -752 J K^{-1} mol^{-1}). There is a significant decrease upon the addition of a hydrophobic spacer region (-752 to -1310 J K^{-1} mol^{-1}) (table 7.1). These values for specific heat capacity are consistent with an increase in the amount of hydrophobic surface being removed from the solution. The loss of the phenylalanine ring in the peptide resulted in an increase in the specific heat capacity of interaction (-698 to -436 J K^{-1} mol^{-1}) (table 7.1). The increase in the specific heat capacity is consistent with the loss of the hydrophobic phenylalanine ring. The values of specific heat capacity vary greatly for SpPCNA and LmPCNA (table 7.1) and are unreliable due to weaker interactions presenting greater variability in the K_d values determined.

7.2.3 Determination of Affinity of Peptide Surfaces with Free PCNA

SA surfaces with different peptides immobilised were used to examine the affinity with Hs, Sp and LmPCNA at different temperatures. Sensorgrams of the interaction between immobilised peptides and HsPCNA at 25°C are shown in figure 7.9.

Peptide2 – Bio-KRRQTSMTDFYH
Peptide3 – Bio-PEG-KRRQTSMTDFYH
Peptide4 – Bio-PEG-PEG-KRRQTSMTDFYH
Peptide5 – Bio-PEG-PEG-Ahx-KRRQTSMTDFYH

PCNA peptide interactions showed no apparent mass transport problems (Karlsson, R. 1999). Interactions of immobilised peptides with HsPCNA at 25°C

gave k_t of 1.64×10^{14} $RU.M^{-1}s^{-1}$, 7.76×10^{15} $RU.M^{-1}s^{-1}$, and 1.37×10^{17} $RU.M^{-1}s^{-1}$ for peptide 3, 4 and 5 respectively. Values of k_t for immobilised peptide 2 with HsPCNA at 25°C could not be determined.

The interactions presented in figure 7.9 have very fast on and off rates (k_{on} and k_{off}). At 25°C, the k_{on} for HsPCNA with immobilised peptides 3, 4 and 5 are ~3.96×10^{3} $M^{-1}s^{-1}$, 3.48×10^{4} $M^{-1}s^{-1}$, 4.91×10^{4} $M^{-1}s^{-1}$ respectively. The k_{off} of HsPCNA interaction for peptides 3, 4 and 5 are ~1.2 s^{-1}, 0.6 s^{-1} and 0.9 s^{-1} respectively. Due to the rapid nature of the PCNA-peptide interactions, the data were analysed by the steady state of interaction (Adamczyk, M. *et al* 2000).

The steady-state analysis of the sensorgrams (figure 7.9) was used to determine the affinity of interaction. Affinity between immobilised peptides 2-5 for HsPCNA and SpPCNA (at different temperatures) are shown in figure 7.10. This data demonstrated that SpPCNA had the strongest affinity for the peptide surfaces, HsPCNA and subsequently LmPCNA had weaker affinities. The interaction of SpPCNA was ~10 fold tighter than that of HsPCNA which was ~14 fold tighter than that of LmPCNA; this is in contradiction with previously determined values (section 8.2 and chapter 6). The different peptide surfaces also presented a differing range of affinities. Peptide 2 gave no data as previously stated, peptide 3 surface gave a weak interaction, surfaces of peptide 4 and 5 gave interactions ~30 fold tighter than those of peptide 3. The K_d of these interactions demonstrated an average shift of 1.5x for an increase of 5K. The thermodynamic values determined from these experiments will be discussed later.

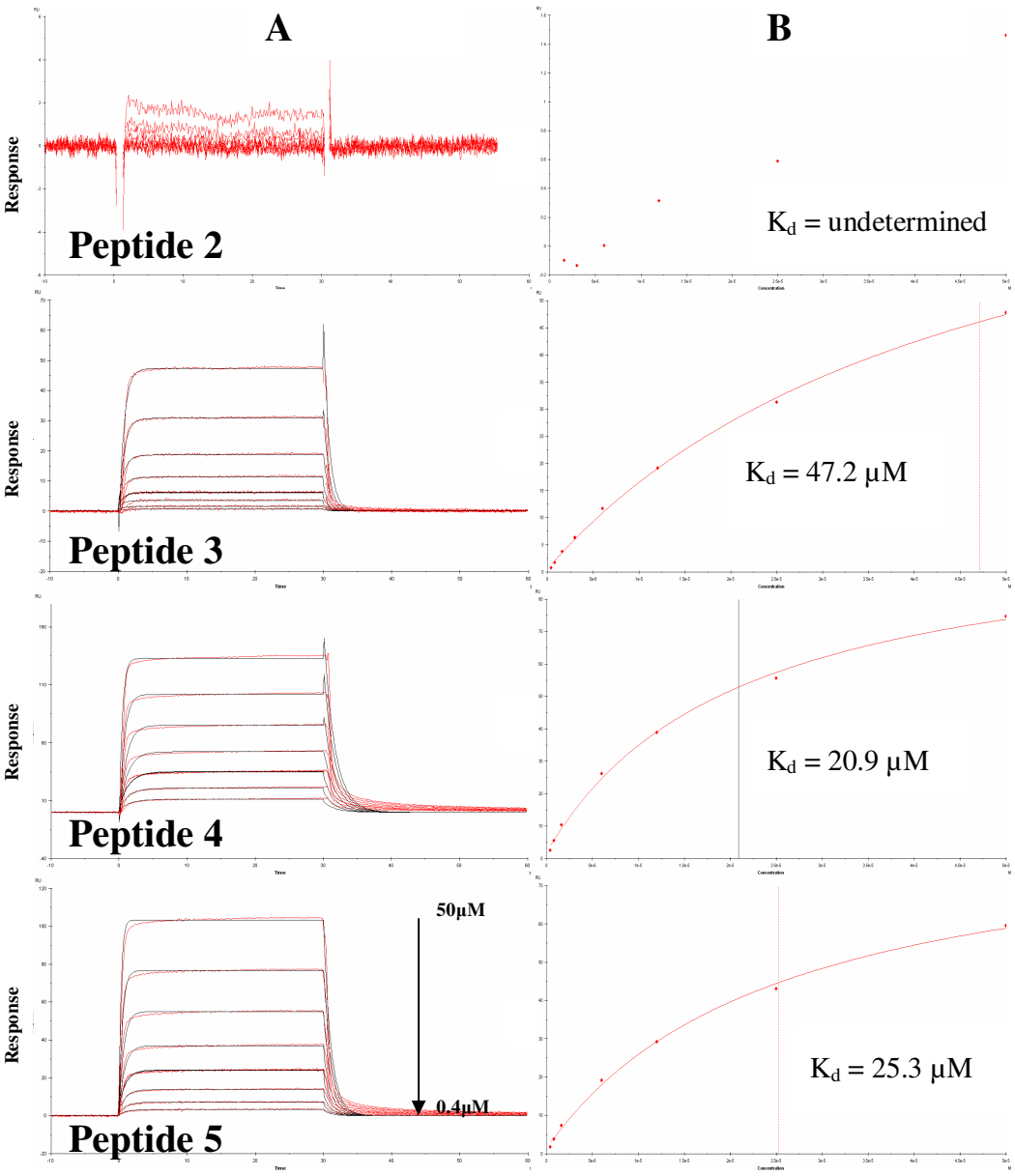
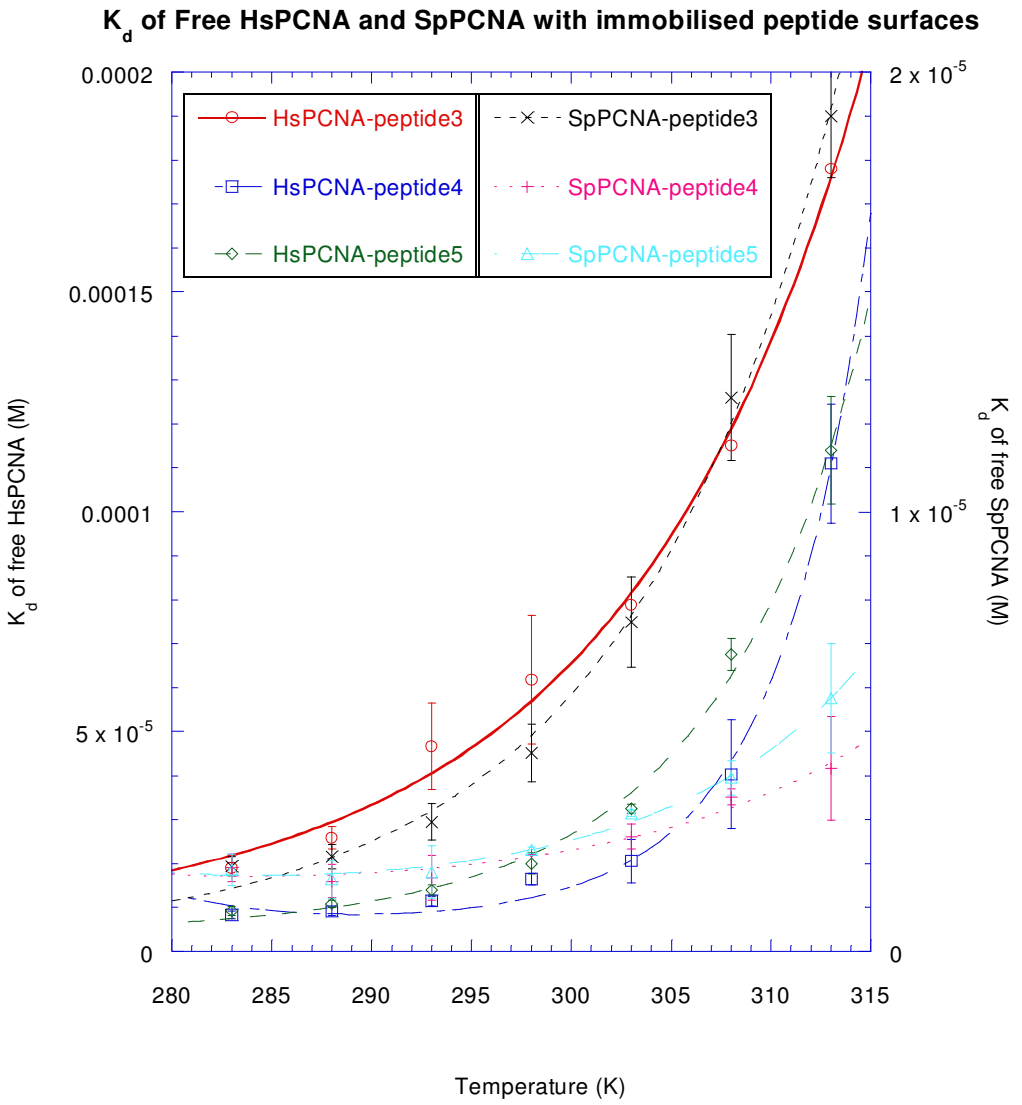


Figure 7-9. **BIAcore data for immobilised peptides for HsPCNA**
The interactions of immobilised peptide for HsPCNA at 298K. **A)** Reference corrected results of BIAcore interaction plotted as RU against time in seconds. Serial dilutions of HsPCNA (50 to 0.4 µM) were passed over peptide. This shows the PCNA peptide interaction is very rapid and repeats of 3 µM of PCNA was passed over the surface twice showing no baseline drift. **B)** Steady state analysis of data plotting RU against peptide concentration was used to give a K_d. This is shown as a black line with the K_d value (in µM) inset. Data shows no obtainable results from surfaces of peptide 2.



Interaction	10 °C	15 °C	20 °C	25 °C	30 °C	35 °C	40 °C
HsPCNA-P3	18.9	25.9	46.7	61.8	78.8	115	178
HsPCNA-P4	8.37	9.18	11.5	16.5	20.6	40.3	111
HsPCNA-P5	9.27	10.9	14.0	20.0	32.5	67.6	114
SpPCNA-P3	1.94	2.16	2.95	4.52	7.49	12.6	19.0
SpPCNA-P4	1.91	1.60	1.67	2.19	2.62	3.52	4.17
SpPCNA-P5	1.86	1.65	1.81	2.31	3.15	3.95	5.76

Figure 7-10. **K_d of PCNA to immobilised peptides across a temperature range**
Figure showing the increase in K_d of immobilised peptides for HsPCNA and SpPCNA with increase in temperature. HsPCNA is shown on the left y-axis while SpPCNA is shown on the right y-axis. The table (below) shows the K_d values (in μM) for each peptide at each temperature, each of these values are the average of three data sets.

7.2.4 Differences in the enthalpy of interaction between HsPCNA and peptides

The non-linear Van't Hoff analysis of this data gives values of ΔH, ΔS and ΔC_p (Yoo, Seung Hyun and Lewis, Marc S. 2002). The linear van't Hoff plot (appendix 1.9) gave similar values of enthalpy and entropy to the non-linear van't Hoff plots (figures 7.6, 7.7 and 7.8). The values derived from the non-linear van't Hoff plots for the three PCNAs with four immobilised peptides are shown in figure 7.11 (table).

7.2.4.1 Differences in the Gibbs free energy of interaction between peptides

The Gibbs free energy of interaction for all immobilised peptides with HsPCNA is negative, indicating that the interaction is favourable. The Gibbs free energy of interaction demonstrates that there is an increase in affinity between HsPCNA and peptide as a result of increasing linker length (-24.1 to -27.6 kJ mol^{-1}) (figure 7.11, table). The corresponding range of binding energies measured with immobilised HsPCNA was between -30.7 and 30.9 kJ mol^{-1} (table 7.1). The small (0.3 kJ mol^{-1}) decrease in Gibbs free energy suggests that increasing linker length does not affect binding affinity. The large (3.5 kJ mol^{-1}) change in free energy of immobilised peptides is most likely the result of the immobilisation method and the reduction in the effect of steric hindrance imposed by Streptavidin. Peptide 5 with an additional hydrophobic linker had similar (however reduced) Gibbs free energy of interaction as peptide 4 (-26.8 and -27.6 kJ mol^{-1} respectively). This is consistent with data obtained with immobilised PCNA.

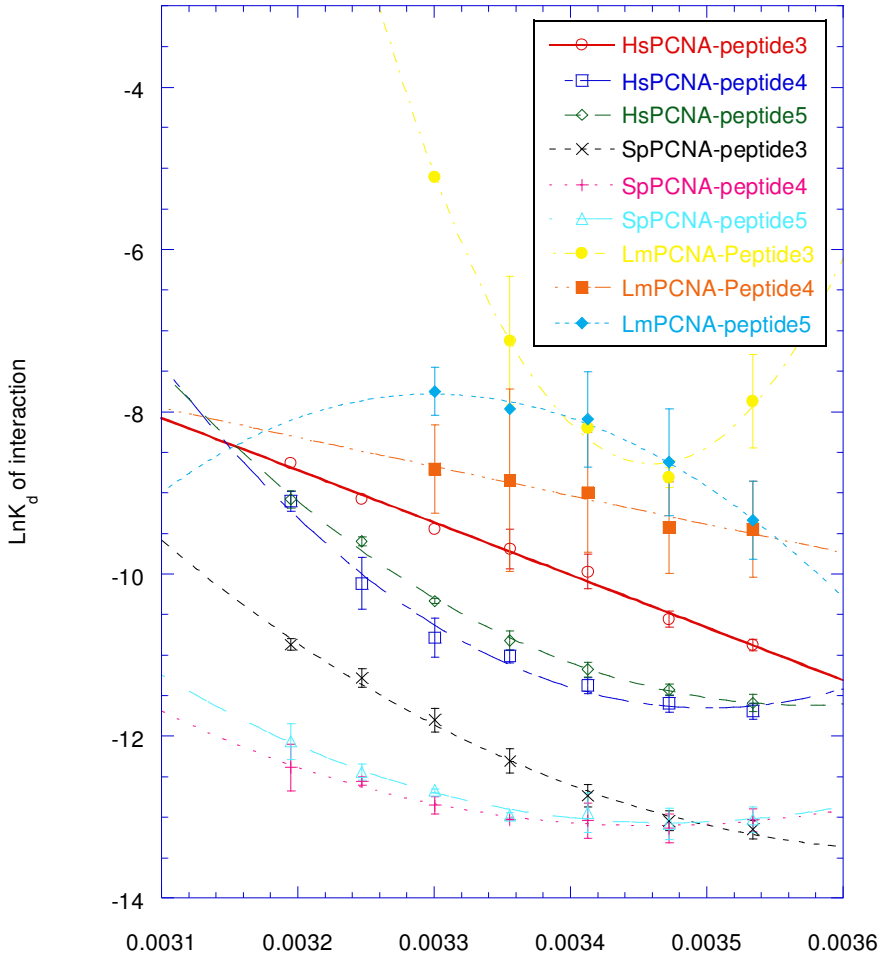
The same trend in increasing affinity of interaction with increasing linker length is observed in both SpPCNA and LmPCNA. SpPCNA has much greater

affinity (smaller Gibbs free energy) to HsPCNA e.g. -32.0 and -26.8 kJ mol^{-1} respectively for peptide 5 (figure 7.11, table). The increased affinity of SpPCNA is not consistent with data shown in section 7.2.1 and could be a result of the method of ligand immobilisation. LmPCNA has a significantly weaker interaction e.g. HsPCNA and LmPCNA to peptide 5 is -26.8 and -19.5 kJ mol^{-1} respectively (figure 7.11, table). The reduced affinity of LmPCNA is consistent with data shown above.

7.2.4.2 Changes in enthalpy entropy and ΔC_p with increasing linker length

All peptides analysed are enthalpically favourable and peptides with increasing linker length give a decrease in enthalpy of interaction for HsPCNA (-53.8 to -65.2 kJ mol^{-1}) (figure 7.11, table). A decrease in enthalpy is consistent with figures derived for immobilised HsPCNA. Peptide binding is entropically unfavourable in all cases and the entropy of interaction decreases with increasing linker length (-99.6 to -129 J $Kmol^{-1}$) (figure 7.11, table). This decreasing entropy is indicative of a longer more flexible peptide having greater disorder and is consistent with thermodynamic values returned for immobilised HsPCNA. The specific heat capacity of interaction with HsPCNA is variable and the figures derived are unreliable (figure 7.11, table). These trends in enthalpy and entropy with increasing linker size are mostly consistent with interactions for SpPCNA and LmPCNA (figure 7.11, table).

Non-linear Van't Hoff analysis of immobilised peptides with free PCNA in solution



PCNA Peptide interactions	Spacer Length (Å)	ΔH^{θ} (kJ mol^{-1})	ΔS^{θ} (J Kmol^{-1})	ΔC_p^{θ} (J K^{-1} mol^{-1})	ΔG^{θ} (kJ mol^{-1})
HsPCNA-P3	20	-53.8	-99.6	17.2	-24.1
HsPCNA-P4	30	-62.2	-116	-4960	-27.6
HsPCNA-P5	40	-65.2	-129	-3570	-26.8
SpPCNA-P3	20	-60.8	-102	-2350	-30.4
SpPCNA-P4	30	-18.4	46.3	-2020	-32.2
SpPCNA-P5	40	-24.8	24.2	-2460	-32.0
LmPCNA-P3	20	-246	-766	-2670	-17.3
LmPCNA-P4	30	-29.7	-25.8	-31.7	-22.0
LmPCNA-P5	40	-26.3	-22.9	5320	-19.5

Figure 7-11. **Non-Linear van't Hoff of LmPCNA for peptides 1-6**
The linear Van't Hoff plot (top) with $\ln K_d$ on the y axis versus 1/T on the x axis plotted with Kaleidagraph. Table (bottom) shows thermodynamic properties, at standard conditions, and change in specific heat capacity of solution determined for each peptide to three significant figures.

7.3 Discussion of the SPR data

7.3.1 Effect of F150A point mutant on thermodynamic values of binding

The Gibbs free energy of interaction between HsPCNA and peptide 1 (-31.6 kJ mol^{-1}) and peptide 6 (-22.3 kJ mol^{-1}) shows a significantly weaker interaction for peptide 6. The reduction in affinity for HsPCNA can be accounted for by the increased entropic cost of interactions with little change in the enthalpy of interaction (discussed below). The reduction in affinity (shown in the values of Gibbs free energy) between peptide 1 and peptide 6 is also observed in SpPCNA (11.0 kJ mol^{-1}) and LmPCNA (8.0 kJ mol^{-1}). This shows remarkably good agreement with Gibbs free energy values for SpPCNA and LmPCNA.

The PIP box sequence forms a helical region at Phe150 as stated above. The Phe150Ala mutation could have been hypothesised to present a lower entropic cost of interaction due to increased propensity for helix formation of alanine and reduced side-chain entropy as shown in (Chellgren, B. W. and Creamer, T. P. 2006; Oneil, K. T. and Degrado, W. F. 1990; Padmanabhan, S. *et al* 1990). Results presented in this work show increased entropic cost upon the Phe150Ala mutation (-32.5 J $Kmol^{-1}$) (table 7.1). With the change in the specific heat capacity of interaction (+162 J K^{-1} mol^{-1}), the increased entropic cost has been demonstrated to be the result of the removal of the phenylalanine solvation shell. A significantly reduced solvation shell prior to binding would result in a greater entropic cost of interaction with no net change in the enthalpy of interaction.

From crystal structures, it is known that the peptide forms a small helical region in the PIP box section of the peptide. Previous literature shows that alanine is a greater promoter of helix formation than phenylalanine (Blaber, M. *et al* 1994;

Blaber, M. *et al* 1993; Chellgren, B. W. and Creamer, T. P. 2006; Creamer, T. P. and Rose, G. D. 1994; Lewis, P. N. *et al* 1970; Oneil, K. T. and Degrado, W. F. 1990; Padmanabhan, S. *et al* 1990; Richardson, J. M. *et al* 2005). The significant drop in affinity of the mutant demonstrates that helix formation is not required as a precursor for peptide binding.

The F150A mutation in peptide 6 resulted in no significant shift in the enthalpy of interaction for HsPCNA compared with peptide 1 (-0.4 kJ mol^{-1}). This result demonstrates that the association of the phenylalanine ring provides little enthalpic contribution to the global interaction. The lack of enthalpic contribution of the phenylalanine side chain may be surprising as crystal structure studies have shown that this mutation has resulted in the loss of van der Waals contacts. The complex of HsPCNA with p21 R143-H152 and HsPCNA with p21 R143-H152 (F150A) have buried surface areas of 1200 and 1160$Å^2$ respectively (figure 7.12). There is no obvious way to explain why the reduction in buried surface area of 40$Å^2$ as a result of the F150A mutation would give such a little effect on binding enthalpy. SpPCNA and LmPCNA had a reduction (+19.4 kJ mol^{-1}) and increased (-25.5 kJ mol^{-1}) changes respectively in the enthalpy of interaction between peptide 1 and peptide 6. The variations could be the result of weaker interactions between immobilised PCNA surfaces and peptide 6 at higher temperatures. SpPCNA and LmPCNA binding to the mutant peptide have a decrease in the buried surface areas of 47.4 and 76.4$Å^2$ (figure 7.12). The increased loss in buried surface area of the F150A mutant may rationalise the increased enthalpy of this interaction.

It has been hypothesised that one of the greatest promoters to helix formation is the burial of hydrophobic residues (Blaber, M. *et al* 1994; Blaber, M. *et al* 1993).

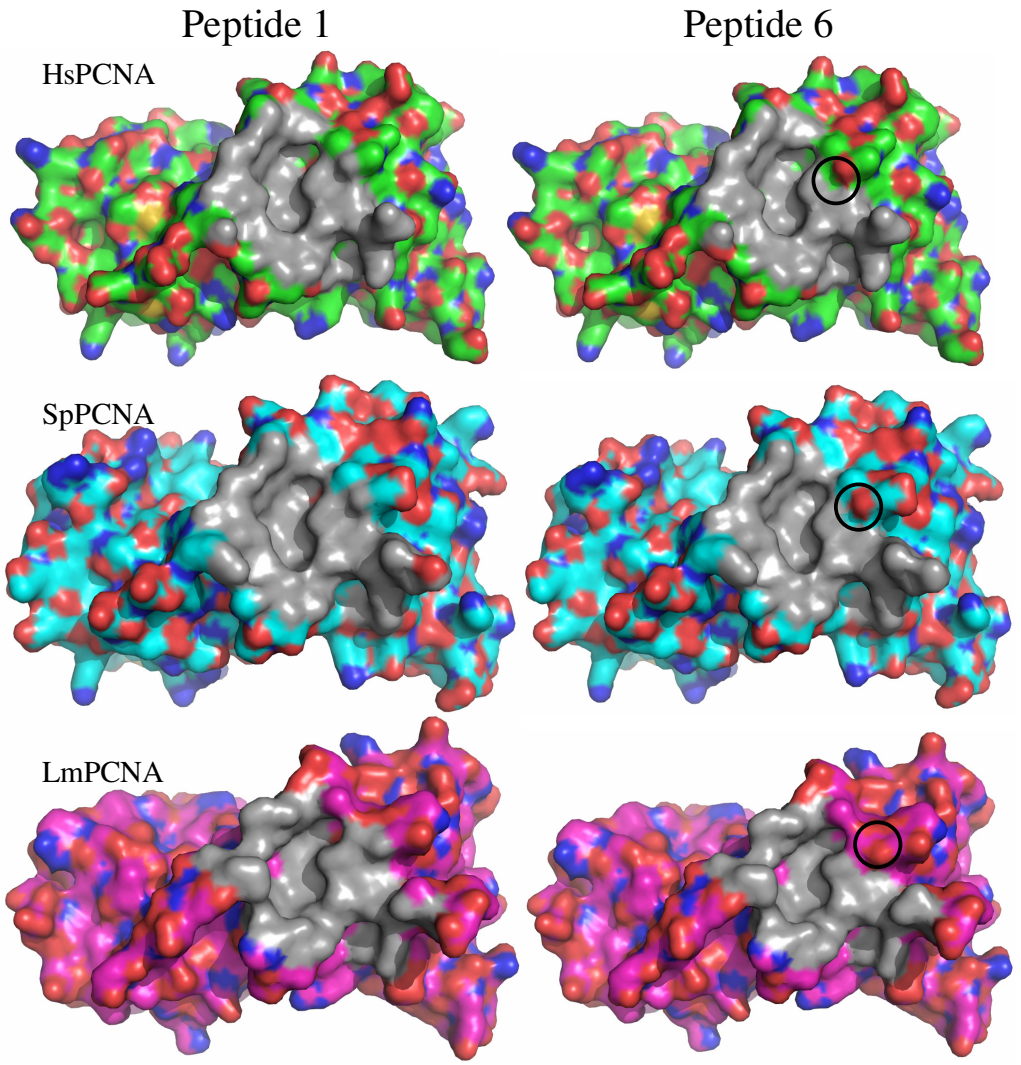


Figure 7-12. **The buried surface areas of Hs, Sp and LmPCNA with peptide 1 and peptide 6**

Figure showing the buried surfaces areas in grey of the HsPCNA (top), SpPCNA (middle) and LmPCNA (bottom). The buried surfaces areas are represented as those atoms within 5Å of peptide 1 (left) and the F150A point mutation peptide 6 (right). This shows a slightly greater buried surface area for SpPCNA. The figure demonstrates the loss of buried surface area due to the point mutation with the position of the phenylalanine ring (circled in black, right).

The importance of the burial of hydrophobics was determined by experimental observation of peptide-protein binding. Computer simulation studies determined side chain entropy loss was the major determinant of helix formation. The simulations were performed in order to determine the increase in entropic cost for each amino acid (Creamer, T. P. and Rose, G. D. 1994). This paper determined that there was a small contribution from ring structures stacking in aiding helix formation, and little contribution was made from the burial of hydrophobics in contradiction to (Blaber, M. *et al* 1994; Blaber, M. *et al* 1993). This apparent contradiction can be resolved using data presented here. The effects of waters upon the system were not incorporated in these publications, and values of the specific heat capacity of interaction aid to explain these apparent contradictions.

7.3.2 Effect on thermodynamic values of an N-terminal biotin tag on binding peptide

The Gibbs free energy of interaction between immobilised HsPCNA and free peptide 1 and peptide 2 was similar, -31.6 and -31.3 kJ mol^{-1} respectively (table 7.1). Little change in the Gibbs free energy of interaction demonstrates no significant change in the affinity of interaction with tagged peptides, due to an enthalpy-entropy compensation effect, discussed below. The enthalpy of interaction showed a decrease for the tagged peptide (from -49.2 to -57.5 J $Kmol^{-1}$) (table 7.1). The decrease in enthalpy of interaction is likely to be a result of the increased number of interactions between HsPCNA and biotin. The entropy of interaction also decreases due to the presence of the biotin tag (-58.9 to -88.0 kJ mol^{-1}) (table 7.1). This is the result of the tag having greater flexibility in solution (figure 7.9) so the immobilisation to HsPCNA will have a greater entropic cost. The specific heat capacity is very similar

for both untagged and tagged peptides, -698 and -668 J K^{-1} mol^{-1} respectively. Little change in specific heat capacity is due to the biotin tag not being either particularly hydrophobic or hydrophilic. Trends in enthalpy and Gibbs free energy are conserved with all PCNAs. The entropy of interaction between peptide 1 and 2 for LmPCNA appears to increase which is in contradiction to data for HsPCNA and SpPCNA. The specific heat capacities for SpPCNA and LmPCNA are unreliable values.

7.3.3 Effect on thermodynamic values of increasing N-terminal biotin linker length

The Gibbs free energy of interaction between HsPCNA and tagged peptides with different linker lengths remains very similar, between -31.3 and -30.7 kJ mol^{-1} (table 7.1). The similar energies of these interactions show that HsPCNA has similar affinities for peptides with any linker length and seems to be a result of enthalpy-entropy compensation discussed below. Interactions for immobilised HsPCNA show a decrease in enthalpy with increasing linker length (from -57.5 to -65.0 kJ mol^{-1}) (table 7.1). The decrease in enthalpy may be due to an increase in the potential contact area of the peptides with longer linkers. The entropy of interaction between tagged peptides and HsPCNA however decreases with increasing linker length (-88.0 to -115 J $Kmol^{-1}$) (table 7.1). The decreasing entropy of interaction with increasing linker length is likely the result of the longer peptide linkers having greater freedom and flexibility. The specific heat capacity of interaction is similar with a slight increase for three tagged peptides (2, 3 and 4) -668, -648 and -752 J K^{-1} mol^{-1} respectively. Peptides 2 and 3 were however linked with large fitting errors (table 7.1, bottom). The small change in specific heat capacity with these linkers suggests that these peptide linkers are not particularly hydrophobic or hydrophilic. The trends

in Gibbs free energy, enthalpy and entropy of interaction with peptides 2, 3 and 4 are well conserved in all three PCNAs (table 7.1).

7.3.4 Effect on thermodynamic values due to a hydrophobic spacer

There is little change in the Gibbs free energy of interaction between peptide 4 and peptide 5. Every interaction (except immobilised HsPCNA) shows a slight increase in the Gibbs free energy of interaction from peptide 4 to peptide 5, from -30.3 to -30.0 kJ mol^{-1} for SpPCNA (+0.3 kJ mol^{-1}) and -25.1 to -24.8 kJ mol^{-1} for LmPCNA (+0.3 kJ mol^{-1}). HsPCNA showed a decrease of 0.2 kJ mol^{-1} in the Gibbs free energy of interaction with the addition of the hydrophobic spacer (table 7.1). The enthalpy of interaction between HsPCNA and peptides 4 and 5 is -65.0 and -63.7 kJ mol^{-1} respectively. The 1.3 kJ mol^{-1} increase in the enthalpy of interaction suggests that HsPCNA can create more non-bonded interactions with peptide directly linked to PEG than those directly linked to the hydrophobic octanoic acid linker. The entropy of interaction between HsPCNA and peptide 4 and 5 is -115 to -110 J $Kmol^{-1}$ respectively. The 5 J $Kmol^{-1}$ increase in the entropy of interaction may suggest that peptide 5 is less flexible than peptide 4 even with increased length or that there is a larger solvation shell which can be released after binding (discussed below). Peptide 5, with the addition of the hydrophobic spacer region, was determined to have a significantly reduced specific heat capacity of interaction (from -752 to -1310 J K^{-1} mol^{-1}). The large decrease (-558 J K^{-1} mol^{-1}) in the specific heat capacity between peptides 4 and 5 is likely the result of the hydrophobic portion of the peptide linker becoming buried. The removal of hydrophobic groups from a system results in the release of waters from a solvation shell, which can give these values of specific heat capacity and potentially explain those for the entropy of interaction. The trends

between the binding of peptides 4 and 5 showing increased enthalpy and entropy of interaction are conserved in SpPCNA and LmPCNA (table 7.1). The values for specific heat capacity are however unreliable.

7.3.5 Difference between Gibbs free energy of free and immobilised peptide

The interaction of immobilised peptides with HsPCNA shows an increased (less favourable) Gibbs free energy of interaction when compared with that of immobilised HsPCNA (table 7.1, figure 7.11). With peptide 2 immobilised to a Streptavidin chip surface, there was only a weak signal and could not be interpreted (figure 7.9). Free peptide 3 when compared to immobilised peptide 3 gave an increased (less favourable) Gibbs free energy of interaction from -30.8 to -24.1 kJ mol^{-1} respectively. Free peptide 4 when compared to immobilised peptide 4 gave an increased Gibbs free energy of interaction from -30.7 to -27.6 kJ mol^{-1} respectively. Free peptide 5 when compared to immobilised peptide 5 gave an increased Gibbs free energy of interaction from -30.9 to -26.8 kJ mol^{-1} respectively. The increase in the Gibbs free energy of interaction may be the result of steric hindrance imposed by the Streptavidin (figure 7.13). The data suggests that the increase in spacer length has little effect on affinity for free peptides (using immobilised PCNA). However, the small decrease in the Gibbs free energy of interaction with increasing linker length of immobilised peptides (-24.1 to -27.6 kJ mol^{-1}) (figure 7.14) may be associated with the method of immobilisation and may also be the result of steric interaction with Streptavidin on the chip.

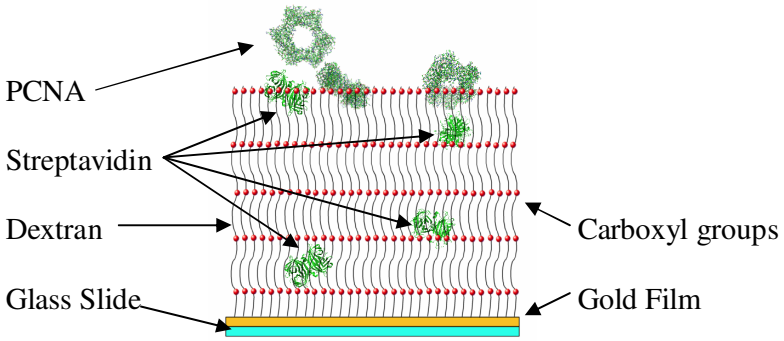


Figure 7-13. **The binding of PCNA to a modified Streptavidin surface**

PCNA bound on the Streptavidin surface showing a slight burial in the dextran matrix which could result in greater entropic cost of the interaction. PCNA is associated to streptavidin and is coupled to the dextran matrix (black strands with red balls). The yellow strip represets the gold film and the blue is the glass slide.

The trends in the Gibbs free energy of immobilised peptides shown for HsPCNA are consistent in both SpPCNA and LmPCNA (table 7.1). Gibbs free energy of interaction between immobilised peptide and SpPCNA showed a decrease with respect to peptide and HsPCNA. For example, immobilised peptide 4 with HsPCNA gave -27.6 kJ mol^{-1} and with SpPCNA gave -32.2 kJ mol^{-1}. This 4.6 kJ mol^{-1} decrease in Gibbs free energy is in contradiction with values determined for immobilised PCNA (which had no significant change between these PCNAs). The greater affinity between SpPCNA and immobilised peptide could be due to some affinity between SpPCNA and the modified Streptavidin.

7.3.6 Difference in enthalpy and entropy between free and immobilised PCNA

The interaction between immobilised HsPCNA and immobilised peptide commonly show similar enthalpy of interactions. The enthalpy of interaction with immobilised HsPCNA for peptides 3-5 is within the range of -57.7 to -65.0 kJ mol^{-1}. The enthalpy of interaction with immobilised peptides 3-5 for HsPCNA is within the

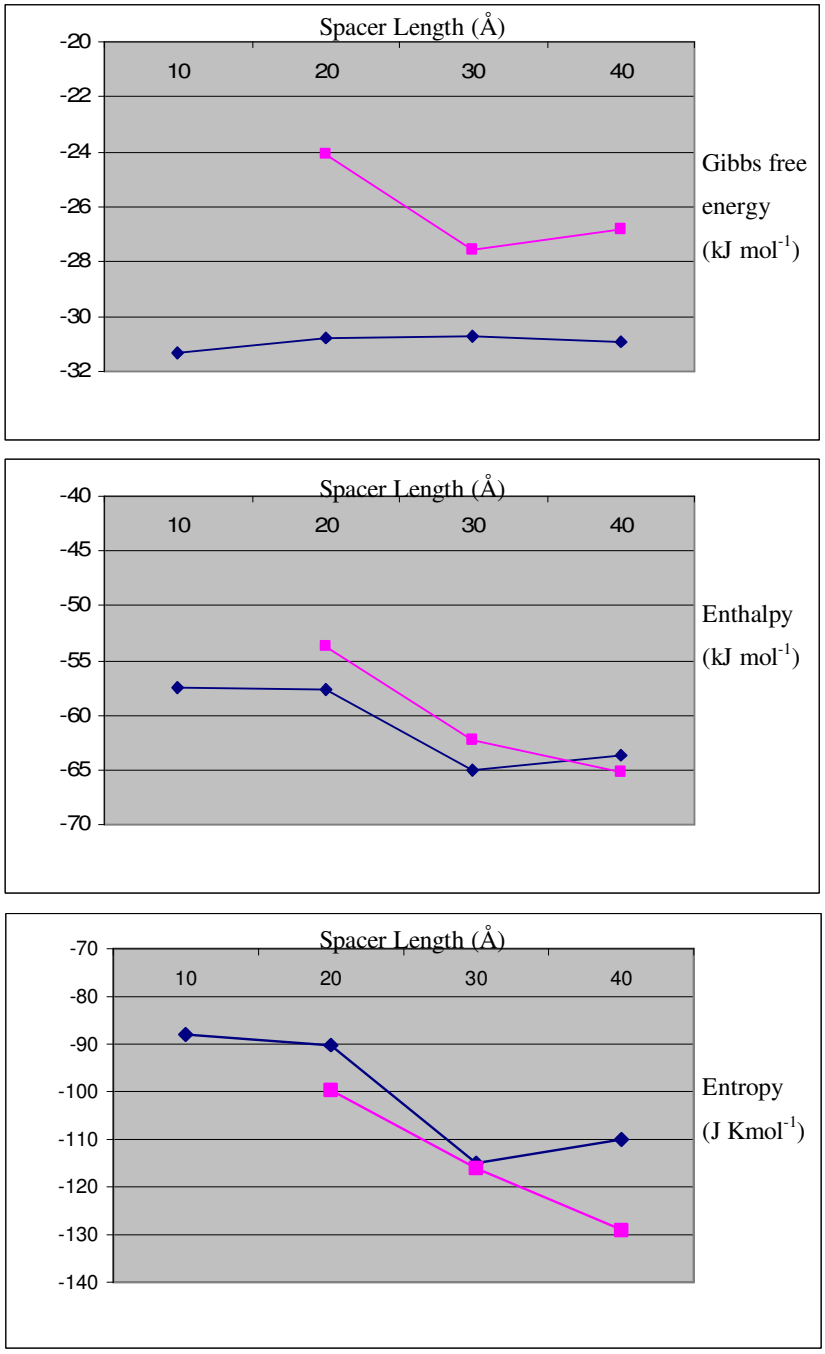


Figure 7-14. **The Gibbs free energy of the HsPCNA peptide interactions**
Thermodynamic properties of HsPCNA with peptides of different linker length (Å). Pink) immobilised peptide, Blue) immobilised PCNA. Top) Gibbs free energy, Middle) enthalpy and Bottom) entropy of interaction. This figure shows little difference in the free energy when PCNA is immobilised however differenaces with peptide length when peptide is immobilised. Immobilised peptides are of lower affinity as a result of a small increase in enthalpy of interaction and a decrease in entropy of interaction.

range of -53.8 to -65.2 kJ mol^{-1}. The greatest deviation (3.9 kJ mol^{-1}) is with interactions of HsPCNA and peptide 3, showing an increased enthalpy of interaction with immobilised peptide. This may be the result of part of the peptide being occluded from the HsPCNA binding site. Similar trends are shown in both SpPCNA and LmPCNA (table 7.1) however values determined for enthalpy are less reliable in these interactions.

The entropy of interaction decreases when peptide is immobilised e.g. HsPCNA to free peptide 5 is -110 J $Kmol^{-1}$ and HsPCNA to immobilised peptide 5 is -129 J $Kmol^{-1}$. The slight increase in entropic cost of the interaction with immobilised peptide could be the result of the large PCNA molecule stabilising some portions of the dextran matrix. The entropy of interaction for immobilised HsPCNA with peptides 3-5 are within the range of -90.2 to -115 J $Kmol^{-1}$ (Δ24.8 J $Kmol^{-1}$) and immobilised peptides 3-5 with free HsPCNA are within the range of -99.6 to -129 J $Kmol^{-1}$ (Δ29.4 J $Kmol^{-1}$). These values demonstrate that the entropy of interaction for all peptides are affected similarly and are not dependent on spacer length (figure 7.14). Similar trends of entropy are shown in both SpPCNA and LmPCNA, however values determined for entropy are less reliable.

The values determined in this chapter will be compared with those determined by ITC and thermal denaturation assays in chapter 8.

CHAPTER 8. Conclusions from the PCNA peptide interaction studied

The work presented in this thesis has been a thorough study of the peptide interactions displayed by PCNA from three different species. These studies have been performed by dynamic scanning fluorimetry, ITC, SPR and observed crystallographically. In this chapter, the results are discussed and differences between the analytical methods are analysed.

The affinities determined in this thesis are put into context and compared to literature values. The thermodynamic analysis ascertained by SPR and the van't Hoff plot are also compared to literature values. The PCNA-peptide interaction can then be broken down to find the contribution of each section of the peptide.

8.1 The PCNA peptide interactions observed in this work show HsPCNA to have the tightest interaction

The affinity data determined from this work has determined a series of K_ds for the PCNA-peptide interactions (tables 8.1). Affinities for the HsPCNA and SpPCNA interactions for peptides 0-5 have been determined at a temperature of 298K using DSF (chapter 5). Affinities for the HsPCNA, SpPCNA and LmPCNA interactions for peptides 0-5 have been determined at two temperatures (283K and 298K) using ITC (chapter 6). Affinities for the HsPCNA, SpPCNA and LmPCNA interactions for peptides 1-6 have been determined at a range of temperatures (283K-313K at 5K intervals) using SPR (chapter 7). These peptides are represented in figure 8.1.

The HsPCNA-peptide interactions studied in this work, as determined by SPR, DSF and ITC, was shown to have the greatest affinity while the LmPCNA interactions have the lowest affinity (table 8.1), see SPR analysis of peptide 2 for

Peptide0 – KKRQTSMTDFYHSKRRLIFS

Peptide1 – KRRQTSMTDFYH

Peptide2 – Bio-KRRQTSMTDFYH

Peptide3 – Bio-PEG-KRRQTSMTDFYH

Peptide4 – Bio-PEG-PEG-KRRQTSMTDFYH

Peptide5 – Bio-PEG-PEG-Ahx-KRRQTSMTDFYH

Peptide6 – KRRQTSMTDAYH

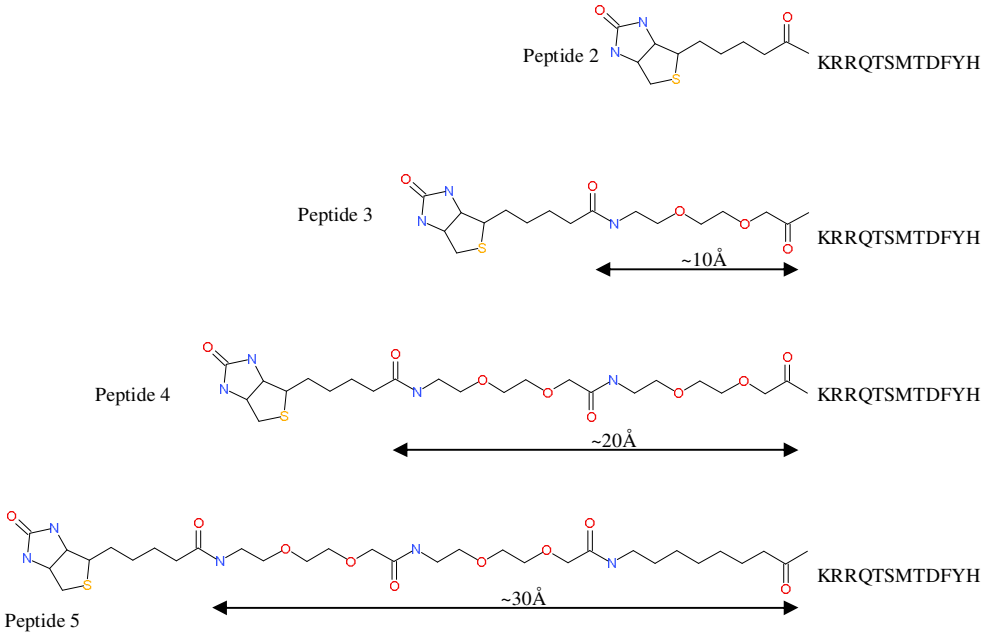


Figure 8-1. **The peptides investigated in this work**
Figure showing the peptides used in SPR experiments carried out in this study.
Top) Peptides 0-6 showing shorthand forms of all 6 peptides. Peptides 1-6 all have the same peptide sequence. Peptide 1 and 6 are untagged and peptide 6 has a F150A point mutation on peptide 6. Peptides 2-5 with biotin tags and linkers made of PEG and a hydrophobic spacer (Ahx).
Bottom) Peptides 2-5, showing the biotin tags and the chemical sequence of the linkers. The additional spacers have their approximate length shown below in Ångström.

HsPCNA

Peptide	Immobilised Protein (SPR)	Immobilised Peptide (SPR)	DSF	ITC
0	NI	NA	0.107μM	0.004μM
1	0.84μM	NA	13.5μM	0.470μM
2	3.16μM	NI	15.0μM	0.375μM
3	3.71μM	61.8μM	16.2μM	0.230μM
4	3.72μM	16.5μM	10.2μM	1.12μM
5	3.46μM	20.0μM	9.68μM	0.670μM
6	128μM	NA	NP	NI

SpPCNA

Peptide	Immobilised Protein (SPR)	Immobilised Peptide (SPR)	DSF	ITC
0	NI	NA	0.012μM	0.010μM
1	1.20μM	NA	1.82μM	0.062μM
2	3.64μM	NI	4.53μM	0.240μM
3	5.07μM	4.52μM	5.10μM	0.410μM
4	4.55μM	2.19μM	3.91μM	NI
5	4.89μM	2.31μM	5.08μM	NI
6	95.9μM	NA	NP	NI

LmPCNA

Peptide	Immobilised Protein (SPR)	Immobilised Peptide (SPR)	DSF	ITC
0	NI	NA	NP	0.426μM
1	10.3μM	NA	NP	1.49μM
2	32.5μM	NI	NP	1.63μM
3	42.5μM	805μM	NP	NI
4	40.2μM	144μM	NP	NI
5	45.9μM	347μM	NP	NI
6	217μM	NA	NP	NI

Tables 8-1. **The K_d of interaction between PCNA and peptide at 298K**
The interaction of peptides with HsPCNA (top), SpPCNA (middle) and LmPCNA (bottom) at 298K. These tables show the affinities determined by the four different methods; SPR (immobilised PCNA), SPR (immobilised peptide), DSF and ITC. The affinities by these different techniques all show reasonable agreement, within ~1 order of magnitude. HsPCNA shows the greatest PCNA peptide affinity, and LmPCNA shows the lowest PCNA peptide affinity.

NA - Not Applicable
NI - Not Interperatable
NP - Not performed

HsPCNA (3.16μM), SpPCNA (3.64μM) and LmPCNA (32.5μM). In all methods the peptide 0 interaction is the strongest with the interaction with peptide 6 being the weakest and peptide 2-5 having similar affinities e.g. DSF analysis of HsPCNA for peptide 0 (107nM), SPR analysis of HsPCNA for peptides 2-5 (~3.5μM) and peptide 6 (128μM) (table 8.1).

These values clearly demonstrate that HsPCNA has the strongest interaction for peptides based on p21 PIP box region. LmPCNA has a significantly weaker interaction with these peptide regions. Peptide 0 had the strongest interaction of all the peptides as a result of the increased peptide length. Peptides with linker regions all showed very similar affinities.

8.1.1 Comparison of the PCNA peptide affinities gave good agreement between different techniques used in this work

PCNA peptide interactions can be compared between these different analytical methods at 298K (table 8.1). This comparison shows reasonable agreement between different methods validating the determined values. In all instances, ITC presented a greater affinity for the PCNA peptide interaction. The values for the K_d of interaction determined by SPR and DSF showed very similar values e.g. HsPCNA for peptide 2 gives 3.16μM by SPR and 15.0μM by DSF; SpPCNA for peptide 2 gives 3.64μM by SPR and 4.53μM by DSF. The similarity of these values validates these methods.

The SPR analysis of immobilised peptides and immobilised PCNA gave different values. As discussed in chapter 7, the peptide, when immobilised, shows a reduced affinity (table 8.1); likely as a result of steric hindrance from Streptavidin. The ITC analysis deviated by the greatest degree from the values determined by SPR

and DSF. In all instances, ITC presented a greater affinity and the thermodynamic values were unreliable for the PCNA-peptide interactions, discussed in chapter 7, as a result of the slight uncertainty in the peptide concentrations. This slight uncertainty has a significant effect on the determined thermodynamic values determined by ITC experiments.

8.1.2 Thermodynamic values of the PCNA peptide interaction gave consistent results from the van't Hoff analysis of SPR data

Thermodynamic values for the PCNA-peptide interaction have been derived from this work. The Gibbs free energy of interaction has been determined by DSF, ITC and SPR, as this can be calculated from K_d of interaction. The Gibbs free energy of interaction can be used to determine the thermodynamic values of the interaction as shown in equation 6.2.

6.2) $\Delta G = \Delta H - T\Delta S$

ITC analysis showed clearly that the interaction was enthalpically favourable and entropically unfavourable. The non-linear van't Hoff analysis of SPR data gave values for the enthalpy and entropy of interaction for HsPCNA with peptides 1-6 (table 8.2). These values showed a logical consistency. There was a decrease (increased favourability) in the enthalpy of interaction with increasing linker length (-49.2 – -65.0 kJ mol^{-1}). The entropy of interaction decreased (decreased favourability) with increasing spacer length (-58.9 – 115 J $Kmol^{-1}$). The non-linear van't Hoff analysis showed a decreasing (greater freedom of water upon interaction) specific heat capacity of interaction with increasing peptide hydrophobicity (-648 – -1310 J K^{-1} mol^{-1}), the greatest ΔC_p associated to peptide with the hydrophobic spacer.

Peptide		1	2	3	4	5	6
HsPCNA	ΔH (kJ mol^{-1})	-49.2	-57.5	-57.7	-65.0	-63.7	-49.6
	ΔS (J $Kmol^{-1}$)	-58.9	-88.0	-90.2	-115	-110	-91.4
	ΔC_p (J K^{-1} mol^{-1})	-698	-668	-648	-752	-1310	-436
	ΔG (kJ mol^{-1})	-31.6	-31.3	-30.8	-30.7	-30.9	-22.3
SpPCNA	ΔH (kJ mol^{-1})	-41.1	-66.5	-67.4	-77.4	-67.8	-66.6
	ΔS (J $Kmol^{-1}$)	-24.6	-120	-125	-158	-127	-147
	ΔC_p (J K^{-1} mol^{-1})	-2110	-1600	-944	-2010	-1070	1290
	ΔG (kJ mol^{-1})	-33.8	-30.7	-30.2	-30.3	-30.0	-22.8
LmPCNA	ΔH (kJ mol^{-1})	-46.6	-34.9	-35.1	-42.0	-35.2	-27.2
	ΔS (J $Kmol^{-1}$)	-61.2	-31.7	-33.7	-56.8	-34.9	-22.7
	ΔC_p (J K^{-1} mol^{-1})	1480	138	419	1390	658	2240
	ΔG (kJ mol^{-1})	-28.4	-25.5	-25.1	-25.1	-24.8	-20.4

Table 8-2. **Thermodynamic values derived from the non-Linear van't Hoff plots of immobilised PCNA interactions with free peptide by SPR**

Table showing the values of the thermodynamic values of PCNA-peptide interactions determined by the non-linear van't Hoff analysis. The enthalpy, entropy, specific heat capacity and Gibbs free energy of interactions are grouped by PCNA species. The values for every peptide are shown. This demonstrates an enthalpically favourable, entropically unfavourable and a Gibbs free energy which demonstrates this to be a spontaneous interaction. HsPCNA demonstrates an increasingly favourable enthalpy and increasingly unfavourable entropy with increasing peptide length. This is consistent with increasing number of non-covalent interactions and removal of more flexible peptides.

The thermodynamic analysis of HsPCNA with peptide 1 and peptide 6 (the point mutant) showed some clear differences. There was little change in the enthalpic contributions, however a substantial increase in entropic cost of interaction with peptide 6. Peptide 1 showed a greater entropy and ΔC_p than peptide 6, due to the increased freedom of waters with peptide 1.

8.2 Literature values for the PCNA peptide interaction have been studied with a range of peptides

The PCNA-peptide interaction has been previously investigated to give the peptide consensus sequence and the affinity of interaction. These have determined the positions of the p21 protein involved in the PCNA interaction (figure 8.2) (Chen, J. J. *et al* 1996b). This work showed that only the C-terminus is required for the binding activity of p21 to PCNA. This early evidence showed the region of p21 which contained the PCNA interacting protein (PIP) box (figure 8.2) (Chen, J. J. *et al* 1996b; Warbrick, E. *et al* 1995).

PCNA from different sources e.g. *H. sapiens*, S.cerevisiae, S.pyogenes, E.coli and RB69 phage; have been analysed for the interaction from a variety of peptides (table 8.3). Peptides have been analysed from a variety of sources from different proteins and different species. *H. sapiens* proteins (p21, p66 and FEN1) have been analysed. S.pyogenes (DNA polymerase C), S.aureus (DNA polymerase C), E.coli (DNA polymerase III, DNA polymerase II) and RB69 (gp43) have also been analysed (table 8.3, figure 8.3).

Point mutants of these peptides have also been studied. The p21 mutants (M147A) and (D149A) (Zheleva, D. I. *et al* 2000) and the p21 mutants (D52A) and (H152A, K154A, R155A and R156A) (Knibiehler, M. *et al* 1996) have been

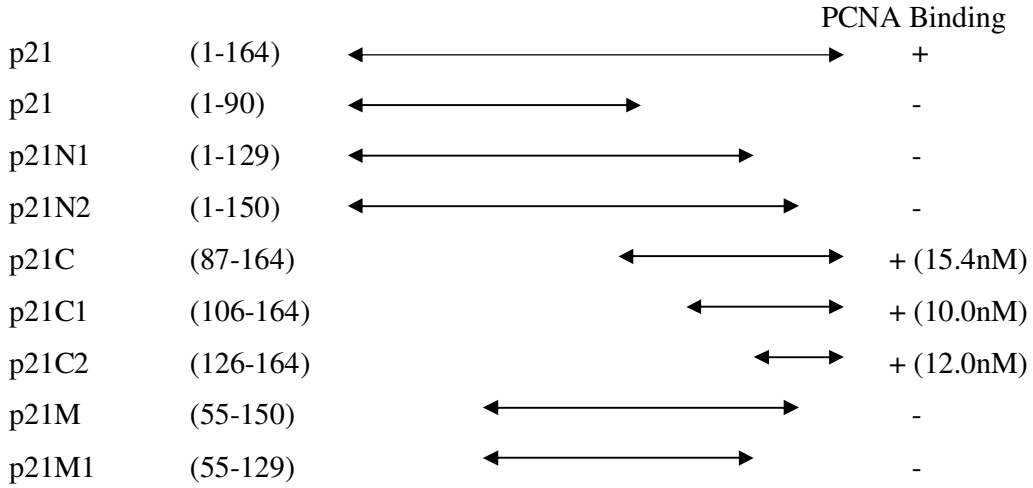


Figure 8-2. **The interaction of HsPCNA with sections of p21**
The interaction of different sections of the p21 protein with HsPCNA as determined in (Chen, J. J. *et al* 1996b). This figure shows the requirement of the C-terminal region for the association to PCNA in *H. sapiens*. It is also demonstrated that there is no other section in the p21 protein which is essential for PCNA association. The K_d of interaction between PCNA - p21C (189-164) was determined to be 15.4 nM and p21C2 (126-164) was 12nM at 37°C (Chen, J. J. *et al* 1996b). Modified from (Chen, J. J. *et al* 1996b)

Hs p21	GRKRRQTSMTDFYHSKRRLIFS	(139-160)
Hs p66	GKANRQVSITGFFQRK	(451-466)
Hs FEN1	RQGSTQGRLDDFFKVTGSLSSA	(332-353)
S.py DNA polC	MPEDNQLSLFDDFF	(1452-1465)
S.au DNA polC	LPDKAQLSIFDM	(1427-1438)
Ec DNA pol III	LIGSEQVELEFD	(1149-1160)
Ec DNA pol II	TLMTGQLGLF	(774-783)
RB69 gp43	ASLDFLFG	(891-898)

Figure 8-3. **PIP box sequences of different PCNA interaction proteins**
A selection of some of the peptide sequences used in the characterisation of the PCNA binding interface. These interacting peptides are from *H. sapiens* proteins (p21, p66 and FEN1), S.pyogenes (DNA polymerase C), S.aureus (DNA polymerase C), E.coli (DNA polymerase III, DNA polymerase II) and RB69 (gp43). The Hs p21 peptide has the residues involved in the PIP box sequence, Q-x-x-h-x-x-a-a, where x represents any amino-acid, h is any mildly hydrophobic residue and a is an aromatic.

PCNA species	Source protein	Peptide Sequence	Affinity	Method	Ref
H. sapiens	Hs p21	Full length (1-164)	2.55nM	SPR	1
	Hs p21	Full length (1-164)	320nM	SPR	2
	Hs p21	Full length (1-164) (D52A)	240nM	SPR	2
	Hs p21	Full length (1-164) (H152A, K154A, R155A, R156A)	39.0µM	SPR	2
	Hs p21	(75-164)	4.10µM	SPR	2
	Hs p21	(87-164)	15.4nM	Rad*	3
	Hs p21	(106-164)	10nM	SPR	4
	Hs p21	(126-164)	12.0nM	Rad*	3
	Hs p21	(139-160) GRKRRQTSMTDFYHSKRRLIFS	2.30nM	SPR	1
	Hs p21	(139-160) GRKRRQTSMTDFYHSKRRLIFS	82.6nM	ITC	5
	Hs p21	(141-160) KRRQTSMTDFYHSKRRLIFS	82.6nM	ITC	5
	Hs p21	(141-160) KRRQTSMTDFYHSKRRLIFS	87.7nM	ITC	6
	Hs p21	(141-160) (M147A) KRRQTSATDFYHSKRRLIFS	-	ITC	6
	Hs p21	(141-160) (D149A) KRRQTSMTAFYHSKRRLIFS	1.28µM	ITC	6
	Hs p66	(451-466) GKANRQVSITGFFQRK	15.6µM	ITC	5
	Hs p66	(451-472) GKANRQVSITGFFQRKRRLIFS	1.54µM	ITC	5
	Hs FEN1	Full length (1-370)	60nM	SPR	4
	Hs FEN1	(- C terminal region) (1-363)	-	SPR	4
	Hs FEN1	(332-353) RQGSTQGRLDDFFKVTGSLSSA	59.9µM	ITC	5
	Synthetic	SAVLQKKITDYFHPKK	100nM	ITC	6
S.cerevisiae	Hs p21	Full length (1-164)	~40nM	SPR	1
	Hs p21	(139-160) GRKRRQTSMTDFYHSKRRLIFS	~200nM	SPR	1
S.pyogenes	Spy Pol C	(1446-1465) MGILGNMPEDNQLSLFDDFF	1.50µM	Fluor	7
	Sa Pol C	DELGSLPNLPDKAQLSIFDM (1419-1438)	2.70µM	Fluor	7
E.coli	Ec Pol III	GATWRVSPSDRLLNDLRGLIGSEQVELEFD (1131-1160)	3.95µM	ITC	7
	Ec Pol III	RLLNDLRGLIGSEQVELEFD (1141-1160)	3.06µM	ITC	7
	Ec Pol III	SEQVELEFD (1152-1160)	910nM	ITC	7
	Ec Pol III	IGQADMFGV (918-926)	≥15µM	ITC	7
	Ec Pol II	TLMTGQLGLF (774-783)	1.70µM	ITC	7
	Synthetic	LQLELELDF	1.80µM	ITC	7
RB69	gp43	Full length (wild type)	1.10µM	ITC	8
	Synthetic	CSLDFLFG	7.00µM	ITC	8

Table 8-3. **Previous affinities determined in the PCNA peptide interaction**

Table showing the PCNA-peptide interaction partners with species of PCNA, the interaction partner species, and the source protein. The sequences of the peptide with PIP box are aligned. The affinity of interaction is shown, as well as the analysis method in each publication. These results are organised by PCNA species and peptide.

The source protein represents the source organism where Hs is *H. sapiens*, Spy is S.pyogenes, Sa is S.aureus, Ec is E.coli

Methods are represented where Rad* is a radioactivity pull-down assay as described in (Chen, J. J. *et al* 1996b) and Fluor is a fluorescence based assay as described in (Georgescu, R. E. *et al* 2008).

Ref are References: 1 is (Gibbs, E. *et al* 1997), 2 is (Knibiehler, M. *et al* 1996), 3 is (Chen, J. J. *et al* 1996b), 4 is (Chen, J. J. *et al* 1996a), 5 is (Bruning, J. B. and Shamoo, Y. 2004), 6 is (Zheleva, D. I. *et al* 2000), 7 is (Georgescu, R. E. *et al* 2008) and 8 is (Alley, S. C. *et al* 2001).

investigated. These mutations allowed the contribution of the residues to be investigated (table 8.3).

8.2.1 A broad range of affinities have been previously determined for the PCNA peptide interaction

Literature values for the analysis of peptide binding to PCNA have been determined using peptides from a diverse number of sources, including synthetic peptides for a broad selection of PCNAs. These affinities have been derived by different techniques e.g. ITC and SPR (table 8.3). The affinities derived from the range of PCNAs and peptides are from 2.3nM (HsPCNA with *H. sapiens* p21 (139-160) as analysed by SPR) (Gibbs, E. *et al* 1997) to 59.9μM (HsPCNA with *H. sapiens* FEN1 (332-353) as analysed by ITC) (Bruning, J. B. and Shamoo, Y. 2004). The table of interactions (table 8.3) shows a range of affinities for S.cerevisiae PCNA of 40-200nM, S.pyogenes 1.5-2.7μM, E.coli 0.91-15μM and RB69 1.1-7.0μM.

8.2.2 Different methods of affinity analysis have given different ranges of the HsPCNA peptide interaction

Early affinity values for the PCNA-peptide interaction were obtained by pull-down assays of radio-labelled PCNA (Chen, J. J. *et al* 1996b). This method involved the incubation of radio-labelled PCNA with peptides (figure 8.2) and the quantity of PCNA pulled down was measured. This technique gave values for the K_d of interaction at 37°C: the affinity of p21 (89-164) was 15.4nM and p21 (126-164) was 12.0nM (figure 8.2 and table 8.3).

ITC has been used to determine affinity data for this interaction. The HsPCNA *H. sapiens* p21 (139-160) (Bruning, J. B. and Shamoo, Y. 2004) and (141-

160) (Zheleva, D. I. *et al* 2000) have been determined in different publications. These different publications gave values for this interaction very similar to each other, 82.6nM and 87.7nM respectively (table 8.3).

SPR analysis of the HsPCNA p21 interaction has given affinity data. Interactions with p21 (106-164) (Chen, J. J. *et al* 1996a) (139-160) (Gibbs, E. *et al* 1997) has given values of 10nM and 2.3nM respectively. This is in close agreement to values previously determined (table 8.3). Within values determined from SPR data, there is variance. The interaction of the HsPCNA and full length p21 had a wide range of affinities from 2.55nM (SPR) (Gibbs, E. *et al* 1997) to 320nM (SPR) (Knibiehler, M. *et al* 1996).

SPR analysis of the PCNA peptide interactions gave some values for the kinetics of the interaction (Knibiehler, M. *et al* 1996). This analysis of full length p21 with HsPCNA gave a k_{on} of 1.1×10^{3} $M^{-1}s^{-1}$ and a k_{off} 3.7×10^{-4} s^{-1}. The kinetics of p21 derivatives are also presented e.g. p21 (D52A) gave a k_{on} of 1.1×10^{3} $M^{-1}s^{-1}$ a k_{off} 2.7×10^{-4} s^{-1}, p21 (H152A, K154A, R155A, R156A) gave a k_{on} of 1.0×10^{2} $M^{-1}s^{-1}$ a k_{off} 3.9×10^{-3} s^{-1} and p21 (75-164) gave a k_{on} of 2.1×10^{3} $M^{-1}s^{-1}$ and a k_{off} 8.7×10^{-3} s^{-1} (Knibiehler, M. *et al* 1996). The affinities given here show relatively slow on and off rates. This slow rate is the result of large peptide lengths and reformations of the PIP box binding sequence.

8.2.3 A broad range of thermodynamic values have been derived in the literature using ITC

HsPCNA-p21 interactions: Thermodynamic values derived from different methods in the literature show that this interaction is enthalpically favourable and entropically unfavourable. The enthalpic contribution has a greater effect upon the Gibbs free

energy than the entropic. Two previous examples have shown a slight entropic favourability of this interaction (Zheleva, D. I. *et al* 2000). The interaction of HsPCNA with p21 (141-160) and with a synthetic consensus sequence have shown entropies of 0.356 and 2.55 J $Kmol^{-1}$ respectively (table 8.4) (Zheleva, D. I. *et al* 2000). These entropically favourable interactions are in contradiction to other interactions outlined in the literature (table 8.4). The literature shows a broad range of thermodynamic values which infers these may not be reliable.

HsPCNA interactions with peptides from other sources: The interaction of HsPCNA with a series of peptides derived from different protein sources have all demonstrated to be enthalpically favourable and entropically unfavourable (Bruning, J. B. and Shamoo, Y. 2004). p21a (139-160), p21b (141-160), p66a (451-466), p66b (451-472) and FEN1 (332-353) have been analysed. Interaction of HsPCNA with all these ranges of peptides give a consistently negative enthalpy and entropy (Bruning, J. B. and Shamoo, Y. 2004) (table 8.4).

E.coli PCNA-peptide interactions: The binding of E.coli DNA polymerase III peptides to the E.coli DNA polymerase β subunit have shown an enthalpically favourable and are entropically unfavourable (Georgescu, R. E. *et al* 2008). The interaction of E.coli PCNA for peptides of different lengths derived from DNA polymerase III (1131-1160), (1141-1160) and (1152-1160) were analysed. Interaction of E.coli PCNA with DNA pol III (1131-1160), (1141-1160) and (1152-1160) shows a negative enthalpy and negative entropy of interaction (table 8.4) (Georgescu, R. E. *et al* 2008). These values show a similar mode of interaction between HsPCNA and peptides and that of E.coli PCNA and peptides.

Peptide	N	K_d	ΔH (kJ mol^{-1})	ΔS (J $Kmol^{-1}$)	ΔG (kJ mol^{-1})	Reference
HsPCNA interacting with *H. sapiens* peptides						
p21 (141-160)	1.0	82.6nM	-122	-267	-42.3	(Bruning, J. B. and Shamoo, Y. 2004)
p21 (141-160)	1.0	87.7nM	-40.8	0.36	-40.9	(Zheleva, D. I. *et al* 2000)
p21 (141-160) (D149A)	0.8	1.28μM	-47.2	-42.8	-34.4	(Zheleva, D. I. *et al* 2000)
p66 (451-466)	1.0	15.6μM	-79.9	-172	-28.0	(Bruning, J. B. and Shamoo, Y. 2004)
p66 (451-472)	1.0	1.54μM	-42.7	-29.7	-33.9	(Bruning, J. B. and Shamoo, Y. 2004)
FEN1 (332-353)	0.9	59.9μM	-83.7	-195	-24.7	(Bruning, J. B. and Shamoo, Y. 2004)
SAVL-QKKITDYF-HPKK	0.7	100nM	-39.8	2.55	-40.6	(Zheleva, D. I. *et al* 2000)
E.coli DNA polymerase β interacting with E.coli peptides						
DNA pol III (1131-1160)	1.3	3.95μM	-37.5	-24.3	-30.3	(Georgescu, R. E. *et al* 2008)
DNA pol III (1141-1160)	1.2	3.06μM	-38.2	-19.9	-32.2	(Georgescu, R. E. *et al* 2008)
DNA pol III (1152-1160)	0.9	0.91μM	-47.4	-47.2	-33.4	(Georgescu, R. E. *et al* 2008)
RB69 phage PCNA interacting with gp43 protein						
gp43	0.8	1.10μM	-58.6	-82.8	-33.9	(Alley, S. C. *et al* 2001)

Table 8-4. **Thermodynamic values of interactions determined by ITC in the literature**

Table showing the previously published thermodynamic values of the PCNA peptide interaction as determined by ITC. This table gives stoichiometry, K_d, enthalpy, entropy, Gibbs free energy and the published reference presenting these values.

The values shown here are split between the human (top), E.coli (middle) and RB69 phage (bottom). All these interactions are enthalpically favourable and commonly entropically unfavourable. This is consistent with the immobilisation of a large flexible peptide.

RB69 PCNA interactions: The interaction of GP45 (W92F/Y165W/W199F) phage PCNA with GP43 (a DNA polymerase) has been analysed by ITC and demonstrated an enthalpically favourable and entropically unfavourable (table 8.4) (Alley, S. C. *et al* 2001). This interaction shows that the PCNA interaction for this, the most structurally divergent PCNA, has a similar mode of interaction.

8.3 The comparison of literature values and those derived in this work showed similar affinities and thermodynamic values

Literature values showed close agreement with values in this work. The well characterised HsPCNA peptide 0 interaction has been determined to have an affinity of 87.7nM (table 8.4) (Zheleva, D. I. *et al* 2000) which is in close agreement to the 107nM affinity determined by DSF in this work (table 8.1).

HsPCNA has been shown to have a 10-fold greater affinity for p21 than lower eukaryotes (Gibbs, E. *et al* 1997). This work has shown a similar affinity of HsPCNA and SpPCNA for p21 peptide with an average of ~0.2 to 0.4 fold weaker for a truncated peptide (p21 (141-152)). Surprisingly, the longer p21 peptide (141-160) demonstrated a greater affinity for SpPCNA than HsPCNA in both DSF and ITC. The shorter peptides were not analysed against HsPCNA, the most similar interaction studied was that of the 9mer peptide for E.coli PCNA. This interaction gave a K_d of 0.91μM (table 8.3) whereas the HsPCNA to p21 (141-152) interaction gave a K_d of 0.84μM (table 8.1).

There is reasonable agreement between the thermodynamic values determined in this work (table 8.2) and those of the literature values (table 8.4). All interactions determined in this work give a negative enthalpy (favourable) and negative entropy (unfavourable) of interaction. Literature values show the same trend

(table 8.4) as determined by ITC. The interaction of the greatest similarity in this work, measuring the thermodynamic values of HsPCNA with peptide 0 by ITC, gave an enthalpy of -88.5 kJ mol^{-1}; entropy of -159 J $Kmol^{-1}$; and Gibbs free energy of -43.7 kJ mol^{-1}. These values are in reasonable agreement with those determined in the literature; with an enthalpy of -122 kJ mol^{-1}; entropy of -267 J $Kmol^{-1}$; and Gibbs free energy of -42.3 kJ mol^{-1} (table 8.4) (Bruning, J. B. and Shamoo, Y. 2004). The shorter peptides were not analysed against HsPCNA, the most similar interaction studied was that of the 9mer peptide for E.coli PCNA. This interaction gave an enthalpy of -47.4 kJ mol^{-1}; entropy of -47.2 J $Kmol^{-1}$; and Gibbs free energy of -33.4 kJ mol^{-1} (table 8.4) (Georgescu, R. E. *et al* 2008). The HsPCNA to p21 (141-152) interaction gave an enthalpy of -49.2 kJ mol^{-1}; entropy of -58.9 J $Kmol^{-1}$; and Gibbs free energy of -31.6 kJ mol^{-1} (table 8.2). The similarity of these values increases the confidence in the data obtained in this work.

8.3.1 Differences between some literature values and those derived in this work may be due to broad range of values given in the literature

Some values of the PCNA peptide interaction (discussed above, section 8.3) agree closely with those determined in this work. Other values shown in the literature (table 8.3) differ e.g. the interaction of the HsPCNA and full length p21 had a wide range of affinities from 2.55nM (SPR) (Gibbs, E. *et al* 1997) to 320nM (SPR) (Knibiehler, M. *et al* 1996). The most similar analysis to those carried out in this work was the interaction of HsPCNA to p21 (141-160) (Zheleva, D. I. *et al* 2000) which returned similar affinities. The interaction of the short peptides were not analysed for HsPCNA; however the interaction of E.coli PCNA for 9mer peptides (table 8.3) gave affinities similar to that determined in this work (table 8.1).

The thermodynamic values derived in this work for the HsPCNA peptide interactions (table 8.2) were self-consistent. These values gave increasing enthalpic return and increasing entropic cost with increasing peptide length (table 8.2) which is the logical progression due to longer more flexible molecules being taken out of solution and forming more intermolecular bonds to stabilise this. Thermodynamic values presented in the literature are less consistent with some interactions showing an entropically favourable interaction (table 8.4) (Zheleva, D. I. *et al* 2000). Values with little similarity to published data may well be the result of published values being poorer.

8.3.2 HsPCNA and PCNA from lower eukaryotes have been shown to have significantly different affinities which is not observed in this work

The interaction of HsPCNA with p21 peptides are in close agreement with the literature values. There has been a previous study comparing the binding affinity of p21 (139-160) to HsPCNA and S.cerevisiae PCNA (Gibbs, E. *et al* 1997). This study demonstrated that the interaction with S.cerevisiae PCNA had an ~100 fold reduction in affinity (2.3nM-200nM) (table 8.3). There is a disagreement with literature values showing a slight increase in affinity for the interaction of peptide 0 with fungal PCNA compared to that of HsPCNA (table 8.1). This may be the result of specific differences in the PIP box binding interface of SpPCNA and ScPCNA.

The SpPCNA interactions for p21 (141-152) in this work only gave a slight reduction in affinity compared to that of HsPCNA (table 8.1). This slight reduction in affinity may be a result of the similarity of the core interactions at this PIP box region. With the increase in the peptide length, literature values give a disproportional increase in affinity for HsPCNA; this may be the result of larger

number of interactions between the peptide and IDCL which are lost when p21 interacts with SpPCNA. This is demonstrated in chapter 3, as the core section of the p21 peptide PIP box (141-152) forms the same number of hydrogen bonds. There has been no previous literature on the Leishmania PCNAs affinity; however, this interaction is clearly weaker than that of eukaryote PCNA. The reduction in affinity could be the result of poorly conserved residues at the PIP box binding interface also discussed in chapter 3.

8.4 Thermodynamic contributions of different sections of the p21 peptide can be determined

A large quantity of binding data and thermodynamic analysis has been carried out on the interaction between PCNA and peptides derived from p21. This information has been obtained with peptides of differing lengths and with point mutations. Thermodynamic analysis of point mutants and peptides of differing lengths have allowed the energetic contribution of different sections to be analysed. The thermodynamic values of the two main sections and two side chains are summarised in this section.

8.4.1 Peptides of different lengths have different thermodynamic contributions to the binding interaction

Thermodynamic values of the large peptides used in this study and in the literature (p21 (141-160)) have been shown (table 8.5). The thermodynamic values of the shorter peptide (p21 (141-152)) have also been shown in this study (table 8.5). The interaction with the weaker binding peptide resulted in a difference in K_d of 776nM, enthalpy of -34.6 kJ mol^{-1}, entropy of -83.1 J $Kmol^{-1}$ and Gibbs free energy of -10.7 kJ mol^{-1}.

The thermodynamic values derived for the different peptides infer that the C-terminal section of the p21 peptide (153-160), which interacts with the IDCL, provides -34.6 kJ mol^{-1} of enthalpy and -83.1 J $Kmol^{-1}$ of entropy. This increases the favourability of the interaction by a Gibbs free energy of -10.7 kJ mol^{-1}.

8.4.2 Different point mutations have been analysed in literature and in this work to give the thermodynamic contribution of these side chains

Thermodynamic values of peptides derived from p21 have been analysed (table 8.5) with point mutations in specific sites. The thermodynamic values of the point mutant peptides (141-160, D149A and 141-152 F150A) have been derived in the literature and in this work (table 8.5). The thermodynamic values of the p21 (141-160, D149A) mutant peptide and that of the average p21 (141-160) peptide have been determined and can be compared (table 8.5). The thermodynamic values of the p21 (141-152, F150A) mutant peptide and that of the p21 (141-152) peptide have been determined and can be compared (table 8.5). The result of the D149A mutation in the thermodynamic values gave a difference in K_d of 1.22 μM, enthalpy of -36.6 kJ mol^{-1}, entropy of -100 J $Kmol^{-1}$ and Gibbs free energy of -7.9 kJ mol^{-1}. The result of the F150A mutation in the thermodynamic values gave a difference in K_d of 127 μM, enthalpy of 0.4 kJ mol^{-1}, entropy of 32.5 J $Kmol^{-1}$ and Gibbs free energy of -9.3 kJ mol^{-1}.

The thermodynamic values derived for the different peptides give information about the D149 side-chain of the p21 peptide (141-160), which exists at the helix region of the p21 peptide. This side-chain provides -36.6 kJ mol^{-1} of enthalpy and -100 J $Kmol^{-1}$ of entropy. The D149 side-chain increases the favourability of the interaction by a Gibbs free energy of -7.9 kJ mol^{-1}. The F150 side chain of the p21

Peptide	N	K_d	ΔH (kJ mol^{-1})	ΔS (J $Kmol^{-1}$)	ΔG (kJ mol^{-1})	Method	Ref.
HsPCNA interacting with *H. sapiens* peptides							
p21 (141-160)	1.0	82.6nM	-122	-267	-42.3	ITC	(Bruning, J. B. and Shamoo, Y. 2004)
p21 (141-160)	1.0	87.7nM	-40.8	0.36	-40.9	ITC	(Zheleva, D. I. *et al* 2000)
p21 (141-160)	0.9	21.9nM	-88.5	-159	-43.7	ITC	This work
p21 (141-160) (D149A)	0.8	1.28μM	-47.2	-42.8	-34.4	ITC	(Zheleva, D. I. *et al* 2000)
p21 (141-152)	n/a	0.84μM	-49.2	-58.9	-31.6	SPR	This work
p21 (141-152) (F150A)	n/a	128μM	-49.6	-91.4	-22.3	SPR	This work

Table 8-5. **Thermodynamic values of interactions in the literature**
Table showing the energetic properties of different peptides based on the *H. sapiens* p21. This table shows the effect of peptide length and point mutations. The energetics of the p21 (141-160) interaction have been published twice and given vastly different values (Bruning, J. B. and Shamoo, Y. 2004; Zheleva, D. I. *et al* 2000). A similar peptide has been studied in this work and has given energetic values which lie between these published values, suggesting that those presented here may be correct. The p21 (141-160) (D149A) mutant is significantly weaker binding than the wild-type (Zheleva, D. I. *et al* 2000) and the differences in the thermodynamic values from the above average are the result of this point mutation. A peptide based on p21 (141-152) investigated in this work has given thermodynamic data and an F150A point mutation has shown differences in the thermodynamic values as a result of this point mutation.

peptide (141-152), associates to a hydrophobic surface at the PIP box binding interface. This side-chain provides 0.4 kJ mol^{-1} of enthalpy and 32.5 J $Kmol^{-1}$ of entropy. The F150 side-chain increases the favourability of the interaction by a Gibbs free energy of -9.3 kJ mol^{-1}.

8.4.3 The energetic contributions of different sections of the p21 peptide have biochemical reasons which affect the overall binding interaction

Contribution of p21 (141-152): The interaction of p21 (141-152) to HsPCNA gives an enthalpy of -49.2 kJ mol^{-1}, entropy of -58.9 J $Kmol^{-1}$ and a Gibbs free energy of -31.6 kJ mol^{-1} (table 8.5) (figure 8.4). The thermodynamic values for this interaction are the result of the structural interactions taking place, which has been shown in chapter 3. The interaction with the section of p21 (143-152) has been studied in HsPCNA, SpPCNA and LmPCNA. The interaction between p21 (143-152) and HsPCNA gives a buried surface area of 1200 sq.Å and allows the formation of seven hydrogen bonds (figure 8.5) (Gulbis, J. M. *et al* 1996).

Contribution of p21 (153-160): The interaction of p21 (153-160) to HsPCNA gives an enthalpy of -34.6 kJ mol^{-1}, entropy of -83.1 J $Kmol^{-1}$ and a Gibbs free energy of -10.7 kJ mol^{-1} (table 8.5) (figure 8.4). The thermodynamic values for this interaction are the result of the structural interactions taking place which has been shown in chapter 3. The interaction with the section of p21 (153-160) has been studied in HsPCNA and SpPCNA. The interaction between p21 (143-152) and HsPCNA gives a buried surface area of 1060 sq.Å and allows the formation of thirteen hydrogen bonds (figure 8.5) (Gulbis, J. M. *et al* 1996). This section of peptide is shorter than p21 (143-152), yet it is however more entropically unfavourable, as a result of the stabilisation of the IDCL as well as the immobilisation of the peptide.

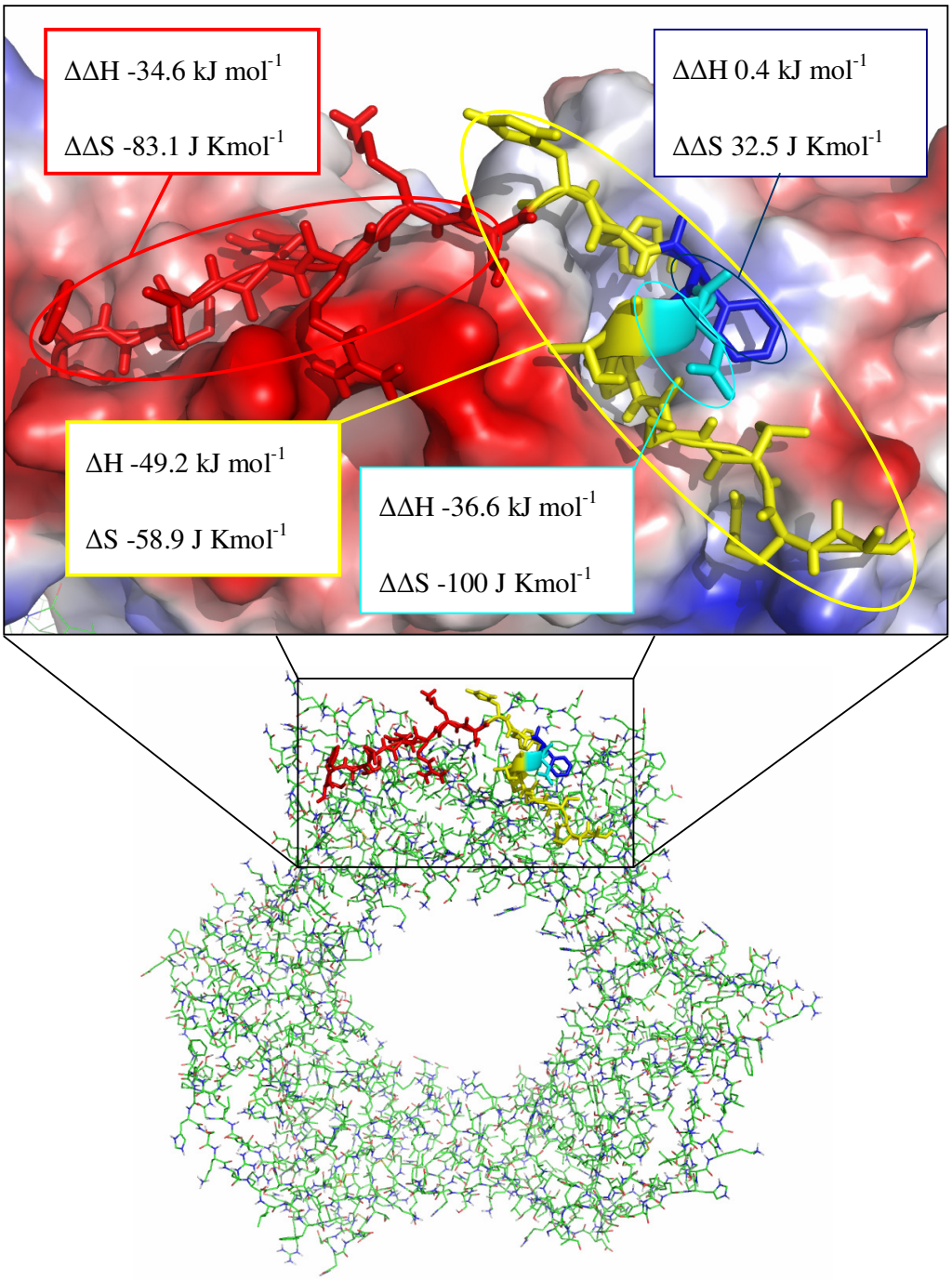


Figure 8-4. **The energetic contributions of the p21 (141-160) peptide**

Figure showing the PCNA trimer (green lines) bound with the p21 (141-160) (cartoon and stick) with the 143-152 section (yellow), D149 (cyan), F150 (blue) and 153-160 section (red). The energetics are shown for these sections in the appropriate colour with the addition of the point mutant side chains as a difference from alanine.

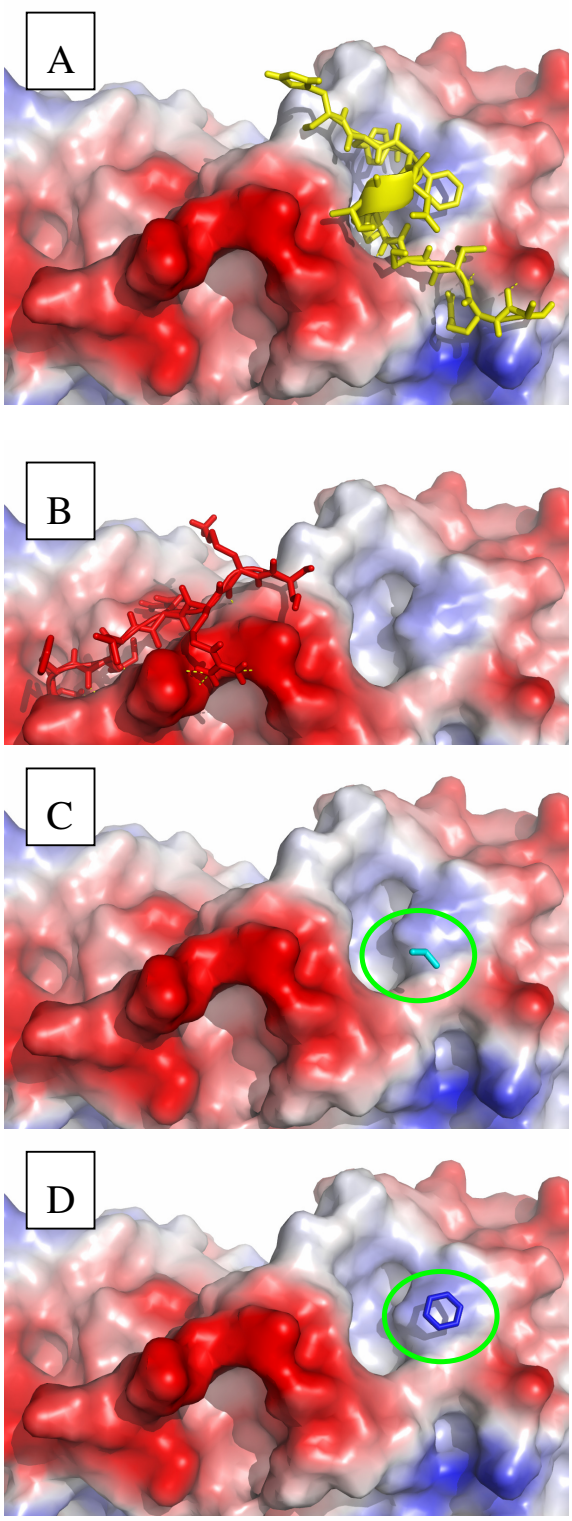


Figure 8-5. **The four sections of the p21 (143-160) peptide with determined thermodynamic values**

Figure showing the different sections of the p21 (143-160) analysed in this work. The hydrogen bonds are represented in yellow for each of these sections which are coloured as in figure 8.4. These peptides are shown in the PIP box binding interface of HsPCNA shown in an electrostatic surface and colours

A) The p21 (143-152) section represented in yellow. Forming seven hydrogen bonds at the interface and gives a buried surface area of 1200 sq.Å.

B) The p21 (153-160) section represented in red. Forming thirteen hydrogen bonds at the interface and gives a buried surface area of 1060 sq.Å.

C) The p21 D149 side chain section represented in cyan. This forms two hydrogen bonds with solvent not shown and results in no additional buried surface area.

D) The p21 F150 side chain section represented in blue. This forms no additional hydrogen bonds however results in an additional buried surface area of 170 sq.Å. The analysis of the interaction of this side-chain with HsPCNA by PISA determined it to be genuine.

Contribution of p21 (D149 side chain): The interaction of p21 (D149 side chain) to HsPCNA gives an enthalpy of -36.6 kJ mol^{-1}, entropy of -100 J $Kmol^{-1}$ and a Gibbs free energy of -7.9 kJ mol^{-1} (table 8.5) (figure 8.4). The thermodynamic values for this interaction are error prone as this is a result observed in only a single publication (Zheleva, D. I. *et al* 2000). The interaction of the D149 side-chain of p21 has been observed in HsPCNA, SpPCNA and LmPCNA. The interaction between the p21 D149 side-chain and HsPCNA results in no additional buried surface areas and forms two hydrogen bonds via waters to other areas in the peptide (figure 8.5) (Gulbis, J. M. *et al* 1996). This residue forms on the helix in the p21 (141-152) peptide, and Alanine is a stronger promoter of helix formation than Aspartic acid (Blaber, M. *et al* 1994; Blaber, M. *et al* 1993; Chellgren, B. W. and Creamer, T. P. 2006; Creamer, T. P. and Rose, G. D. 1994; Lewis, P. N. *et al* 1970; Oneil, K. T. and Degrado, W. F. 1990; Padmanabhan, S. *et al* 1990; Richardson, J. M. *et al* 2005). Mutation to A150 may increase the propensity for the peptide to be formed as a helix, increasing favourability of the interaction by reducing the entropic cost.

Contribution of p21 (F150 side chain): The interaction of p21 F150 side-chain to HsPCNA gives an enthalpy of 0.4 kJ mol^{-1}, entropy of 32.5 J $Kmol^{-1}$ and a Gibbs free energy of -9.3 kJ mol^{-1} (table 8.5) (figure 8.4). The thermodynamic values for this interaction are the result of the structural interactions taking place, which have been shown in chapter 3. The interaction with the section of p21 F150 side-chain has been studied in HsPCNA, SpPCNA and LmPCNA. The interaction between p21 F150 side-chain with HsPCNA gives a buried surface area of 170 sq.Å and allows the formation of no hydrogen bonds (figure 8.5) (Gulbis, J. M. *et al* 1996). The hydrophobic interaction of the F150 ring to the HsPCNA would be expected to

provide some favourable enthalpic contribution to the interaction. It can be argued that the solvation shell that surrounds the F150 ring would be held together by hydrogen bonds and impose an enthalpic cost to the complex formation. The release of waters can allow a favourable entropic contribution to the complex formation and may explain the small enthalpic change and large entropic cost of the p21 F150A mutant.

CHAPTER 9. Conclusions and future work

9.1 Overall coclusions of this work

In this work PCNA from these three species have been prepared in milligram quantities for biochemical and biophysical studies. The previously unknown structure of LmPCNA has been solved in an uncomplexed form and also complexed with a dodecapeptide to a resolution of 3.0Å. A comparison of PCNA structures and their peptide complexes for the three species identifies structural differences which may be relevant in analysing thermodynamic contributions of binding.

The binding affinities for these three PCNAs have been determined with a selection of peptides derived from the Hs p21 protein. This work has shown, using a selection of different techniques including Surface Plasmon Resonance (SPR), Isothermal Titration Calorimetry (ITC) and Dynamic Scanning Fluorimetry (DSF); that HsPCNA and SpPCNA have similar affinities for a 12mer peptide (K_d of ~1μM) however LmPCNA shows significantly weaker interactions (K_d of ~10μM). This is most likely due to divergence in the sequence and structure of LmPCNA.

A systematic investigation by SPR on the effect of peptide linker length on binding has been carried out using a series of synthesised peptides with different lengths of chemical spacer. The series of streptavidin immobilised peptides show that longer spacers are required for the recovery of the PCNA peptide binding affinity. The results presented in this work indicate that a linker length of at least 20Å is required for measurable protein binding activity. This interaction is improved with longer peptide spacers.

9.2 Possible improvements to the crystal structure solutions

The crystal structure described in chapter 3 had a low resolution and low completeness. It is likely that this is the result of the high (~70%) solvent content of these crystals. The loose crystal packing may be the result of the flexible loop domains holding the PCNA rings apart and not allowing a tight packing form.

The problems associated to this flexible loop region may be alleviated by the limited proteolysis of the LmPCNA prior to crystallisation. Alternatively, a deletion mutant can be created to produce LmPCNA without this region and improve the crystal packing, and by extension, the crystal structure.

9.2.1 Benefits of the improved structure of LmPCNA

With a significantly more accurate structure solved, including the position of immobilised waters, modification to the PIP box binding interface upon peptide binding can be more accurately observed. These changes can reveal information regarding the mode of binding and account for the thermodynamic values determined in this work.

The accurate protein structure of the *L. major* PCNA can also be used for structure based drug design. This can be used with the PIP box binding interface in both the complexed and uncomplexed forms. A study of this type may create new novel small molecule inhibitors which may lead to potential cures for Leishmaniasis detailed in chapter 1.

9.3 Possible improvements to the thermodynamic values determined

The thermodynamic values determined in this work appear to be genuine for the HsPCNA peptide interaction. This has, however, been solved by the van't Hoff analysis of SPR affinity data. It is preferable for thermodynamic values to be

determined by ITC. ITC was attempted in this work, however slight uncertainties in the peptide concentrations and the high demand on the available peptide resulted in no thermodynamic values being determined by this method. Accurate peptide concentration can be obtained by quantitative amino acid analysis, allowing these experiments to be repeated with greater accuracy.

9.3.1 Benefits of improved ITC data

The ITC data cannot only give accurate determination of the enthalpy and entropy of an interaction, but also, when performed over a temperature range, the specific heat capacity of interaction. The values that may be determined using this method can validate those determined by SPR. This may also allow data to be determined for the SpPCNA and LmPCNA peptide interactions. Thermodynamic information on all three PCNAs can allow differences in the mode of interaction to be ascertained. These possible differences can then be exploited in the process of designing small molecule inhibitors.

9.4 Further investigation into the effect of spacers in SPR experiments

The effect of spacers on peptide affinity when immobilised to a sensor surface is apparent as shown in this work. The shorter peptides result in the PCNA peptide affinity to be drastically diminished. It has been clearly demonstrated that biotinylated peptides require spacers of >20Å in order to give reasonable affinities. This has however only been demonstrated in this peptide PCNA-interaction.

The effect of spacers on the SPR data determined may fluctuate for protein-DNA interactions or interactions with other peptides or other proteins. It is possible to study the effect of spacers in other interactions and determine if these effects are consistent in different interactions.

9.4.1 Benefits of further investigation into the effect of spacers in SPR experiments

The immobilisation of ligand by Streptavidin in SPR experiments is simple and does not damage the interacting molecule. As a result, this method has significant advantages over other methods e.g. amine coupling. The requirement of the spacer is commonly assumed, however, values for the minimum spacer length for different interactions have not been determined. If this information were resolved, the method of SPR can be made more accessible to other divisions of sciences by the increased simplicity of ligand immobilisation and increased reliability of the determined values.

Adamczyk, M., Moore, J. A., and Yu, Z. G.: **Application of surface plasmon resonance toward studies of low-molecular-weight antigen-antibody binding interactions**. Methods-A Companion to Methods in Enzymology 2000, 20:319-328

Aiyar, A.: **The use of CLUSTAL W and CLUSTAL X for multiple sequence alignment**. Methods Mol Biol 2000, 132:221-241

Alley, S. C., Trakselis, M. A., Mayer, M. U., Ishmael, F. T., Jones, A. D., and Benkovic, S. J.: **Building a replisome solution structure by elucidation of protein-protein interactions in the bacteriophage T4 DNA polymerise holoenzyme**. Journal of Biological Chemistry 2001, 276:39340-39349

Anderson, E. H.: **Growth Requirements of Virus-Resistant Mutants of Escherichia-Coli Strain-B**. Proceedings of the National Academy of Sciences of the United States of America 1946, 32:120-128

Ayyagari, R., Impellizzeri, K. J., Yoder, B. L., Gary, S. L., and Burgers, P. M. J.: **A Mutational Analysis of the Yeast Proliferating Cell Nuclear Antigen Indicates Distinct Roles in Dna-Replication and Dna-Repair**. Molecular and Cellular Biology 1995, 15:4420-4429

Baca, A. M. and Hol, W. G. J.: **Overcoming codon bias: A method for high-level overexpression of Plasmodium and other AT-rich parasite genes in Escherichia coli**. International Journal for Parasitology 2000, 30:113-118

Bambara, R. A., Murante, R. S., and Henricksen, L. A.: **Enzymes and reactions at the eukaryotic DNA replication fork**. Journal of Biological Chemistry 1997, 272:4647-4650

Bates, P. A., Kelley, L. A., MacCallum, R. M., and Sternberg, M. J. E.: **Enhancement of protein modeling by human intervention in applying the automatic programs 3D-JIGSAW and 3D-PSSM**. Proteins-Structure Function and Genetics 2001,39-46

Becker, P. B. and Horz, W.: **ATP-dependent nucleosomere modeling**. Annual Review of Biochemistry 2002, 71:247-273

Beckman, R. A. and Loeb, L. A.: **Multistage Proofreading in Dna-Replication**. Quarterly Reviews of Biophysics 1993, 26:225-331

Blaber, M., Zhang, X. J., Lindstrom, J. D., Pepiot, S. D., Baase, W. A., and Matthews, B. W.: **Determination of Alpha-Helix Propensity Within the Context of A Folded Protein - Sites 44 and 131 in Bacteriophage-T4 Lysozyme**. Journal of Molecular Biology 1994, 235:600-624

Blaber, M., Zhang, X. J., and Matthews, B. W.: **Structural Basis of Amino-Acid Alpha-Helix Propensity**. Science 1993, 260:1637-1640

Boeckler, F. M., Joerger, A. C., Jaggi, G., Rutherford, T. J., Veprintsev, D. B., and Fersht, A. R.: **Targeted rescue of a destabilized mutant of p53 by an in silico screened drug**. Proceedings of the National Academy of Sciences of the United States of America 2008, 105:10360-10365

Bomar, M. G., Pai, M. T., Tzeng, S. R., Li, S. S. C., and Zhou, P.: **Structure of the ubiquitin-binding zinc finger domain of human DNA Y-polymerase eta**. Embo Reports 2007, 8:247-251

Bond, C. S.: **TopDraw: a sketchpad for protein structure topology cartoons**. Bioinformatics 2003, 19:311-312

Boulton, S. J. and Jackson, S. P.: **Saccharomyces cerevisiae Ku70 potentiates illegitimate DNA double-strand break repair and serves as a barrier to error-prone DNA repair pathways**. EMBO Journal 1996, 15:5093-5103

Bowman, G. D., O'Donnell, M., and Kuriyan, J.: **Structural analysis of a eukaryotic sliding DNA clamp-clamp loader complex**. Nature 2004, 429:724-730

Boysen, R. I., Wang, Y., Keah, H. H., and Hearn, M. T. W.: **Observations on the origin of the non-linear van't Hoff behaviour of polypeptides in hydrophobic environments**. Biophysical Chemistry 1999, 77:79-97

Bozhenok, L., Wade, P. A., and Varga-Weisz, P.: **WSTF-ISWI chromatin remodeling complex targets heterochromatic replication foci**. EMBO Journal 2002, 21:2231-2241

Brokx, R. D., Lopez, M. M., Vogel, H. J., and Makhatadze, G. I.: **Energetics of target peptide binding by calmodulin reveals different modes of binding**. Journal of Biological Chemistry 2001, 276:14083-14091

Bruck, I. and O'Donnell, M.: **The ring-type polymerase sliding clamp family**. Genome Biol 2001, 2:REVIEWS3001

Bruning, J. B. and 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

Burgers, P. M. J., Koonin, E. V., Bruford, E., Blanco, L., Burtis, K. C., Christman, M. F., Copeland, W. C., Friedberg, E. C., Hanaoka, F., Hinkle, D. C., Lawrence, C. W., Nakanishi, M., Ohmori, H., Prakash, L., Prakash, S., Reynaud, C. A., Sugino, A., Todo, T., Wang, Z. G., Weill, J. C., and Woodgate, R.: **Eukaryotic DNA polymerases: Proposal for a revised nomenclature**. Journal of Biological Chemistry 2001, 276:43487-43490

Burgers, P. M. J. and Yoder, B. L.: **ATP-Independent Loading of the Proliferating Cell Nuclear Antigen Requires DNA Ends**. Journal of Biological Chemistry 1993, 268:19923-19926

Burnett, J. C., Ruthel, G., Stegmann, C. M., Panchal, R. G., Nguyen, T. L., Hermone, A. R., Stafford, R. G., Lane, D. J., Kenny, T. A., McGrath, C. F., Wipf, P., Stahl, A. M., Schmidt, J. J., Gussio, R., Brunger, A. T., and Bavari, S.: **Inhibition of metalloprotease botulinum serotype A from a pseudo-peptide binding mode to a small molecule that is active in primary neurons**. Journal of Biological Chemistry 2007, 282:5004-5014

Chang, A. C. Y. and Cohen, S. N.: **Construction and Characterization of Amplifiable Multicopy Dna Cloning Vehicles Derived from P15A Cryptic Miniplasmid**. Journal of Bacteriology 1978, 134:1141-1156

Chellgren, B. W. and Creamer, T. P.: **Side-chain entropy effects on protein secondary structure formation**. Proteins-Structure Function and Bioinformatics 2006, 62:411-420

Chen, J. J., Chen, S., Saha, P., and Dutta, A.: **p21(Cip1/Waf1) disrupts the recruitment of human Fen1 by proliferating-cell nuclear antigen into the DNA replication complex**. Proceedings of the National Academy of Sciences of the United States of America 1996a, 93:11597-11602

Chen, J. J., Peters, R., Saha, P., Lee, P., Theodoras, A., Pagano, M., Wagner, G., and Dutta, A.: **A 39 amino acid fragment of the cell cycle regulator p21 is sufficient to bind PCNA and partially inhibit DNA replication in vivo**. Nucleic Acids Research 1996b, 24:1727-1733

Chen, Y. H., Yang, J. T., and Chau, K. H.: **Determination of Helix and Beta-Form of Proteins in Aqueous-Solution by Circular-Dichroism**. Biochemistry 1974, 13:3350-3359

Chotiyarnwong, P., Stewart-Jones, G. B., Tarry, M. J., Dejnirattisai, W., Siebold, C., Koch, M., Stuart, D. I., Harlos, K., Malasit, P., Screaton, G., Mongkolsapaya, J., and Jones, E. Y.: **Humidity control as a strategy for lattice optimization applied to crystals of HLA-A*1101 complexed with variant peptides from dengue virus**. Acta Crystallographica Section F-Structural Biology and Crystallization Communications 2007, 63:386-392

Coen, C. J., Blanch, H. W., and Prausnitz, J. M.: **Salting-Out of Aqueous Proteins - Phase-Equilibria and Intermolecular Potentials**. Aiche Journal 1995, 41:996-1004

Collaborative Computational Project, Number: **The CCP4 Suite: Programs for protein crystallography**. Acta Crystallographica Section D Biological Crystallography 1994, 50:760-763

Corry, J. E. L.: **Relationships of Water Activity to Fungal Growth**. Beuchat, Larry R.Food and Beverage Mycology.Xi+527P.Illus.Avi Publishing Co., Inc.: Westport, Conn., Usa.Isbn 0-87055-247-3 1978,45-82

Creamer, T. P. and Rose, G. D.: **Alpha-Helix-Forming Propensities in Peptides and Proteins**. Proteins-Structure Function and Genetics 1994, 19:85-97

Cromer, D. T. and Liberman, D.: **Relativistic Calculation of Anomalous Scattering Factors for X-Rays**. Journal of Chemical Physics 1970, 53:1891-&

Dai, H., Dubin, P. L., and Andersson, T.: **Permeation of small molecules in aqueous size-exclusion chromatography vis-a-vis models for separation**. Analytical Chemistry 1998, 70:1576-1580

Dalmas, B. and Bannister, W. H.: **Prediction of Protein Secondary Structure from Circular-Dichroism Spectra - An Attempt to Solve the Problem of the Best-Fitting Reference Protein Subsets**. Analytical Biochemistry 1995, 225:39-48

Datta, S. and Grant, D. J. W.: **Crystal structures of drugs: Advances in determination, prediction and engineering**. Nature Reviews Drug Discovery 2004, 3:42-57

DeLano, W. L.: **The PyMOL Molecular Graphics System**.2002,

Delcueto, J. A. and Shevchik, N. J.: **Synchrotron-Like Intensities with A Rotating Anode and A Focusing Dispersing Crystal**. Journal of Physics E-Scientific Instruments 1978, 11:616-617

Diguan, C., Li, P., Riggs, P. D., and Inouye, H.: **Vectors That Facilitate the Expression and Purification of Foreign Peptides in Escherichia-Coli by Fusion to Maltose-Binding Protein**. Gene 1988, 67:21-30

Ellison, V. and Stillman, B.: **Biochemical characterization of DNA damage checkpoint complexes: Clamp loader and clamp complexes with specificity for 5 ' recessed DNA**. Plos Biology 2003, 1:231-243

Emsley, P. and Cowtan, K.: **Coot: model-building tools for molecular graphics**. Acta Crystallographica Section D-Biological Crystallography 2004, 60:2126-2132

Engel, M., Hindie, V., Lopez-Garcia, L. A., Stroba, A., Schaeffer, F., Adrian, I., Imig, J., Idrissova, L., Nastainczyk, W., Zeuzem, S., Alzari, P. M., Hartmann, R. W., Piiper, A., and Biondi, R. M.: **Allosteric activation of the protein kinase PDK1 with low molecular weight compounds**. EMBO Journal 2006, 25:5469-5480

Ericsson, U. B., Hallberg, B. M., DeTitta, G. T., Dekker, N., and Nordlund, P.: **Thermofluor-based high-throughput stability optimization of proteins for structural studies**. Analytical Biochemistry 2006, 357:289-298

Esteve, P. O., Chin, H. G., Smallwood, A., Feehery, G. R., Gangisetty, O., Karpf, A. R., Carey, M. F., and Pradhan, S.: **Direct interaction between DNMT1 and G9a coordinates DNA and histone methylation during replication**. Genes & Development 2006, 20:3089-3103

Freire, E., Mayorga, O. L., and Straume, M.: **Isothermal Titration Calorimetry**. Analytical Chemistry 1990, 62:A950-A959

Fujita, N. and Wade, P. A.: **Use of bifunctional cross-linking reagents in mapping genomic distribution of chromatin remodeling complexes**. Methods 2004, 33:81-85

Fukuda, K., Morioka, H., Imajou, S., Ikeda, S., Ohtsuka, E., and Tsurimoto, T.: **Structure-Function Relationship of the Eukaryotic Dna-Replication Factor, Proliferating Cell Nuclear Antigen**. Journal of Biological Chemistry 1995, 270:22527-22534

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., and Bairoch, A.: **ExPASy: the proteomics server for in-depth protein knowledge and analysis**. Nucleic Acids Research 2003, 31:3784-3788

Georgescu, R. E., Yurieva, O., Kim, S. S., Kuriyan, J., Kong, X. P., and O'Donnell, M.: **Structure of a small-molecule inhibitor of a DNA polymerase sliding clamp**. Proceedings of the National Academy of Sciences of the United States of America 2008, 105:11116-11121

Gibbs, E., Kelman, Z., Gulbis, J. M., ODonnell, M., Kuriyan, J., Burgers, P. M. J., and Hurwitz, J.: **The influence of the proliferating cell nuclear antigen-interacting domain of p21(CIP1) on DNA synthesis catalyzed by the human and Saccharomyces cerevisiae polymerase delta holoenzymes**. Journal of Biological Chemistry 1997, 272:2373-2381

Gogol, E. P., Young, M. C., Kubasek, W. L., Jarvis, T. C., and Vonhippel, P. H.: **Cryoelectron Microscopic Visualization of Functional Subassemblies of the Bacteriophage-T4 Dna-Replication Complex**. Journal of Molecular Biology 1992, 224:395-412

Goldfarb, A. R. and Saidel, L. J.: **Ultraviolet Absorption Spectra of Proteins**. Science 1951, 114:156-157

Gonzalez, Urba, Pinart, Mariona, Reveiz, Ludovic, and Alvar, Jorge: **Interventions for Old World cutaneous leishmaniasis**. Cochrane Database Syst Rev 2008,CD005067

Green, N. M., Anfinsen, C. B., John T.Edsall, and Frederic M.Richards: **Avidin**. Advances in Protein Chemistry 1975, 29:85-133

Grosse-Kunstleve, R. W. and Adams, P. D.: **Patterson correlation methods: a review of molecular replacement with CNS**. Acta Crystallographica Section D-Biological Crystallography 2001, 57:1390-1396

Grossman, T. H., Kawasaki, E. S., Punreddy, S. R., and Osburne, M. S.: **Spontaneous cAMP-dependent derepression of gene expression in stationary phase plays a role in recombinant expression instability**. Gene 1998, 209:95-103

Grover, P. K. and Ryall, R. L.: **Critical appraisal of salting-out and its implications for chemical and biological sciences**. Chemical Reviews 2005, 105:1-10

Gu, Y., Parker, A., Wilson, T. M., Bai, H., Chang, D. Y., and Lu, A. L.: **Human MutY homolog, a DNA glycosylase involved in base excision repair, physically and functionally interacts with mismatch repair proteins human MutS homolog 2/human MutS homolog 6**. Journal of Biological Chemistry 2002, 277:11135-11142

Gulbis, J. M., Kelman, Z., Hurwitz, J., ODonnell, M., and Kuriyan, J.: **Structure of the C-terminal region of p21(WAF1/CIP1) complexed with human PCNA**. Cell 1996, 87:297-306

Harding, S. E.: **On the Hydrodynamic Analysis of Macromolecular Conformation**. Biophysical Chemistry 1995, 55:69-93

Harker, D.: **The Determination of the Phases of the Structure Factors of Non-Centrosymmetric Crystals by the Method of Double Isomorphous Replacement**. Acta Crystallographica 1956, 9:1-9

Henderson, D. S., Wiegand, U. K., Norman, D. G., and Glover, D. M.: **Mutual correction of faulty PCNA subunits in temperature-sensitive lethal mus209 mutants of Drosophila melanogaster**. Genetics 2000, 154:1721-1733

Hennessey, J. P. and Johnson, W. C.: **Information-Content in the Circular-Dichroism of Proteins**. Biochemistry 1981, 20:1085-1094

Heras, B. and Martin, J. L.: **Post-crystallization treatments for improving diffraction quality of protein crystals**. Acta Crystallographica Section D-Biological Crystallography 2005, 61:1173-1180

Heras, Bego a, Edeling, Melissa A., Byriel, Karl A., Jones, Alun, Raina, Satish, and Martin, Jennifer L.: **Dehydration Converts DsbG Crystal Diffraction from Low to High Resolution**. Structure 2003, 11:139-145

Hermann, A., Gowher, H., and Jeltsch, A.: **Biochemistry and biology of mammalian DNA methyltransferases**. Cellular and Molecular Life Sciences 2004, 61:2571-2587

Hochuli, E., Dobeli, H., and Schacher, A.: **New Metal Chelate Adsorbent Selective for Proteins and Peptides Containing Neighboring Histidine-Residues**. Journal of Chromatography 1987, 411:177-184

Huang, L., Kim, Y., Turchi, J. J., and Bambara, R. A.: **Structure-Specific Cleavage of the Rna Primer from Okazaki Fragments by Calf Thymus Rnase Hi**. Journal of Biological Chemistry 1994, 269:25922-25927

Huang, Shir Ly, Lin, Fu Yung, and Yang, Chih Ping: **Microcalorimetric studies of the effects on the interactions of human recombinant interferon-[alpha]2a**. European Journal of Pharmaceutical Sciences 2005, 24:545-552

Iida, T., Suetake, I., Tajima, S., Morioka, H., Ohta, S., Obuse, C., and Tsurimoto, T.: **PCNA clamp facilitates action of DNA cytosine methyltransferase 1 on hemimethylated DNA**. Genes to Cells 2002, 7:997-1007

Ikemura, T.: **Correlation Between the Abundance of Escherichia-Coli Transfer-Rnas and the Occurrence of the Respective Codons in Its Protein Genes**. Journal of Molecular Biology 1981, 146:1-21

Iwai, T., Kurosawa, N., Itoh, Y. H., and Horiuchi, T.: **Phylogenetic analysis of archaeal PCNA homologues**. Extremophiles 2000, 4:357-364

Izasrd, Tina, Sarfaty, Steve, Westphal, Adrie, De Kok, Arie, and Hol, Wim G.: **Improvement of diffraction quality upon rehydration of dehydrated icosahedral Enterococcus faecalis pyruvate dehydrogenase core crystals**. Protein Science 1997, 6:913-915

Jachimska, B., Wasilewska, M., and Adamczyk, Z.: **Characterization of globular protein solutions by dynamic light scattering, electrophoretic mobility, and viscosity measurements**. Langmuir 2008, 24:6866-6872

Jelesarov, I. and Bosshard, H. R.: **Isothermal titration calorimetry and differential scanning calorimetry as complementary tools to investigate the energetics of biomolecular recognition**. Journal of Molecular Recognition 1999, 12:3-18

Jin, R. S., Rummel, A., Binz, T., and Brunger, A. T.: **Botulinum neurotoxin B recognizes its protein receptor with high affinity and specificity**. Nature 2006, 444:1092-1095

Jiricny, J.: **The multifaceted mismatch-repair system**. Nature Reviews Molecular Cell Biology 2006, 7:335-346

Johnson, W. C.: **Analyzing protein circular dichroism spectra for accurate secondary structures**. Proteins-Structure Function and Genetics 1999, 35:307-312

Jones, S. and Thornton, J. M.: **Principles of protein-protein interactions**. Proceedings of the National Academy of Sciences of the United States of America 1996, 93:13-20

Kantardjieff, K. A. and Rupp, B.: **Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals**. Protein Science 2003, 12:1865-1871

Karlsson, R.: **Affinity analysis of non-steady-state data obtained under mass transport limited conditions using BIAcore technology**. Journal of Molecular Recognition 1999, 12:285-292

Kelly, S. M., Jess, T. J., and Price, N. C.: **How to study proteins by circular dichroism**. Biochimica et Biophysica Acta-Proteins and Proteomics 2005, 1751:119-139

Kelman, Z.: **PCNA: Structure, functions and interactions**. Oncogene 1997, 14:629-640

Kelman, Z. and ODonnell, M.: **Dna-Polymerase-Iii Holoenzyme - Structure and Function of A Chromosomal Replicating Machine**. Annual Review of Biochemistry 1995, 64:171-200

Knibiehler, M., Goubin, F., Escalas, N., Jonsson, Z. O., Mazarguil, H., Hubscher, U., and Ducommun, B.: **Interaction studies between the p21Cip1/Waf1 cyclin-dependent kinase inhibitor and proliferating cell nuclear antigen (PCNA) by surface plasmon resonance**. FEBS Letters 1996, 391:66-70

Knipscheer, P., van Dijk, W. J., Olsen, J. V., Mann, M., and Sixma, T. K.: **Noncovalent interaction between Ubc9 and SUMO promotes SUMO chain formation**. EMBO Journal 2007, 26:2797-2807

Kong, X. P., Onrust, R., O'Donnell, M., and Kuriyan, J.: **Three-dimensional structure of the beta subunit of E. coli DNA polymerase III holoenzyme: A sliding DNA clamp**. Cell 1992, 69:425-437

Kontopidis, G., Wu, S. Y., Zheleva, D. I., Taylor, P., McInnes, C., Lane, D. P., Fischer, P. M., and Walkinshaw, M. D.: **Structural and biochemical studies of human proliferating cell nuclear antigen complexes provide a rationale for cyclin association and inhibitor design**. Proceedings of the National Academy of Sciences of the United States of America 2005, 102:1871-1876

Kostanski, L. K., Keller, D. M., and Hamielec, A. E.: **Size-exclusion chromatography - a review of calibration methodologies**. Journal of Biochemical and Biophysical Methods 2004, 58:159-186

Krishna, T. S. R., Kong, X. P., Gary, S., Burgers, P. M., and Kuriyan, J.: **Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA**. Cell 1994, 79:1233-1243

Krissinel, E. and Henrick, K.: **Inference of macromolecular assemblies from crystalline state**. Journal of Molecular Biology 2007b, 372:774-797

Krissinel, E. and Henrick, K.: **Protein interfaces, surfaces and assemblies service PISA at European Bioinformatics Institute version 1.13 accessed 08/08/07**.2007a,

Kronig, R. D. L.: **On the theory of dispersion of x-rays**. Journal of the Optical Society of America and Review of Scientific Instruments 1926, 12:547-557

Laskay, T., van Zandbergen, G., and Solbach, W.: **Neutrophil granulocytes - Trojan horses for Leishmania major and other intracellular microbes?** Trends in Microbiology 2003, 11:210-214

Laskowski, R. A., Macarthur, M. W., Moss, D. S., and Thornton, J. M.: **Procheck - A Program to Check the Stereochemical Quality of Protein Structures**. Journal of Applied Crystallography 1993, 26:283-291

Lee, J. H. and Paull, T. T.: **ATM activation by DNA double-strand breaks through the Mre11-Rad50-Nbs1 complex**. Science 2005, 308:551-554

Lee, W. C. and Lee, K. H.: **Applications of affinity chromatography in proteomics**. Analytical Biochemistry 2004, 324:1-10

Leslie, A. G. W.: **Integration of macromolecular diffraction data**. Acta Crystallographica Section D-Biological Crystallography 1999, 55:1696-1702

Lewis, P. N., Go, N., Go, M., Kotelchu, D., and Scheraga, H. A.: **Helix Probability Profiles of Denatured Proteins and Their Correlation with Native Structures**. Proceedings of the National Academy of Sciences of the United States of America 1970, 65:810-&

Li, H., Koshiba, S., Hayashi, F., Tochio, N., Tomizawa, T., Kasai, T., Yabuki, T., Motoda, Y., Harada, T., Watanabe, S., Inoue, M., Hayashizaki, Y., Tanaka, A., Kigawa, T., and Yokoyama, S.: **Structure of the C-terminal phosphotyrosine interaction domain of Fe65L1 complexed with the cytoplasmic tail of amyloid precursor protein reveals a novel peptide binding mode**. Journal of Biological Chemistry 2008, 283:27165-27178

Liang, Y.: **Applications of isothermal titration calorimetry in protein science**. Acta Biochimica et Biophysica Sinica 2008, 40:565-576

Liang, Y., Li, J., Chen, J., and Wang, C. C.: **Thermodynamics of the folding of D-glyceraidehyde-3-phosphate dehydrogenase assisted by protein disulfide isomerase studied by microcalorimetry**. European Journal of Biochemistry 2001, 268:4183-4189

Lim, N. S., Kozlov, G., Chang, T. C., Groover, O., Siddiqui, N., Volpon, L., De Crescenzo, G., Shyu, A. B., and Gehring, K.: **Comparative peptide binding studies of the PABC domains from the ubiquitin-protein isopeptide ligase HYD and poly(A)-binding protein - Implications for HYD function**. Journal of Biological Chemistry 2006, 281:14376-14382

Liu, Y. Y., Kvaratskhelia, M., Hess, S., Qu, Y. X., and Zou, Y.: **Modulation of replication protein A function by its hyperphosphorylation-induced conformational change involving DNA binding domain B**. Journal of Biological Chemistry 2005, 280:32775-32783

Lo, M. C., Aulabaugh, A., Jin, G. X., Cowling, R., Bard, J., Malamas, M., and Ellestad, G.: **Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery**. Analytical Biochemistry 2004, 332:153-159

Loladze, V. V., Ermolenko, D. N., and Makhatadze, G. I.: **Heat capacity changes upon burial of polar and nonpolar groups in proteins**. Protein Science 2001, 10:1343-1352

Lopez, P. J., Marchand, I., Joyce, S. A., and Dreyfus, M.: **The C-terminal half of RNase E, which organizes the Escherichia coli degradosome, participates in mRNA degradation but not rRNA processing in vivo**. Molecular Microbiology 1999, 33:188-199

Luo, R. G. and Hsu, J. T.: **Optimization of gradient profiles in ion-exchange chromatography for protein purification**. Industrial & Engineering Chemistry Research 1997, 36:444-450

Machold, C., Deinhofer, K., Hahn, R., and Jungbauer, A.: **Hydrophobic interaction chromatography of proteins - I. Comparison of selectivity**. Journal of Chromatography A 2002, 972:3-19

Maga, G. and Hubscher, U.: **Proliferating cell nuclear antigen (PCNA): a dancer with many partners**. Journal of Cell Science 2003, 116:3051-3060

Matsuoka, S., Ballif, B. A., Smogorzewska, A., McDonald, E. R., Hurov, K. E., Luo, J., Bakalarski, C. E., Zhao, Z. M., Solimini, N., Lerenthal, Y., Shiloh, Y., Gygi, S. P., and Elledge, S. J.: **ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage**. Science 2007, 316:1160-1166

Matsuoka, S., Yamaguchi, M., and Matsukage, A.: **D-Type Cyclin-Binding Regions of Proliferating Cell Nuclear Antigen**. Journal of Biological Chemistry 1994, 269:11030-11036

Matthews, B. W.: **X-Ray Crystallographic Studies of Proteins**. Annual Review of Physical Chemistry 1976, 27:493-523

Matthews, B. W.: **Solvent Content of Protein Crystals**. Journal of Molecular Biology 1968, 33:491-&

Mattock, H., Lane, D. P., and Warbrick, E.: **Inhibition of cell proliferation by the PCNA-binding region of p21 expressed as a GFP miniprotein**. Experimental Cell Research 2001, 265:234-241

Mcpherson, A.: **Crystallization of Macromolecules - General-Principles**. Methods in Enzymology 1985, 114:112-120

Mechanic, L. E., Frankel, B. A., and Matson, S. W.: **Escherichia coli MutL loads DNA helicase II onto DNA**. Journal of Biological Chemistry 2000, 275:38337-38346

Melcher, K.: **New chemical Crosslinking methods for the identification of transient protein-protein interactions with multiprotein complexes**. Current Protein & Peptide Science 2004, 5:287-296

Messing, J., Gronenborn, B., Mullerhill, B., and Hofschneider, P. H.: **Filamentous Coliphage M13 As A Cloning Vehicle - Insertion of A Hindii Fragment of Lac Regulatory Region in M13 Replicative Form Invitro - (Single-Stranded-Dna Phage-Blunt End Ligation Lactose Operon Alpha-Complementation)**. Proceedings of the National Academy of Sciences of the United States of America 1977, 74:3642-3646

MicroCal, LLC: **VP-ITC MicroCalorimeter User's Manual**.20-11-2002,

Miller, P., Kendall, F., and Nicolini, C.: **Thermal-Denaturation of Sheared and Unsheared Chromatin by Absorption and Circular-Dichroism Measurements**. Nucleic Acids Research 1976, 3:1875-1881

Moffatt, B. A. and Studier, F. W.: **T7 Lysozyme Inhibits Transcription by T7 Rna-Polymerase**. Cell 1987, 49:221-227

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

Moore, J. K. and Haber, J. E.: **Cell cycle and genetic requirements of two pathways of nonhomologous end-Joining repair of double-strand breaks in Saccharomyces cerevisiae**. Molecular and Cellular Biology 1996, 16:2164-2173

Moreira, R. F. and Noren, C. J.: **Minimum Duplex Requirements for Restriction Enzyme Cleavage Near the Termini of Linear Dna Fragments**. Biotechniques 1995, 19:56-59

Morris, A. L., Macarthur, M. W., Hutchinson, E. G., and Thornton, J. M.: **Stereochemical Quality of Protein-Structure Coordinates**. Proteins-Structure Function and Genetics 1992, 12:345-364

Morton, C. J. and Ladbury, J. E.: **Water mediated protein-DNA interactions: The relationship of thermodynamics to structural detail**. Protein Science 1996, 5:2115-2118

Mortusewicz, O., Schermelleh, L., Walter, J., Cardoso, M. C., and Leonhardt, H.: **Recruitment of DNA methyltransferase I to DNA repair sites**. Proceedings of the National Academy of Sciences of the United States of America 2005, 102:8905-8909

Mottram, J. C., Kinnaird, J. H., Shiels, B. R., Tait, A., and Barry, J. D.: **A Novel Cdc2-Related Protein-Kinase from Leishmania-Mexicana, Lmmcrk1, Is Posttranslationally Regulated During the Life-Cycle**. Journal of Biological Chemistry 1993, 268:21044-21052

Murphy, K. P., Privalov, P. L., and Gill, S. J.: **Common Features of Protein Unfolding and Dissolution of Hydrophobic Compounds**. Science 1990, 247:559-561

Murshudov, G. N., Vagin, A. A., and Dodson, E. J.: **Refinement of macromolecular structures by the maximum-likelihood method**. Acta Crystallographica Section D-Biological Crystallography 1997, 53:240-255

Myers, E. W. and Miller, W.: **Optimal Alignments in Linear-Space**. Computer Applications in the Biosciences 1988, 4:11-17

Naktinis, V., Turner, J., and ODonnell, M.: **Molecular switch in a replication machine defined by an internal competition for protein rings**. Cell 1996, 84:137-145

Naryzhny, S. N., Zhao, H., and Lee, H.: **Proliferating cell nuclear antigen (PCNA) may function as a double homotrimer complex in the mammalian cell**. Journal of Biological Chemistry 2005, 280:13888-13894

Nieba, L., NiebaAxmann, S. E., Persson, A., Hamalainen, M., Edebratt, F., Hansson, A., Lidholm, J., Magnusson, K., Karlsson, A. F., and Pluckthun, A.: **BIACORE analysis of histidine-tagged proteins using a chelating NTA sensor chip**. Analytical Biochemistry 1997, 252:217-228

Niesen, F. H., Berglund, H., and Vedadi, M.: **The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability**. Nature Protocols 2007, 2:2212-2221

Nishimiya, Y., Tsumoto, K., Shiroishi, M., Yutani, K., and Kumagai, I.: **Thermodynamic consequences of grafting enhanced affinity toward the mutated antigen onto an antibody - The case of anti-lysozyme antibody, HyHEL-10**. Journal of Biological Chemistry 2000, 275:12813-12820

Norrander, J., Kempe, T., and Messing, J.: **Construction of Improved M13-Vectors Using Oligodeoxynucleotide-Directed Mutagenesis**. Gene 1983, 26:101-106

Ohmori, H., Friedberg, E. C., Fuchs, R. P. P., Goodman, M. F., Hanaoka, F., Hinkle, D., Kunkel, T. A., Lawrence, C. W., Livneh, Z., Nohmi, T., Prakash, L., Prakash, S., Todo, T., Walker, G. C., Wang, Z. G., and Woodgate, R.: **The Y-family of DNA polymerases**. Molecular Cell 2001, 8:7-8

Okhrimenko, O. and Jelesarov, M.: **A survey of the year 2006 literature on applications of isothermal titration calorimetry**. Journal of Molecular Recognition 2008, 21:1-19

Oku, T., Ikeda, S., Sasaki, H., Fukuda, K., Morioka, H., Ohtsuka, E., Yoshikawa, H., and Tsurimoto, T.: **Functional sites of human PCNA which interact with p21 (Cip1/Waf1), DNA polymerase delta and replication factor C**. Genes to Cells 1998, 3:357-369

Olivier, M., Badaro, R., Medrano, F., and Moreno, J.: **The pathogenesis of Leishmania/HIV co-infection: Cellular and immunological mechanisms**. Annals of Tropical Medicine & Parasitology 2003, 97:S79-S98

Oneil, K. T. and Degrado, W. F.: **A Thermodynamic Scale for the Helix-Forming Tendencies of the Commonly Occurring Amino-Acids**. Science 1990, 250:646-651

Orlando, V.: **Mapping chromosomal proteins in vivo by formaldehyde-crosslinked-chromatin immunoprecipitation**. Trends in Biochemical Sciences 2000, 25:99-104

Oshannessy, D. J., Brighamburke, M., Soneson, K. K., Hensley, P., and Brooks, I.: **Determination of Rate and Equilibrium Binding Constants for Macromolecular Interactions Using Surface-Plasmon Resonance - Use of Nonlinear Least-Squares Analysis-Methods**. Analytical Biochemistry 1993, 212:457-468

Padmanabhan, S., Marqusee, S., Ridgeway, T., Laue, T. M., and Baldwin, R. L.: **Relative Helix-Forming Tendencies of Nonpolar Amino-Acids**. Nature 1990, 344:268-270

Pascal, J. M., O'Brien, P. J., Tomkinson, A. E., and Ellenberger, T.: **Human DNA ligase I completely encircles and partially unwinds nicked DNA**. Nature 2004, 432:473-478

Paunesku, T., Mittal, S., Protic?, M., Oryhon, J., Korolev, S. V., Joachimiak, A., and Woloschak, G. E.: **Proliferating cell nuclear antigen (PCNA): Ringmaster of the genome**. International Journal of Radiation Biology 2001, 77:1007-1021

Peterson, E. A. and Sober, H. A.: **Chromatography of Proteins .1. Cellulose Ion-Exchange Adsorbents**. Journal of the American Chemical Society 1956, 78:751-755

Polo, S. E. and Almouzni, G.: **Chromatin assembly: a basic recipe with various flavours**. Current Opinion in Genetics & Development 2006, 16:104-111

Poot, R. A., Bozhenok, L., van den Berg, D. L. C., Steffensen, S., Ferreira, F., Grimaldi, M., Gilbert, N., Ferreira, J., and Varga-Weisz, P. D.: **The Williams syndrome transcription factor interacts with PCNA to target chromatin remodelling by ISWI to replication foci**. Nature Cell Biology 2004, 6:1236-1244

Porath, J. and Flodin, P.: **Gel Filtration - Method for Desalting and Group Separation**. Nature 1959, 183:1657-1659

Potzschke, H., Barnikol, W. K. R., Domack, U., and Kirste, R. G.: **A novel procedure for the determination of molar masses of polymers with broad molar mass distribution by means of gel-permeation chromatography and viscosimetry for haemoglobin-hyperpolymers**. Macromolecular Chemistry and Physics 1996, 197:3229-3250

Powell, H. R.: **The Rossmann Fourier autoindexing algorithm in MOSFLM**. Acta Crystallographica Section D-Biological Crystallography 1999, 55:1690-1695

Prakash, S., Johnson, R. E., and Prakash, L.: **Eukaryotic translesion synthesis DNA polymerases: Specificity of structure and function**. Annual Review of Biochemistry 2005, 74:317-353

Privalov, P. L. and Gill, S. J.: **Stability of Protein-Structure and Hydrophobic Interaction**. Advances in Protein Chemistry 1988, 39:191-234

Privalov, Peter L. and Potekhin, Sergey A.: **Scanning microcalorimetry in studying temperature-induced changes in proteins**.1986, Volume 131:4-51

Radhakrishnan, S. K., Gierut, J., and Gartel, A. L.: **Multiple alternate p21 transcripts are regulated by p53 in human cells**. Oncogene 2006, 25:1812-1815

Rainaldi, M., Yamniuk, A. P., Murase, T., and Vogel, H. J.: **Calcium-dependent and -independent binding of soybean calmodulin isoforms to the calmodulin binding domain of tobacco MAPK phosphatase-1**. Journal of Biological Chemistry 2007, 282:6031-6042

Remaut, E., Tsao, H., and Fiers, W.: **Improved Plasmid Vectors with A Thermoinducible Expression and Temperature-Regulated Runaway Replication**. Gene 1983, 22:103-113

Richardson, J. M., Lopez, M. M., and Makhatadze, G. I.: **Enthalpy of helix-coil transition: Missing link in rationalizing the thermodynamics of helix-forming propensities of the amino acid residues**. Proceedings of the National Academy of Sciences of the United States of America 2005, 102:1413-1418

Riekel, C., Burghammer, M., and Schertler, G.: **Protein crystallography microdiffraction**. Current Opinion in Structural Biology 2005, 15:556-562

Rossmann, M. G. and Blow, D. M.: **Detection of Sub-Units Within Crystallographic Asymmetric Unit**. Acta Crystallographica 1962, 15:24-&

Rost, B., Yachdav, G., and Liu, J. F.: **The PredictProtein server**. Nucleic Acids Research 2004, 32:W321-W326

Sadqi, M., Hernandez, F., Pan, U. M., Perez, M., Schaeberle, M. D., Avila, J., and Munoz, V.: **alpha-helix structure in Alzheimer's disease aggregates of tau-protein**. Biochemistry 2002, 41:7150-7155

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., and Erlich, H. A.: **Primer-Directed Enzymatic Amplification of Dna with A Thermostable Dna-Polymerase**. Science 1988, 239:487-491

Sakakura, C., Hagiwara, A., Tsujimoto, H., Ozaki, K., Sakakibara, T., Oyama, T., Ogaki, M., and Takahashi, T.: **Inhibition of Gastric-Cancer Cell-Proliferation by Antisense Oligonucleotides Targeting the Messenger-RNA Encoding Proliferating Cell Nuclear Antigen**. British Journal of Cancer 1994, 70:1060-1066

Sali, A., Glaeser, R., Earnest, T., and Baumeister, W.: **From words to literature in structural proteomics**. Nature 2003, 422:216-225

Sambrook, J., Fritsch, E. F., and Maniatis, T.: **Molecular Cloning A Laboratory Manual Second Edition Vols. 1 2 and 3**. Sambrook, J., E.F.Fritsch and T.Maniatis.Molecular Cloning: A Laboratory Manual, Second Edition, Vols.1, 2 and 3.Xxxix+Pagination Varies(Vol.1); Xxxiii+Pagination Varies(Vol.2): Xxxii+Pagination Varies(Vol.3) Cold Spring Harbor Laboratory Press: 1989,

Sancar, A., Lindsey-Boltz, L. A., Unsal-Kacmaz, K., and Linn, S.: **Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints**. Annual Review of Biochemistry 2004, 73:39-85

Sarraf, S. A. and Stancheva, I.: **Methyl-CpG binding protein MBD1 couples histone H3 methylation at lysine 9 by SETDB1 to DNA replication and chromatin assembly**. Molecular Cell 2004, 15:595-605

Schein, C. H.: **Production of Soluble Recombinant Proteins in Bacteria**. Bio-Technology 1989, 7:1141-1147

Shamoo, Y. and Steitz, T. A.: **Building a replisome from interacting pieces: Sliding clamp complexed to a peptide from DNA polymerase and a polymerase editing complex**. Cell 1999, 99:155-166

Shiroishi, M., Yokota, A., Tsumoto, K., Kondo, H., Nishimiya, Y., Horii, K., Matsushima, M., Ogasahara, K., Yutani, K., and Kumagai, I.: **Structural evidence for entropic contribution of salt bridge formation to a protein antigen-antibody interaction - The case of hen lysozyme-HyHEL-10 Fv complex**. Journal of Biological Chemistry 2001, 276:23042-23050

Shore, D., Langowski, J., and Baldwin, R. L.: **Dna Flexibility Studied by Covalent Closure of Short Fragments Into Circles**. Proceedings of the National Academy of Sciences of the United States of America 1981, 78:4833-4837

Siligardi, G., Hu, B., Panaretou, B., Piper, P. W., Pearl, L. H., and Prodromou, C.: **Co-chaperone regulation of conformational switching in the Hsp90 ATPase cycle**. Journal of Biological Chemistry 2004, 279:51989-51998

Sleigh, S. H., Seavers, P. R., Wilkinson, A. J., Ladbury, J. E., and Tame, J. R. H.: **Crystallographic and calorimetric analysis of peptide binding to OppA protein**. Journal of Molecular Biology 1999, 291:393-415

Smith, D. B. and Johnson, K. S.: **Single-Step Purification of Polypeptides Expressed in Escherichia-Coli As Fusions with Glutathione S-Transferase**. Gene 1988, 67:31-40

Smith, S. and Stillman, B.: **Purification and Characterization of Caf-I, A Human Cell Factor Required for Chromatin Assembly During Dna-Replication In-vitro**. Cell 1989, 58:15-25

Smyth, M. and Martin, J.: **X-Ray crystallography**. Molecular Pathology 2000, 53:8-14

Soares, A. S., Caspar, D. L. D., Weckert, E., Heroux, A., Holzer, K., Schroer, K., Zellner, J., Schneider, D., Nolan, W., and Sweet, R. M.: **Three-beam interference is a sensitive measure of the efficacy of macromolecular refinement techniques**. Acta Crystallographica Section D-Biological Crystallography 2003, 59:1716-1724

Sorensen, T. L. M., McAuley, K. E., Flaig, R., and Duke, E. M. H.: **New light for science: synchrotron radiation in structural medicine**. Trends in Biotechnology 2006, 24:500-508

Souza, D. H. F., Selistre-de-Araujo, H. S., and Garratt, R. C.: **Determination of the three-dimensional structure of toxins by protein crystallography**. Toxicon 2000, 38:1307-1353

Spolar, R. S. and Record, M. T.: **Coupling of Local Folding to Site-Specific Binding of Proteins to Dna**. Science 1994, 263:777-784

Sreerama, N., Venyaminov, S. Y., and Woody, R. W.: **Estimation of the number of alpha-helical and beta-strand segments in proteins using circular dichroism spectroscopy**. Protein Science 1999, 8:370-380

Sreerama, N. and Woody, R. W.: **A Self-Consistent Method for the Analysis of Protein Secondary Structure from Circular-Dichroism**. Analytical Biochemistry 1993, 209:32-44

Stryer, L.: **Biochemistry Fourth Edition**. Stryer, L.Biochemistry, Third Edition.Xxxiii+1089P.W.H.Freeman and Co.: New York, New York, Usa.Illus 1995,802-817

Studier, F. W. and Moffatt, B. A.: **Use of Bacteriophage-T7 Rna-Polymerase to Direct Selective High-Level Expression of Cloned Genes**. Journal of Molecular Biology 1986, 189:113-130

Stukenberg, P. T., Studwellvaughan, P. S., and ODonnell, M.: **Mechanism of the Sliding Beta-Clamp of Dna Polymerase-Iii Holoenzyme**. Journal of Biological Chemistry 1991, 266:11328-11334

Tellinghuisen, J.: **Calibration in isothermal titration calorimetry: Heat and cell volume from heat of dilution of NaCl(aq)**. Analytical Biochemistry 2007, 360:47-55

Towbin, H., Staehelin, T., and Gordon, J.: **Electrophoretic Transfer of Proteins from Polyacrylamide Gels to Nitrocellulose Sheets - Procedure and Some Applications**. Proceedings of the National Academy of Sciences of the United States of America 1979, 76:4350-4354

Tsurimoto, T.: **PCNA, a multifunctional ring on DNA**. Biochimica et Biophysica Acta-Gene Structure and Expression 1998, 1443:23-39

Tyler, J. K., Adams, C. R., Chen, S. R., Kobayashi, R., Kamakaka, R. T., and Kadonaga, J. T.: **The RCAF complex mediates chromatin assembly during DNA replication and repair**. Nature 1999, 402:555-560

Tyler, J. K., Collins, K. A., Prasad-Sinha, J., Amiott, E., Bulger, M., Harte, P. J., Kobayashi, R., and Kadonaga, J. T.: **Interaction between the Drosophila CAF-1 and ASF1 chromatin assembly factors**. Molecular and Cellular Biology 2001, 21:6574-6584

Vagin, A. and Teplyakov, A.: **MOLREP: an automated program for molecular replacement**. Journal of Applied Crystallography 1997, 30:1022-1025

Vailaya, A. and Horvath, C.: **Retention thermodynamics in hydrophobic interaction chromatography**. Industrial & Engineering Chemistry Research 1996, 35:2964-2981

Vanstokkum, I. H. M., Spoelder, H. J. W., Bloemendal, M., Vangrondelle, R., and Groen, F. C. A.: **Estimation of Protein Secondary Structure and Error Analysis from Circular-Dichroism Spectra**. Analytical Biochemistry 1990, 191:110-118

Vedadi, M., Niesen, F. H., lali-Hassani, A., Fedorov, O. Y., Finerty, P. J., Wasney, G. A., Yeung, R., Arrowsmith, C., Ball, L. J., Berglund, H., Hui, R., Marsden, B. D., Nordlund, P., Sundstrom, M., Weigelt, J., and Edwards, A. M.: **Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination**. Proceedings of the National Academy of Sciences of the United States of America 2006, 103:15835-15840

Vieira, J. and Messing, J.: **The Puc Plasmids, An M13Mp7-Derived System for Insertion Mutagenesis and Sequencing with Synthetic Universal Primers**. Gene 1982, 19:259-268

Vijayakumar, S., Chapados, B. R., Schmidt, K. H., Kolodner, R. D., Tainer, J. A., and Tomkinson, A. E.: **The C-terminal domain of yeast PCNA is required for physical and functional interactions with Cdc9 DNA ligase**. Nucleic Acids Research 2007, 35:1624-1637

von Bergen, M., Friedhoff, P., Biernat, J., Heberle, J., Mandelkow, E. M., and Mandelkow, E.: **Assembly of tau protein into Alzheimer paired helical filaments depends on a local sequence motif ((306)VQIVYK(311)) forming beta structure**. Proceedings of the National Academy of Sciences of the United States of America 2000, 97:5129-5134

Waga, S. and Stillman, B.: **The DNA replication fork in eukaryotic cells**. Annual Review of Biochemistry 1998, 67:721-751

Warbrick, E.: **The puzzle of PCNA's many partners**. Bioessays 2000, 22:997-1006

Warbrick, E.: **A functional analysis of PCNA-binding peptides derived from protein sequence, interaction screening and rational design**. Oncogene 2006, 25:2850-2859

Warbrick, E., Heatherington, W., Lane, D. P., and Glover, D. N.: **PCNA binding proteins in Drosophila melanogaster: the analysis of a conserved PCNA binding domain**. Nucleic Acids Research 1998, 26:3925-3932

Warbrick, E., Lane, D. P., Glover, D. M., and Cox, L. S.: **A small peptide inhibitor of DNA replication defines the site of interaction between the cyclin-dependent kinase inhibitor p21WAF1 and proliferating cell nuclear antigen**. Current Biology 1995, 5:275-282

Wear, M. A., Patterson, A., Malone, K., Dunsmore, C., Turner, N. J., and Walkinshaw, M. D.: **A surface plasmon resonance-based assay for small molecule inhibitors of human cyclophilin A**. Analytical Biochemistry 2005, 345:214-226

Weckert, E. and Hummer, K.: **On the Quantitative-Determination of Triplet Phases by X-Ray 3-Beam Diffraction**. Acta Crystallographica Section A 1990, 46:387-393

Weik, M., Ravelli, R. B. G., Silman, I., Sussman, J. L., Gros, P., and Kroon, J.: **Specific protein dynamics near the solvent glass transition assayed by radiation-induced structural changes**. Protein Science 2001, 10:1953-1961

Weiss, M. S. and Hilgenfeld, R.: **On the use of the merging R factor as a quality indicator for X-ray data**. Journal of Applied Crystallography 1997, 30:203-205

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

Wiencek, J. M.: **New strategies for protein crystal growth**. Annual Review of Biomedical Engineering 1999, 1:505-534

Wilson, T. E., Grawunder, U., and Lieber, M. R.: **Yeast DNA ligase IV mediates non-homologous DNA end joining**. Nature 1997, 388:495-498

Wong, M. L. and Medrano, J. F.: **Real-time PCR for mRNA quantitation**. Biotechniques 2005, 39:75-85

Wood, R. D., Mitchell, M., Sgouros, J., and Lindahl, T.: **Human DNA repair genes**. Science 2001, 291:1284-1289

World Health Organisation (WHO): Official Website 1-7-2009,

Xu, J. and Morris, G. F.: **p53-mediated regulation of proliferating cell nuclear antigen expression in cells exposed to ionizing radiation**. Molecular and Cellular Biology 1999, 19:12-20

Yamniuk, A. P. and Vogel, H. J.: **Structurally homologous binding of plant calmodulin Isoforms to the calmodulin-binding domain of vacuolar calcium-ATPase**. Journal of Biological Chemistry 2004, 279:7698-7707

Yang, F., Zhou, B. R., Zhang, P., Zhao, Y. F., Chen, J., and Liang, Y.: **Binding of ferulic acid to cytochrome c enhances stability of the protein at physiological pH and inhibits cytochrome c-induced apoptosis**. Chemico-Biological Interactions 2007, 170:231-243

Yao, N., Turner, J., Kelman, Z., Stukenberg, P. T., Dean, F., Shechter, D., Pan, Z. Q., Hurwitz, J., and ODonnell, M.: **Clamp loading, unloading and intrinsic stability of the PCNA, beta and gp45 sliding clamps of human, E-coli and T4 replicases**. Genes to Cells 1996, 1:101-113

Yokota, A., Tsumoto, K., Shiroishi, M., Kondo, H., and Kumagai, I.: **The role of hydrogen bonding via interfacial water molecules in antigen-antibody complexation - The HyHEL-10-HEL interaction**. Journal of Biological Chemistry 2003, 278:5410-5418

Yoo, Seung Hyun and Lewis, Marc S.: **Thermodynamic Study of the pH-Dependent Interaction of Chromogranin A with an Intraluminal Loop Peptide of the Inositol 1,4,5-Trisphosphate Receptor**. Biochemistry 1-5-2002, 34:632-638

Yoon, J. H., Iwai, S., O'Connor, T. R., and Pfeifer, G. P.: **Human thymine DNA glycosylase (TDG) and methyl-CpG-binding protein 4 (MBD4) excise thymine glycol (Tg) from a Tg : G mispair**. Nucleic Acids Research 2003, 31:5399-5404

Young, M. E., Carroad, P. A., and Bell, R. L.: **Estimation of Diffusion-Coefficients of Proteins**. Biotechnology and Bioengineering 1980, 22:947-955

Zeng, K., Rose, J. P., Chen, H. C., Strickland, C. L., Tu, C. P. D., and Wang, B. C.: **A Surface Mutant (G82R) of A Human Alpha-Glutathione S-Transferase Shows Decreased Thermal-Stability and A New Mode of Molecular**

Association in the Crystal. Proteins-Structure Function and Genetics 1994, 20:259-263

Zheleva, D. I., Zhelev, N. Z., Fischer, P. M., Duff, S. V., Warbrick, E., Blake, D. G., and 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

Zhou, Y. L., Liao, J. M., Du, F., and Liang, Y.: **Thermodynamics of the interaction of xanthine oxidase with superoxide dismutase studied by isothermal titration calorimetry and fluorescence spectroscopy**. Thermochimica Acta 2005, 426:173-178

Zimm, B. H.: **The Scattering of Light and the Radial Distribution Function of High Polymer Solutions**. Journal of Chemical Physics 1948, 16:1093-1099

Zylicz, M., Ang, D., Liberek, K., and Georgopoulos, C.: **Initiation of Lambda-Dna Replication with Purified Host-Encoded and Bacteriophage-Encoded Proteins - the Role of the Dnak, Dnaj and Grpe Heat-Shock Proteins**. EMBO Journal 1989, 8:1601-1608

1 Appendix

1.1 DNA and protein sequence details of PCNA

1.2 DNA and protein sequence details of HsPCNA

DNA sequence of HsPCNA

ATGTTCGAGGCGCGCCTGGTCCAGGGCTCCATCCTCAAGAAGGTGTTGGAGGCACTCAAGGACCTCAT
CAACGAGGCCTGCTGGGATATTAGCTCCAGCGGTGTAAACCTGCAGAGCATGGACTCGTCCCACGTCT
CTTTGGTGCAGCTCACCCTGCGGTCTGAGGGCTTCGACACCTACCGCTGCGACCGCAACCTGGCCATG
GGCGTGAACCTCACCAGTATGTCCAAAATACTAAAATGCGCCGGCAATGAAGATATCATTACACTAAG
GGCCGAAGATAACGCGGATACCTTGGCGCTAGTATTTGAAGCACCAAACCAGGAGAAAGTTTCAGACT
ATGAAATGAAGTTGATGGATTTAGATGTTGAACAACTTGGAATTCCAGAACAGGAGTACAGCTGTGTA
GTAAAGATGCCTTCTGGTGAATTTGCACGTATATGCCGAGATCTCAGCCATATTGGAGATGCTGTTGT
AATTTCCTGTGCAAAAGACGGAGTGAAATTTTCTGCAAGTGGAGAACTTGGAAATGGAAACATTAAAT
TGTCACAGACAAGTAATGTCGATAAAGAGGAGGAAGCTGTTACCATAGAGATGAATGAACCAGTTCAA
CTAACTTTTGCACTGAGGTACCTGAACTTCTTTACAAAAGCCACTCCACTCTCTTCAACGGTGACACT
CAGTATGTCTGCAGATGTACCCCTTGTTGTAGAGTATAAAATTGCGGATATGGGACACTTAAAATACT
ACTTGGCTCCCAAGATCGAGGATGAAGAAGGATCTTAGGCATTCTTAAAATTCAAGAAAATAAAACTA
AGCTCTTTGAGAACTGCTTCTAAGATGCCAGCATATACTGAAGTCTTTTCTGTCACCAAATTTGTACC
TCTAAGTACATATG

Protein sequence of HsPCNA construct used in this work

MGSSHHHHHH SSGLVPRGSH MLEMFEARLV QGSILKKVLE ALKDLINEAC
WDISSSGVNL QSMDSSHVSL VQLTLRSEGF DTYRCDRNLA MGVNLTSMSK
ILKCAGNEDI ITLRAEDNAD TLALVFEAPN QEKVSDYEMK LMDLDVEQLG
IPEQEYSCVV KMPSGEFARI CRDLSHIGDA VVISCAKDGV KFSASGELGN
GNIKLSQTSN VDKEEEAVTI EMNEPVQLTF ALRYLNFFTK ATPLSSTVTL
SMSADVPLVV EYKIADMGHL KYYLAPKIED EEGS*

1.3 Protein sequence details of SpPCNA

Protein sequence of SpPCNA construct used in this work

MRGSHHHHHH GIPMLEARFQ QAALLKKLLD AIKELVTDAN FDCNDNGISL
QAMDSSHVAL VSMLIKSDGF EPYRCDRNIA LGINLNALSK VLRCAQNEDL
VTLKAEDTPE VLNLVFESEK NDRISDYDVK LMDIDQEHLG IPDIEYDATI
TMPAAEFQRI TRDLLTLSDS VTINASKEGV RFSCKGDIGN GSTTLKQHTD
LSDQDQSIEI SLTQAVTLTF SLKYLAQFTK ATPLATRVTL SMSNDVPLLV
EYKMESGFLR FYLAPKIGEE DEE*

1.4 DNA and protein sequence details of LmPCNA

DNA sequence of LmPCNA

atgctggaggctcaggtccagtacgccagcctgtggaagcgcctggtcgagtgcatcaacgggttggtgaatgaggcg
aacttcgactgcaacccgggaggtctgtccattcaggccatggactccagccacgtcgccctcgtgcacatgttgctgcgt
gacgactgctttgtcaagtaccagtgcgagcgcaacataatccttggcctcaaccttgcctccctctccaaggtgctgaag
atcgtcgacggcaacgacagcctcagtctgcgtcacgacgatgactcggacgtggtcaccctcacctcagagaacccg

gagaagacgcgcaagtgcgagtaccagctgaagctgctcgaaatcgaggcggagtcgatgggcattcctgagatggac
taccgcagcactgtcacactcaactccgccgagttcgcgaagattgtgcgcgatatgcaggtcttcggcgatactgtcacc
atcgccatcagcaaggagggcgtgaagttcagctcatcgggcgatgttgggcagggctacacattccttcaggccgccg
gcgtgtcggatcgcagcgccaagtcggaggtgaaatcagaggtgaaggcggaggcccgggacgaggacgagcacg
agcctatctctcgccggtacaacaaggcggagggcggcaacggcgccatcggggtcgaggtggcgatggaggagcc
catcacgctctcttttgcgctccgcttcatgggcatctttgccaagggctcgacgctcagcgagcgcgtcacgctcaagttt
gccaaggacagtccttgcatggtggaatacggcatcgacaacgtcggttatcttcgctattacctggcgcccaaggtgga
cgatgcggag

Protein sequence of LmPCNA construct used in this work

MGSSHHHHHH SSGLVPRGSH MLEAQVQYAS LWKRLVECIN GLVNEANFDC
NPGGLSIQAM DSSHVALVHM LLRDDCFVKY QCERNIILGL NLASLSKVLK
IVDGNDSLSL RHDDDSDVVT LTSENPEKTR KCEYQLKLLE IEAESMGIPE
MDYRSTVTLN SAEFAKIVRD MQVFGDTVTI AISKEGVKFS SSGDVGQGYT
FLQAAGVSDR SAKSEVKSEV KAEARDEDEH EPISRRYNKA EGGNGAIGVE
VAMEEPITLS FALRFMGIFA KGSTLSERVT LKFAKDSPCM VEYGIDNVGY
LRYYLAPKVD DAE*

1.5 ITC values derived for PCNA

1.6 ITC values derived for HsPCNA

[Peptide] (μM)	N	K_d (μM)	ΔH (kJ mol^{-1})	ΔS (J $Kmol^{-1}$)	TΔS (kJ mol^{-1})	ΔG (kJ mol^{-1})
Peptide0						
[87.75]	2.03	0.0213	-38.46	11.09	3.140	-41.60
[33.81]	0.795	0.011	-96.99	-190.0	-53.76	-43.23
[37.37]	0.882	0.0092	-88.53	-158.6	-44.88	-43.66
Peptide1						
[154.5]	2.12	3.31	-46.48	-59.41	-16.81	-29.67
[60.51]	0.829	1.3	-118.9	-307.1	-86.91	-32.00
[66.88]	0.916	1.43	-107.6	-268.2	-75.90	-31.67
Peptide2						
[145.9]	2.03	2.81	-26.18	13.81	3.9100	-30.09
[47.63]	0.734	0.92	-80.17	-167.8	-47.48	-32.68
[52.65]	0.664	1.01	-72.55	-141.4	-40.02	-32.53
Peptide3						
[125.3]	3.9	-0.154	-9.270			-9.270
[52.90]	0.995	5.29	-78.37	-175.7	-49.73	-28.64
[58.46]	1.1	5.85	-70.88	-150.2	-42.51	-28.37
Peptide4						
[113.2]	1.49	3.46	-37.37	-27.45	-7.770	-29.60
[54.65]	0.72	1.68	-77.40	-162.8	-46.06	-31.34
[60.41]	0.797	1.85	-70.00	-137.7	-38.96	-31.04
Peptide5						
[103.7]	2.21	4.37	-26.22	10.00	2.830	-29.05
[51.84]	1.1	2.18	-52.43	-76.99	-21.79	-30.64
[57.30]	1.22	2.41	-47.45	-59.83	-16.93	-30.51

Stoichiometry closest to 1.00

[Peptide] (μM)	N	K_d (μM)	ΔH (kJ mol^{-1})	ΔS (J $Kmol^{-1}$)	TΔS (kJ mol^{-1})	ΔG (kJ mol^{-1})
P0 [37.37]	0.882	0.0092	-88.53	-158.6	-44.88	-43.66
P1 [66.88]	0.916	1.43	-107.6	-268.2	-75.90	-31.67
P2 [47.63]	0.734	0.92	-80.17	-167.8	-47.48	-32.68
P3 [52.90]	0.995	5.29	-78.37	-175.7	-49.73	-28.64
P4 [60.41]	0.797	1.85	-70.00	-137.7	-38.96	-31.04
P5 [51.84]	1.1	2.18	-52.43	-76.99	-21.79	-30.64

Biophysical characteristics of HsPCNA-peptide interactions at 283K. Each ITC result was solved for the concentrations based on dissolved mass (first), +5% A_{280} (middle) and -5% A_{280} (bottom). The results for each peptide showing a stochiometery closest to 1.0 is shown in the lower table. This data demonstrates at 283K the K_d of peptide0 in the low nM region and the shorter peptides in the μM, also a Gibbs free energy demonstrating a spontaneous interaction.

[Peptide] (μM)	N	K_d (μM)	ΔH (kJ mol^{-1})	ΔS (J $Kmol^{-1}$)	TΔS (kJ mol^{-1})	ΔG (kJ mol^{-1})
Peptide0						
[88.54]	1.61	0.018	-28.74	46.86	13.97	-42.72
[26.29]	0.29	0.0054	-90.79	-162.34	-48.37	-42.43
[29.05]	0.322	0.0034	-80.92	-123.43	-36.78	-44.14
Peptide1						
[129.0]	4.09	1.36	-13.35	65.27	19.46	-32.80
[42.36]	1.34	0.45	-40.63	-21.97	-6.53	-34.10
[46.82]	1.48	0.49	-36.78	-9.12	-2.7196	-34.06
Peptide2						
[148.1]	1.01	1.18	-30.92	13.89	4.14	-35.06
[58.19]	0.396	0.47	-71.80	-132.21	-39.41	-32.38
[64.31]	0.502	0.28	-50.04	-51.46	-15.36	-34.69
Peptide3						
[120.9]	1.12	1.47	-19.87	41.51	12.38	-32.26
[63.63]	0.663	0.22	-24.85	39.71	11.84	-36.69
[70.33]	0.733	0.24	-22.47	47.28	14.10	-36.57
Peptide4						
[114.1]	2.22	2.46	-21.55	31.21	9.29	-30.84
[45.54]	0.888	0.98	-54.02	-75.73	-22.55	-31.46
[50.34]	0.992	1.25	-51.71	-69.45	-20.71	-31.00
Peptide5						
[104.2]	2.75	0.41	-11.17			
[47.52]	0.817	0.27	-28.87	23.68	7.07	-35.94
[52.52]	1.2	1.07	-39.08	-23.72	-7.07	-32.01

Stoichiometry closest to 1.00

[Peptide] (μM)	N	K_d (μM)	ΔH (kJ mol^{-1})	ΔS (J $Kmol^{-1}$)	TΔS (kJ mol^{-1})	ΔG (kJ mol^{-1})
P0 [88.54]	1.61	0.018	-28.74	46.86	13.97	-42.72
P1 [42.36]	1.34	0.45	-40.63	-21.97	-6.53	-34.10
P2 [148.1]	1.01	1.18	-30.92	13.89	4.14	-35.06
P3 [120.9]	1.12	1.47	-19.87	41.51	12.38	-32.26
P4 [50.34]	0.992	1.25	-51.71	-69.45	-20.71	-31.00
P5 [52.52]	1.2	1.07	-39.08	-23.72	-7.07	-32.01

Biophysical characteristics of HsPCNA-peptide interactions at 298K. Each ITC result was solved for the concentrations based on dissolved mass (first), +5% A_{280} (middle) and -5% A_{280} (bottom). The results for each peptide showing a stochiometery closest to 1.0 is shown in the lower table. This data demonstrates at 298K the K_d of peptide0 greater than that at 283K, the shorter peptides consistently in the μM region, with the untagged peptide showing a greater affinity, also a Gibbs free energy demonstrating a spontaneous interaction.

1.7 ITC values derived for SpPCNA

[Peptide] (μM)	N	K_d (μM)	ΔH (kJ mol^{-1})	ΔS (J $Kmol^{-1}$)	TΔS (kJ mol^{-1})	ΔG (kJ mol^{-1})
Peptide0						
[88.54]	1.05	0.046	-30.88	31.46	8.912	-39.79
[30.04]	0.355	0.016	-90.04	-168.6	-47.70	-42.34
[33.20]	0.395	0.021	-83.55	-148.1	-41.92	-41.63
Peptide1						
[147.4]	0.673	0.133	-50.63	-47.28	-13.39	-37.24
[77.15]	0.352	0.070	-96.78	-205.0	-58.03	-38.74
[85.27]	0.390	0.085	-88.78	-178.2	-50.46	-38.33
Peptide2						
[150.2]	0.647	0.68	-38.33	-17.28	-4.895	-33.43
[68.76]	0.290	0.34	-85.27	-177.4	-50.21	-35.06
[76.00]	0.313	0.46	-83.14	-172.4	-48.79	-34.35
Peptide3						
[116.7]	1.70	1.75	-31.80	-2.050	-0.5858	-31.21
[58.73]	0.856	0.88	-63.22	-107.1	-30.29	-32.93
[64.91]	0.946	0.97	-57.20	-87.03	-24.64	-32.55
Peptide4						
[114.1]	1.80	1.07	-24.69	27.20	7.699	-32.38
[45.54]	0.711	0.45	-60.63	-92.47	-26.15	-34.48
[50.34]	0.791	0.56	-56.99	-81.59	-23.10	-33.89
Peptide5						
[104.5]	1.75	0.89	-22.89	35.02	9.916	-32.80
[56.16]	0.948	0.69	-45.65	-43.51	-12.30	-33.35
[62.07]	1.03	0.56	-37.61	-13.26	-3.766	-33.85

Stoichiometry closest to 1.00

[Peptide] (μM)	N	K_d (μM)	ΔH (kJ mol^{-1})	ΔS (J $Kmol^{-1}$)	TΔS (kJ mol^{-1})	ΔG (kJ mol^{-1})
P0 [88.54]	1.05	0.046	-30.88	31.46	8.912	-39.79
P1 [147.4]	0.673	0.13	-50.63	-47.28	-13.39	-37.24
P2 [150.2]	0.647	0.68	-38.33	-17.28	-4.895	-33.43
P3 [64.91]	0.946	0.97	-57.20	-87.03	-24.64	-32.55
P4 [50.34]	0.791	0.56	-56.99	-81.59	-23.10	-33.89
P5 [62.07]	1.03	0.56	-37.61	-13.26	-3.766	-33.85

Biophysical characteristics of SpPCNA-peptide interactions at 283K. Each ITC result was solved for the concentrations based on dissolved mass (first), +5% A_{280} (middle) and -5% A_{280} (bottom). The results for each peptide showing a stochiometery closest to 1.0 is shown in the lower table. This data demonstrates at 283K the K_d of peptide0 greater than that for HsPCNA, the shorter peptides in the μM range, with the untagged peptide showing a greater affinity, also a Gibbs free energy demonstrating a spontaneous interaction.

[Peptide] (μM)	N	K_d (μM)	ΔH (kJ mol^{-1})	ΔS (J $Kmol^{-1}$)	TΔS (kJ mol^{-1})	ΔG (kJ mol^{-1})
Peptide0						
[88.54]	1.32	0.043	-30.04	34.81	10.38	-40.42
[22.63]	0.334	0.007	-109.5	-230.1	-68.58	-40.96
[25.02]	0.373	0.012	-106.0	-223.0	-66.44	-39.58
Peptide1						
[148.2]	0.874	0.154	-52.47	-54.81	-16.32	-36.15
[72.61]	0.424	0.056	-102.5	-223.0	-66.44	-36.02
[80.25]	0.470	0.068	-93.89	-194.1	-57.86	-36.02
Peptide2						
[150.3]	0.917	0.50	-34.94	-2.812	-0.837	-34.10
[68.76]	0.419	0.23	-76.40	-142.7	-42.51	-33.89
[76.00]	0.464	0.25	-69.12	-118.0	-35.15	-33.97
Peptide3						
[116.6]	1.27	0.61	-23.81	34.85	10.38	-34.18
[48.94]	0.542	0.36	-62.05	-95.81	-28.53	-33.51
[54.10]	0.603	0.46	-57.99	-83.26	-24.81	-33.18
Peptide4						
[]						
[]						
[]						
Peptide5						
[]						
[]						
[]						

Stoichiometry closest to 1.00

[Peptide] (μM)	N	K_d (μM)	ΔH (kJ mol^{-1})	ΔS (J $Kmol^{-1}$)	TΔS (kJ mol^{-1})	ΔG (kJ mol^{-1})
P0 [88.54]	1.32	0.043	-30.04	34.81	10.38	-40.42
P1 [148.2]	0.874	0.154	-52.47	-54.81	-16.32	-36.15
P2 [150.3]	0.917	0.50	-34.94	-2.812	-0.837	-34.10
P3 [116.6]	1.27	0.61	-23.81	34.85	10.38	-34.18
P4 []						
P5 []						

Biophysical characteristics of SpPCNA-peptide interactions at 298K. Each ITC result was solved for the concentrations based on dissolved mass (first), +5% A_{280} (middle) and -5% A_{280} (bottom). The results for each peptide showing a stochiometery closest to 1.0 is shown in the lower table. This data demonstrates at 298K the K_d of all peptides to be lower than that for 283K, peptide0 is the strongest interactor however this is anomalous data. The Gibbs free energy still demonstrates a spontaneous interaction.

1.8 ITC values derived for LmPCNA

[Peptide] (μM)	N	K_d (μM)	ΔH (kJ mol^{-1})	ΔS (J $Kmol^{-1}$)	TΔS (kJ mol^{-1})	ΔG (kJ mol^{-1})
Peptide0						
[88.54]	2.51	0.990	-17.41	53.56	15.15	-32.55
[26.29]	0.808	0.386	-57.53	-80.33	-22.72	-34.81
[29.05]	0.894	0.429	-52.17	-62.34	-17.66	-34.52
Peptide1						
[142.4]	2.05	3.91	-18.74	37.32	10.54	-29.29
[84.71]	1.22	2.33	-31.55	-3.60	-1.004	-30.54
[93.63]	1.35	2.58	-28.53	6.19	1.757	-30.29
Peptide2						
[300.2]	3.03		-4.351			
[142.8]	1.31	3.72	-17.91	40.67	11.51	-29.41
[157.9]	1.36	1.27	-24.64	6.611	1.883	-26.53
Peptide3						
[]						
[]						
[]						
Peptide4						
[]						
[]						
[]						
Peptide5						
[]						
[]						
[]						

Stoichiometry closest to 1.00

[Peptide] (μM)	N	K_d (μM)	ΔH (kJ mol^{-1})	ΔS (J $Kmol^{-1}$)	TΔS (kJ mol^{-1})	ΔG (kJ mol^{-1})
P0 [29.05]	0.894	0.429	-52.17	-62.34	-17.66	-34.52
P1 [84.71]	1.22	2.33	-31.55	-3.60	-1.004	-30.54
P2 [142.8]	1.31	3.72	-17.91	40.67	11.51	-29.41
P3 []						
P4 []						
P5 []						

Biophysical characteristics of LmPCNA-peptide interactions at 283K. Each ITC result was solved for the concentrations based on dissolved mass (first), +5% A_{280} (middle) and -5% A_{280} (bottom). The results for each peptide showing a stochiometery closest to 1.0 is shown in the lower table. This data demonstrates at 283K the K_d of peptide0 greater than that for both HsPCNA and SpPCNA, in the sub μM range, the shorter peptides in the μM range, with the untagged peptide showing a greater affinity, also a Gibbs free energy demonstrating a spontaneous interaction.

[Peptide] (μM)	N	K_d (μM)	ΔH (kJ mol^{-1})	ΔS (J $Kmol^{-1}$)	TΔS (kJ mol^{-1})	ΔG (kJ mol^{-1})
Peptide0						
[177.1]	1.75	0.794	-16.32	58.99	17.57	-33.89
[90.22]	0.89	0.405	-32.05	9.163	2.720	-34.77
[99.72]	0.98	0.446	-29.00	19.12	5.690	-34.69
Peptide1						
[294.8]	1.19	1.99	-19.12	41.55	12.38	-31.51
[208.5]	0.84	1.41	-27.07	16.44	4.895	-31.97
[230.4]	0.93	1.56	-24.48	24.77	7.364	-31.84
Peptide2						
[300.5]	0.68	25.4	-34.77	-34.77	-10.38	-24.39
[179.2]	0.37	15.5	-64.85	-137.2	-40.88	-23.97
[198.0]	0.42	17.0	-56.07	-106.7	-31.80	-24.27
Peptide3						
[]						
[]						
[]						
Peptide4						
[]						
[]						
[]						
Peptide5						
[]						
[]						
[]						

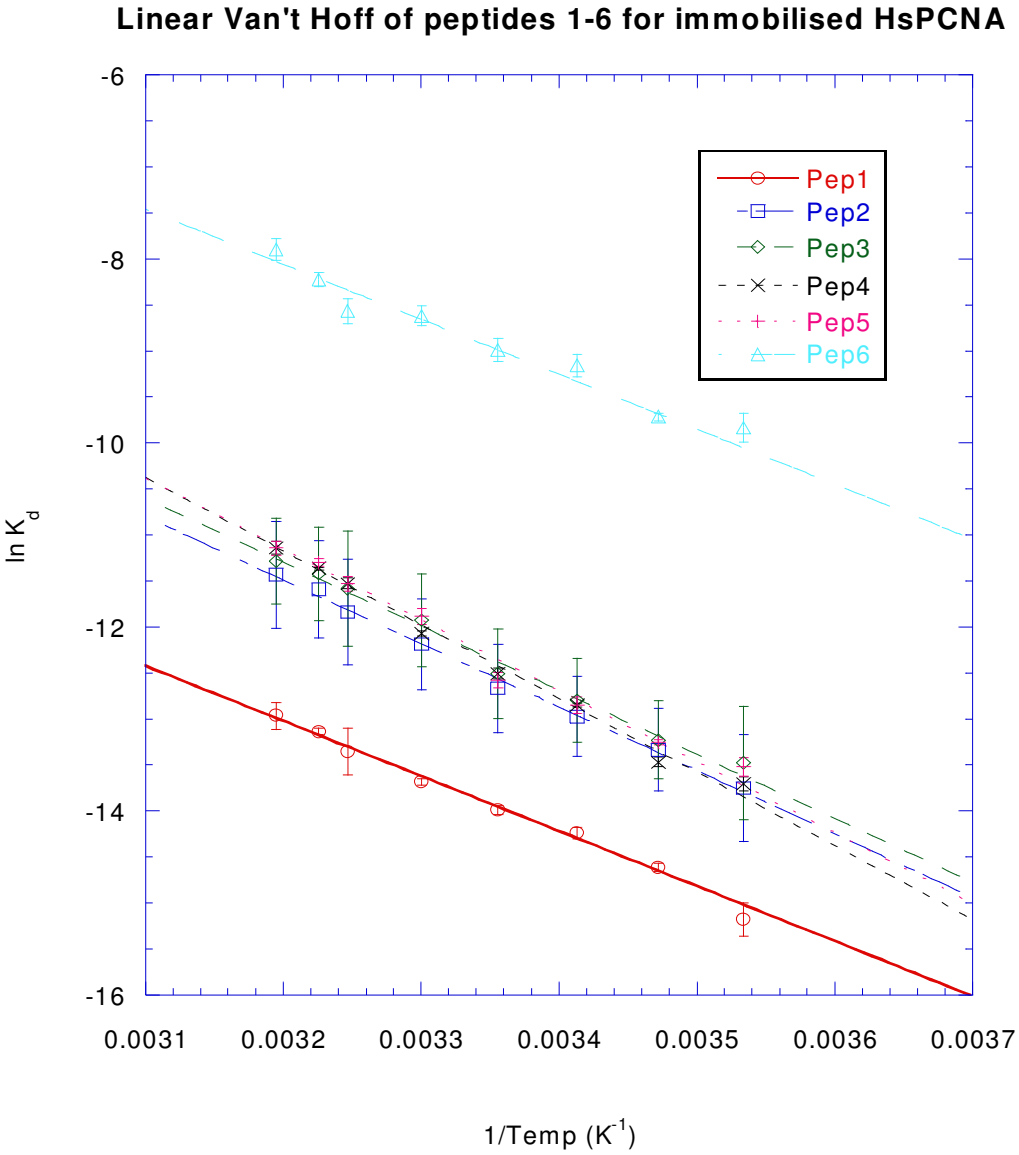
Stoichiometry closest to 1.00

[Peptide] (μM)	N	K_d (μM)	ΔH (kJ mol^{-1})	ΔS (J $Kmol^{-1}$)	TΔS (kJ mol^{-1})	ΔG (kJ mol^{-1})
P0 [99.72]	0.98	0.446	-29.00	19.12	5.690	-34.69
P1 [230.4]	0.93	1.56	-24.48	24.77	7.364	-31.84
P2 [300.5]	0.68	25.4	-34.75	-34.77	-10.38	-24.37
P3 []						
P4 []						
P5 []						

Biophysical characteristics of LmPCNA-peptide interactions at 298K. Each ITC result was solved for the concentrations based on dissolved mass (first), +5% A_{280} (middle) and -5% A_{280} (bottom). The results for each peptide showing a stochiometery closest to 1.0 is shown in the lower table. This data demonstrates at 298K the K_d is lowest for peptide 0 also a Gibbs free energy demonstrating a spontaneous interaction.

1.9 Linear van't Hoff analysis of SPR data

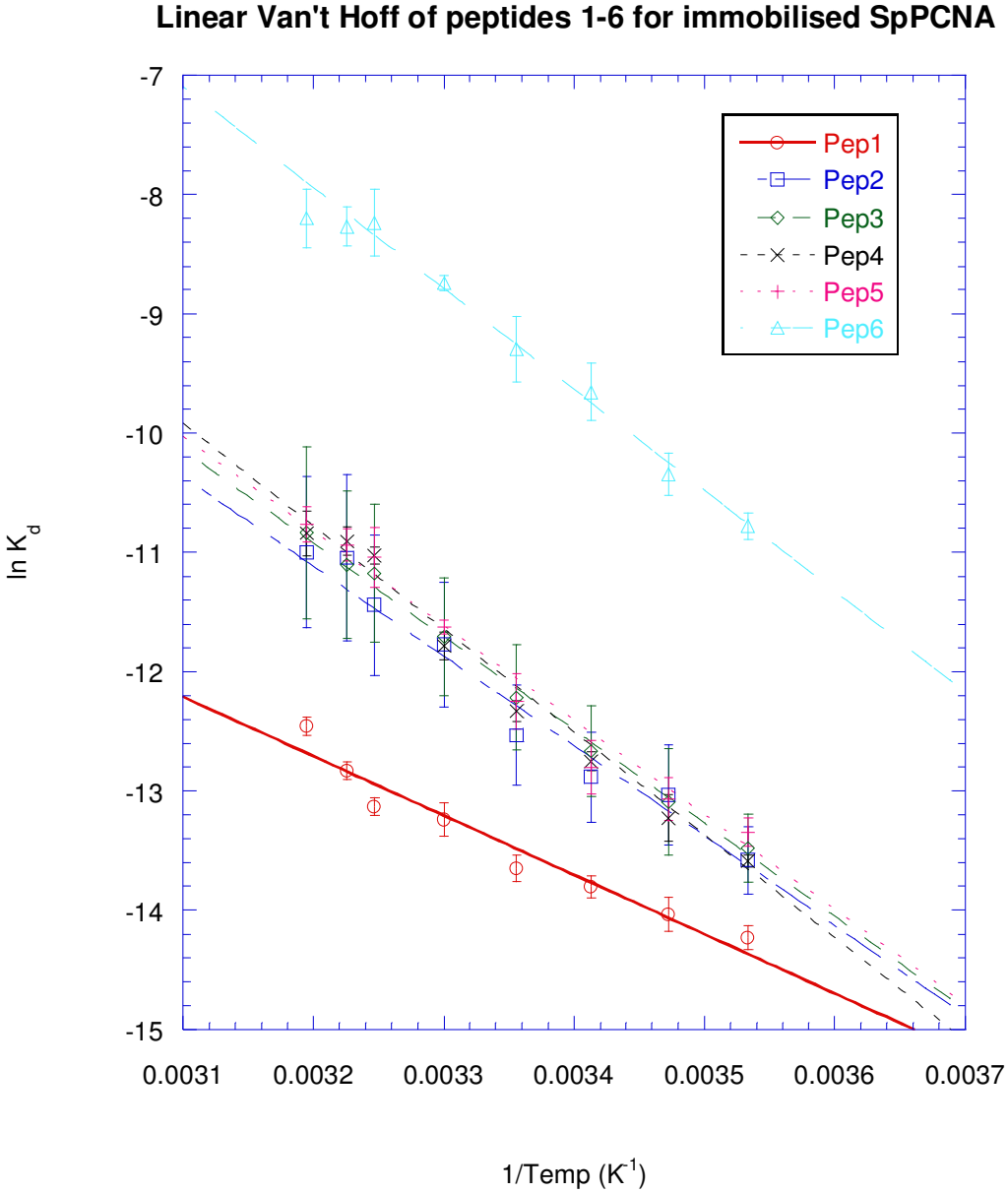
1.10 Linear van't Hoff analysis of HsPCNA



Peptide	Spacer Length (Å)	ΔH (kJ mol^{-1})	ΔS (J Kmol^{-1})	ΔG (kJ mol^{-1})
1	-	-49.8	-50.9	-34.6
2	10	-57.3	-87.9	-31.1
3	20	-57.9	-91.2	-30.7
4	30	-66.6	-120	-30.8
5	40	-64.1	-113	-30.6
6	-	-49.7	-92.0	-22.3

The linear Van't Hoff plot (top) with lnK_d on the y axis verses 1/T on the x axis plotted with Kaleidagraph. Table (bottom) shows thermodynamic properties determined for each peptide to three significant figures.

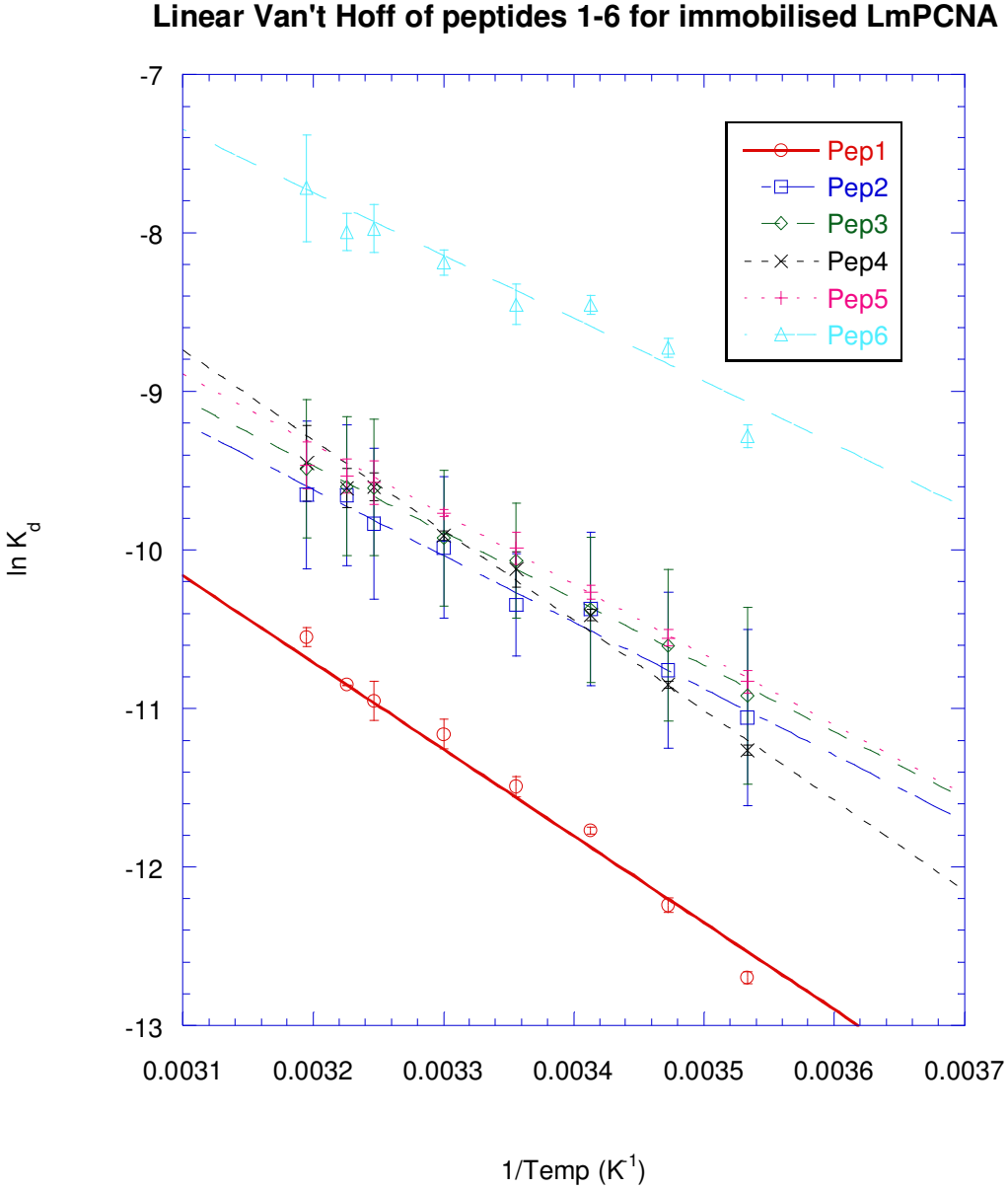
1.11 Linear van't Hoff analysis of SpPCNA



Peptide	Spacer Length (Å)	ΔH (kJ mol^{-1})	ΔS (J Kmol^{-1})	ΔG (kJ mol^{-1})
1	-	-41.3	-26.6	-33.4
2	10	-62.6	-108	-30.5
3	20	-65.0	-117	-30.1
4	30	-71.7	-140	-30.0
5	40	-65.9	-121	-29.9
6	-	-70.3	-159	-22.9

The linear Van't Hoff plot of SPR data (top) with lnK_d on the y axis verses 1/T on the x axis plotted with Kaleidagraph. Table (bottom) shows thermodynamic properties determined for each peptide to three significant figures.

1.12 Linear van't Hoff analysis of LmPCNA

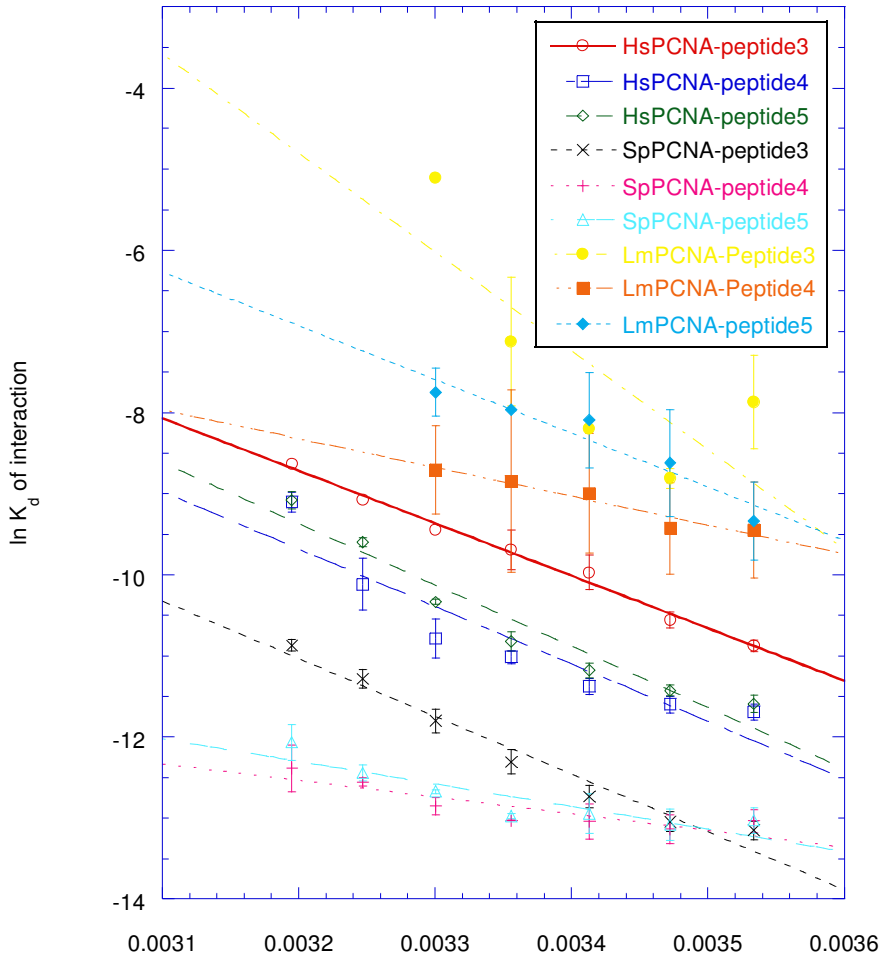


Peptide	Spacer Length (Å)	ΔH (kJ mol^{-1})	ΔS (J Kmol^{-1})	ΔG (kJ mol^{-1})
1	-	-45.5	-56.5	-28.6
2	10	-34.8	-31.5	-25.5
3	20	-34.9	-33.0	-25.1
4	30	-47.2	-73.7	-25.3
5	40	-36.8	-40.1	-24.8
6	-	-33.0	-41.2	-20.7

The linear Van't Hoff plot of SPR data (top) with lnK_d on the y axis verses 1/T on the x axis plotted with Kaleidagraph. Table (bottom) shows thermodynamic properties determined for each peptide to three significant figures.

1.13 Linear van't Hoff analysis of immobilised peptides

Linear Van't Hoff analysis of immobilised peptides with free PCNA in solution



PCNA Peptide interactions	Spacer Length (Å)	ΔH (kJ mol^{-1})	ΔS (J Kmol^{-1})	ΔG (kJ mol^{-1})
HsPCNA-P3	20	-53.8	-99.6	-24.1
HsPCNA-P4	30	-58.8	-108	-26.7
HsPCNA-P5	40	-62.8	-123	-26.1
SpPCNA-P3	20	-59.2	-97.6	-30.1
SpPCNA-P4	30	-17.1	49.6	-31.9
SpPCNA-P5	40	-23.1	28.3	-31.6
LmPCNA-P3	20	-101	-284	-16.5
LmPCNA-P4	30	-29.5	-25.3	-22.0
LmPCNA-P5	40	-55.1	-119	-19.7

The linear Van't Hoff plot of SPR data (top) with lnK_d on the y axis verses 1/T on the x axis plotted with Kaleidagraph. Table (bottom) shows thermodynamic properties determined for each peptide to three significant figures.