Structural and biochemical studies of PCNA and its molecular interactions

A Thesis

Submitted for the Degree of

Doctor of Philosophy

by

Cornelia Ludwig, B.Sc.

Structural Biochemistry Group

Institute of Structural and Molecular Biology

University of Edinburgh

August 2008

Summary of Contents

Abstract

Proliferating cell nuclear antigen (PCNA) plays a key role in DNA replication and repair. It also interacts directly or indirectly with proteins involved in cell-cycle control. These proteins include p21, a p53-regulated association control factor with check point function, and p53-regulated Gadd45, which plays an important role in the cell cycle G2-M checkpoint. There are a number of structures currently available for the trimeric ring-forming PCNA family of proteins found in eukaryotes and archaea as well as those related PCNA-like sliding clamps in prokaryotes and viruses. However, despite the great similarity between the PCNA and sliding clamps it is important to increase the understanding about these proteins from other organisms, in particular model organisms such as *Schizosaccharomyces pombe.* Structural comparisons highlight a well conserved hydrophobic binding pocket involved in protein-protein interactions, which is found in all PCNAs and other sliding clamps. Previously this pocket has been identified as a potential druggable target for the development of new anti-cancer drug leads due to the central role of PCNA in the cell-cycle. Proteins that bind to this pocket contain all or part of the conserved PIP-box motif (Q-x-x-h-x-x-a-a).

In this thesis I have determined the crystal structure of *S. pombe* PCNA on its own and in complex with a peptide derived from human p21. This peptide was found previously to bind to SpPCNA despite there being no homologue of p21 in fission yeast. These crystal structures of SpPCNA confirmed structural similarities in comparison with human PCNA, but also hints at differences between the two species in terms of the hydrophobic pocket and ligand binding modes.

Using the p21 peptide as a template, chemical libraries were screened for small drug-like molecules that are able to mimic the protein-protein interaction between human PCNA and p21 and were tested for binding using various fluorescence-based methods. Binding of selected compounds could not be observed, however, which may have been due to poor solubility of the compounds as well as a lack of assays sensitive enough to pick up on ligands with low affinity.

Not all direct PCNA binding proteins bind to the above mentioned pocket and PCNA/protein co-crystal structures of these examples without the PIP-box motif are not available so far, so information on the mode of binding in those cases is not accessible yet. One objective was, therefore, to narrow down and characterize the PCNA binding site of Gadd45. This protein does not contain a PIP-box motif, but previously was shown using yeast-two hybrid technology to bind directly to PCNA *via* its C-terminus. *In vitro* binding of the two full-length proteins could not be confirmed, due to the inherent instability of Gadd45, which also has been reported by others. GST-tagged truncations of the C-terminus of Gadd45 were expressed and purified and binding to PCNA was studied using SPR. These result are at odds with the published binding data and suggest that either the interaction between PCNA and Gadd45 is not direct and needs to mediated by a third factor. Alternatively, Gadd45 binds to PCNA *via* a different part than the C-terminus (as also suggested in the literature).

Declaration

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

Cornelia Ludwig

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 in his group and the School of Biological Sciences for funding.

I am very grateful to the members of the Structural Biochemistry Group, in particular Dr. Martin Wear and Dr. Iain McNae who have provided significant help and assistance.

In addition I would like to thank Dr. Simon Harding and Dr. Mike Greaney for their collaboration on the structure-based drug design project and Dr. Emma Warbrick for providing various materials.

Furthermore I would like to mention Dr. David Dryden, Dr. Julia Richardson, Dr. Hugh Morgan, Dr. Chris Brown, Dr. Majorie Harding, Dr. Paul Taylor and Peter Brown for help in a variety of things during the course of my studies. My thanks also goes to everyone on the third floor of the Michael Swann Building for providing a friendly and stimulating work environment which has made all the difference to me.

Lastly I would like to thank my family - Walburga and Andrea - and my friends for their support and encouragement over the past few years and I am dedicating this thesis to Dad.

List of Abbreviations

AP	apurinic/apyrimidinic
Apn2	AP-endonuclease 2
BER	Base Excision Repair
CAF-1	Chromatin assembly factor 1
CCP4	Collaborative Computational Project Number 4
DLS	Dynamic Light Scattering
DMSO	Dimethyl sulfoxide
DNMT1	DNA cytosine methyltransferase 1
DTT	Dithiothreitol
EDTA	ethylenediaminetetraacetic acid
FP	Fluorescence Polarization
GADD45	Growth Arrest and DNA Damage gene 45 (Gadd45 = protein product of GADD45)
HIC	hydrophobic interaction
hPCNA	human PCNA
IDCL	interdomain connector loop
IPTG	isopropyl-β-D-thiogalactopyranoside
ITC	Isothermal titration calorimetry
logP	octanol-water partition coefficient
MMR	Mismatch Repair
NER	Nucleotide Excision Repair
NTA	nitrilotriacetic acid
OD_{600}	optical cell density measured at 600nm
ORF	open reading frame
PBS	Phosphate Buffered Saline
PCNA	proliferating cell nuclear antigen
PCR	polymerase chain reaction
PDB	Protein Data Bank
PEG	polyethyleneglycol
PfuPCNA	Pyrococcus furiosus
PIP box/motif	PCNA interacting protein **box/motif**
PIP score	Pose Interaction Profile **score**
PL	Pogo/Ligase I
PMSF	phenylmethanesulphonyl fluoride
RFC	replication factor C
RU	response units
ScPCNA	Saccharomyces cerevisiae PCNA
SDS PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SpPCNA	Schizosaccharomyces pombe PCNA
SPR	surface plasmon resonance
TLS pathway	translesion synthesis pathway
TLS refinement	translation-libration-screw refinement
T_m	melting temperature
Ungl	uracil DNA N-glycosylase

Table of Contents

List of Figures

List of Tables

CHAPTER 1. Introduction: The structure of PCNA and its homologues

1.1 GENERAL STRUCTURE

The ring-shaped PCNA (proliferating cell nuclear antigen) is probably best known as the processivity factor of the DNA polymerase δ in eukaryotes, however, it is involved in a number of processes of DNA metabolism, such as DNA damage repair (Maga and Hubscher, 2003; Paunesku et al., 2001). Eukaryotic PCNA encircles DNA as a homotrimer. Its ability to slide along DNA explains why it is also called a 'sliding clamp' (Figure 1-1). However, proteins with related functional and structural similarities are also found in prokaryotes and viruses, but are usually only referred to as sliding clamps, rather than PCNA (Table 1-1).

The first evidence for the structure of sliding clamps came from studies in *Escherichia coli*. Stukenberg and colleagues (Stukenberg et al., 1991) carried out experiments with β sliding clamps on both circular and linear DNA molecules and found that the protein bound stably to circular DNA, but dissociated readily from the linear form. This finding shed new light on the way the protein binds to DNA by encircling the double helix. Further evidence for the overall structure of PCNA-like proteins came from T4 phage sliding clamp g45 (Gogol et al., 1992) and electron microscope pictures suggested that DNA was threaded through this ring-shaped molecule.

Indirect evidence for the structure of PCNA came from Burgers and Yoder (Burgers and Yoder, 1993), who showed that in the absence of Replication Factor C (RFC), the clamp loader protein, DNA synthesis could only take place when linear DNA

was used as the substrate. Tinker et al. (Tinker et al., 1994) found that hPCNA could only be cross-linked to circular DNA. Previously, Kong et al. (Kong et al., 1992) solved the crystal structure of the β sliding clamp of *E. coli* which proved the validity of previous claims about the structure-function relationship in these proteins. Their structure showed a homodimeric ring with an inner diameter of ~35Å, big enough to let double stranded DNA pass through.

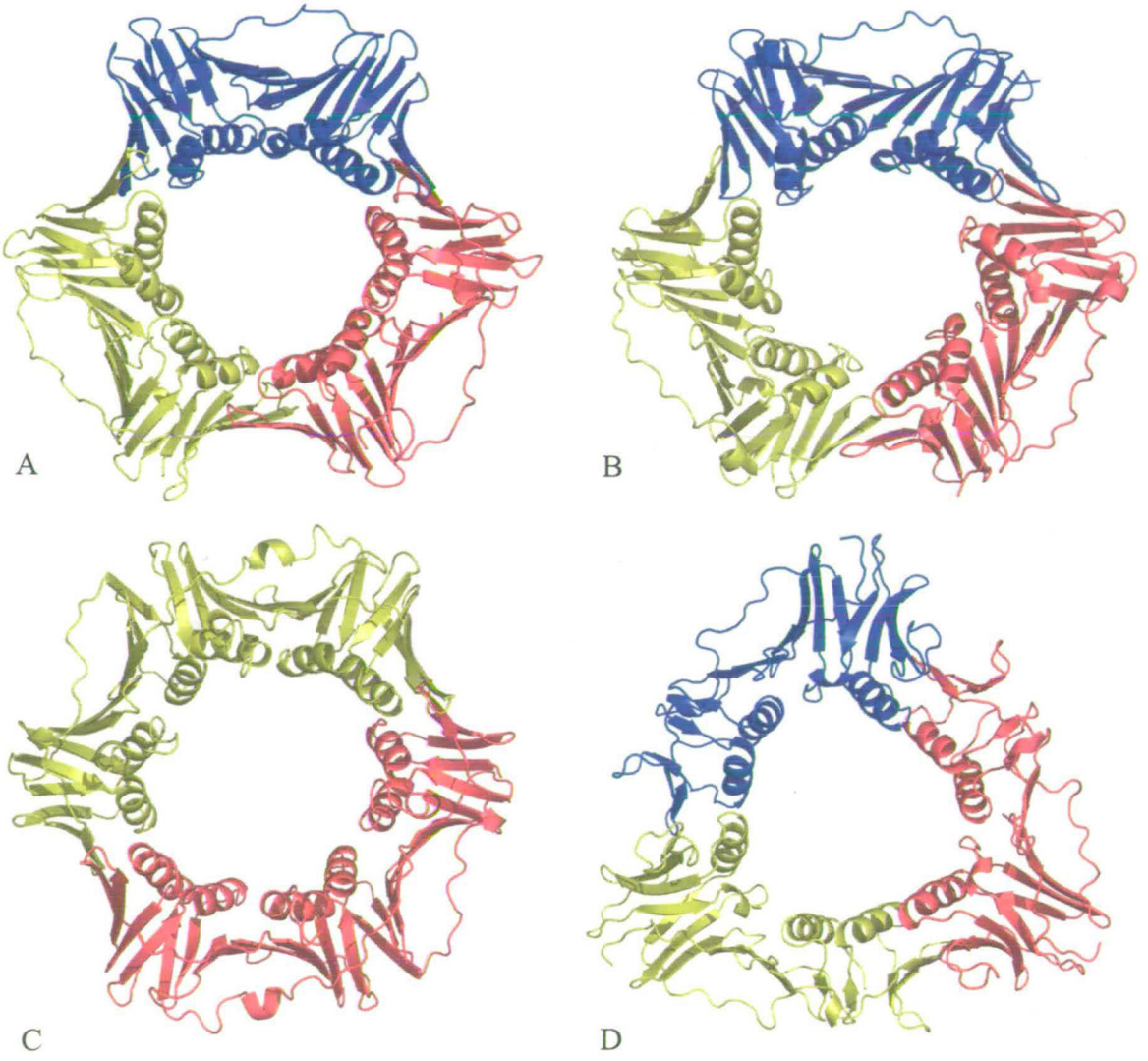


Figure 1-1 Crystal structures of human PCNA and its homologues.
A) human PCNA, PDB ID: 1VYM (Kontopidis et al., 2005) B) *Archaeoglobus fulgidus* PCNA, PDB ID: 1RWZ (Chapados et al., 2004) C) *Escherichia coli* β subunit of DNA polymerase III, PDB ID: 2POL (Kong et al., 1992) D) bacteriophage RB69 gene 45 protein, PDB ID: 1B77 (Shamoo and Steitz, 1999) Apart from the *Escherichia coli* β subunit all sliding clamps show the trimeric configuration. The overall 3D structure, however, remains similar in all four examples from across all domains of life including viruses.

An alignment of phage, *E. coli*, archaeal and eukaryotic PCNA secondary structure elements showed a high degree of structural conservation (Figure 1-2), but almost no similarity in sequence.

Species	PDB Code	Number of amino acids in monomer	Resolution in Å	Complexed with	Author
human	1VYM	261	2.3	naive	(Kontopidis et al., 2005)
human	1AXC		2.6	p21(141-160)	(Gulbis et al., 1996)
human	1U76		2.6	Pol δ (453-465)	(Bruning and Shamoo, 2004)
human	1U7B		1.88	Fen-1 (336-348)	(Bruning and Shamoo, 2004)
human	1VYJ		2.8	PL (1-16)	(Kontopidis et al., 2005)
Saccharomyces cerevisiae	1PLQ	258	2.3	native	(Krishna et al., 1994)
Saccharomyces cerevisiae	1SXJ		2.85	RFC hetero-pentamer	(Bowman et al., 2004)
Archaeoglobus fulgidus	1RWZ	245	2.0	native	(Chapados et al., 2004)
Archaeoglobus fulgidus	1RXZ		1.8	Fen-1 (326-336)	(Chapados et al., 2004)
Pyrococcus furiosus	1GE8	249	2.1	native	(Matsumiya et al., 2001)
Pyrococcus furiosus	1ISQ		2.3	RFC (469-476)	(Matsumiya et al., 2002)
Escherichia coli	2POL	366	2.5	native	(Kong et al., 1992)
Escherichia coli	1OK7		1.65	pol IV (345-351)	(Burnouf et al., 2004)
T4	1CZD	228	2.45	native	(Moarefi et al., 2000)
RB69	1B77	228	2.1	native	(Shamoo and Steitz, 1999)
RB69	1B8H		3.0	pol (893-903)	(Shamoo and Steitz, 1999)

Table 1-1 List of selected sliding clamp X-ray structures deposited in the PDB till March 2006.

Figure 1-2 Alignment of the sequences of the domains of human PCNA, *A. fulgidus* PCNA, *E. coli* β subunit and RB69 gene 45 protein.
Although these proteins come from different domains of life (as well as viruses) and bear no significant sequence conservation, the arrangement of secondary structure elements (red = α helix, green = β strand) is highly conserved among sliding clamps. The most important loops have been labelled in blue, where C-loop = Centre Loop and b.s. loop designates loops placed on the back side of PCNA.

Since then it has been shown by various groups that the structure of sliding clamp proteins is indeed highly conserved in all three domains of life as well as viruses. In all these different organisms the function is similar and in each case a ring shaped trimer (or dimer) formed by smaller monomers provides the structural motif (Figure 1-1) (Chapados et al., 2004; Kong et al., 1992; Kontopidis et al., 2005; Shamoo and Steitz, 1999). In eukaryotes the sliding clamps are homotrimeric doughnut-shaped molecules with each subunit in human PCNA weighing about 29kDa. In *E. coli* a larger gene product leads to the formation of a dimeric ring with each monomer contributing 41kDa.

The PCNA family has an overall excess of acidic residues, which gives them a low pI (Kelman and Odonnell, 1995a) (Figure 1-3). The distribution of charge is asymmetric, with strong negative potential on the outer surface of the protein, which might aid the prevention of non-specific interactions with DNA. There is a net positive electrostatic potential at the central cavity (Kong et al., 1992; Krishna et al., 1994).

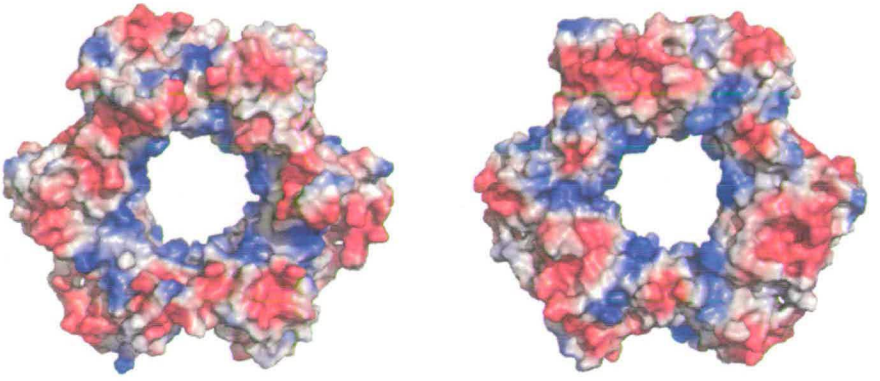


Figure 1-3 Surface charge of human PCNA.
Right: C-side, left: back side. The central cavity is predominantly lined by basic residues, where the remaining surface is rather covered in acidic residues. Overall a negative electrostatic potential predominates.

PCNA trimers show an approximate six-fold symmetry like the *E. coli* dimer. Each monomer contains two globular domains which are made up of two pseudosymmetric domains following a βαβββ motif. Hence each domain of the trimeric subunit consists of two α helices, which coat the central cavity (Figure 1-1) and a continuous layer of nine anti-parallel β-sheets and leads to the formation of a two-layer structure. All together, twelve helices line the central cavity (Figure 1-1), which has a diameter of 34Å in hPCNA wide enough for B-DNA to fit through (Gulbis et al., 1996; Krishna et al., 1994).

In a number of DNA binding proteins α helices are placed in such a way as to fit into the groves of DNA (Warren, 2002). In PCNA the helices are arranged perpendicular to the phosphate back bone, which makes this an unusual recognition mode (Kuriyan and Odonnell, 1993; Odonnell et al., 1993), and basic residues on these helices are used to recognise and bind DNA (Krishna et al., 1994). The importance of these residues has been shown by mutation studies, which showed that removal of any of these basic side chains leads to down regulation of DNA synthesis by the appropriate polymerase. The proper distribution of positive charge seems to be rather crucial (Fukuda et al., 1995).

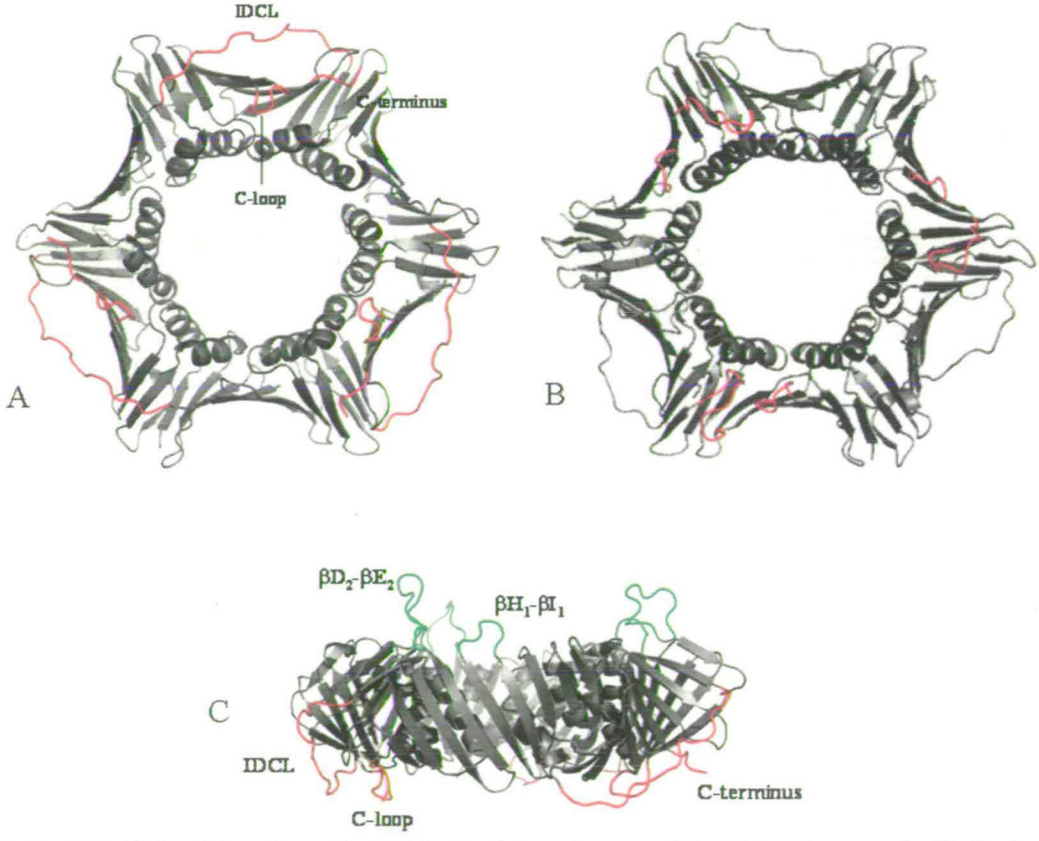
Trimeric PCNA forms a stable complex with DNA in solution down to a concentration of 20nM. At lower concentrations it dissociates into monomers and dimers (Podust et al., 1995). As one can see in Figure 1-1 the basic architecture of the sliding clamps is the same, which also holds for bacteriophage clamp gene 45 protein. In RB69 and T4 the ring is also a trimer with an internal diameter of about 35Å, but has rather a triangular appearance. The structural irregularities make it less stable than ScPCNA or the β clamp and it only manages to stay bound to DNA in the presence of the polymerase in contrast to hPCNA, which independently binds DNA once it is loaded (Yao et al., 1996).

1.2 THE TWO FACES OF PCNA

As mentioned above PCNA and its homologues are rather acidic proteins with pIs of 4.59 in budding yeast (Krishna et al., 1994), a pI of 5.38 in the β subunit of *E. coli* (Kong et al., 1992) and a pI of 5.05 in the T4 bacteriophage gp45 (Moarefi et al., 2000), which has a 77% sequence identity with the closely related bacteriophage

RB69 gp45 (Shamoo and Steitz, 1999). A colour-coded representation of the surface residues in Figure 1-3 shows clearly how both front and back side of the ring are covered mainly covered in acidic residues save for the central cavity where a few lysines and arginines create the positive charge. Despite the similar charge distribution the front and backside are very different in terms of structure due to the head-to-tail conformation of the monomers. A number of loops extend from the core fold as can be seen in Figure 1-4. The back side has a number of prominent loops which connect adjacent anti-parallel β strands sticking out into the solvent. The two main ones here are βD_2-βE_2 and βH_1-βI_1 (Figure 1-2).

Figure 1-4 The two faces of human PCNA. (Kontopidis et al., 2005)



A) C-side: Highlighted in red are the IDCL, the Centre Loop and the C-terminus on the B) Backside: Marked in red are the two long loops in each monomer C) Side view of PCNA: The loops on the backside marked in green are protruding into the solvent, while the loops in red on the C-side stay fairly close.

The other side of PCNA, termed the C-side due to the featured C-terminus, is especially marked by the interdomain connector loop (IDCL, residues 121-132 in hPCNA). In ScPCNA the βD_2-βE_2 loop and IDCL are highly antigenic and already in the early nineties were predicted to be involved in interactions with other proteins (Brand et al., 1994; Roos et al., 1993). Interestingly, a range of proteins in hyperthermophilic archaea like *Pyrococcus furiosus* and *A. fulgidus* have been found to contain shorter loops in comparison with homologues from eukaryotes (Chapados et al., 2004; Jaenicke and Bohm, 1998; Matsumiya et al., 2001; Vogt and Argos, 1997).

This difference is thought to increase heat stability of those molecules and is also observed in PCNA. In *P. furiosus* particularly loops βD_2-βE_2, βH_1-βI_1 as well as the C-terminus are significantly shorter.

1.3 THE INTERFACE CONTACT BETWEEN MOLECULES IN THE TRIMER

In sliding clamps the monomers are joined head to tail to form a ring. The internal interface consists of two β-sheets of two domains from different monomers (βD_2 and βI_2). Furthermore there is a hydrogen bond network which might be important for the orientation of the monomers at appropriate angles. For example, hydrogen bonds in hPCNA between Glu109 - Ser183, Val111 - Lys181, Asp113 - Asn179, Glu115 - Asn177 and Lys117 - Leu17 link two anti-parallel β-sheets of two adjacent monomers (Figure 1-5). Between the backbone of analogous residues seven (ScPCNA) or eight (hPCNA) clustered hydrogen bonds are formed as can be seen in

both yeast and human crystal structures (Gulbis et al., 1996; Krishna et al., 1994) (Table 1-2). Mutation studies with those residues have shown that they are very important for the stability of the trimer (Amin and Holm, 1996; Arroyo et al., 1996; Ayyagari et al., 1995; Jonsson et al., 1995).

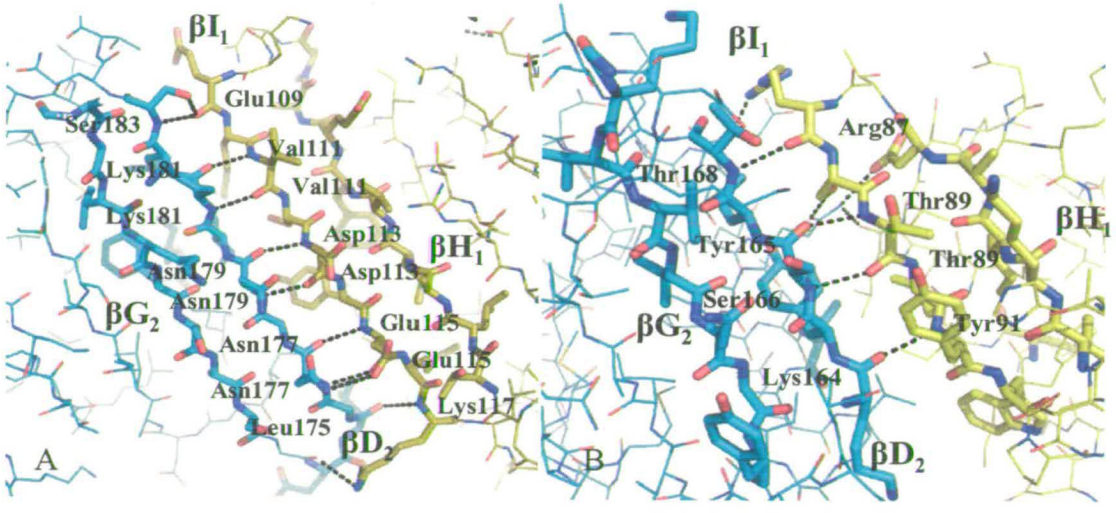


Figure 1-5 The intermolecular interface in the PCNA/sliding clamp trimer.
A) human PCNA (1VYM) (Kontopidis et al., 2005)B) RB69 gp45 (1B77) (Shamoo and Steitz, 1999). Two monomers are held together by main chain-main chain hydrogen bonds between β-strands of adjacent monomers, βD_2-βI_1. The interface of human PCNA is far more extended than in the viral sliding clamp and has eight hydrogen bonds.

Additionally, van der Waals forces within a hydrophobic core at the interface along with ion pairs hold the molecules together, although not all interactions are equally important as shown in mutation studies (Kelman and Odonnell, 1995a; Krishna et al., 1994). For example the mutation of a Y114A in hPCNA prevents the formation of trimers all together (Jonsson et al., 1995), S115P in ScPCNA has the same effect (Ayyagari et al., 1995), since the Pro residue is not usually found in β-sheets due to its conformation. Similarly, the mutation L108P in *E. coli* also leads to the disruption of dimerization (Yao et al., 1996).

As mentioned above, the T4 sliding clamp is less stable on DNA than is PCNA. This is due to the fact that the interface contacts are less exten sive compared to ScPCNA. Only four hydrogen bonds link two monomers instead of eight in the eukaryotic sliding clamp and the total surface area that is buried at the interface is only about two thirds of hPCNA (Table 1-2).

Characteristics / Protein	Number of hydrogen bonds at interface	Buried surface area in Å^2
Human PCNA	8	1014
***S. cerevisiae* PCNA**	7	888
***P. furiosus* PCNA**	5	1274
***E. coli* β subunit**	4.5	1285
RB69	4	692
T4	4	772

Table 1-2 Characteristics of the interface in some sliding clamps.
Distance for donor/acceptor < 3.9Å, probe size = 1.4Å.

In *E. coli* there is a head-to-tail arrangement of the monomers in the dimer. The average surface buried at the interface is 1285Å^2 which is similar to the surface buried at the interface in PfuPCNA. At least four strong hydrogen bonds are formed between adjacent β-sheets of two monomers. At the centre, as in the other structures, is a small hydrophobic core consisting of closely packed side chains Phe106 and Ile278 from one monomer and Ile272 and Leu273 from the other. Positioned around this core are six potential intermolecular ion pairs, while only a handful of ion pairs are found at the interfaces in the yeast trimer. Interestingly, all positively charged residues are contributed by one monomer and all negative ones by the second. This electrostatic complementarity might help with the head-to-tail assembly of the dimer.

Looking at different domains of life, the archaeon *Pyrococcus furiosus* grows optimally at 100C°. Proteins from thermophiles have been found to share certain features like increased numbers of ion pairs as surveys have shown (Jaenicke and Bohm, 1998; Vogt and Argos, 1997). Again, main chain amide-to-carbonyl hydrogen bonds are formed between anti-parallel β-strands βI_1 and βD_2 (Matsumiya et al., 2001). However, the number of bonds is reduced to only four in contrast to the eukaryotic clamps. Loops adjacent to βI_1 and βD_2 are somewhat shorter than in human or yeast PCNA as mentioned above, which might cause the differences in the hydrogen bond network. On the other hand the buried surface area is even more extended than in hPCNA. Additional stability within and between PCNA molecules, which is needed for a life in such a hot environment, is probably enhanced by a number of short ion pairs found between oppositely charged groups (Matsumiya et al., 2001). Ten of those ion pairs are located at a single intermolecular surface of PfuPCNA. In contrast to that, the trimeric conformation of hPCNA is supported by only five ion pairs per primer and even by fewer in yeast (Gulbis et al., 1996; Krishna et al., 1994).

1.4 THE HYDROPHOBIC POCKET

1.4.1 The interdomain connector loop and the hydrophobic pocket

The domain topology of each PCNA monomer places the last strand of the N-terminal domain remote from the first strand in the C-terminal domain and so a

linker strand (residues 121-132 in hPCNA) is an important feature, which additionally forms a part of a well defined hydrophobic pocket (Figure 1-1, Figure 1-2, Figure 1-4). As shown in Figure 1-6 in ScPCNA a number of hydrophobic residues from the linker strand reach out towards the hydrophobic pocket underneath to maximize van der Waals interactions. Additionall to those hydrophobic interactions hydrogen bonds are formed between polar side chains of the linker and amino acids in the β-sheet underneath. But not all the hydrogen bonding potential of the backbone is tied up in intramolecular interactions. The remaining hydrogen bond capacity is available to the binding of proteins, which is shown by complexes of PCNA with peptides of partner proteins. One example for that is the complex of PCNA with p21 (141-160) (Figure 1-7 A1 and A2). In all structures of sliding clamps this linker is found and the conservation of this feature might actually show a functional requirement (Figure 1-1).

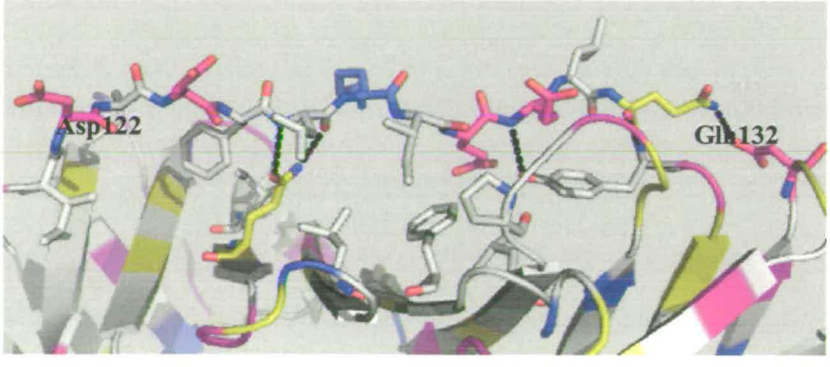


Figure 1-6 Intramolecular interactions of the IDCL (121-IDADFLKIEELQ-132) in ScPCNA. Acidic (magenta: E130, D200), hydrophobic (white: L126, Y133) and polar (yellow: Q38, Q132) side chains are involved in four hydrogen bonds. Furthermore a number of hydrophobic residues (e.g. L47, L126, I128, L236, F249) are engaged in van der Waals contacts.

A number of proteins have been shown to interact with this loop upon complex formation with sliding clamps (Eissenberg et al., 1997; Gulbis et al., 1996; Warbrick et al., 1995). DNA

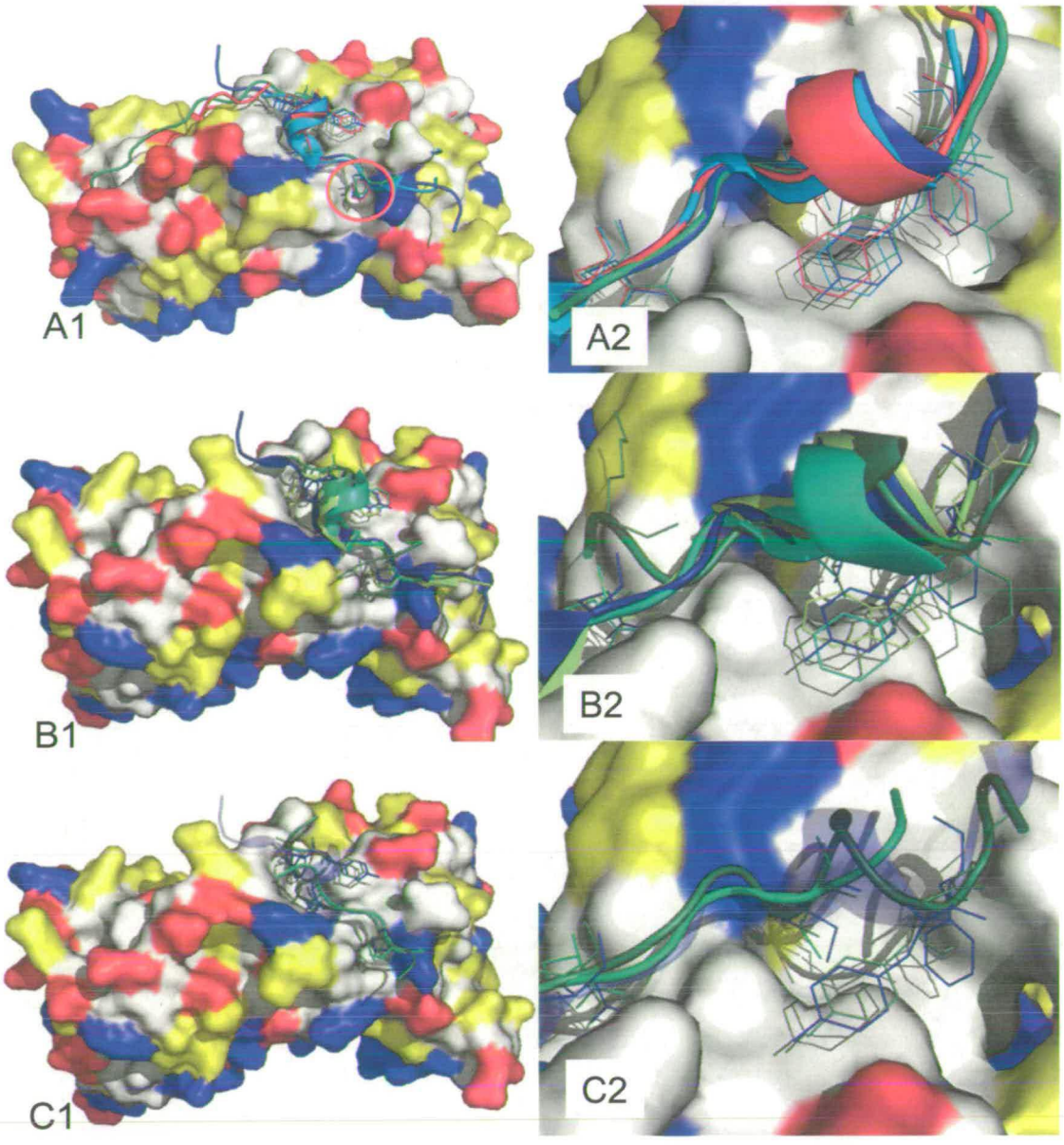


Figure 1-7 Peptide binding to the hydrophobic pocket in PCNA/sliding clamps is conserved throughout the domains of life.

The surface shown is that of a human PCNA monomer and viewed in A-C from the front/above. All peptide sequences are given in Figure 1-8. A1 and A2 show human peptides bound human PCNA: PL= deep blue (1VYJ) (Kontopidis et al., 2005), Fen-1 = red (1U7B) (Bruning and Shamoo, 2004), DNA polymerase δ = cyan (1U76) (Bruning and Shamoo, 2004), p21 = green (1AXC) (Gulbis et al., 1996). The cartoon backbone with the α-helical turn is placed at the same position in the pocket. Especially the glutamine side chain is fixed in the same position (Q-pocket, red circle in A1), a small groove a little further away from the hydrophobic pocket, in all four complexes. The hydrophobic residues, on the other hand, that grip into the hydrophobic pocket underneath the IDCL, vary a bit more (Figure 1-8). B1 and B2: Peptides of non human proteins complexed with the respective sliding clamp have been superimposed with hPCNA: *A. fulgidus* Fen-1 = yellow (1RXZ) (Chapados et al., 2004), *P.furiosus* RFC = lime (1ISQ) (Matsumiya et al., 2002), RB69 polymerase = forest green (1B8H) (Shamoo and Steitz, 1999). PL was added for comparison = deep blue. Despite their origin from both archaea and virus, the displayed peptides have retained the overall fit to the hydrophobic pocket including the α-helical turn. The side chains still grip into the their respective pockets, apart from RB69 pol the N-terminal end does not contain a glutamine and it also bend upwards compared to the other three peptides. C1 and C2: Peptides from *E.coli* polymerase DNA polymerase IV (1OK7) (Burnouf et al., 2004) (light green) and *S. cerevisiae* RFC (1SXJ) (Bowman et al., 2004) (forest green)

are shown with PL (shadow). Neither of them contains the typical α-helical turn, but still place two (*E.coli* pol IV) or four (ScRFC) residues in the pockets. It should be noted that N398 in ScRFC is placed in the appropriate pocket for the usual Q at that position, which can be explained by the similar chemical character of the two amino acids.

Polymerase δ (Eissenberg et al., 1997; Fukuda et al., 1995; Jonsson et al., 1998; Oku et al., 1998; Zhang et al., 1998) binds to this loop as well as to the hydrophobic pocket underneath (Figure 1-7 A1 and A2). Also for binding of p21 and Fen-1 the recognition of the loop is crucial (Eissenberg et al., 1997; Gulbis et al., 1996; Jonsson et al., 1998; Oku et al., 1998). It is probably also important for the maintenance of the ring structure and adds flexibility at opening of the ring (Tsurimoto, 1998).

Due to its high mobility the IDCL is rather unstructured and often difficult or impossible to model in most crystal structures and its conformation depends on the ligand binding to that area. However, both the IDCL and the C-terminus, which also tends to be flexible in unliganded structures, are rigidified upon complexing with a ligand (Kontopidis et al., 2005).

As the IDCL is conserved among organisms of all three domains of life including phages, so is the hydrophobic pocket underneath the loop. Most similar are the pockets in eukaryotic PCNA, as not only the fold is conserved but also a substantial part of the sequence as can be seen in Figure 1-2. A superimposition of the complex structures of RB69 sliding clamp with a DNA polymerase peptide and hPCNA with the PL peptide (Figure 1-7 B1 and B2, Figure 1-8) (Bruning and Shamoo, 2004; Kontopidis et al., 2005; Shamoo and Steitz, 1999) shows that also here the mode of interaction is more or less conserved. The PL peptide was designed from sequences of human DNA ligase 1 and *D. melanogaster* Pogo transposase. The aim was to find a tight binding peptide which did not contain the cyclin-binding motif of p21.

Among eukaryotes the pockets and hence modes of interactions with peptide are in fact so similar, that a peptide from one species like human p21 (141-160) can also bind to fission yeast PCNA (Reynolds et al., 2000).

141-KRRQTSMTDFYHSKRRLIFS-160	human p21	1AXC
1-SAVLQKKITDYFHPKK-16	Pogo/Ligase (PL)	1VYJ
453-ANRQVSITGFFQR-465	human DNA pol δ	1U76
336-TQGRLDDFFKVTG-348	human Fen1	1U7B
396-LDNMSVVGYFKH-407	yeast RFC	1SXJ
469-KQATLFDF--476	P.fur RFC	1ISQ
326-KSTQATLERWF-336	A.ful Fen-1	1RXZ
345-RQLVLGL-351	Ecoli pol	1OK7
893-KKASLFDMFDF-903	RB69 pol	1B8H

Figure 1-8 Sequence alignment of selected peptides and proteins that bind to sliding clamps in crystal structures with PDB codes.
In red are marked the residues that fit the PIP-box motif. The M in the RB69 polymerase is highlighted in green since it is not an aromatic, but does bind into the same pocket as Y, W or F.

1.4.2 Partner peptides binding to the hydrophobic pocket

Previously it has been shown that a number of proteins that bind to hPCNA contain a certain motif, termed the PIP box. The sequence is Q-x-x-h-x-x-a-a where 'h' stands for a moderately hydrophobic residue like methionine or leucine and 'a' means an aromatic residue (Warbrick, 2000). These four residues have been found to bind in the hydrophobic pocket (Figure 1-7). It is conserved in human Fen-1, DNA polymerase δ and p21 (Figure 1-8); all three of them bind to the hydrophobic pocket as well as making connections with the IDCL (Figure 1-7A1 and A2). This binding motif seems to be also conserved to some extent among other domains of life (Figure 1-7 B and C). The peptide of RB69 DNA polymerase in the complex with the phage

sliding clamp has the sequence, KKAS**L**FD**MF**DF, and even contains a single helical turn, just as it is the case in the complex of hPCNA with the human DNA polymerase δ peptide (ANRQVSITGFFQ) (Figure 1-7). Residues L, M and F the final F are all found to bind in the pocket, however, the glutamate, which is so important for the binding in hPCNA is not present and in its place K is just sticking out, not making any important contact in or near the pocket.

1.5 THE C-TERMINUS AND OTHER BINDING SITES FOR INTERACTING PROTEINS

There are two distinct faces on the trimer and there are loops protruding at each side. The C-terminus sticks out from the C-side (Jonsson and Hubscher, 1997) and in unliganded structures of hPCNA the C-terminal residues 255-261 tend to be disordered (

Figure 1-9) in the crystal structures. If the C-terminus is then deleted the clamp loader protein RFC cannot be stimulated to load PCNA onto DNA (Fukuda et al., 1995).

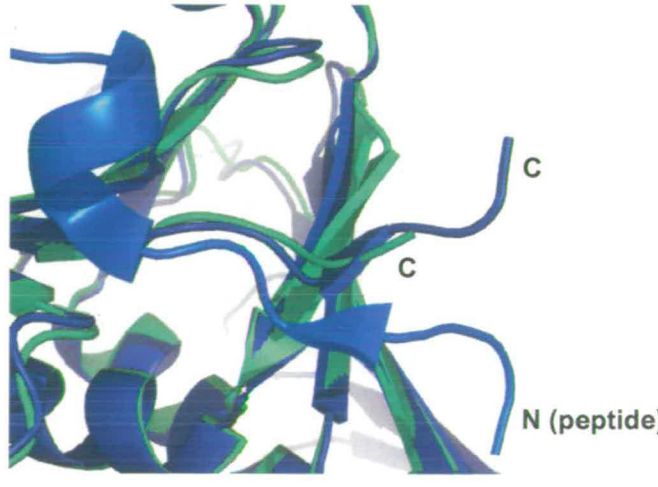


Figure 1-9 Superimposition of uncomplexed human PCNA and with PL.

The C-terminus is more ordered in the complex structure (1VYJ) (Kontopidis et al., 2005) (dark blue with peptide in light blue) and two more residues can be modelled compared to the unliganded structure (1VYM) (Kontopidis et al., 2005) (green).

This suggests that the C-terminal region is important for the interaction of RFC with the clamp (Naktinis et al., 1996). In *E. coli* it has also been shown that the C-terminus of the sliding clamp interacts with polymerase III (Naktinis et al., 1996), which is similar to DNA polymerase ε, but not like DNA polymerase δ in eukaryotes (Arroyo et al., 1996; Eissenberg et al., 1997; Fukuda et al., 1995). Different studies have also shown that the C-terminus is recognized as well by p21, Fen-1 and Gadd45 (Eissenberg et al., 1997; Gulbis et al., 1996; Hall et al., 1995b; Oku et al., 1998). In budding yeast residues Asp41 to His43 form the Centre Loop (Figure 1-2, Figure 1-4) at the C-side between the two domains in a monomer. Asp97 seems to be important for the interaction with RFC as mutation studies illustrate (Fukuda et al., 1995). Cold sensitive mutants of budding yeast showed further the importance of the residues Asp41 andAsp42 in the Centre Loop (Ayyagari et al., 1995) for RFC interaction, whereas residues 41-44 are needed for the binding of DNA polymerase δ(Oku et al., 1998).

In PfuPCNA the C-terminus is also shortened, just as some other loops in that structure, and this is thought to increase heat stability. Additionally, in the complex structure of PfuPCNA with a peptide of RFC it seems to play an important role in the interaction of both proteins (Matsumiya et al., 2002). In the complex of *A. fulgidus* PCNA with a peptide derived from Fen-1 the C-terminal open flexible loop is engaging in the interaction with the peptide and forms with it an anti-parallel intermolecular β-sheet (Chapados et al., 2004). This fact is thought to provide a control mechanism for mismatch repair.

Further residues that are important for protein interactions are 231-233 for the complex with p21 and 1-68 & 195-261 for binding to cyclin D (Tsurimoto, 1998).

1.6 POTENTIAL DOUBLE HOMOTRIMER COMPLEX FORMATION

Although many proteins with various functions have been found so far to interact with PCNA (Maga and Hubscher, 2003), none of them proved to bind to the back side of the trimer.

Crosslinking and SDS-PAGE experiments showed the presence of particles twice the size of a trimer, suggesting the existence of double homotrimers, although this complex seems to be of a rather dynamic, loose nature (Naryzhny et al., 2005). We should also note that this complex has never shown up in any of the crystal structures so far (Gulbis et al., 1996; Krishna et al., 1994). Deletion studies led to the conclusion that Arg5 and Lys110 are crucial for the formation of the double homotrimer. Also R5A/K110A double mutant CHO cells showed clearly that these residues are vital for survival. If PCNA does exist as double homotrimer in the cell, this could explain why no proteins seem to be binding to the back of the ring. However, it has been proposed before that the back side of PCNA is involved in regulation of the protein's functions, though no evidence has been found for that so far (Fukuda et al., 1995). Two front sides, on the other hand, would allow PCNA to work in both directions on DNA and help with the coordination of multiple functions as repair, replication and cell-cycle control.

DNA replication is coupled with chromatin assembly, which requires the association of PCNA and chromatin assembly factor-1, CAF-1, (Moggs et al., 2000; Shibahara and Stillman, 1999; Zhang et al., 2000). Unsurprisingly, the largest subunit of CAF-

1 has been shown to bind directly to the processivity factor (Zhang et al., 2000). Naryzhny et al. (Naryzhny et al., 2005) carried out a number of studies on interaction of PCNA with DNA polymerase δ and CAF-1, a complex of three subunits, p150, p60 and p48 (Smith and Stillman, 1989).

They found that the R5A/K110A PCNA mutant can only bind CAF-1 or DNA polymerase δ, while the wild-type protein is able to bind both simultaneously. This might come as a surprise since in theory there should be three binding sites, one on each monomer. But it appears that binding of one protein can cover up a large part of the trimer surface, as found in the complex of DNA ligase I (Pascal et al., 2004) and PCNA, and so exclude other proteins by competition/steric hindrance.

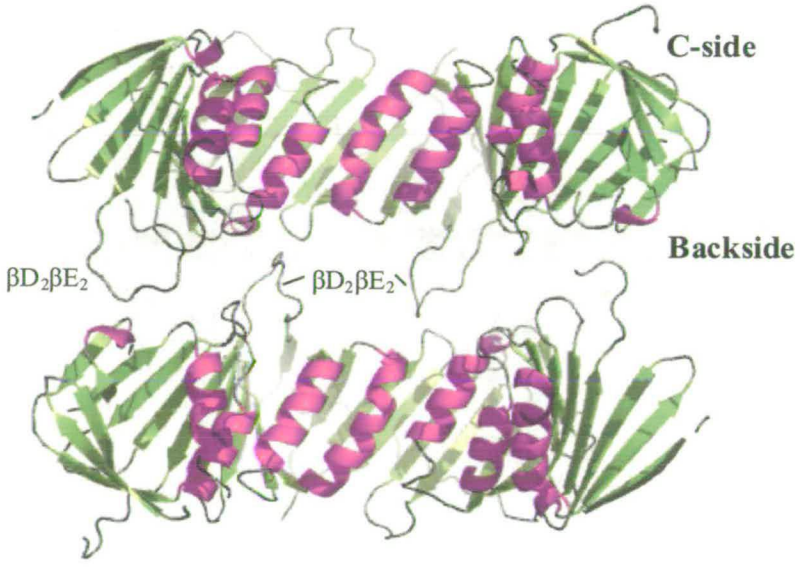


Figure 1-10 Model of the possible back-to-back arrangement of the double homotrimer complex of hPCNA. (Naryzhny et al., 2005)
The two rings of human PCNA (1VYM) are twisted against each other, this way the larger loop ($\beta D_2\beta E_2$) on the backside might fit into a grove on the other ring around αA_1 (residues 9-20), αB_2 (residues 209-211), βH_2 (residues 235-241) and βI_2 (residues 254-260). For clarity the structures are shown as a cross-section.

PCNA crystal structures have shown that the $\beta D_2\beta E_2$ loop sticks out from the back of the ring, so a direct interaction between R5 of one trimer and K110 of another might only be possible if the rings are twisted against each other to make the loops fit into

groves on the opposite trimer. The model proposed by Naryzhny et al.(Naryzhny et al., 2005) (Figure 1-10) suggests hence that a "locking-in" might occur by fitting the $\beta D_2\beta E_2$ loop into a grove on the other ring around αA_1 (residues 9-20), αB_2 (residues 209-211), βH_2 (residues 235-241) and βI_2 (residues 254-260).

Another group working on PCNA in *Drosophila melanogaster* also found that the results of their genetic studies only made sense if two PCNA trimers formed a back-to back complex (Henderson et al., 2000), so there seems to be a growing body of evidence to support this hypothesis.

1.7 ALTERNATIVE PCNA ARCHITECTURES

Dimeric PCNA structures like the *E. coli* sliding clamp have been also found in eukaryotes. In carrot two distinct PCNA genes have been found in somatic embryogenesis (Hata et al., 1992). One shows the typical 264 amino acids in trimeric PCNA monomers, the other consists of 365 residues and is 40kDa in size, which is very much like the *E. coli* monomer. Furthermore, in *Xenopus laevis* a 43kDa gene product has been found (Leibovici et al., 1990). The reason for the existence of alternative architectures might be the need of rapid chromosomal replication during embryogenesis (Kelman and Odonnell, 1995b).

On the other hand, an alternative PCNA gene product has been found in *E. coli*. Its internal promoter is UV inducible and is located within the *dnaN* gene which leads to the formation of a 26kDa peptide. Crosslinking and size exclusion experiments have shown that this alternative form exists as a homotrimer and stimulates DNA

polymerase III. It has been suggested that it confers increased UV resistance for DNA repair (Skaliter et al., 1996).

A different alternative to the common homotrimer comes from the archaeon *Sulfolobus solfataricus*. Dionne et al. (Dionne et al., 2003) detected three PCNA homologues in the genome sequence: PCNA 1, 2 and 3. Though PCNA1 and 2 can form dimers together, PCNA3 is unable to interact with either of the proteins individually and only a mix of all three homologues leads to the formation of a heterotrimer with 1:1:1 stoichiometry. Their work furthermore suggested that Fen-1, DNA ligase I and DNA polymerase B1 can associate with the trimer simultaneously and act possibly as a preassembled processing complex. This, of course, would facilitate the coupling of both DNA synthesis and Okazaki fragment processing. Also it could be shown that the homotetramer of the small subunit of RFC is binding only to the PCNA1-2 dimer and that the single large subunit of RFC is contacting PCNA3, which is the last of the three to join the ring (Dionne et al., 2003). Other archaea seem to possess several genes for PCNA homologues as well, though it is so far unknown, if these form heterotrimers (Daimon et al., 2002). It is widely accepted that hyperthermophilic archaea are the most ancient of this domain of life, which makes them closer to the common ancestor of both eukarya and archaea. Hence it is thought, that the heterotrimer might be an evolutionary fossil of original sliding clamp before it specialized into homotrimeric PCNA for replication and the heterotrimeric 911 complex in *S. pombe*, which is denoted for DNA repair (Parrilla-Castellar et al., 2004).

1.8 OPEN CLAMP STRUCTURE IN THE CLAMP-LOADING COMPLEX

The RFC is essential for loading PCNA onto DNA at template-primer junctions by multiple stepwise ATP-binding events. This heteropentameric complex belongs to the AAA^+ family and consists of Rfc1, Rfc2, Rfc3, Rfc4, and Rfc5 subunits. Although crystal structures so far show PCNA always in the planar ring structure, latest electron microscopy technology has enabled us to have a look at the archaeal RFC-PCNA complex of *Pyrococcus furiosus* at 12Å resolution (Miyata et al., 2005). It shows a possible intermediate state in the process of clamp-loading. In this state PCNA is kept open in the complex by RFC and assumes a helical shape. The intrinsic planarity found in crystal structures has been deformed (Figure 1-11), giving PCNA the opportunity for a broader contact interface to RFC binding possibly all five subunits of the clamp loader. This contrasts with the ScPCNA-RFC complex where the sliding-clamp only makes contact to three subunits of RFC (Bowman et al., 2004).

This alternative mode of interaction between these two proteins may be essential for the opening of the PCNA ring. A possible model was proposed by Miyata et al. that PCNA is converted to the open form upon RFC binding. Binding of RFC to DNA may induce a conformational change in this protein and consequently, under ATP hydrolysis, PCNA will reduce its contact surface to RFC and return into the planar ring structure encircling DNA.

Molecular dynamics simulations on PCNA have furthermore suggested that PCNA is oscillating between planar and non-planar conformations as long as it is not

constrained to the ring structure (Kazmirski et al., 2005). These simulations also show that when choosing the non-planar state the molecule distinctly seems to prefer a right-handed spiral distortion. The authors set up their simulations by removing one of the three subunits of the PCNA trimer in yeast, human and archaeal PCNA. They assumed stability of the open form as the T4 bacteriophage clamp has been shown to be predominantly an open trimer in solution (Millar et al., 2004). Also the bacteriophage clamp appears to be open at one interface in the presence of the clamp-loader complex suggesting stability of the open ring.

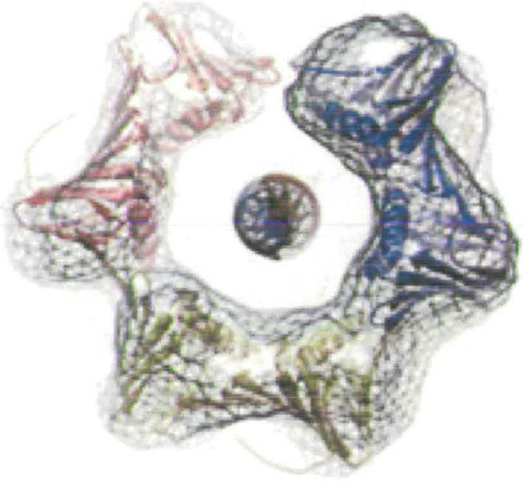


Figure 1-11 The open clamp structure of PfuPCNA observed with EM. (image taken from (Miyata et al., 2005))
The image shows that in the complex with RF-C (not shown) the ring structure is opened in order to be able to encircle DNA.

The results showed spontaneous lateral relaxation of the subunits and transient adoption of spiral conformations. The flexibility of conformations lies in the interface of the two PCNA subunits where the hydrogen bonds between the β-strands

are maintained. The right-handed twist in the anti-parallel β-sheets seems to lead to the preferred formation of a right-handed spiral.

1.9 PCNA'S ROLE IN THE CELL CYCLE

1.9.1 General

As mentioned above, PCNA is involved in various aspects of DNA metabolism. In eukaryotes replication occurs at about 500 too 5000 bases per minute while accuracy is very important. DNA lesions and other obstacles can lead to replication failure or even broken chromosomes with fatal consequences for cells. Therefore cells have a need for numerous safeguard mechanisms that are directly linked to DNA synthesis. This setup allows not only for replication to proceed despite of problematic regions, but also facilitates repair of those areas or induces cell-cycle arrest through check point pathway in case of serious damage. The following section explains how PCNA is involved in the various aspects of DNA metabolism.

1.9.2 DNA synthesis, mode of action and coordination of replication events

In DNA replication PCNA is loaded onto primed DNA by Replication Factor C (RFC) with the C side facing the growing end. This ensures polymerases and other enzymes binding to this side are close to the 3' end of the elongating DNA. The next step is the association of polymerases Pol δ or Pol ε. Enhanced by PCNA these two proofreading enzymes carry out most of the DNA synthesis. On the lagging strand each Okazaki fragment is left with a flap structure that needs to be removed and the

fragments to be joined up. For these steps Fen-1 and DNA ligase I associate with PCNA.

Having no enzymatic activity on its own, PCNA is able to act as a moving platform for various factors that need to act sequentially for highest efficiency. As mentioned above, most proteins that are known to bind to PCNA contain the PIP-box motif and are hence competing with each other for the three binding sites on the sliding clamp. Various studies (Chapados et al., 2004; Pascal et al., 2004), suggest that PCNA coordinates Okazaki fragment processing and joining in a stepwise fashion. Of all proteins tested for affinity to PCNA p21 is the one which binds most tightly, which makes sense as this interaction is most important when DNA replication has to stop upon DNA damage.

In order to prevent rereplication at the origin once replication has been initiated, several mechanisms are in place. PCNA is contributing by its interaction with Cdt1, a licensing cofactor which is essential for initiating DNA replication at the origin. By binding to PCNA *via* its PIP-box motif this factor is marked for ubiquitylation and therefore degradation (Arias and Walter, 2006).

1.9.3 Bypass replication and ubiquitylation of PCNA

Although accurate replication is most desirable for the cell, it is also important that synthesis continues despite obstacles on the DNA. When DNA damage is detected bypass mechanisms are triggered that lead to ubiquitylation of PCNA. The sliding clamp and ubiquitin are linked *via* the conserved K164 in PCNA and K63 in ubiquitin. This modification can be a single ubiquitin moiety or a polyubiquitin chain. While the first kind triggers the error-prone translesion synthesis (TLS)

pathway, the latter one leads to an error-free bypass mode. This mechanism has been coined the "ubiquitin-PCNA switch" (Hoege et al., 2002) and allows both passing of lesions as well as a choice between bypass modes.

1.9.4 Maintaining genomic stability by prevention of sister-chromatid recombination

Even in the absence of DNA damage a fraction of PCNA remains modified in S phase by the ubiquitin-related SUMO *via* K164 as observed in ubiquitylation (Arakawa et al., 2006; Hoege et al., 2002). SUMOylation of PCNA has been shown to recruit Srs2, a helicase with antirecombinogenic properties. This interaction leads to the disruption of Rad51 nucleoprotein filaments that are essential for recombination (Pfander et al., 2005). This mechanism is there to control sister-chromatid recombination as a salvage pathway for stalled replication forks which can lead to gross chromosomal rearrangements.

1.9.5 DNA Repair Systems

1.9.5.1 Mismatch Repair (MMR)

The DNA Mismatch Repair system (reviewed by (Jiricny, 2006)) recognises and corrects base-base mismatches, but also fixes small insertions or deletions. The erroneous parts of the newly synthesised DNA strand are removed which is followed by targeting the DNA-synthesis machinery to the single-stranded gap. The MMR components are sensor complexes MSH2-MSH6 or MSH2-MSH3 for recognising

smaller or bigger loops/deletions respectively, the transducer complex MHL1-PMS2, exonucleases like EXO1 that remove the flawed DNA stretch, single-stranded binding protein RPA, the nonhistone chromatin protein HMGB1 (only in higher eukaryotes) and the DNA synthesis machinery.

PCNA is required for the early steps in the MMR process (Umar et al., 1996) and interacts directly with MSH6, MSH3, MLH1 and EXO1 while the first three proteins also contain PIP boxes. Since MMR is strand-specific, i.e. it always works on the newly synthesised strand, PCNA is thought to act as some sort of marker along with a gap or nick in the DNA originating for example from Okazaki fragments (Modrich, 2006). Furthermore, PCNA is likely to recruit EXO1 to the site followed by its direct part in DNA synthesis. As binding of sensors, transducers and any of the excision or DNA synthesis proteins to PCNA is mutually exclusive, this arrangement seems to ensure the ordered course of MMR.

1.9.5.2 Base Excision Repair (BER)

The Base Excision Repair system (reviewed by (Sancar et al., 2004)) is responsible for the correction of chemical alterations of DNA bases. Degenerative influences leading to such alterations include alkylation, oxidation, reduction and deamination.

BER starts with the recognition of damaged sites by specific DNA glycosylases. These enzymes catalyse the removal of bases from the nucleotides. The remaining phosphodiester bond at the 5' end of such abasic sites is cleaved by apurinic/apyrimidinic (AP) endonucleases. Some enzymes like NTH1 combine both activities in one molecule, while most other enzymes carry out either the one or the

other reaction. After base removal the single nucleotide gab is filled by Polβ followed by sealing of the nick through XRCC1-DNA ligase 3. Alternatively, for bigger lesions the RFC-PCNA-Polδ/Polε complex is recruited to the site to perform DNA synthesis involving Fen-1 and DNA ligase I, which makes this process analogous to lagging strand synthesis.

As in MMR PCNA seems to be involved both in early and later steps of this system. Uracil-DNA glycosylase UNG2 contains PIP-box motif and interacts directly with PCNA during the initial steps of BER. Later AP endonucleases bind to the sliding clamp. NTH1 has also been found to bind directly to PCNA, but does not contain the PIP-box residues (Oyama et al., 2004). It is worth pointing out that in this repair pathway PCNA is not only colocalizing with the mentioned repair enzymes, but is also able to stimulate enzymatic activities in its binding partners (Tsuchimoto et al., 2001; Unk et al., 2002).

1.9.5.3 Nucleotide Excision Repair (NER)

Nucleotide Excision Repair is a system for removal of bulky DNA lesions originating from exposure to radiation or chemicals (reviewed by (Sancar et al., 2004)). The process includes detection by XPC, which is followed by unwinding and opening up the damaged DNA stretch by helicases XPD and XPB. Endonucleases XPG and XPF cut out the lesion generously and the single stranded gap is then closed through DNA synthesis and ligation. XPG contains a PIP-box motif and binds to PCNA *via* that region. In this interaction PCNA is thought to coordinate DNA excision and synthesis as has been suggested for the other repair systems.

1.9.6 Chromatin assembly

PCNA's function in nucleosome assembly is mediated by its interaction with heterotrimeric chromatin assembly factor CAF-1. CAF-1 is targeted to sites of DNA synthesis by PCNA *via* its PIP-box motif (Krawitz et al., 2002; Zhang et al., 2000). There it is responsible for delivering histones H3 and H4 to the replication fork for chromatin assembly (Smith and Stillman, 1989). PCNA-mediated CAF-1 recruitment is also observed at sites of NER where it is thought to aid chromatin assembly after the repair process is finished (Gaillard et al., 1996).

1.9.7 Epigenetic inheritance

The maintenance of epigenetic information encoded as modification on DNA and chromatin is important for cellular differentiation. This information is copied with high fidelity over rounds of cell division by coupling the process of DNA replication directly to enzymes required for preservation of epigenetic patterns.

Cytosine methylation on C5 is found mainly in CpG islands where it induces gene silencing in promoter regions and is controlled by PCNA (Bird, 2002). Maintenance of the methylation pattern is carried out by DNA cytosine methyltransferase 1 (DNMT1) (Hermann et al., 2004). PCNA targets this protein to replication forks by binding to its PIP-box motif located in the N-terminal region. It has been shown that PCNA not only anchors DNMT1 to these sites, but is also responsible for stimulating enzymatic activity (Chuang et al., 1997; Iida et al., 2002). Unsurprisingly, PCNA-mediated recruitment of DNMT1 is also found at site of DNA repair to restore

epigenetic information to newly synthesized DNA sequences (Mortusewicz et al., 2005).

Furthermore, PCNA might also be involved in gene silencing as it has been found to interact with histone deacetylation enzyme 1 (HDAC1) *in vitro* and both proteins colocalize at replication foci *in vivo* (Kovacevic et al., 2002).

1.9.8 Cell-cycle control and survival

As previously mentioned, binding of p21 to PCNA is effectively blocking out many other proteins containing the PIP-box motif. Exclusion of proteins (including Polδ, Fen-1, Ctd1 and DNMT1) (Arias and Walter, 2006; Chuang et al., 1997; Ducoux et al., 2001; Warbrick et al., 1997) prevents cell-cycle progression until all damage is repaired.

PCNA is also involved in processes that determine cell survival or death. Binding of the tumour suppressor protein ING1b to PCNA *via* a PIP-box motif is 10-fold increased after UV radiation leading to apoptosis (Scott et al., 2001). In contrast to that PCNA performs also an antiapoptotic role through interaction with proteins of the Gadd45 family by inhibiting their activities (Azam et al., 2001; Vairapandi et al., 1996). As described in further detail in Chapter 4 these proteins are implicated in growth control, apoptosis and DNA repair. They are very similar in sequence, but do not contain the specific PCNA-binding motif.

1.9.9 Regulation of PCNA functions

Since PCNA is involved in so many crucial processes of cell cycle progression and regulation it is important to control its ability to associate with the multitude of partner proteins in order to prevent confusion. PCNA's partner proteins could be ordered according to their binding affinities with p21 right at the top, virtually jamming the binding pocket. On the other hand, this hierarchy can be altered through phosphorylation of each factor. Fen-1 for example dissociated from PCNA after phosphorylation by the CDK2-PCNA complex (Henneke et al., 2003).

As PCNA provides three binding sites for PIP-box containing proteins, it is possible to imagine that the sliding clamp can bind three different proteins at once. An organization like that would be very handy in stepped processes like the repair systems. That this may be a possible scenario has been shown in studies with heterotrimeric PCNA from the archeon *Sulfolobus solfataricus*. Here each of the subunits binds specifically the polymerase, Fen-1 or DNA ligase I (Dionne et al., 2003).

Another form of controlling PCNA interactions is through modifying the clamp itself. As described above ubiquitinylated or SUMOylated PCNA attracts specific binding partners like Srs2 (Papouli et al., 2005) involved in prevention of sister-chromatid recombination or translesion synthesis polymerases (Bienko et al., 2005). The addition of a SUMO moiety to PCNA has been shown to repress binding of PIP-box containing proteins (Moldovan et al., 2006).

Since only a small fraction of PCNA molecules in a cell are actually modified at any given time, the hypothesis is that either only certain PCNA molecules are selected for modification or that this alteration is generally of a transient nature.

Another way of PCNA regulation is the unloading of the protein from DNA and subsequent degradation. Unloading is mediated by RFC, which is also needed for loading PCNA into DNA in the first place. Human PCNA degradation seems to be triggered by PCNA dephosphorylation. The phosphate group is added by the nuclear version of epidermal growth factor receptor (EGFR) (Wang et al., 2006) and phosphorylated PCNA is stably associated with nucleosomes. Dephosphorylated PCNA and those molecules which cannot be phosphorylated are marked by ubiquitylation for degradation by the proteasome.

Despite the numerous interactions PCNA engages in, the sliding clamp does not seem to bind directly to members of the checkpoint cascade that are responsible for cell-cycle arrest in case of DNA damage. This part is covered in eukaryotes by the PCNA-related 9-1-1 complex and features association of certain 9-1-1 subunits with defined checkpoint proteins (Parrilla-Castellar et al., 2004). It is assumed that PCNA and 9-1-1 have evolved from the same ancestor to separate replisome and checkpoint components to allow ongoing replication without any hindrance due to bound checkpoint signalling proteins.

1.9.10 PCNA and diseases

It has been shown that tumours of a similar pathologic state can result in very different clinical outcomes. Being such a central protein to cell-cycle control, dysregulation of PCNA has been an attractive parameter as a biomarker for cancer prognosis.

Studies have found that in esophageal squamous-cell carcinoma the main mechanism of progression after invasion of carcinoma is due to cell proliferation (Kuwano et al.,

1998), hence PCNA has been suggested as a profiling marker in cancer prognosis. PCNA expression as a biomarker in characterizing the degree of proliferative disregulation has also been suggested after a study on head and neck tumorigenesis (Shin et al., 1993). No significant connection could be established, however, for colorectal cancer and clinical outcome (Lyall et al., 2006).

Furthermore, the link between DNA damage and repair systems in neurodegenerative processes has been examined. It seems that both p53 and MSH2 are involved in neurodegenerative processes as well as part of the recognized cancer prevention machinery. Many human pathological conditions that exhibit genetic defects in DNA damage control also show neurological deficits (Rolig and McKinnon, 2000), which implies an important role for DNA repair proteins neurological development.

The deficiency in p53 is also associated with learning difficulties and has an effect on behaviour in mice (Amson et al., 2000). While p53 levels were elevated in non-proliferating cultures of cortical neurons upon beta-amyloid injury associated with Alzheimer's disease (Copani et al., 2001), MSH2 and PCNA translocation to the nucleus was observed in differentiated human neuroblastoma cell lines after oxidative insult with H_2O_2 (Uberti et al., 2003).

It is likely that PCNA is involved in neurological process by dint of p53 and MSH2 association, either directly or indirectly.

1.10 APOPTOSIS-LIKE CELL DEATH IN YEASTS SIMILAR TO THAT OBSERVED IN METAZOANS

Despite the fact that neither p53 nor p53-upregulated p21 homologues are present in yeasts, they have proved to be ideal model systems for studying function of mammalian proteins involved in apoptosis. Both budding and fission yeast display an apoptotic-like cell death induced by DNA damage or other stimuli suggesting similarities with mammalian p53-independent apoptotic pathways. This programmed cell death is accompanied by specific symptoms that are also presented by metazoan cell undergoing the process (Ink et al., 1997; Jurgensmeier et al., 1997; Madeo et al., 1997; Tao et al., 1997; Watanabe et al., 2002). These markers include the production of reactive oxygen species (ROS), chromatin condensation, DNA fragmentation and the release of cytochrome *c*.

Induction of this apoptosis in yeasts requires homologues of cell cycle checkpoint proteins that have been shown to be involved in metazoan cell death such as kinases ATM and ATR, 9-1-1 complex and Chk1 and Chk (Qin and Li, 2003). In fact, initial genetic studies to understand the DNA damage responses were carried out in yeasts. Expression of human and mammalian pro-apoptotic proteins in yeasts led to cell death (Ink et al., 1997; Jurgensmeier et al., 1997; Ryser et al., 1999), but these unicellular organisms also harbour an intrinsic apoptotic pathway (Fröhlich and Madeo, 2000), which can be induced by exposure to low concentrations of acetic acid or UV radiation for example. This programmed cell death is distinctive from a passive cell death, also called necrosis, upon exposure to high concentrations of acetic acid or UV radiation (Del Carratore et al., 2002; Ludovico et al., 2001).

1.11 PROJECT OUTLINES

1.11.1 Crystallization and structure determination of *S. pombe* PCNA

As described above both budding and fission yeast are model organisms which have been used to determine the genetic background for many complex processes in mammalians. As both human and budding yeast PCNA structures have been solved it is just a logical progress to examine the fission yeast counterpart. In many ways fission yeast is more closely related to humans than budding yeast and comparison between the eukaryotic structures was carried out to gain further insight in conserved protein-protein interactions.

1.11.2 Screening for and testing peptidomimetic anti-cancer drug leads to block the hydrophobic pocket of human PCNA for DNA polymerase

Experiments with antisense oligonucleotides that caused a suppression of PCNA expression showed selective inhibition of gastric cancer cell proliferation both *in vitro* and *in vivo* (Sakakura et al., 1994). Based on these findings it seems reasonable to suggest that a similar inhibitory effect could be achieved, if PCNA was kept from interacting with partner proteins involved in DNA replication as has been previously suggested (Kontopidis et al., 2005). A number of different proteins bind to the hydrophobic pocket beneath the IDCL. A small molecule binding to that place by

mimicking these peptides should inhibit PCNA action, stop cell proliferation and hence pave the way to a new class of anti-cancer drugs.

Chemical data bases were screened for suitable compounds employing the programme LIDAEUS and selected molecules were analyzed for binding to human PCNA using a thermal shift assay and a competitive assay employing labelled PL peptide.

1.11.3 Analysis of the non-PIP-box-mediated interaction between human PCNA and human Gadd45

Although most proteins that have found to bind directly to PCNA exhibit the conserved PIP-box binding motif, some binding partners like Gadd45 do not contain this sequence. The third objective of this thesis was to determine exactly the section of Gadd45 that mediates direct contact with PCNA and to analyze the kinetics underlying this interaction using SPR (Surface Plasmon Resonance).

CHAPTER 2. Crystal structure and biophysical characterization of *Schizosaccharomyces pombe* PCNA unliganded and in complex with a human p21 peptide

2.1 INTRODUCTION

Schizosaccharomyces pombe is also referred to as fission yeast and is used as a model organism in molecular and cell biology. Its genome with 13.8 million base pairs is very short for an eukaryote which makes this organism attractive to work with. Its genome has been fully sequenced and many genes homologous to human disease genes, including diabetes and cystic fibrosis, have been identified (Wood et al., 2002). Furthermore, many fission yeast proteins involved in cell cycle regulation and DNA synthesis are equivalent to the ones present in human pathways, PCNA being one of them.

Knowledge of the structure of SpPCNA will complement what is known about eukaryotic sliding clamps. In particular the conservation of the hydrophobic pocket, as described in Chapter 1, is of interest. This is the site where numerous PCNA partner proteins that contain the PIP-box motif (e.g. p21) bind to the sliding clamp. Despite the fact that no homologue of p21 has been found in fission yeast, the human p21 20-mer that has been complex with hPCNA by Gulbis and his colleagues (Gulbis et al., 1996) was found to bind SpPCNA (Reynolds et al., 2000). A search of the fission yeast proteome for PIP-box containing proteins found plenty of hits as discussed in section 2.3.6.

In order to gain a deeper understanding of the mechanisms of protein:protein and protein:ligand recognition/binding, SpPCNA was crystallized on its own and in

complex with the human p21 peptide. The interaction between SpPCNA and p21 is unlikely to be biologically relevant, but shows how small differences in the target protein can change the binding mode of a particular ligand.

To characterize the interaction between SpPCNA and the p21 peptide surface plasmon resonance (SPR) and a thermal shift assay were employed.

SPR allows monitoring the interaction between molecules in real time. The technology behind it involves attaching one interacting partner, called ligand in SPR terminology, to the surface of a sensor chip. The analyte or second interaction partner is passed over the surface. Binding of the analyte to the sensor surface generates a response which is proportional to the bound mass, and changes in the amount of bound analyte can be detected down to a few picograms or less per mm^2.

Due to the ability to observe interactions in real time, SPR can provide information about the specificity, the kinetics and affinity of a particular interaction.

The fluorescence-based thermal shift assay is a general method for identification of ligands for target proteins. Mostly it is used for isolating inhibitors of the protein of interest from compound libraries. Using an environmentally sensitive fluorescent dye to monitor protein thermal unfolding, the ligand-binding affinity can be assessed from the shift of the unfolding temperature (ΔT_m) obtained in the presence of ligands relative to that obtained in the absence of ligands. The background to this method is further described in section 3.2.1.

2.2 MATERIAL AND METHODS

2.2.1 Peptides for crystallization and assays

The human p21(141-160) peptide was provided by Dr. Emma Warbrick (University of Dundee) as a 10 mg/ml solution in 0.1% trichloroacetic acid. Aliquots were stored at -20°C.

IGHHHH, HHHHHH and p21 (141-152) peptides were synthesised by Peptide Protein Research Ltd at >90% purity. DMSO stock solutions were produced at 50mM for IGHHHH and p21(141-152) and at 25mM for HHHHHH, and stored at -20°C.

2.2.2 Expression and purification of SpPCNA

The N-terminally MetArgSer(6xHis)GlyIleProMet-tagged fission yeast PCNA (30.4kDa) was expressed in *Escherichia coli* M15 [pREP4] (Qiagen) from a pQE32-Pcn1 expression vector (Qiagen), which was kindly provided by Dr. Stuart A. MacNeill (University of Copenhagen). The tag added by this vector to simplify the purification process is uncleavable, but does not interfere with trimerization or functionality in *S. pombe* (Reynolds et al., 2000). Expression and purification basically followed an established protocol (private communication from Dr. Stuart MacNeill). Essentially it was carried out as follows.

Cell cultures (1L) were grown shaking in LB medium at 30°C to an OD_{600} = 0.7 in the presence of 100 µg/ml ampicillin and 25 µg/ml kanamycin. Expression was then induced with 1 mM IPTG and the cells were incubated for another 6 hours. Upon

harvest, cell pellets were suspended in 40 ml Lysis Buffer A (50 mM NaH_2PO_4, 150 mM NaCl, 20 mM imidazole pH 8.0, 1 mM PMSF and 0.1% Triton-X) and sonicated with 4 x 15 second bursts. Cell lysate was cleared by centrifugation for 25 minutes at 10000 x g and the supernatant was filtered through a 0.2µm filter. The lysate was applied to a 5 ml Ni-NTA agarose K9/15 column (Amersham) that had been equilibrated previously with Buffer X (50 mM NaH_2PO4, 300 mM NaCl and 20 mM imidazole pH 8.0). PCNA was eluted with a gradient of Buffer X and Y (50 mM NaH_2PO_4, 150 mM NaCl and 250 mM imidazole pH 8.0) at 135-145 mM imidazole and pooled fractions were dialysed against 20 mM Tris pH7.5, 1 mM DTT, 0.5 mM EDTA, 100 mM NaCl (Figure 2-1). The protein at an estimated purity of >95% was then concentrated to 4-5 mg/ml and stored at 4°C.

Figure 2-1 Purification of *S. pombe* PCNA using a Ni-NTA column.
A 12% SDS-Polyacrylamide gel was run and subsequently stained with Coomassie. The total yield of cell pellets derived from 1L liquid culture was 4 mg of protein at high purity. Key: lane 1 = marker, lane 2 = lysate, lane 3 = flow through, lanes 4 - 6 = wash steps at 40 mM, 61.5 mM and 100 mM respectively, lanes 7 -19 = consecutive elution fractions from a gradient between 100 mM and 150 mM imidazole. His-tagged SpPCNA runs slightly larger than the calculated 30.4kDa, which is coherent with the literature (Reynolds et al., 2000).

2.2.3 Expression and Purification of hPCNA

The wild type human PCNA expression plasmid (pT7-hPCNA) was kindly provided by Emma Warbrick (Fien and Stillman, 1992). The protein was over-expressed in *E.*

coli BL21 (DE3) cells grown in LB medium containing 100 μg/ml ampicillin at 37 °C to an OD_{600} of 0.7. Over-expression was induced by addition of 1 mM IPTG and cells were grown for another 3 hours before harvesting. Cell pellets were either resuspended in ice-cold 10% sucrose and 25 mM Tris pH 8.0, snap-frozen in liquid nitrogen and then stored at -80 °C, or processed immediately by adding 40 ml of ice-cold Lysis Buffer B (25 mM Tris pH 8.0, 1 mM EDTA, 25 mM NaCl, 2 mM benzamidine, 1 mM PMSF, 1 mM DTT). Lysis was performed by ultra sonication for 8 x 20 seconds with 40 second breaks in between bursts. All manipulations were carried out on ice. Following sonication, the lysate was then centrifuged for 30 minutes at 10000 x g and the supernatant filtered through 0.2 μm syringe filters prior to application onto a 40 ml Q sepharose (Amersham) XK 16/20 column pre-equilibrated with Buffer A (25 mM Tris-HCl pH 8.0, 1 mM EDTA, 10% glycerol, 2 mM benzamidine, 1 mM PMSF, 1 mM DTT, 100 mM NaCl). The protein was eluted in a two-step process using Buffer A supplemented with 700 mM NaCl. The column was washed with 310 mM NaCl and protein eluted with 400 mM NaCl. Fractions were analysed by 12% acrylamide SDS-PAGE, pooled according to PCNA content and dialyzed against 5L of Buffer B1 (20 mM Na-acetate pH 7.5, 2 mM DTT, 0.5 mM EDTA and 10% glycerol) over night, with two buffer changes. The dialysate was filtered through 0.2 μm syringe filters prior to application onto a 30 ml Heparin sepharose (Amersham) XK 16/20 column pre-equilibrated with Buffer B1. PCNA was eluted with a NaCl gradient from 0 – 500 mM in Buffer B1. PCNA elutes between 100 mM and 120 mM NaCl. The peak fractions were then pooled and concentrated using a spin concentrator (molecular weight cut-off: 10 kDa) to approximately 2ml. Concentrated PCNA was then applied to a Sephacryl S-200 HR

gel filtration column (Amersham) ($V_t \approx 120$ ml; 16 cm x 60 cm) pre-equilibrated with Buffer B1 supplemented with 100 mM NaCl. This step was selectively repeated to achieve higher purity. The protein was concentrated to 10-20 mg/ml and stored at 4 °C in Storage Buffer (25 mM Tris-HCl, pH 7.5, 1 mM EDTA, 0.01% NP-40, 10% glycerol, 1 mM DTT, 2 mM NaCl).

2.2.4 Crystallization of *S. pombe* PCNA

2.2.4.1 Crystallization of monoclinic form unliganded SpPCNA crystals

Crystallization of unliganded SpPCNA and the PCNA/p21(141-160) complex yielded crystals in various conditions. The best unliganded crystals were obtained by the hanging-drop, vapour diffusion method at 18°C in a condition from Structure Screen 1 (200 mM Lithium sulphate, 100 mM Tris pH 8.5, 30%(w/v) PEG-4000) (Molecular Dimensions Limited). This was optimized to 200 mM Lithium sulphate, 100 mM Tris pH 8.5 and 25% w/v PEG-4000. Hanging drops were obtained by mixing 2 µl of protein 4 mg/ml and 2 µl of well solution on siliconized cover-slips. Crystals appeared after 1-2 days.

2.2.4.2 Crystallization of hexagonal form SpPCNA crystals in complex with human p21(141-160)

Crystals of PCNA complexed with the p21 peptide were obtained using the same method, but grew from peptide/protein mixtures with the ratio 3:1 and a protein concentration of 2-5 mg/ml. Screening for crystal growth conditions was based around mixtures that had yielded unliganded SpPCNA crystals before. The best

condition was found to be 100 mM succinic acid pH 6.2 and 47% ammonium sulphate and 0.1% glycerol. Crystals appeared after 2 weeks.

2.2.5 Data collection, Data analysis, Molecular Replacement and Refinement

A crystal of unliganded SpPCNA with the dimensions of 0.5 mm x 0.4 mm x 0.2 mm was collected in a Cryo-loop, dipped briefly into freezing solution (30% PEG-4000 in mother liquor) and flash frozen in liquid nitrogen. Diffraction data were recorded using the synchrotron source in Daresbury station 14.1 (Table 2-1). The crystal was found to be in the monoclinic space group $P2_1$ with a trimer present in the asymmetric unit. X-ray data were processed to 2.4 Å using MOSFLM (CCP4) (Leslie, 1992) and scaled with SCALA (CCP4) (Evans, 1997). An mtz file was produced using 165 frames taken at 1 degree increments. The search model for molecular replacement was obtained by use of the programme 3DPSSM (Kelley et al., 2000). This programme takes the sequence of protein of interest and attempts to predict its 3-dimensional structure by calculating a score for matches between residues in the probe and residues in the library sequence (here hPCNA) using the global dynamic programming algorithm.

Molecular replacement was carried out by using the program MOLREP (Vagin and Teplyakov, 1997). For refinement the programs REFMAC (CCP4) (Murshudov et al., 1997) and COOT (Emsley and Cowtan, 2004) were used.

A data set of the trigonal crystals of the SpPCNA/p21 peptide (141-KRR**Q**TS**M**TD**FY**HSKRRLIFS-160) complex was collected at the Daresbury synchrotron.

The crystal of the SpPCNA/p21 peptide complex was 1 mm long. Crystals were flash frozen by dipping into liquid nitrogen. No freezing solution was used. The space group was found to be $H3_2$ with one peptide and two PCNA monomers present in the asymmetric unit, where each of them belongs to a different trimeric ring. X-ray data were processed to 2.2 Å using the programs mentioned above and frames 29-178 while omitting frames 83 and 84 due to elevated R_{merge} values.

2.2.6 Binding Assays of SpPCNA and hPCNA

2.2.6.1 Thermal shift assay of hPCNA and the His-tags

The assay was carried out using thin-wall PCR 96-well plates (Bio-Rad) covered with Optical-Quality Sealing Tape (Bio-Rad) using the iCycler iQ Real Time PCR Detection System (Bio-Rad).

Each 50 µl sample contained 0.25 µM recombinant human PCNA or SpPCNA, a 1:1000 dilution of Sypro Orange (Molecular Probes) dye and increasing concentrations of either IGHHHH or HHHHHH peptide in thermal denaturation buffer 1 (25 mM Tris pH 7.5, 100 mM NaCl, 2 mM DTT). Each sample was carried out in triplicate and averages were calculated. The wavelengths for excitation and emission were 490 nm and 575 nm, respectively. Fluorescent readings were taken between 20 °C and 88 °C in 1 °C increments after each temperature had been maintained for 90 seconds.

The fluorescent intensities were plotted against temperature in the programme Kaleida-Graph 4.03 (Synergy Software, Reading, PA) and the T_m for each sample was calculated using the

$$F(T) = F_{post} + [(F_{pre} - F_{post}) / (1 + e^{(Tm - T)/C})] \quad (1)$$

where F(T) is the fluorescence intensity at temperature T, F_{pre} and F_{post} are the intensities for completely folded and completely unfolded protein, respectively and C is the slope factor. Here the slope factor acts as a scaling factor for the total increase of fluorescence, which may vary for different samples.

2.2.6.2 Surface Plasmon Resonance of SpPCNA and p21 peptides

Pure hexa His-SpPCNA was immobilized on the surface of a NTA sensor chip essentially as described (Wear and Walkinshaw, 2006). Experiments were performed with the help of Dr. Martin Wear.

Briefly, hexa His-SpPCNA was first captured and orientated, *via* its N-terminal His-tag, under neutral physiological conditions (pH 7.4, 150 mM NaCl) on an Ni^{2+}-NTA surface, followed by brief covalent stabilization using standard primary amine coupling chemistry. The final amount of protein immobilised on the surface was 860 response units (RU), which gave an 85%-90% active surface. A concentration series of p21(141-152) from 50 nM to 25 μM, in running buffer (10 mM Hepes pH 7.4, 150 mM NaCl, 0.005% surfactant P20, and 2% ethanol), was injected over the surface at a flow rate of 50 μl/min at 25 °C for 60 seconds. The surface was regenerated after each cycle by letting any complex dissociate in running buffer.

Kinetic data were obtained by carrying out the experiment with same concentration range at 10 °C – 45 °C in 5 °C increments.

2.3 RESULTS AND DISCUSSION

2.3.1 Purification of recombinant untagged human PCNA

The purification was carried out in three steps using ion-exchange (Q sepharose chromatography), affinity chromatography (Heparin sepharose resin) and finally gel filtration. The average purity was estimated to be >95% as shown by SDS - polyacrylamide gels of a typical purification (Figure 2-2).

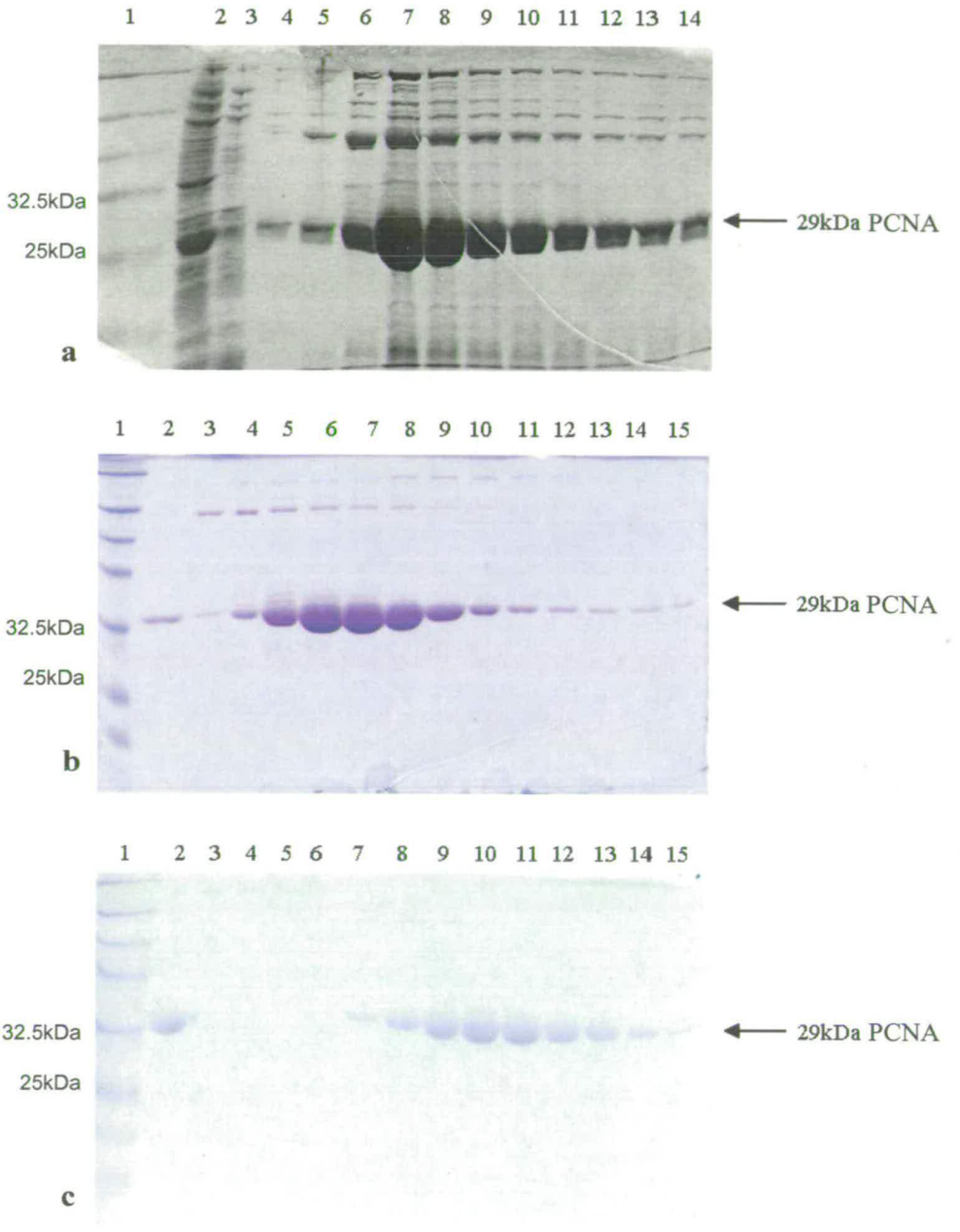


Figure 2-2 Purification of untagged human PCNA.
12% SDS-polyacrylamide gels were run with equal volumes of elution fractions and subsequently stained with coomassie. a) first purification step on Q sepharose; lane 1 = marker, lane 2= lysate, lane 3 = flow through, lanes 4-14 = alternating elution fractions b) second purification step on Heparin sepharose; lane 1 = marker, lane 2 = pooled fractions from previous step, lanes 3-15 = alternating elution fractions c) third purification step using a S200 gel filtration column; lane 1 = marker, lane 2 = pool from previous step, lanes 3-15 = alternating elution fractions. As in fission yeast PCNA, the human protein runs at a higher molecular weight as calculated.

2.3.2 Sequence comparison of hexa His-SpPCNA with human and budding yeast PCNA

As already discussed in Chapter 1, the overall secondary and tertiary structure (Figure 1-1, Figure 1-2) is well conserved among sliding clamps. Two eukaryotic PCNA crystal structures from humans and budding yeast have been published. The sequence alignment (Figure 2-3) by ClustalW (Thompson et al., 1994) shows that SpPCNA and hPCNA have 50% residues in common, while the two different yeast species share only 44% of their amino acids. hPCNA and ScPCNA are even less similar with a score of 35%. Nevertheless, the amino acids that make up the conserved hydrophobic pocket in PCNA seem to be highly conserved in all three species. The prediction would be that the hydrophobic pocket and the way p21(141-160) binds to these PCNA molecules will be very similar.

```
6His-SpPCNA MRSHHHHHHGIPMLEARFQQAALLKKLLDAIKELVTDANFDCNDNGISLQAMDSSHVALV 60
hPCNA       ------------MFEARLVQGSILKKVLEALKDLINEACWDISSSGVNLQSMDSSHVSLV 48
ScPCNA      ------------MLEAKFEEASLFKRIIDGFKDCVQLVNFQCKEDGIIAQAVDDSRVLLV 48
                        *:**:: :.:::*::::.:*: :  . :: ...*:  *::*.*:* **

6His-SpPCNA SMLIKSDGFEPYRCDRNIALGINLNALSKVLRCAQNEDLVTLKAEDTPEVLNLVFESEKN 120
hPCNA       QLTLRSEGFDTYRCDRNLAMGVNLTSMSKILKCAGNEDIITLRAEDNADTLALVFEAPNQ 108
ScPCNA      SLEIGVEAFQEYRCDHPVTLGMDLTSLSKILRCGNNTDTLTLIADNTPDSIILLFEDTKK 108
            .: :  :.*: ****: :::*::*.::**:*:*. * * :** *::..: : *:**  ::

6His-SpPCNA DRISDYDVKLMDIDQEHLGIPDIEYDATITMPAAEFQRITRDLLTLSDSVTINASKEGVR 180
hPCNA       EKVSDYEMKLMDLDVEQLGIPEQEYSCVVKMPSGEFARICRDLSHIGDAVVISCAKDGVK 168
ScPCNA      DRIAEYSLKLMDIDADFLKIEELQYDSTLSLPSSEFSKIVRDLSQLSDSINIMITKETIK 168
            :::::*.:****:* : * * : :*...:.:*:.** :* ***  :.*:: *  :*: ::

6His-SpPCNA FSCKGDIGNGSTTLKQHTDLSDQDQSIEISLTQAVTLTFSLKYLAQFTKATPLATRVTLS 240
hPCNA       FSASGELGNGNIKLSQTSNVDKEEEAVTIEMNEPVQLTFALRYLNFFTKATPLSSTVTLS 228
ScPCNA      FVADGDIGSGSVIIKPFVDMEHPETSIKLEMDQPVDLTFGAKYLLDIIKGSSLSDRVGIR 228
            * ..*::*.*.  :.   ::.. : :: :.: :.* ***. :**  : *.:.*:  * :

6His-SpPCNA MSNDVPLLVEYKMES-GFLRFYLAPKIGEEDEE 272
hPCNA       MSADVPLVVEYKIADMGHLKYYLAPKIEDEEGS 261
ScPCNA      LSSEAPALFQFDLKS-GFLQFFLAPKFNDEE-- 258
            :* :.* :.:::.: . *.*:::****: :*:
```

Figure 2-3 Sequence alignment of hexa His-SpPCNA, hPCNA and *Saccharomyces cerevisiae* PCNA.
Purple marks the residues that form hydrophobic pocket in hPCNA to which p21(141-160) binds in the complex structure. Most of them are well conserved in all three eukaryotic proteins: 56% are identical, while 19% are conserved in terms of charge or hydrophobicity.

2.3.3 Crystallization of unliganded SpPCNA and in complex with p21(141-160)

After optimization of a condition from Structure Screen 1 unliganded SpPCNA crystals grew diffracting up to 2.2 Å (a). The best SpPCNA/p21 co-crystals under optimized conditions diffracted to 2.2 Å as well (Figure 2-4c). SpPCNA crystallized also in various other conditions (Figure 2-4b+d).

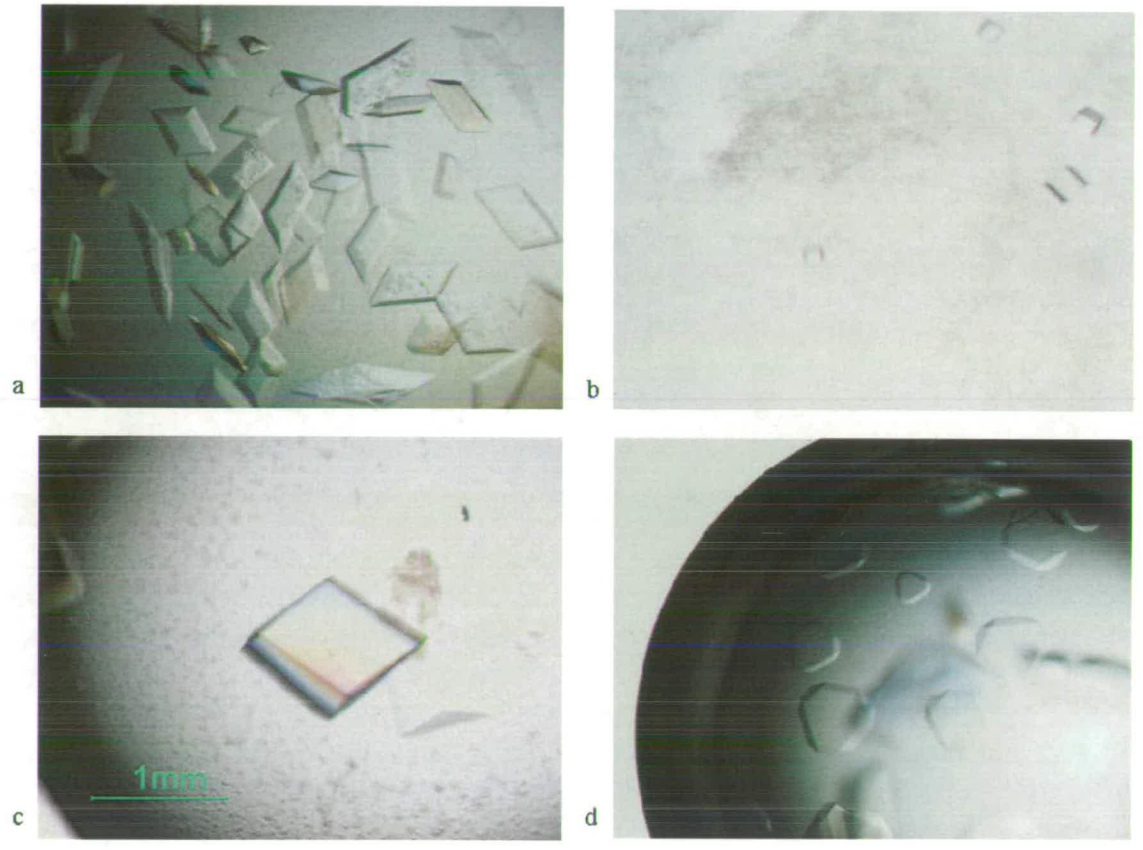


Figure 2-4 Protein crystals of SpPCNA.
a) Unliganded SpPCNA, Space group: $P2_1$, resolution: 2.4 Å b) Unliganded SpPCNA, Space group: unknown, resolution: 3.6 Å, conditions: 45% AS, pH8.0 (100mM Tris), 10 mM CoCl2, 1% PEG-400(50%), 2.7 mg/ml protein at 4 °C for 4 weeks c) SpPCNA in complex with p21(141-160), Space group: $H3_2$, resolution: 2.2 Å d) SpPCNA in complex with p21(141-160), Space group: $H3_2$, resolution: 3.5 Å, conditions: 45% AS, 100 mM succinic acid pH 6.0, 4 – 6 mg/ml, peptide: protein ratio = 1.5:1, 18 °C for 2 days

2.3.4 X-ray data analysis

2.3.4.1 Structure solution of monoclinic unliganded hexa His-SpPCNA crystals

The best crystals of the 30.4 kDa His-tagged fission yeast protein (see section 2.3.3) diffracted to 2.2 Å (Figure 2-5a), but the structure was solved using data up to 2.4 Å, as data in higher resolution shells were of lesser quality. Statistics of the data are shown in Table 2-1.

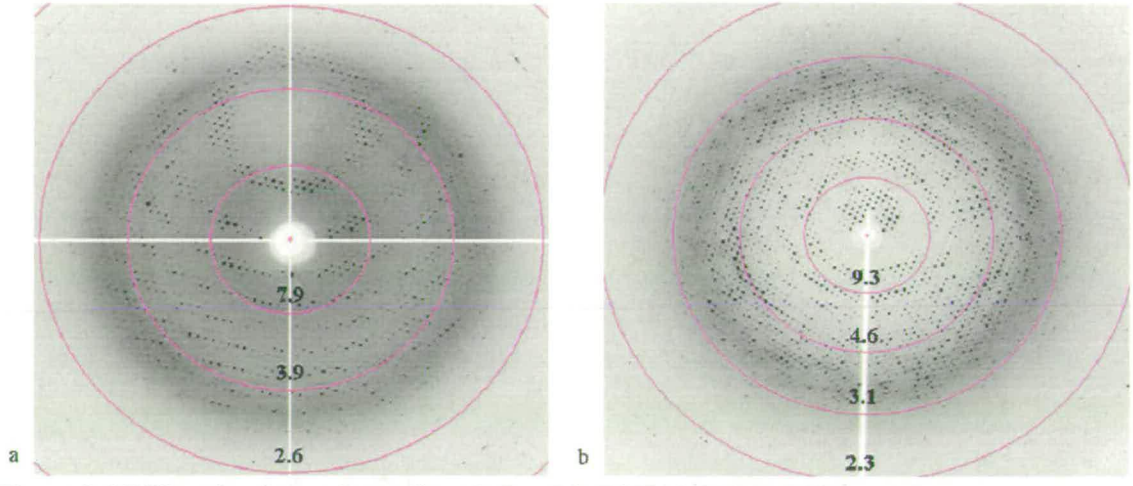


Figure 2-5 Diffraction image for unliganded and SpPCNA/p21(141-160) crystals.
a) Diffraction image for a crystal of unliganded SpPCNA b) Diffraction for a crystal of SpPCNA in complex with p21(141-160)

2.3.4.2 Structure solution of hexagonal SpPCNA/p21(141-160) crystals

The best crystals of the His-tagged fission yeast protein in complex with p21(141-160) diffracted to 2.2 Å (Figure 2-5b) and the structure was solved using data up to that limit. Statistics of the data are shown in Table 2-1.

Unit-cell parameters	Unliganded (outer shell)	PCNA/p21(141-160) complex (outer shell)
a (Å)	89.29	112.79
b (Å)	70.76	112.79
c (Å)	89.39	417.63
alpha	90	90
beta	119.69	90
gamma	90	120
Space group	$P2_1$	$H3_2$
Mean (I) / sd (I)	11.7 (2.4)	14. 3 (2.6)
No of independent reflections	37113	51476
Resolution range used in refinement	38.84-2.40	42.14-2.20
Highest resolution shell	2.53-2.40	2.32-2.20
overall completeness	97.5 (98.1)	98.2(91.4)
R_{sym} (%)	8.3 (42.5)	8.0 (38.6)
Multiplicity	2.7 (2.7)	5.6 (2.7)

Table 2-1 Statistics for SpPCNA crystal structures.

2.3.4.3 Space group determination

2.3.4.3.1 Unliganded SpPCNA

Unliganded SpPCNA diffracted to 2.2 Å and indexing in MOSFLM suggested P2 or $P2_1$ with unit-cell dimensions a = 89.29 Å, b = 70.76 Å, c = 89.39 Å and $\beta = 119.69°$ as a possibility. The space group was confirmed as monoclinic $P2_1$ by the inspection of systematic absences of reflections in the SCALA output file (0k0 reflections with

k ≠ 2n were absent). Data were successfully indexed and processed to an R_{sym} of 8.3%. The unit-cell parameters and data reduction statistics are shown in Table 2-1. The calculated unit cell volume is 490621.8 $Å^3$ and gave a calculated Matthew's Coefficient (2.69 $Å^3$/Da) (Kantardjieff and Rupp, 2003) corresponding to 3 monomers per asymmetric unit for a protein with the molecular weight of 30433 Da when packing in $P2_1$ with 54.25% solvent content.

A self-rotation map calculated using data processed in space group $P2_1$ confirms the outcome from the Matthew's coefficient and reveals a small degree of rotational non-crystallographic symmetry (Figure 2-6). Two threefold peaks appear at $\kappa = 120°$ marking non-crystallographic symmetry (

Figure 2-6).

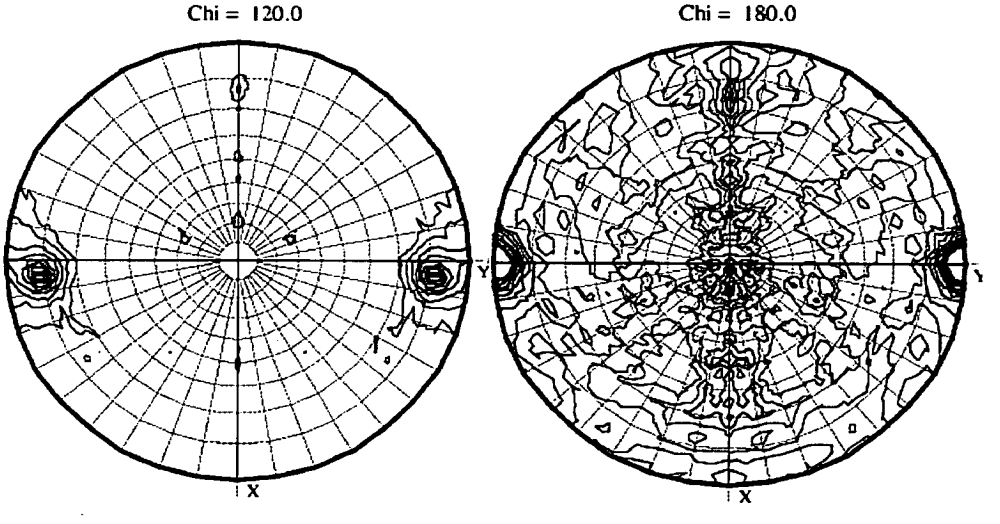


Figure 2-6 Stereographic projection plot of the κ= 120° and κ= 180° sections of the self-rotation function of the unliganded SpPCNA data set.

2.3.4.3.2 Space group determination (complex)

SpPCNA/p21(141-160) complex crystals diffracted to 2.2 Å and indexing in MOSFLM suggested H3 or $H3_2$ initially. Both variants were tried, but $H3_2$ was the only one that gave a solution. The unit-cell dimensions are a = b = 112.791 Å and c = 417.628 Å. The systematic absences of reflections in the SCALA output file give further evidence for a 3_2 three-fold screw symmetry. Data were successfully indexed and processed with an R_{sym} of 8.0%. The unit-cell parameters and data reduction statistics are shown in Table 2-1.

The Matthews Coefficient (2.1 $Å^3$/Da) gives the highest probability for two monomers of hexa His-SpPCNA present in the asymmetric unit (V= 4601190 $Å^3$) with 41.46% solvent content.

A self-rotation Patterson map calculated using data processed in space group $H3_2$ confirms two molecules per asymmetric unit.

2.3.4.4 Model building and Refinement

2.3.4.4.1 Unliganded SpPCNA

3DPSSM attempts to predict a protein's 3-dimensional structure from its amino acid sequence by comparing it with structures of similar proteins. In this case the closest crystal structure available was hPCNA.

Using the trimer as a model MOLREP found three peaks for the rotation function. The Rf/sigma scores were 14.64, 12.60 and 12.36 respectively, while the fourth solution was just a third of the value of the third rotation peak. For each rotation peak there was one solution of the translation function with correlation factors of 0.756, 0.638 and 0.627.

The electron density maps resulting from molecular replacement with the 3DPSSM output file, which included the first 254 residues of SpPCNA, as the search model were of sufficient quality to begin model building. One run of rigid body refinement in Refmac for 20 cycles yielded an R/R_{free} factor of 45.5%/45% at a weighting factor of 0.3. In restrained refinement both the X-ray residual (reflecting the agreement between the observed and the calculated Fs), and the geometric residual (reflecting the fit between the expected and the observed geometry) are minimised at the same time. The weighting factor determines relative weighting of these two terms. A value of 0.3 is considered to be tight and is in the recommended range for structures with a low resolution limit (>1.5 Å).

The data were used up to 2.5 Å and each chain was regarded as separate unit. This procedure was followed with a 35 cycle run of restrained refinement which yielded R/R_{free} of 28.7%/36.2% at a weighting of 0.05. A first manual refinement step in O (Jones et al., 1991) led to an R/R_{free} of 27.2%/34.4% after 10 cycles in restrained refinement. Loops (residues 101-108 and 186-189) in areas with particularly poor density were initially removed, neither were any residues of the hexa His-tag modelled. Building the loops and the hexa His-tags, which were visible for both chain A and B, but not for chain C, led to a drop in R/R_{free} to 25.3%/33.5%. Various

rounds of adding water molecules using ARP/wARP as part of the CCP4 suite in addition to manual refinement steps led to a final R/R_{free} factor of 22.2%/29.7%.

In an attempt to gain a higher resolution structure data from a second crystal from the same drop were obtained. 135 single frames in 1 degree increments were used for data processing. The model for use in molecular replacement was the final result of the previous refinement with all water molecules removed. A single rigid body refinement run with 20 cycles and a weighting factor of 0.3 yielded an R/R_{free} of 32.8%/33.1%. Data were initially used up to 2.2Å. A round of restrained refinement for 35 cycles and a weighting factor of 0.05 lowered the R/R_{free} to 25.2%/30.5%. The addition of water molecules and manual refinement using COOT led to an R/R_{free} of 22.8%/30.1%. Since it seemed to be impossible to improve the statistics any further the frames used for the structure solution were reviewed and batches number 18-20 were excluded as their R_{merge} was elevated in comparison with the remaining frames. Refinement was continued with a new SCALA output file (while retaining the R_{free} data from the first SCALA output) to 22.7%/30.7% R/R_{free}. As the R/R_{free} statistics were essentially unaffected by omitting those frames, TLS (thermal –libration-screw) refinement was employed for further improvement. Initially one group consisted of the entire PCNA monomer.

When refinement stagnated at 21.7%/29.3% R/R_{free} the TLS grouping was reviewed. The number of TLS groups to include in the refinement was assessed with the TLSMD server (Painter and Merritt, 2006) and different subdivisions were tested. The best result was obtained with five groups per monomer (1-19 plus His-tag if visible, 20-80, 81-124, 125-213, 214-254 or the last residue that was modelled). Throughout the refinement water was added in COOT and questionable loops were

attempted to be remodelled by superimposition with hPCNA. This led to an R/R_{free} of 21.3%/28.7%. Final statistics for refinement and structural details can be found in Table 2-2.

The quality of the model was assessed with PROCHECK (Laskowski et al., 1993) and analysis of the Ramachandran plot (Figure 2-7) showed 91.1% of non-glycine and non-proline residues in the most favoured regions, 7.9 % of all residues were in additionally allowed regions, 0.7% of the residues lie in generously allowed areas and 0.3% of residues are located in disallowed regions of the plot. This latter fraction consists of Glu106 and Asn108 in chain B and both of them are located in the same loop element of the protein structure. Many of PCNA's loops are very flexible with high B factors, which indicates disorder of those regions.

Structure Parameters	unliganded	PCNA/p21(141-160) complex
R factor (%)	21.3	19.2
R free (%)	28.7	24.4
B values ($Å^2$)		
Main chain	40.41	37.13
Side chain and waters	41.13	41.783
Ligand (peptide)	-	43.91
Succinic acid	-	58.25
Observed r.m.s. from ideal geometry in Å		
Bond lengths	0.010	0.026
Bond angles	1.469	2.272
Dihedrals	Chiral 0.11	Chiral 0.17
Ramachandran plot, residues in %		
Most favoured regions	91.1	93.6
Additionally allowed regions	7.9	6.5
Generously allowed regions	0.7	-
Disallowed regions	0.3	-
No. of water molecules	302	398

Table 2-2 Structure Refinement Statistics for SpPCNA crystal structures.

A fully consistent model of these regions could not be built, however, for completeness they were included in the PDB file in modelled positions. While the residues in the generously allowed and disallowed regions coincide with loop elements, the residues in allowed regions are located in well ordered sheets and helices.

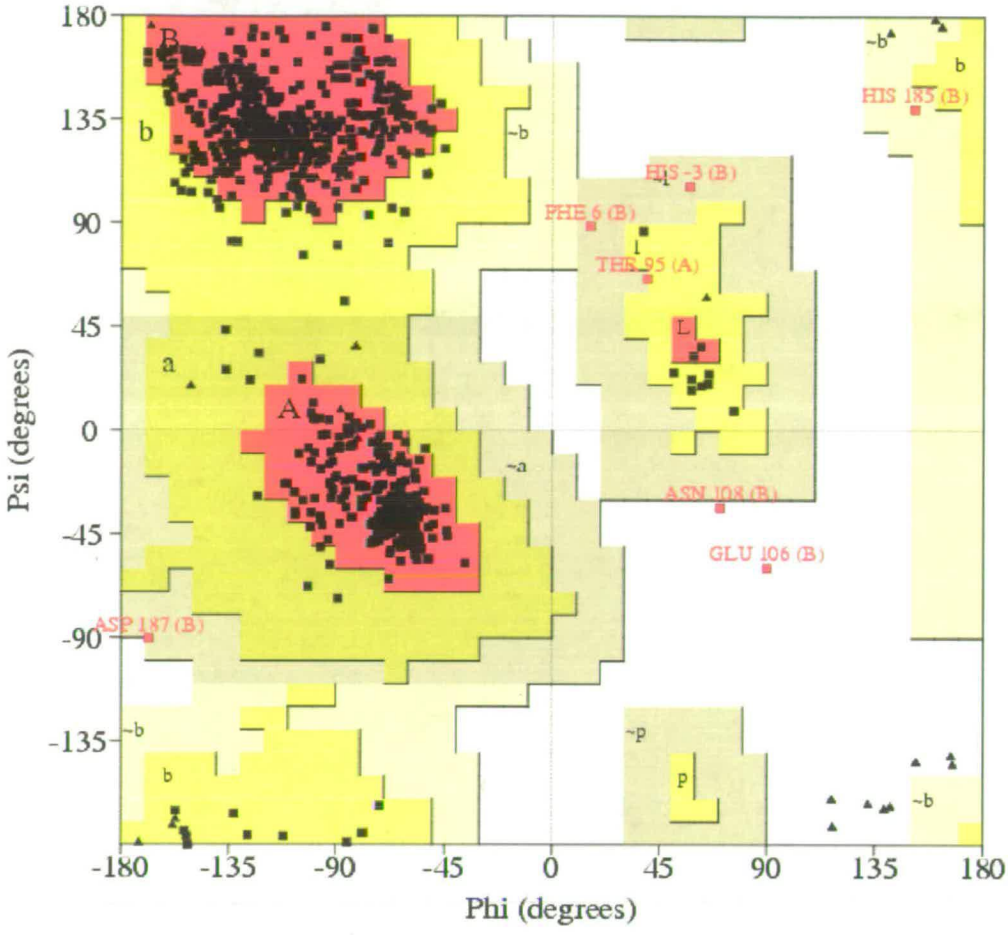


Figure 2-7 Ramachandran plot for the structure of unliganded SpPCNA.
Most residues have allowed conformations, but 7 residues (red) in difficult to model loops are outside the ideal areas. For details see Table 2-2.

2.3.4.4.2 SpPCNA/p21(141-160) complex

The data that were obtained from a first crystal diffracted to 3.45Å. Various SCALA runs were attempted for both space groups H3 and $H3_2$ as the determination of the correct space group was not straight forward.

MOLREP runs with different number of molecules in the asymmetric unit were carried out. The only correct solution, however, was with 2 monomers in the asymmetric unit using data processed in $H3_2$. Using chain A from unliganded SpPCNA as a model MOLREP found one peak for the rotation function. The

Rf/sigma score was 9.43, while the second solution was just half of the value of the first rotation peak. For the single rotation peak there was one solution of the translation function for the first monomer with a correlation factor of 0.350. The solution of the translation function for the second monomer yielded a correlation factor of 0.518.

A number of restrained refinement steps with a weighting of 0.5 for 20 cycles yielded an R/R_{free} of 19.3%/28.0%.

A second crystal yielded useable data of up to 2.2Å. A SCALA output file was obtained using frames 29-178 while omitting frames 83 and 84 due to elevated R_{merge} values. The best model from the initial low quality crystal was used for molecular replacement. A single step of rigid body refinement for 20 cycles and a weighting factor of 0.05 gave an R/R_{free} of 38%/37.9%. Further restrained refinement steps led to an improvement of statistics to 23.6%/28.4%. For further decrease of the R/R_{free} values TLS refinement was employed using one group for each monomer. This procedure yielded an R/R_{free} of 19.7%/24.3%. The use of 5 groups per monomer as described for the unliganded SpPCNA refinement procedure led to an R/R_{free} of 19.2%/24.4% which was a slight deterioration, but helped to bring the remaining residues of disordered regions into allowed conformations. Final refinement statistics are displayed in Table 2-2.

The quality of the model was assessed with PROCHECK (Laskowski et al., 1993) and analysis of the Ramachandran plot (Figure 2-8) showed 93.5 % of non-glycine and non-proline residues were in the most favoured regions while the remaining 6.5% were found in the additionally allowed areas. As in the unliganded SpPCNA

structure most loop regions are very flexible with high B factors, which resulted in regions that were difficult to model.

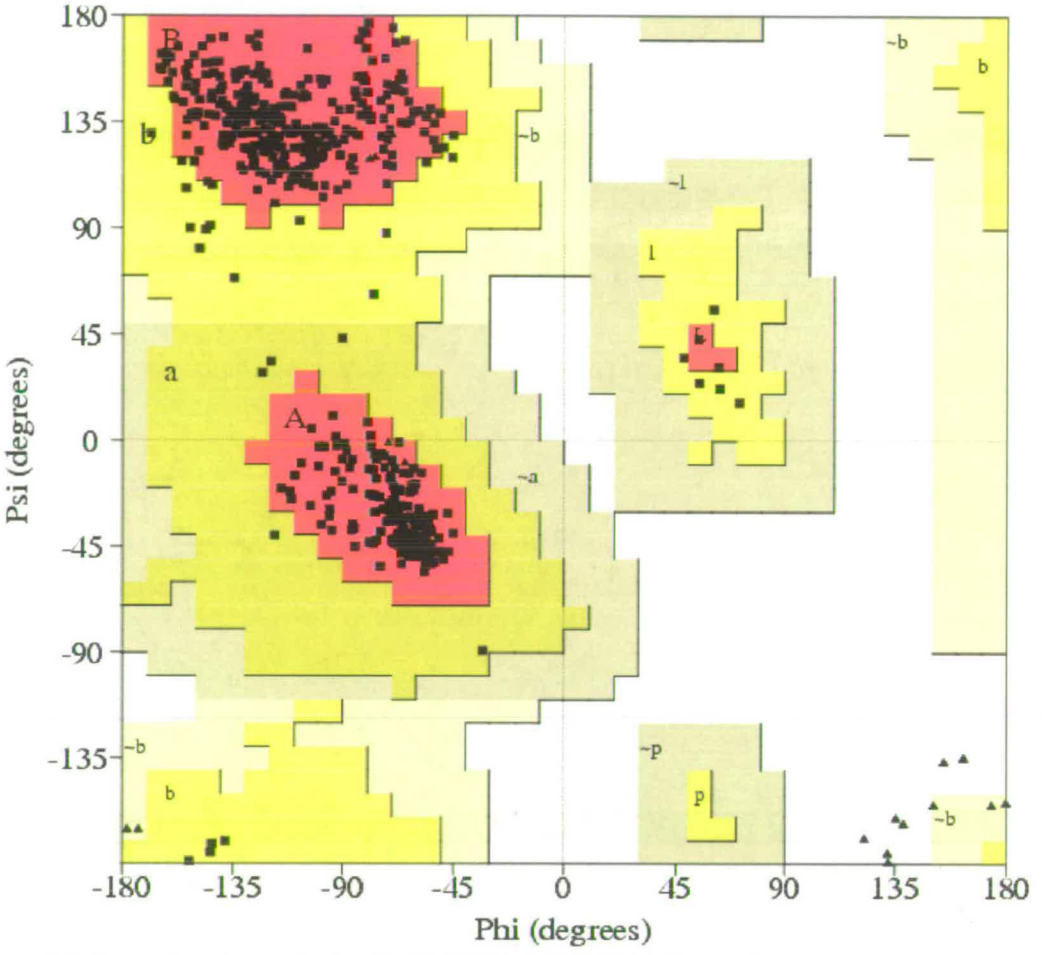


Figure 2-8 Ramachandran plot for the SpPCNA/p21(141-160) complex structure.
All residues have allowed conformations as the modelling was facilitated by a better data quality/slightly higher resolution compared to the unliganded SpPCNA structure. For details see Table 2-2.

2.3.5 Description of the unliganded SpPCNA crystal structure

2.3.5.1 Conservation of the general structure

The model of the PCNA trimer comprises the first 254 residues of 260 in chain A and C and 253 in chain B, 302 water molecules and 5 or 8 residues of the His-tag in

chain A or B, respectively. No density for a His-tag in chain C was detected. The amino acid sequence is shown in Figure 2-3, where it is aligned with human and budding yeast PCNA.

As in other eukaryotic loading clamps, functional SpPCNA consists of a set of three PCNA monomers that assemble to form a ring with a hole wide enough for a DNA double helix to pass through (about 35Å) (Figure 2-9). A superimposition of the C_α backbone of all three chains of the SpPCNA model shows that each chain is folded very similarly, with a maximum rmsd fit of 0.338 Å based on all 253 C_α atoms (Figure 2-10a).

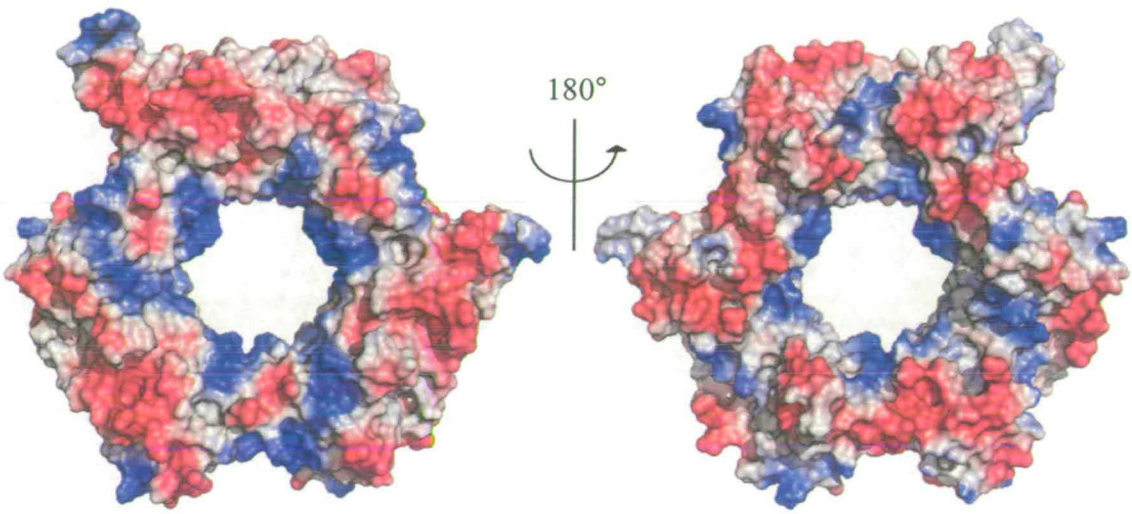


Figure 2-9 Surface charge of unliganded SpPCNA.
Left = front = C-side, right = back side. The overall surface charge is negative with a positively charged cavity to accommodate DNA. The same pattern is observed in all sliding clamps.

Unsurprisingly, when superimposing SpPCNA, hPCNA and ScPCNA, the 3D structure of these proteins is very similar (Figure 2-10b).

SpPCNA seems to have a single helical turn in monomer A and C formed by residues 54-56 in a loop connecting βD_1 and βE_1 (Figure 2-11). Assembly into the ring- structure *via* βD_2 and βI_1 of adjacent monomers leads to the formation of a contiguous surface across each intermolecular boundary as the anti-parallel nine-stranded β-sheets of each monomer connect. The central hole is lined by all together 12 α helices, 4 from each monomer, which is a conserved structural feature in PCNA. These are held in place by a scaffold formed by the already mentioned 3 β-sheets plus another three sheets located underneath the connector loop that are formed by βD_1 and βI_2 of each monomer.

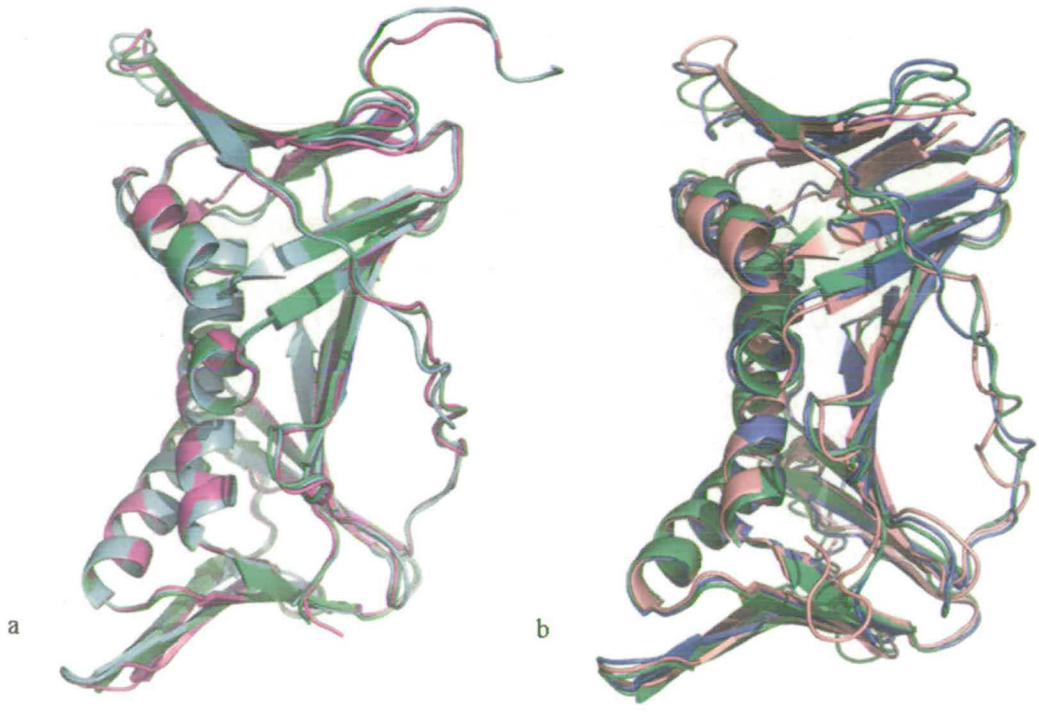


Figure 2-10 Comparison of the backbone structure of SpPCNA, hPCNA and ScPCNA.
a) Superimpositions of the C_α backbone of the three monomers of SpPCNA trimer. RMSD: chain B to chain A: 0.338Å, chain C to chain A: 0.327 Å; Key: green = A, blue = B, magenta = C; b) Superimposition of C_α backbone of the three eukaryotic PCNA structures. RMSD is 0.834Å for hPCNA and SpPCNA and 1.132Å for ScPCNA and SpPCNA_A. Key: purple = hPCNA, green = SpPCNA, pink = ScPCNA

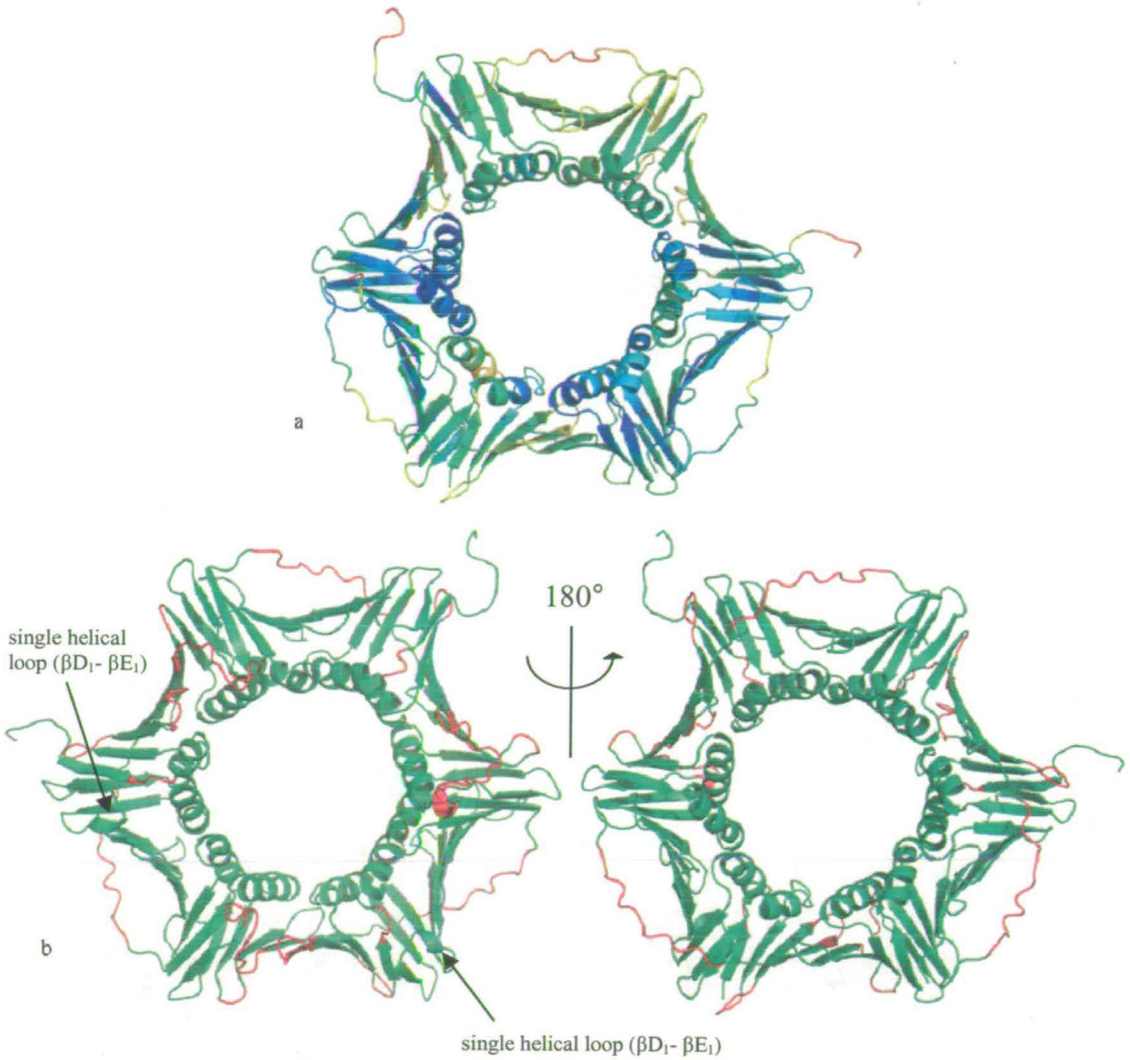


Figure 2-11 Regions of varying B factors and disorder in unliganded SpPCNA.
a) Backbone structure of unliganded SpPCNA with colouring according to B factors. Red shows high and blue signifies low values. b) The red sections show areas that were difficult to model and that fit the density less well than the remaining areas.

2.3.5.2 Connector loop and conservation of hydrophobic pocket

As in the structure of unliganded human PCNA, the final residues 255-260 of SpPCNA are not visible as the C-terminus seems to be coordinated upon binding to $p21^{WAF}$ or a peptide derived from this checkpoint protein, which is an association control factor for PCNA in the cell (Waga et al., 1994). Residues 121-132 form the

interdomain connector loop (IDCL), which joins the N-terminal domain with the C-terminal domain. The IDCL is flexible in unliganded PCNA structures and difficult to model. In addition to the IDCL the unliganded structure of a SpPCNA monomer contains, as in the hPCNA, zones with high temperature factors that correspond to loop regions between adjacent β-strands. Loops that are particularly affected include residues 6-8 (βA_1-αA_1), 93-97 (βG_1- βH_1), 105-111 (βG_1- βI_1), 164-165 (βB_2- βC_2) and 182-193 (βD_2- βE_2) (Figure 2-11) (Figure 1-2).

As to be expected, the His-tag residues that appear in density also have high B-factors.

Some of these loop regions are also disordered and difficult to model, which leads to a small number of residues to be in disallowed regions of the Ramachandran plot (Figure 2-7).

Furthermore, the sequence alignment of SpPCNA and hPCNA shows that the residues involved in the formation of the hydrophobic pocket are nearly identical in *S. pombe* and human PCNA (Figure 2-3). The only exceptions of those 16 residues are Cys27 vs. Asn27, Leu121 vs. Ile 121 and Val123 vs. Gln123. Ile and Leu are very similar in their chemical properties, but not in size.

In the crystal structure one can see that indeed most residues forming the pocket are located at the same positions except for hydrophobic residues from the connector loop, which is slightly distorted anyway (Figure 2-12). The residues deep in the pocket (Met40, Val45, Leu/Ile47, Tyr133, Pro234, Tyr249/50, Ala251/2 and Pro252/3) under the loop are very much in the same position.

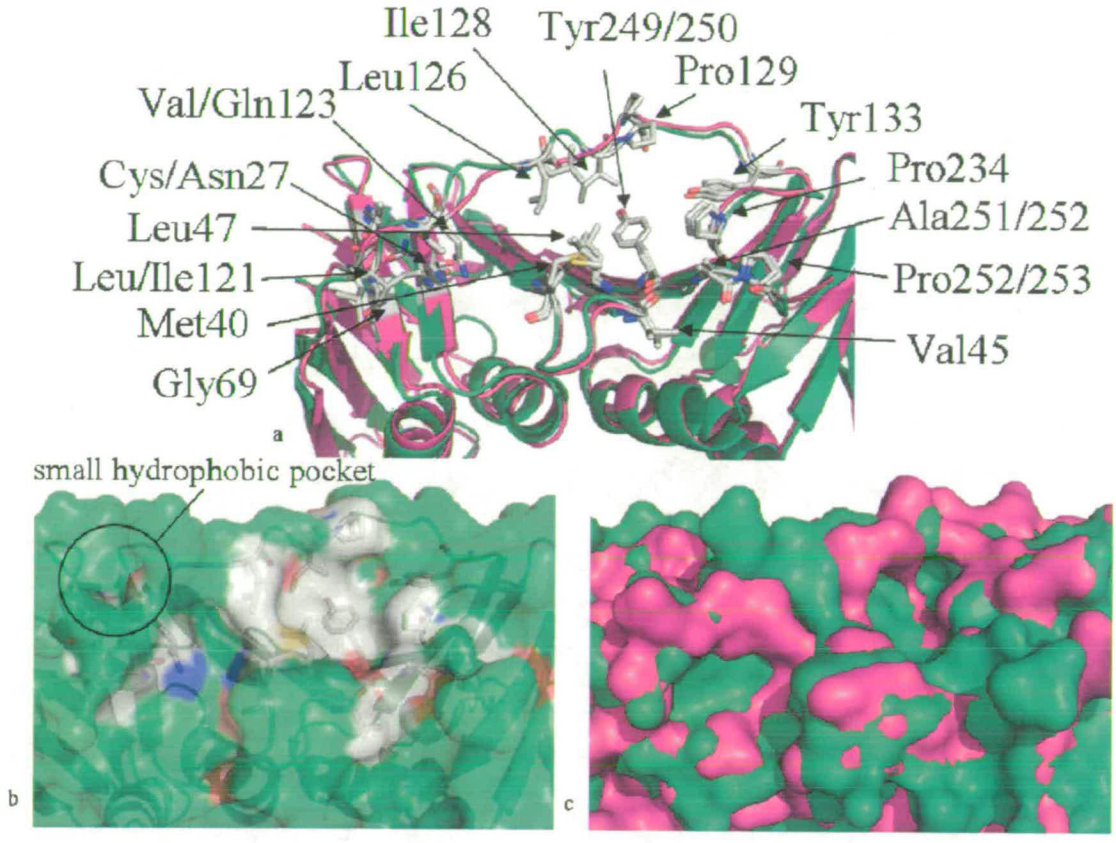


Figure 2-12 PCNA's hydrophobic pocket.
a) Overlay of unliganded hPCNA (magenta) and unliganded SpPCNA (green). The residues in grey show the hydrophobic pocket which accommodates the p21 peptide in the complex crystal structure with hPCNA. In cases where name or position of an amino acid is not conserved between human and fission yeast proteins the first name or number stands for the human variant, the second one for yeast. As one can see the amino acids deep in the pocket are orientated in the same position in both structures. The residues in the IDCL that contribute to the hydrophobic pocket are in different positions, which is to be expected as that linker is very flexible and disordered in the absence of a peptide ligand. b) Depictured is the semitransparent surface of SpPCNA. The grey areas show residues of the hydrophobic pocket. c) Overlay of the surface of hPCNA and SpPCNA. The larges variations are contributed by the flexible IDCL. On the whole the core pocket of fission yeast PCNA seems to be slightly larger than in hPCNA.

2.3.5.3 Crystal packing showing His-tags act as a p21 mimic

When looking at the crystal packing of the SpPCNA trimers (Figure 2-13), it becomes clear that both His-tags of chain A and B in one ring make interactions with the hydrophobic pocket/IDCL of an adjacent trimer. This contact makes the tag become more rigid and can hence be detected in the electron density maps (Figure

2-14). Crystal structures containing clear density for His-tags are fairly rare, since this feature is normally very flexible.

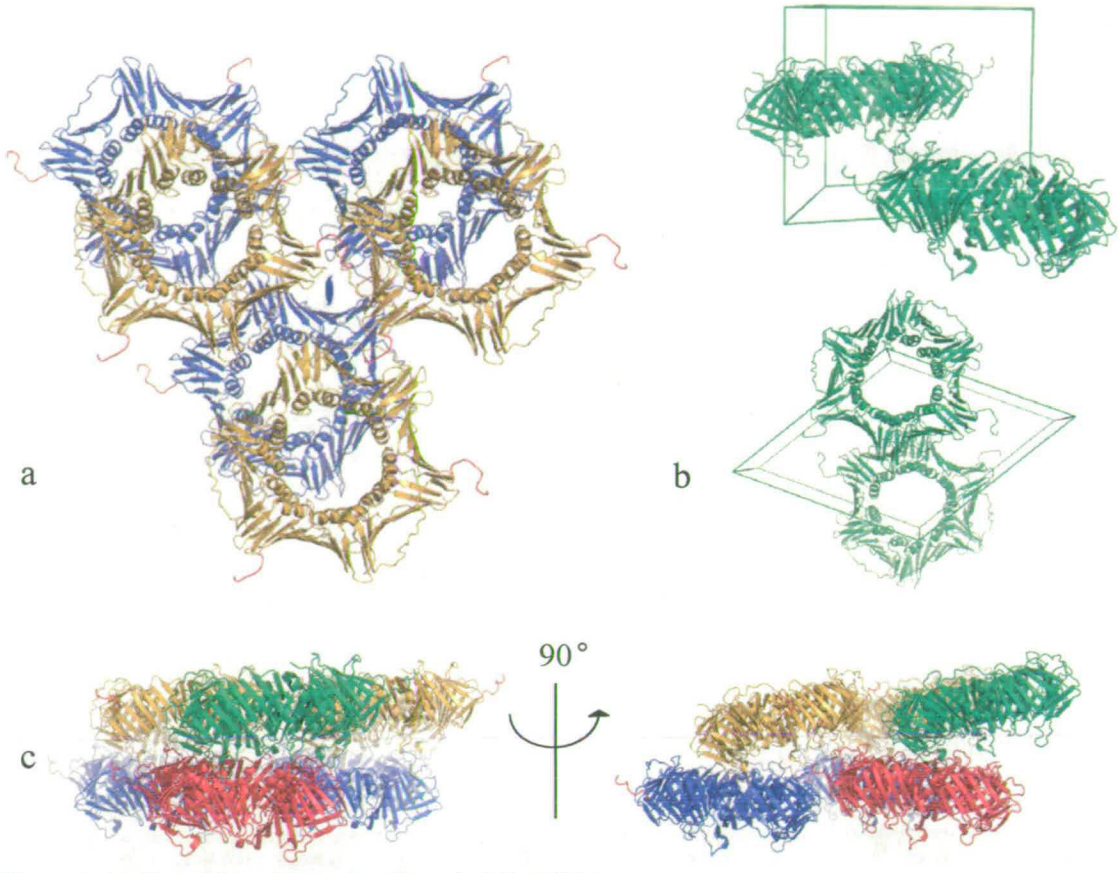


Figure 2-13 Crystal packing of unliganded SpPCNA.
a) This picture shows clearly the 2-fold screw symmetry. The His-tags of chain A and B (red) make neat interactions with adjacent rings. There is no density for the third tag. Trimers of different layers are staggered b) This picture shows the unit cell outline and the positioning of the trimeric ring and second ring related by the two-fold screw symmetry. c) The trimers are not neatly stacked, but are tilted against each other. While chains A → C and B → A in adjacent rings are close together, the tilt in the layering means that the N-terminus from one ring is too far away from the hydrophobic pocket of the neighbouring chain B. Therefore the His-tag of chain C and cannot make the stabilizing contact with a hydrophobic pocket.

The interactions between the tags and the hydrophobic pocket are both hydrophobic and hydrophilic. The hydrophobic contributions come from Ile-1 and His residues that have no hydrogen-bonding potential left (Figure 2-15). Hydrophilic interactions are formed by direct and indirect hydrogen bonds: Chain A → His-4/NE2 ··· Leu126 (Chain C)/ O, His-3/NE2 ··· Water 14; Chain B → His-4/NE2 ··· Leu126 (Chain

A)/O). Furthermore, the interaction between two rings is stabilized by hydrogen bonds between non-His tag residues and the hydrophobic pocket. For Chain A: Asp63/OD2 ··· Asp130/N, Arg64/NH1 and Arg64/NE ··· Gly127/O. For chain B: Asp63/OD2 ··· Water 84 and Water 136, Arg64/NH1 ··· Leu126/O and Gly127/O. A list of all distances of up to 4Å between atoms of two adjacent rings can be found in Table 2-3.

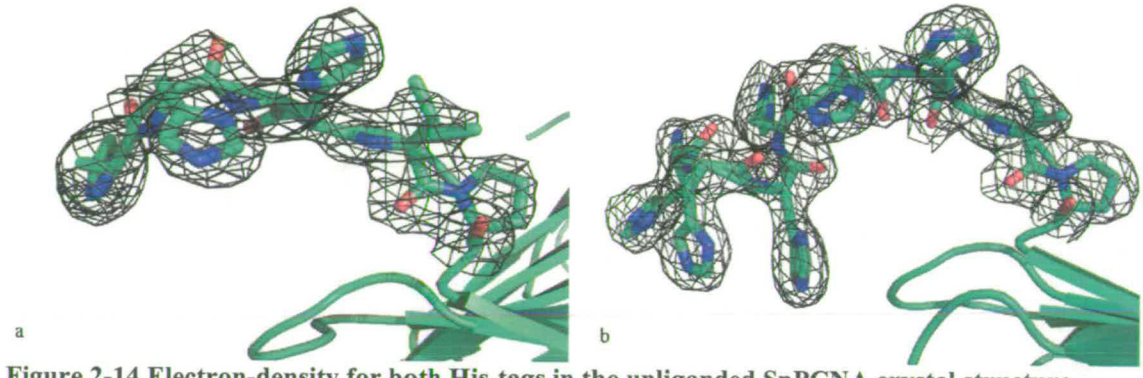


Figure 2-14 Electron-density for both His-tags in the unliganded SpPCNA crystal structure. Density maps show clearly the location for up to 8 residues of the His-tags. a) SpPCNA Chain A b) SpPCNA Chain B

His -4 and His -3 bind directly to the hydrophobic pocket (Figure 2-16). While His -3 is located in the same place as Tyr151 in the hPCNA/p21(141-160) complex, His-4 is situated near where Met147 in the human complex structure can be found.

Both His-residues have an overall neutral charge as they use up their hydrogen bonding potential by forming a hydrogen bond with each other and with nearby water molecules (Figure 2-15). Hence they seem to be able to assume the position of the hydrophobic/aromatic residues of the p21 peptide, especially at neutral pH.

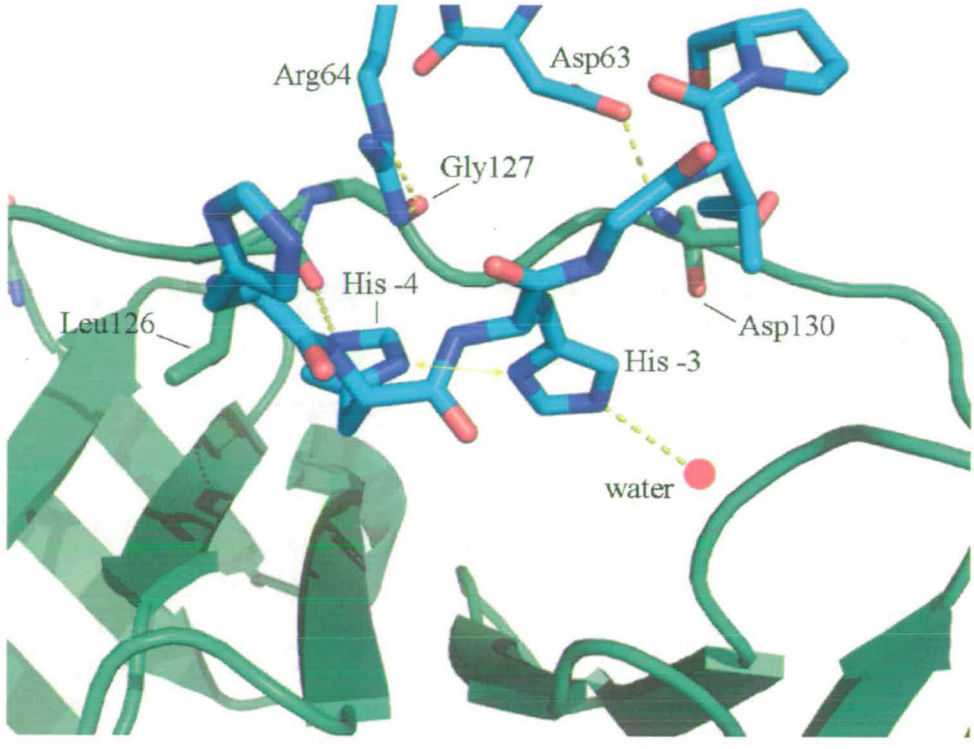


Figure 2-15 Hydrogen bonds that facilitate binding of the His-tag in the hydrophobic pocket. Both His -4 and His -4 are forming two hydrogen bonds with their side chains. Partners are either the other histidine residue or other residues/water molecules.

Despite the seeming specificity of the hydrophobic pocket, which accommodates normally residues of the PIP-box binding motif, it surprisingly seems to be able to accept any residues which mimic aromatic amino acids (as found in the His-tags), if their freedom of movement is restricted. This notion is further reinforced by findings in the SpPCNA/p21 co-crystal structure described below.

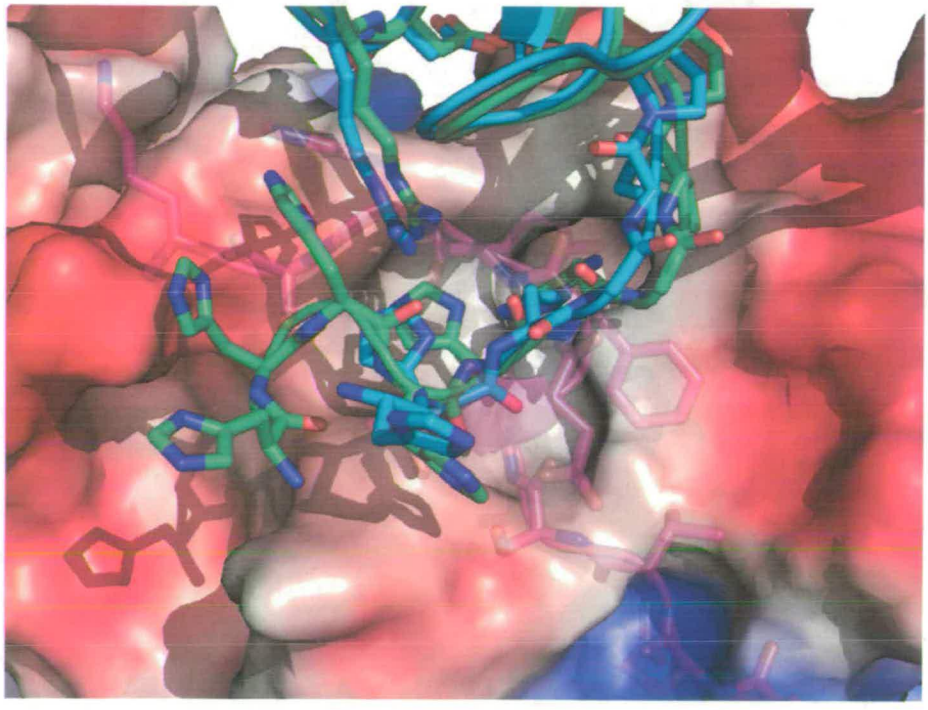


Figure 2-16 Orientation of the His-tags in the hydrophobic pocket.
For both chain A and B residues His-4 and His-3 are able to bind in the hydrophobic pocket due to overall neutral charge. Key: blue = residues -5 – 0 of chain A, green = residues -8 – 0 chain of B, magenta = p21 of complex with hPCNA.

His -tag	Chain A	Chain C	Distance
O HIS -8	SD MET 40		3.37
	CE MET 40		3.51
CB HIS -7	CG GLU 124		3.85
	O HIS 125		3.67
CG HIS -7	CG GLU 124		3.4
	O HIS 125		3.61
ND1 HIS -7	CG GLU 124		3.58
	C HIS 125		3.9
	O HIS 125		2.74
CE1 HIS -7	CG GLU 124		3.86
	O HIS 125		3.76
NE2 HIS -7	CG GLU 124		3.86
CD2 HIS -7	CG GLU 124		3.58
O HIS -7	O LEU 126		3.35
	CA LEU 126		3.54
	C LEU 126		3.9
	CD1 LEU 126		3.49
ND1 HIS -6	O HIS 125		3.9
	C LEU 126		3.87

	O LEU 126		3.1
CE1 HIS -6	O HIS 125		3.73
	O LEU 126		3.57
C HIS -5	CE MET 40		3.68
O HIS -5	CE MET 40		3.19
	O HIS 44		3.88
CB HIS -4		SD MET 40	3.79
		CE MET 40	3.8
	CD2 LEU 47		3.95
CG HIS -4	CD2 LEU 47		3.67
ND1 HIS -4	CD1 LEU 47		3.8
CE1 HIS -4		CD2 LEU 47	3.86
	C LEU 126		3.98
	O LEU 126		3.25
	O GLY 127		3.8
NE2 HIS -4		CD2 LEU 47	3.82
	C LEU 126	C LEU 126	3.76/3.91
	O LEU 126	O LEU 126	3.03/3.4
	CB LEU 126	CB LEU 126	3.87/3.2
		CG LEU 126	3.3
CD2 HIS -4		SD MET 40	3.99
	CE MET 40		3.86
		CB LEU 126	3.59
		CG LEU 126	3.45
		CD2 LEU 126	3.73
	CD2 LEU 47		3.65
CB HIS -3	CG PRO 129		3.98
CG HIS -3	CG PRO 129	CG PRO 129	3.56/3.95
	CD PRO 129		3.81
ND1 HIS -3	CG PRO 129		3.36
	CD PRO 129	CD PRO 129	3.94/4.0
	CG PRO 234		3.81
	CD PRO 234		3.8
CE1 HIS -3	CG PRO 129		3.78
	CB PRO 234		4
	CG PRO 234		3.73
	O HOH W 5		3.23
NE2 HIS -3	CG PRO 234	CG PRO 234	3.87/3.74
		CD PRO 234	3.87
CD2 HIS -3		CG PRO 129	3.81
		CG PRO 234	3.62
		CD PRO 234	3.5
	CD PRO 129		3.83
CG1 ILE -1		CG PRO 129	3.99
CD1 ILE -1		CB PRO 129	3.73
		CG PRO 129	3.83
		CG1 ILE 131	3.6
	CD1 ILE 131	CD1 ILE 131	3.78/3.53
	CG2 ILE 131		3.62
CG2 ILE -1		CB VAL 233	3.99

		CG1 VAL 233	3.69
CE MET 1	CG PRO 129	CG PRO 129	3.83
	CD PRO 129	CD PRO 129	3.8
CD ARG 61	O ASP 130	O ASP 130	3.36/3.44
		O HOH W 84	3.82
NE ARG 61	O ASP 130	O ASP 130	3.52/3.53
CZ ARG 61		O ASP 130	3.95
NH2 ARG 61	CB PRO 129		3.86
CB ASP 63		CA PRO 129	3.99
CG ASP 63	CA PRO 129	CA PRO 129	3.86/3.83
		N ASP 130	3.88
OD1 ASP 63	CA PRO 129		3.29
	CB PRO 129		3.68
	CD PRO 129		3.7
OD2 ASP 63		CA PRO 129	3.26
		N ASP 130	3.01
		C PRO 129	3.62
O ASP 63		CA GLY 127	3.78
		CB PRO 129	3.95
CG ARG 64		O GLY 127	3.96
NE ARG 64	O LEU 126	O LEU 126	3.92/2.94
	O GLY 127	O GLY 127	3.66/3.9
CZ ARG 64	O LEU 126		3.84
	O GLY 127	O GLY 127	3.92/3.51
NH1 ARG 64	O LEU 126	O GLY 127	3.11/3.19
	O GLY 127		3.27

Table 2-3 Distances between two adjacent rings *via* the affinity tags and other atoms at the interface.
The yellow highlighted selection shows hydrogen bonds (or possible ones) between two chains. There are three monomers in the asymmetric unit, defined as chains A, B and C. Only chain A and C have contacts to affinity tag residues. In a few instances the same atoms in chain A and C are able to form hydrogen bond contacts with affinity tag residues.

Many proteins are purified *via* an affinity tag of which the hexa His-tag is the most popular one. Often it is cleaved prior crystallization, but sometimes it is left on as it has been shown previously that the effect of the His-tag on the protein structure is generally not significant (Chant et al., 2005). A recent review of the impact of His-tags on protein structures (Carson et al., 2007) found that less than 10% of the solved structures actually contain a His-tag and of those less than 6% (65 structures) have main-chain coordinates for four or more consecutive histidine residues. Only 29

structures have five or more consecutive histidine residues with complete side-chain coordinates and an average B factor for those residues of 40 $Å^2$ or less. A closer examination of the latter group showed that almost all of those tags take part in making crystal contacts, where a third of them are bound to the unliganded protein and the remainder extends into the solvent to bind into pockets on adjacent protein molecules. As one expects those pockets are mainly of an acidic nature to provide hydrogen-bond acceptors. Sometimes metal ions like nickel, cadmium or zinc are associated with the tags. In a few cases the electrostatic surface of the adjacent molecule shows an overall basic charge.

Histidine has a special status among amino acids as its protonation state is pH dependent. Its pK_a lies between pH 6 and 7 for the side chain and is highly dependent of the environment. At a neutral pH histidine is "singly protonated" and is uncharged (Figure 2-17). In this state it is able to interact with positively charged residues acting as a hydrogen-bond acceptor. Therefore histidine will be harder to protonate in a positively charged environment which leads to a lower effective pK_a. In the neutral form histidine has also aromatic properties and can be found stacked between or head-on with tyrosine and phenylalanine.

At pH 6 the amino acid is "doubly protonated" and hence positively charged. In this form it is likely to seek interactions with acidic residues.

What is the importance of this special amino acid in the formation of protein-protein interactions? Upon comparison of homodimers (e.g. cytochrome c') and heterocomplexes (e.g. antibody-antigen, enzyme-inhibitor) histidine residues have been found to have a relatively high interface propensity in heterocomplexes. Propensity describes the likelihood of a residue to be involved in the formation of a

protein-protein interaction interface over the rest of the protein surface (Jones and Thornton, 1996). Hence it is in accordance with naturally occurring protein-protein complexes that His-tags make contact with adjacent protein molecules in crystals. Of course the space restrictions in crystal packing can promote the non-specific binding of histidine residues to protein pockets or surfaces due to reduced degrees of freedom.

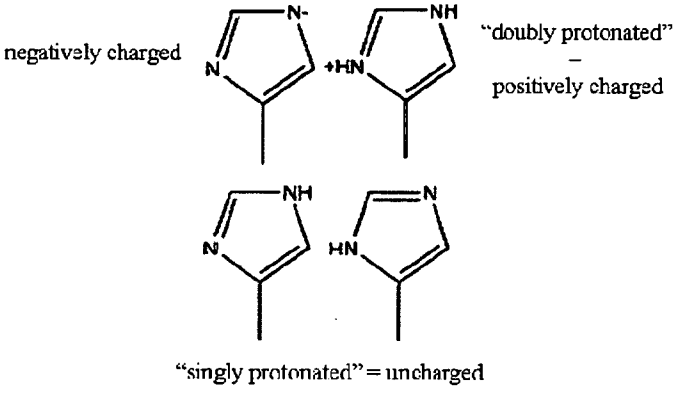


Figure 2-17 Protonation states of histidine.

2.3.6 Thermal denaturation assay of SpPCNA and hPCNA to test for His-tag binding

The fact that the His-tag is located in the hydrophobic pocket of PCNA is explainable, but nevertheless intriguing since the drugability of this pocket in hPCNA has been discussed recently (Kontopidis et al., 2005). If anything with a remotely aromatic character can be found to bind there, then the specificity of this pocket may be questioned.

To test if the binding of the His-tag to the pocket is a mere artefact of crystal packing induced by spatial restriction and to confirm the specificity of the pocket, a thermal denaturation assay was carried out with both SpPCNA and hPCNA. In this assay the unfolding of both proteins is monitored in the absence and under increasing concentrations of two hexameric peptides (100 μM, 200 μM, 500 μM and 1 mM for SpPCNA and 500 μM and 1 mM for hPCNA): IGHHHH and HHHHHH. The crystal structure of unliganded SpPCNA shows that residues Ile-1 and Gly-2 which are part of the linker region between protein and the six histidine residues make an interaction with certain residues at the pocket. As Figure 2-12 shows the comparison of the human and fission yeast binding pocket shows minor differences, which might be just significant. It has been shown that ligand binding has a stabilizing effect on proteins which can be visualized with the thermal denaturation assay as a T_m shift towards a higher temperature (Ericsson et al., 2006; Lo et al., 2004). Figure 2-18 shows the melting curves of SpPCNA in the presence of the two peptides. It is difficult to distinguish any significant differences between the melting curves by eye, so it is important to apply a suitable equation which calculates the melting or transition midpoint of each sample at which 50% of the protein is unfolded.

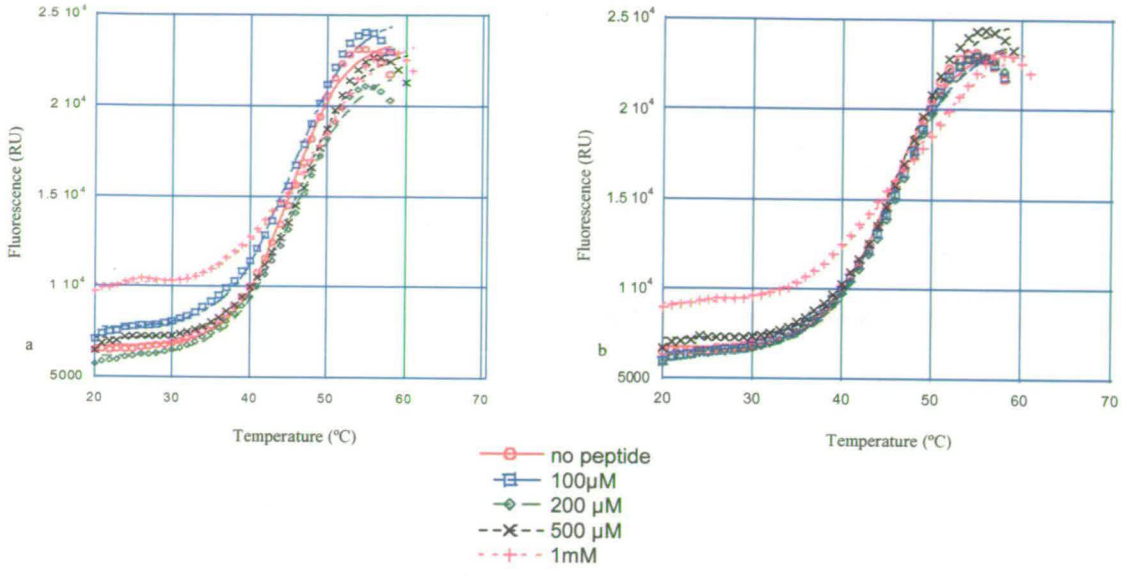


Figure 2-18 Thermal denaturation curves of SpPCNA in the absence and presence of IGHHHH and HHHHHH.
a) Fission yeast PCNA under the influence of IGHHHH b) Fission yeast PCNA under the influence of HHHHHH. Key: red = no peptide, dark blue = 100 µM peptide, green = 200 µM, black = 500 µM peptide, pink = 1 mM. When analysing the curves just by eye the difference between them, i.e. the T_m, seems small if at all present. Only by fitting the data to a function, significant changes in T_m may become apparent.

Table 2-4 summarizes the T_m values for both SpPCNA and hPCNA. While the transition midpoint shifts at 0.1 mM and 0.2 mM lie within the error compared to protein in the absence of any potential ligand at 0.5 mM of both peptides seem to have a stabilizing effect on fission yeast PCNA. With increasing peptide concentrations the amount of DMSO in the sample increases as well as this is the solvent of the stock solutions (25 mM for HHHHHH and 50 mM of IGHHHH). DMSO has been shown to have a destabilizing effect on PCNA as is to be expected (the extent will be quantified in section 3.4.3). At same concentration of each peptide HHHHHH will contain hence twice the amount of solvent. In order to correct for the effect of DMSO controls have been carried out that include just protein and DMSO at the equivalent peptide concentration. At 1 mM concentrations of both peptides it seems that HHHHHH has a slightly stronger stabilizing effect on SpPCNA than

IGHHHH when judging by the ΔT_m shift between control and experiment, which implies a higher affinity for the protein. Human PCNA was only exposed to peptides at 0.5mM and 1mM.

Concentrations of peptides (mM)	T_m (°C) for SpPCNA		T_m (°C) for hPCNA	
	IGHHHH	HHHHHH	IGHHHH	HHHHHH
0	45.22 ± 0.21		52.92 ± 0.20	
0.1	45.32 ± 0.22	45.56 ± 0.22	-	-
0.2	45.68 ± 0.22	45.90 ± 0.20	-	-
0.5	46.27 ± 0.21	45.98 ± 0.17 (control = 45.35 ± 0.36)	52.86 ± 0.15	52.18 ± 0.15 (control = 51.67 ± 0.15)
1	48.08 ± 0.59 (control = 45.35 ± 0.36)	47.42 ± 0.52 (control = 43.78 ± 0.39)	50.04 ± 0.15 (control = 51.67 ± 0.15)	49.16 ± 0.17 (control = 50.69 ± 0.15)

Table 2-4 Comparison of T_m values for SpPCNA and hPCNA for various peptide concentrations.
Control: To correct for the destabilizing effect of DMSO at higher peptide concentrations equivalent volume of just DMSO is added to the protein to observe the effect on T_m of solvent alone.

To sum up, both peptides have a stabilizing effect on SpPCNA, while HHHHHH is slightly more stabilizing than IGHHHH, suggesting a higher affinity for the sliding clamp. The calculated K_ds of IGHHHH and HHHHHH for SpPCNA at the melting temperature are 893μM and 917μM, respectively.

In contrast to SpPCNA the human sliding clamp experiences no significant stabilization upon complexing with either of the hexapeptides; in fact at a concentration of 1mM for both peptides, the T_m value of the DMSO control is higher than for the T_m value for either of the peptides. This suggests a destabilizing effect of the peptides which is observed when a ligand has a higher affinity for the unfolded

protein than for the folded form. Of course this experiment is carried out under non-natural conditions. Neither DMSO nor this kind of elevated temperature would be experienced in a cell.

Possible explanations for the difference of the thermal shift results

On the whole hPCNA seems to be more stable than the yeast version of the sliding clamp as the transition midpoint difference between hPCNA (T_m= 52.92°C ± 0.20) and SpPCNA (T_m= 45.22°C ± 0.21) in the absence of peptides and DMSO suggests. As already mentioned in Chapter 1 sliding clamps differ in the number of hydrogen bonds that are located at the interface of monomers. The higher developed an organism is the more bonds are located at this interface which is increasing the stability of the trimer/dimer. When calculating hydrogen bonds of up to 3.5Å in SpPCNA there are 13 (7 for main chain only) between chain A and chain B and 11 between both A and C (6 for main chain only) and B and C (8 for main chain only) (Figure 2-19). In comparison hPCNA contains for the same parameters the same number of hydrogen bonds, but two more main chain to main chain bonds. These two contacts offer an explanation for the increased T_m of hPCNA in comparison to SpPCNA.

Despite the high degree of conservation of the hydrophobic pocket, the data from the thermal shift assay suggest that there are certain differences between both sliding clamps that are not obvious, but may be able to influence the way proteins recognize each other in both species. Proteins containing the PIP-box binding motif have been found also in the fission yeast genome. Despite the lack of an obvious homologue to p21, other proteins like DNA ligase I/cdc17 (Martin et al., 2004), DNA polymerase

δ/Cdc27 (Reynolds et al., 2000), Fen-1 (Tanaka et al., 2004), AP-endonuclease Apn2 (Dianova et al., 2001), uracil DNA N-glycosylase Ung1 (Ko and Bennett, 2005) and MSH6/MutS (Clark et al., 2000) protein have been found to either bind directly to SpPCNA *via* the PIP-box motif or at least contain the sequence (Table 2-5).

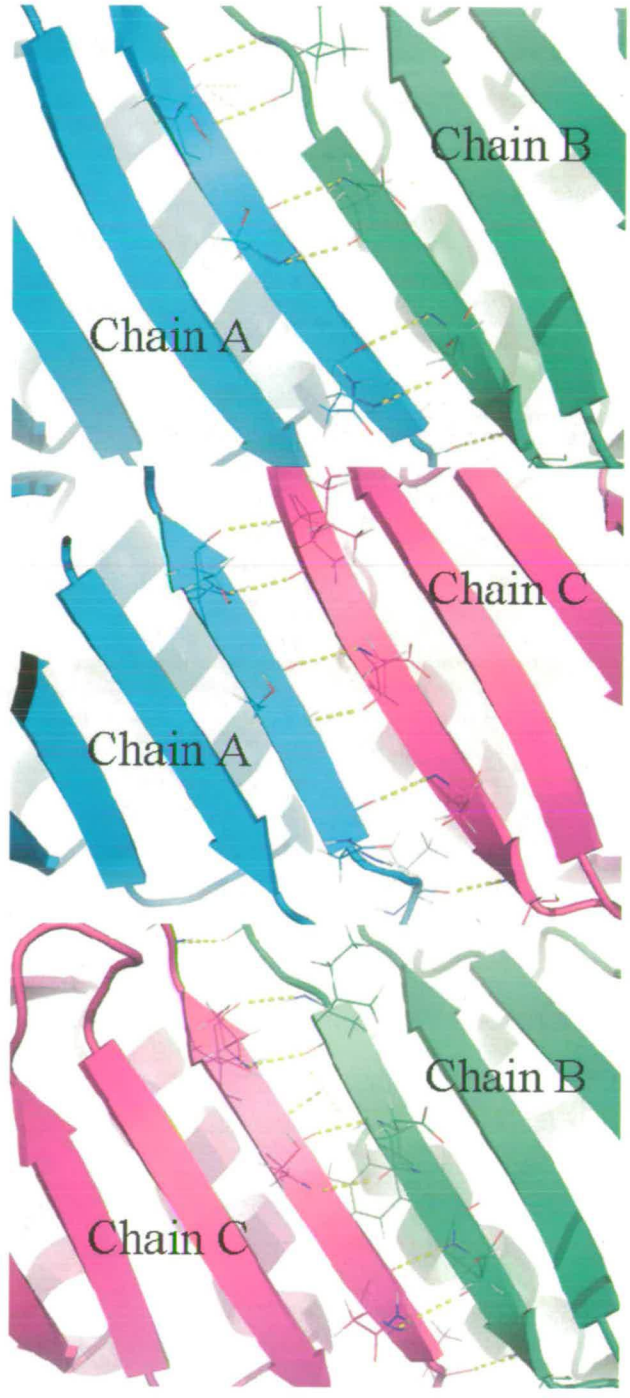


Figure 2-19 Hydrogen bonds between monomers in SpPCNA trimer.

The PIP-motif is mainly located near the termini of the proteins which makes it more likely to be solvent accessible and available for protein-protein interaction. Numerous other proteins have been found to contain the motif (140 with Q-x-x-I/M/L-x-x-Y/F-Y/F as search template); some have the sequence near a terminus, others have it in the middle of the sequence, which does not necessarily occlude the site. However, other proteins have shown to contain only a fraction of the binding motif.

An example is the adenine DNA glycosylase MutY (Table 2-5). This protein contains the full motif in human and mouse, but only L and F are present in fission yeast. Nevertheless, the remainder of the motif is essential for binding to PCNA (Chang and Lu, 2002). A search with parts of the motif will therefore undoubtedly uncover many more proteins that may bind to that pocket.

***S. pombe* proteins with homologues in humans proven to bind to hPCNA**	**Motif**	**Total length of PIP-motif containing proteins in *S. pombe***
AP-endonuclease Apn2	396-QSKLLSFF-403	523
uracil DNA N-glycosylase Ung1	25-QPRLDNFF-32	332
MutS protein homolog (mismatch repair, msh6)	22-QKTLFGFF-29	1254
Fen1	338-QGRLDSFF-345	380
Cdc27/DNA Pol δ	362-QKSIMSFF-369	372
cdc17/DNA ligase I	24-QSDISNFF-31	768
MutY	438-VTSLSNFK-445	461

Table 2-5 PIP-box motifs in fission yeast proteins.
Human homologues of these proteins have been found to bind to hPCNA directly.

2.3.7 Description of the hexagonal SpPCNA/p21(141-160) crystal structure

2.3.7.1 General structure and crystal packing

The asymmetric unit contains two monomers, which belong to two different trimeric rings of PCNA and one peptide molecule, which is shared between the two PCNA monomers (Figure 2-20, Figure 2-21).

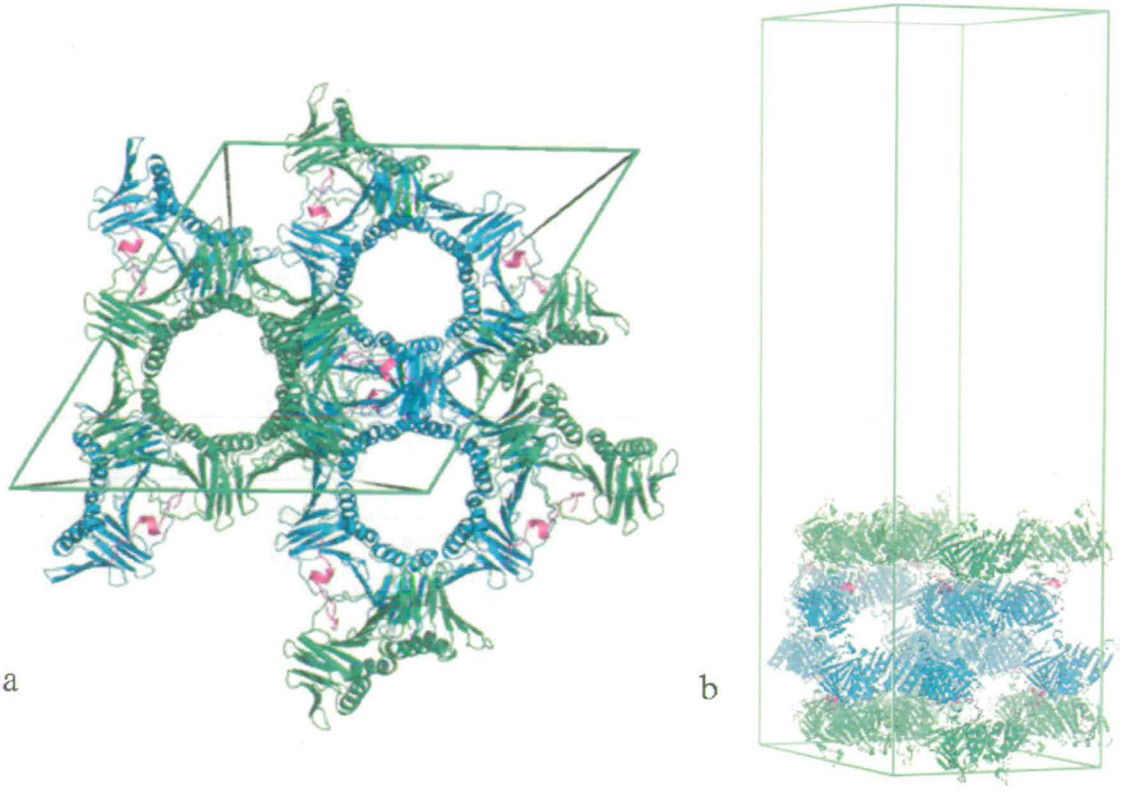


Figure 2-20 Crystal packing in hexagonal SpPCNA/p21(141-160) complex crystals.
a) View along axis c b) Side view of the unit cell. The two PCNA monomers of one asymmetric unit are shown in blue and green, while the shared peptide is highlighted as magenta. The green monomer contains the N-terminal fraction of the peptide which includes the PIP-box motif with its α-helical turn.

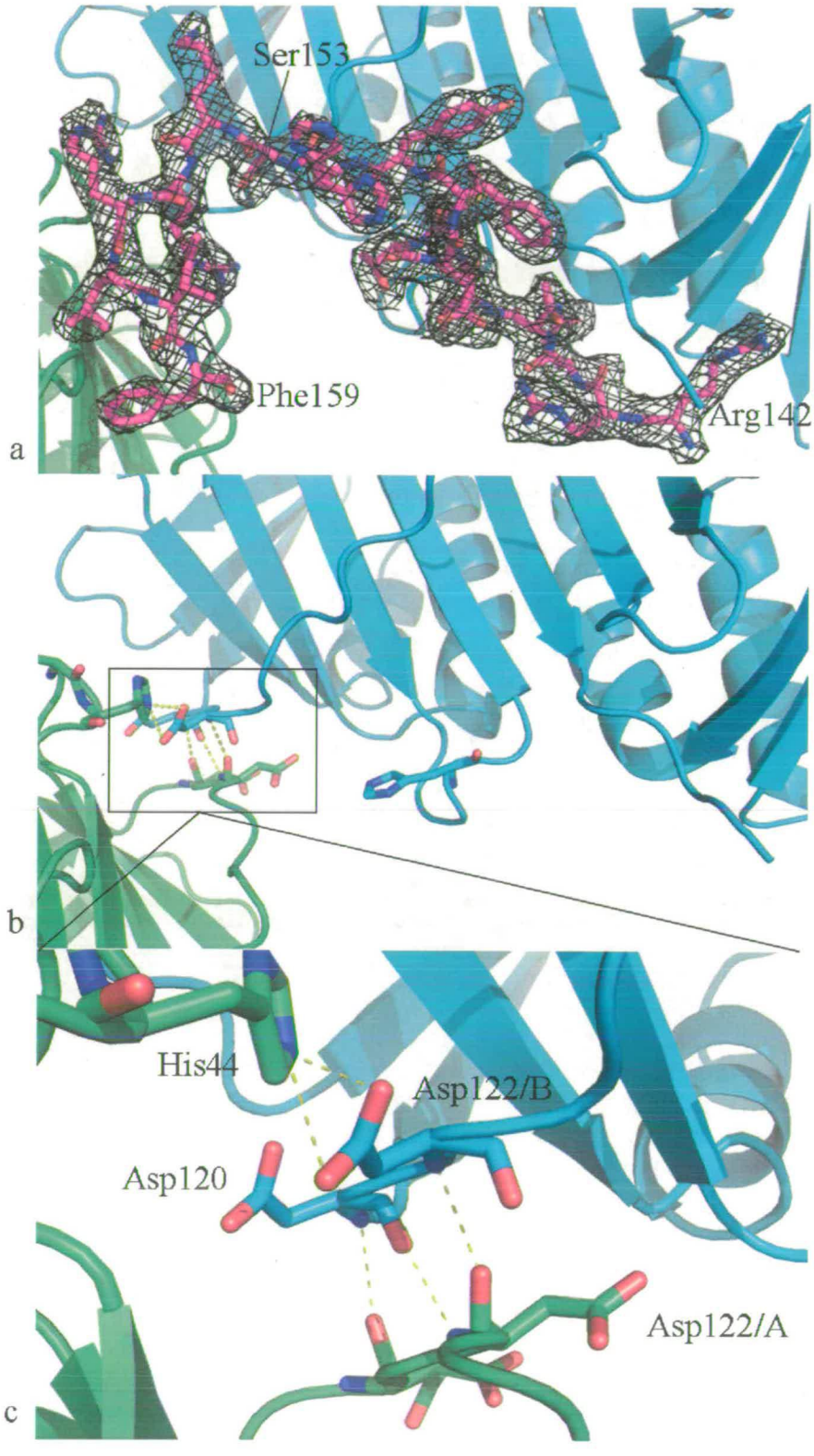


Figure 2-21 The complex of SpPCNA with p21 (140-160).
a) Electron-density for p21(141-160) in the SpPCNA/p21 complex structure. The map clearly shows the location of residues 141-159. While amino acids 142-153 are bound to Chain B (blue), Chain A (green) accommodates the remainder of the peptide. b) + c) Direct hydrogen bonds between the two monomers of SpPCNA.

In order for the unit to repeat, the two joined rings have to be fitted back to front with other pairs. The two molecules touch each other head to head. Apart from having the p21 peptide bound across, the two PCNA monomers are linked directly by a few interactions *via* their IDCLs (Figure 2-20, Figure 2-21).

Table 2-6 gives a list of distances <4.0 Å and possible contacts between the peptide and both PCNA monomers.

Residue in SpPCNA	Peptide/Chain C	Chain A Distance in Å	Chain B Distance in Å
CE MET 40	CA MET C 147		3.66
	CA MET C 147		3.23
	C MET C 147		3.98
	N THR C 148		3.82
	CG2 THR C 148		3.97
O SER 43	CZ ARG C 156	3.48	
	NH1 ARG C 156	3.3	
	NH2 ARG C 156	3.4	
CB HIS 44	CG ARG C 155	3.73	
O HIS 44	CA SER C 146		3.61
	CB SER C 146		3.71
	OG SER C 146		3.89
	C SER C 146		3.7
	N MET C 147		2.84
	CA MET C 147		3.71
	CB MET C 147		3.54
	N ARG C 156	3.72	
	NH1 ARG C 156	3.46	
	N LEU C 157	3.34	
	CB LEU C 157	3.58	
CG1 VAL 45	CD GLN C 144		3.62
	OE1 GLN C 144		3.95
	NE2 GLN C 144		3.5
	CD ARG C 156	3.68	
	NE ARG C 156	3.81	
	CZ ARG C 156	3.98	
	NH1 ARG C 156	3.4	
CG2 VAL 45	CG GLN C 144		3.72
	O THR C 145		3.98
C VAL 45	CG MET C 147		3.78
O VAL 45	CG MET C 147		3.51
	SD MET C 147		3.69
	CB LEU C 157	3.92	

	CG LEU C 157	3.88	
	CD2 LEU C 157	3.8	
C ALA 46	CG MET C 147		3.93
CB LEU 47	CG MET C 147		3.94
	CD1 LEU C 157	3.82	
O HIS 125	CA SER C 153		3.36
	CB SER C 153		3.26
CA LEU 126	C AHIS C 152		3.92
	C BHIS C 152		3.86
	O AHIS C 152		3.53
	O BHIS C 152		3.47
CD1 LEU 126	CB MET C 147		3.96
	O TYR C 151		3.82
	CD1 LEU C 157	3.54	
	O LEU C 157	3.79	
C LEU 126	O AHIS C 152		3.65
	O BHIS C 152		3.59
N GLY 127	C AHIS C 152		3.9
	C BHIS C 152		3.85
	O AHIS C 152		2.85
	O BHIS C 152		2.79
CA GLY 127	O AHIS C 152		3.81
	O BHIS C 152		3.77
O GLY 127	CB TYR C 151		3.38
	CG TYR C 151		3.82
	CD2 TYR C 151		3.84
	CB PHE C 159	3.81	
	CD1 PHE C 159	4.0	
CG1 ILE 128	CD1 TYR C 151		3.92
CD1 ILE 128	CE1 PHE C 159	3.54	
CG2 ILE 128	CZ PHE C 159	3.91	
CD PRO 129	CE2 TYR C 151		3.89
	CD2 TYR C 151		3.83
	CB PHE C 159	3.65	
	CG PHE C 159	3.74	
	CD2 PHE C 159	3.73	
CD1 ILE 131	OH TYR C 151		3.3
CB THR 206	CD ARG C 142		3.84
	NE ARG C 142		3.93
OG1 THR 206	NE ARG C 142		3.81
	CZ ARG C 142		3.45
	NH1 ARG C 142		3.24
CG2 THR 206	CD ARG C 142		3.75
O ASP 232	CE2 PHE C 150		3.13
	CD2 PHE C 150		3.69
CG1 VAL 233	OH TYR C 151		3.74
CB PRO 234	SD MET C 147		3.9
	CE MET C 147		3.92
	CD2 LEU C 157	3.78	

CG PRO 234	SD MET C 147		3.85
	CE MET C 147		3.76
	CG PHE C 150		3.75
	CD1 PHE C 150		3.6
	CE1 TYR C 151		3.92
	CD2 LEU C 157	3.78	
	O ILE C 158	3.71	
	CE2 PHE C 159	3.99	
CD PRO 234	CB PHE C 150		3.87
	CG PHE C 150		3.51
	CD1 PHE C 150		3.84
	CD2 PHE C 150		3.74
	OH TYR C 151		3.7
	O ILE C 158	3.67	
CB TYR 249	SD MET C 147		3.94
	CE MET C 147		3.63
	CD2 LEU C 157	3.89	
CG TYR 249	CZ PHE C 159	3.95	
CE2 TYR 249	CZ PHE C 159	3.94	
CD2 TYR 249	CZ PHE C 159	3.52	
N ALA 251	SD MET C 147		3.77
N ALA 251	CD2 LEU C 157	3.82	
CB ALA 251	NE2 GLN C 144		3.84
	CG2 THR C 145		3.32
	C THR C 145		3.72
	O THR C 145		3.31
	C SER C 146		3.58
	O SER C 146		3.79
	N MET C 147		3.7
	CB ARG C 156	3.93	
	C ARG C 156	3.79	
	O ARG C 156	3.68	
	N LEU C 157	3.92	
C ALA 251	NE2 GLN C 144		3.54
O ALA 251	CD GLN C 144		3.92
	NE2 GLN C 144		2.91
N PRO 252	CG2 THR C 145		3.74
CD PRO 252	CG2 THR C 145		3.54
	CE1 PHE C 150		3.99
	CZ PHE C 150		3.98
C PRO 252	NE2 GLN C 144		3.92
O PRO 252	O ARG C 143		3.77
	NE2 GLN C 144		3.51
	N THR C 145		3.14
	CA THR C 145		3.74
	CB THR C 145		2.97
	OG1 THR C 145		3.74
	CG2 THR C 145		3.22
CA LYS 253	O ARG C 143		3.15

CB LYS 253	O ARG C 143		3.88
CG LYS 253	OE1 GLN C 144		3.4
CD LYS 253	OE1 GLN C 144		3.87
C LYS 253	O ARG C 143		3.44
O LYS 253	CD ARG C 142		3.81
	NE ARG C 142		3.55
N ILE 254	C ARG C 143		3.88
	O ARG C 143		2.82
CA ILE 254	O ARG C 143		3.98
CG1 ILE B 254	O ARG C 143		3.88
	OG1 THR C 145		3.83
CD1 ILE 254	OG1 THR C 145		3.99
C ILE 254	N ARG C 143		3.96
O ILE 254	CA ARG C 142		3.14
	CB ARG C 142		3.75
	C ARG C 142		3.33
	N ARG C 143		2.75
	CA ARG C 143		3.74
	CG ARG C 143		3.74
	O ARG C 143		3.58

Table 2-6 Distances between p21(141-160) and both monomers in the SpPCNA complex structure.
The yellow highlighted selection shows hydrogen bonds (or possible ones) between the peptide (KRRQTSMTDFYHSKRRLIFS) and either of the two monomers in the asymmetric unit.

2.3.7.2 Comparison of unliganded and complexed SpPCNA

The N and C terminal residues of the peptide p21 (141-160) were not visible and only electron density for residues 142-159 was detectable. No His-tags are visible in this structure. When superimposing a monomer of unliganded SpPCNA and one from the complex with p21 peptide (ligand removed) (Figure 2-22), it is obvious that the secondary features remain the same. Loops, including the IDCL, are the only regions with significant backbone changes, but as they are generally flexible and assume different conformations in each single monomer, this is to be expected.

Particularly in the area of the hydrophobic pocket the superimposition shows identical conformation. Interestingly the three β-strands (βI_1, βC_2 and βD_2) are more extended in the unliganded structure than in the complex: residues 109-117 vs. residues 111-117, residues 167-172 vs. residues 166-170, residues 176-182 vs. residues 178-182.

As mentioned above no density was observed for His-tags (Figure 2-20). A reason might be the unavailability of connector loop and hydrophobic pocket for coordination of the histidine residues in this structure (see section 2.3.5.3 for crystal packing of unliganded SpPCNA) as connection between the two monomers is established *via* these structural features (Figure 2-20, Figure 2-21).

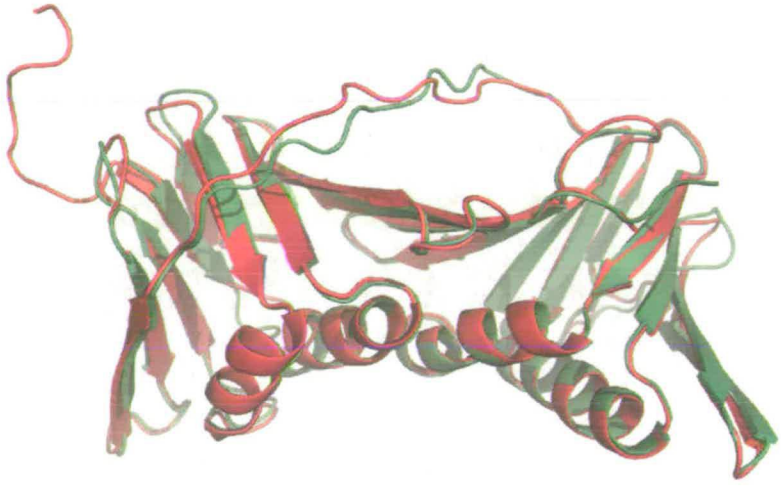


Figure 2-22 Superimposition of the backbone structure of unliganded and complexed SpPCNA (peptide omitted).
Key: blue = unliganded, green = complexed, RMSD: 0.67Å

Apart from the shared peptide which may act a bit as a connector holding the two separate PCNA rings together, five hydrogen bonds are formed between the monomers directly (His44/NE2 (chain A) ··· Asp122/OD1 and OD2 (chain B),

Asp122/N (chain A) ··· Asp120/O (chain B)) or by bridging water molecules (Figure 2-21b+c).

Some electron density is displayed for the C-terminus of Chain A (1-254), though it was not possible to model the positions of the side chains correctly. Chain B is modelled up to residue 255. However, this additional density for the C-terminus agrees with previous finding that the C-terminus is coordinated upon complex formation with p21 peptide derived as the N-terminus of the peptide pairs up with the C-terminus of the sliding clamp.

Furthermore, the IDCL is more rigidified in the complex with p21, shown by a reduction in B factors, which was also observed as previously in the human PCNA complex structure, although the difference is more obvious in the human case (Table 2-7).

However, while the peptide binds along the IDCL in the human structure (Figure 2-23), in the fission yeast example the C-terminal half (Ser153-Phe159) bends away from the monomer in which it occupies the hydrophobic pocket with the PIP-box motif. So while the N-terminus with PIP-box motif is bound to monomer B, the C-terminus is coordinated into the hydrophobic pocket of an adjacent monomer, which belongs to a second trimer, in this structure called monomer A (Figure 2-20, Figure 2-21, Figure 2-23).

	Unliganded SpPCNA				SpPCNA/p21(141-160)			Unliganded hPCNA	hPCNA/ p21(141-160)
IDCL Residue Number	Chain A	Chain B	Chain C	**Average**	Chain A	Chain B	**Average**	Chain A	Chain A
121	45.63	43.08	39.53	**42.75**	27.61	30.23	**28.92**	119.06	16.45
122	48.79	46.14	41.08	**45.34**	25.99	33.83	**29.91**	126.10	31.29
123	51.50	48.75	42.74	**47.66**	30.94	37.54	**34.24**	129.83	29.38
124	52.57	50.35	44.28	**49.07**	34.07	36.62	**35.35**	132.17	27.67
125	53.22	51.11	44.35	**49.56**	37.5	39.44	**38.47**	129.24	21.09
126	51.81	50.67	44.40	**48.96**	38.97	35.31	**37.14**	128.65	16.47
127	49.02	50.47	42.16	**47.22**	40.6	33.32	**36.96**	125.14	18.10
128	45.84	50.08	39.58	**45.17**	42.85	35.63	**39.24**	127.98	15.66
129	45.08	49.15	38.19	**44.14**	47.87	38.96	**43.42**	124.98	34.41
130	44.43	47.56	37.45	**43.15**	52.34	42.71	**47.53**	111.91	36.66
131	44.21	45.54	36.44	**42.06**	47.89	40.17	**44.3**	91.93	41.35
132	43.53	42.98	35.67	**40.73**	40.03	39.43	**39.73**	126.97	45.42

Table 2-7 B factors for the IDCL residues in both unliganded and complex crystal structures of SpPCNA and hPCNA. The overall B factor for main chains atoms is 40.41Å^2 for the unliganded structure and 37.13Å^2 for SpPCNA in complex with p21(141-160).

In the hPCNA structure the peptide forms hydrogen bonds with the IDCL (Table 2-8). Seven out of eleven of those bonds involve main-chain to main-chain contacts and should be available for complex formation in the yeast case as well. Residues Asp120 and Glu124 are conserved between both fission yeast and human PCNA, so the side chain contacts provided by these amino acids should be possible for both proteins.

Table 2-6 shows distances of 4 Å or less between the peptide and either of the two chains in the asymmetric unit.

As the superimposition of unliganded hPCNA and SpPCNA shows (Figure 2-12a, Figure 2-23) there are only minor differences in the backbone and surface.

IDCL residues in hPCNA involved in hydrogen bonds with p21 (141-160)	Equivalent IDCL residues in *S. pombe* PCNA
Met119 (O)	Met119
Asp120 (OD1)	Asp120
Leu121 (N and O)	Ile121
Asp122 (O)	Asp122
Val123 (N and O)	Gln123
Glu124 (OD2)	Glu124
Gln125 (N and O)	His125
Gly127 (N)	Gly127

Table 2-8 Residues involved in bonding between the IDCL in hPCNA and p21(141-160).
The highlighted amino acids contribute entirely towards main-main chain contacts. Most of those are also identical or similar in both proteins. The two remaining residues Asp120 and Glu124 form contacts *via* side chain atoms, which are conserved in hPCNA and SpPCNA.

Even the residues of the small second pocket (human/SpPCNA: Cys/Asn27, Met40, Ala67, Gly69, Leu/Ile 121 and Val/Gln123) (Figure 2-12b) that accommodates the C-terminus of the peptide in the human complex structure are largely conserved. Most of the residues are conserved with very similar orientation of side chains.

However, a few small changes on the surface, mainly in the linker region, do exist (Figure 2-12c) and residues Cys27 and Val123 are replaced by Asn and Gln respectively in fission yeast (Figure 2-12a). So there is the possibility that these variations between hPCNA and SpPCNA lead to the observed difference in p21 binding.

When superimposing the complex of human PCNA and p21 with SpPCNA and the same peptide, then one can see that the PIP-Box residues (Gln144, Met 147, Phe150 and Tyr 151) are in both structures at the same position (Figure 2-23, Figure 2-24a), which supports the idea of the conserved pocket/binding motif that has been observed in previous structures of sliding clamps/PCNA (Chapter 1).

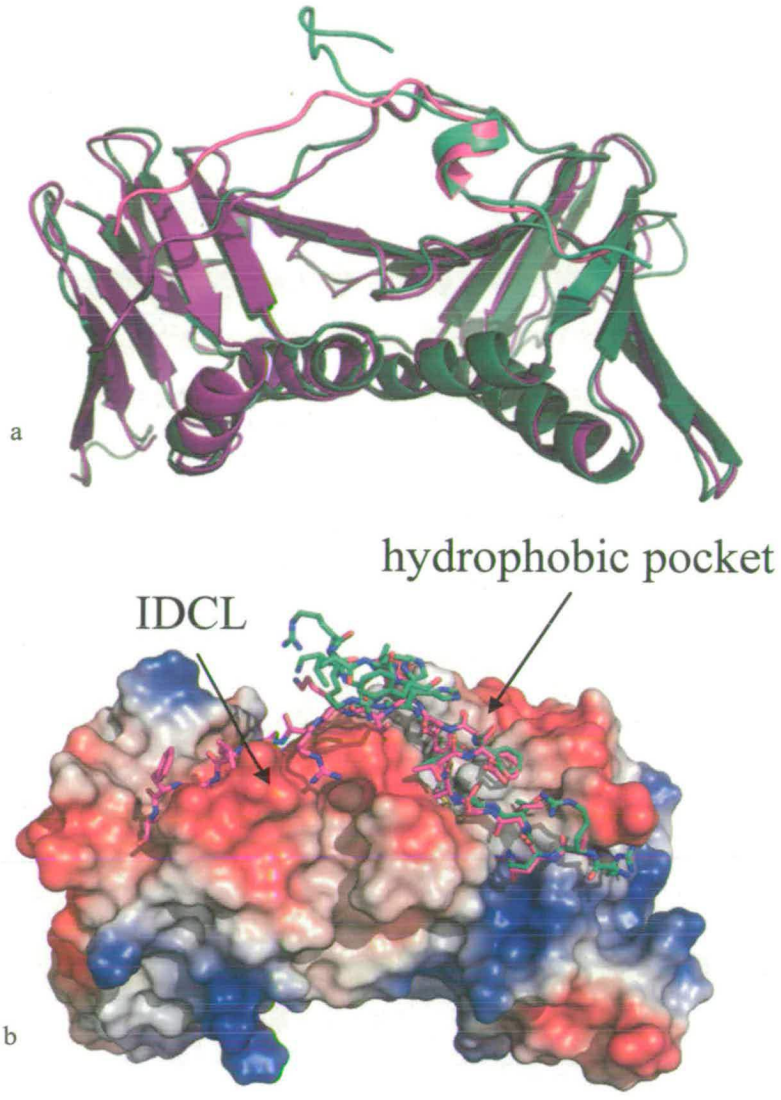


Figure 2-23 Comparison of the complexes of human and fission yeast p21 complex structures. a) Superimposition of SpPCNA (dark green) with p21 peptide (light green) bound and hPCNA (purple) with p21 peptide (magenta); b) Surface charge of SpPCNA with stick models of p21peptides as it binds to hPCNA (magenta) and to SpPCNA (light green).

As Gulbis and colleagues (Gulbis et al., 1996) found out, the second part of the peptide binds along the connector loop in the human complex structure, where it forms a 2-stranded β-sheet.

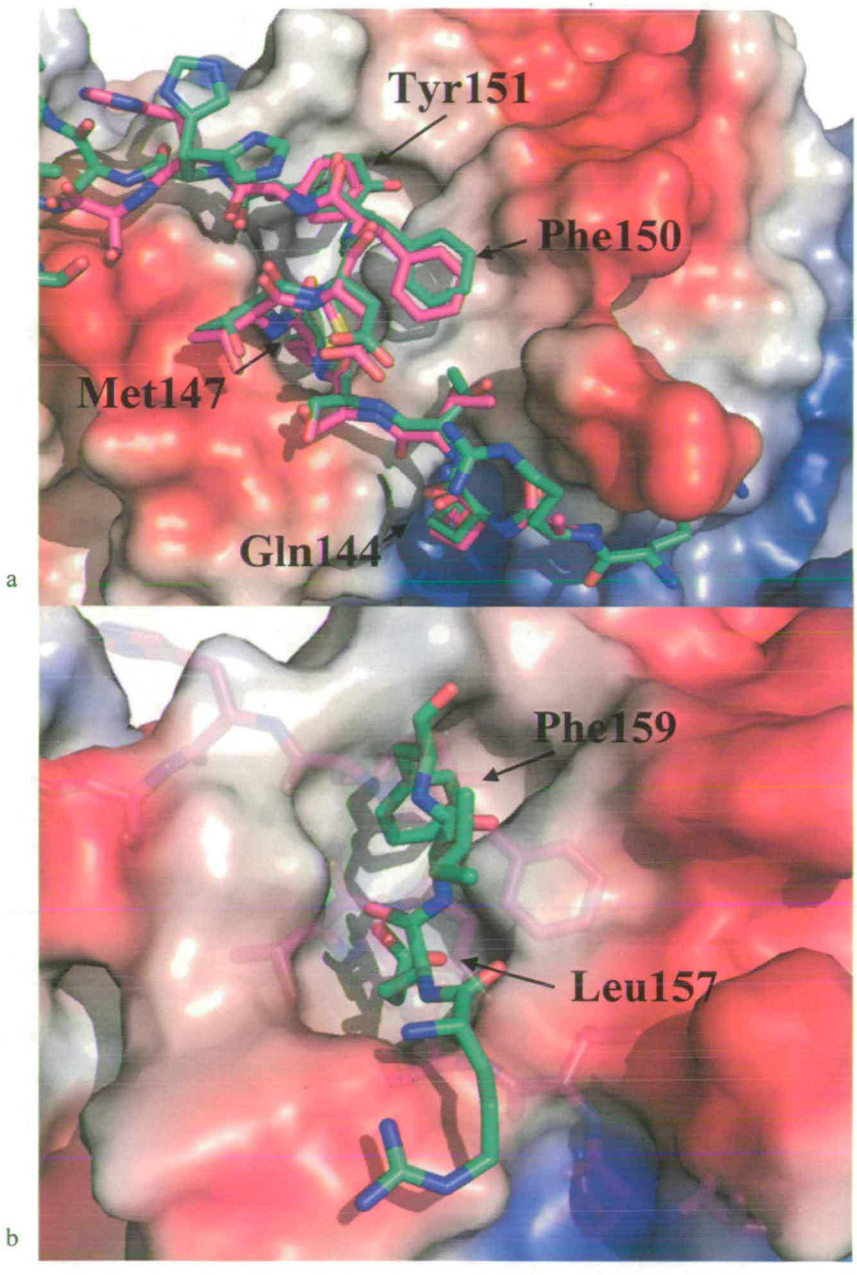


Figure 2-24 Binding of p21(141-160) to SpPCNA in comparison to hPCNA.
a) The PIP-box part of the peptide seems to bind in a very similar way to the hydrophobic pocket in both complex structures. Gln144, Met147, Phe150 and Tyr151 are exactly in the same place with same side chain orientations in the human (magenta) and the fission yeast (light green) structure. b) The very C-terminal portion of the peptide is coordinated into the hydrophobic pocket of the second monomer in the asymmetric unit. In magenta is shown the p21 peptide as it binds to the pocket in hPCNA. Only the residues 156-159 of the peptide in the SpPCNA are shown for clarity. Here Phe159 acts as Tyr151 and Leu157 assumes roughly the position of Met147.

In contrast with the SpPCNA complex structure, the second half of the peptide does not form a β-sheet with the IDCL strand as it does in the hPCNA structure. Instead it dissociates from the one subunit at Ser153 (Figure 2-20, Figure 2-21, Figure 2-23) and is coordinated by the hydrophobic pocket of a neighbouring ring. While Phe159 assumes the position of Tyr151 in the PIP-box motif (Figure 2-24b), which agrees with the hydrophobic interaction of the PIP-box motif on that site, Leu157 is found at the same site where Met147 is positioned in the PIP-box motif.

2.3.8 SPR binding experiments of SpPCNA and human p21 peptides

The crystal structure of SpPCNA in complex with human p21(141-160) suggests that the portion of the peptide which binds along the linker strand in the human crystal complex has a low affinity for the yeast linker as it appears to prefer to bind to the hydrophobic pocket of an adjacent ring *via* aromatic/hydrophobic. The possible reasons for that have been discussed above. In order to find out if a peptide version missing this portion [p21 (141-152)] has a similar affinity for SpPCNA as the full-length one, SPR studies have been carried out. If the second portion is freely mobile in solution then the affinity of the full length peptide and the shorter version should be similar, otherwise the longer peptide should bind more tightly to the sliding clamp.

Globally fitting a kinetic model where a 1:1 complex is formed between His-SpPCNA and p21(141-152) to data similar to those illustrated in Figure 2-25 gave very good fits. Kinetic data suggest a reaction with a fast on rate at $1.9*10^6$/Ms and a

fast off rate of 0.47/s at 25 °C for the shorter version of the p21 peptide (Figure 2-25, bottom). A steady-state affinity of about 500 nM (Figure 2-25, top) for p21(141-152) was calculated, which is close to the ITC derived 307nM (see section 3.3.3) for the interaction of p21(141-152) with hPCNA. These values increase with temperature. The K_d of p21(141-160) for the interaction with hPCNA has been determined as 114 nM by ITC (Zheleva et al., 2000), which agrees with the fact that the second portion of the full-length peptide is contributing to the binding by making interactions with the linker strand as seen in the crystal structure. In the fission yeast crystal complex this second part of the interaction is not observed, so the K_d for p21(141-160) is likely to be similar to that of p21(141-152) for SpPCNA.

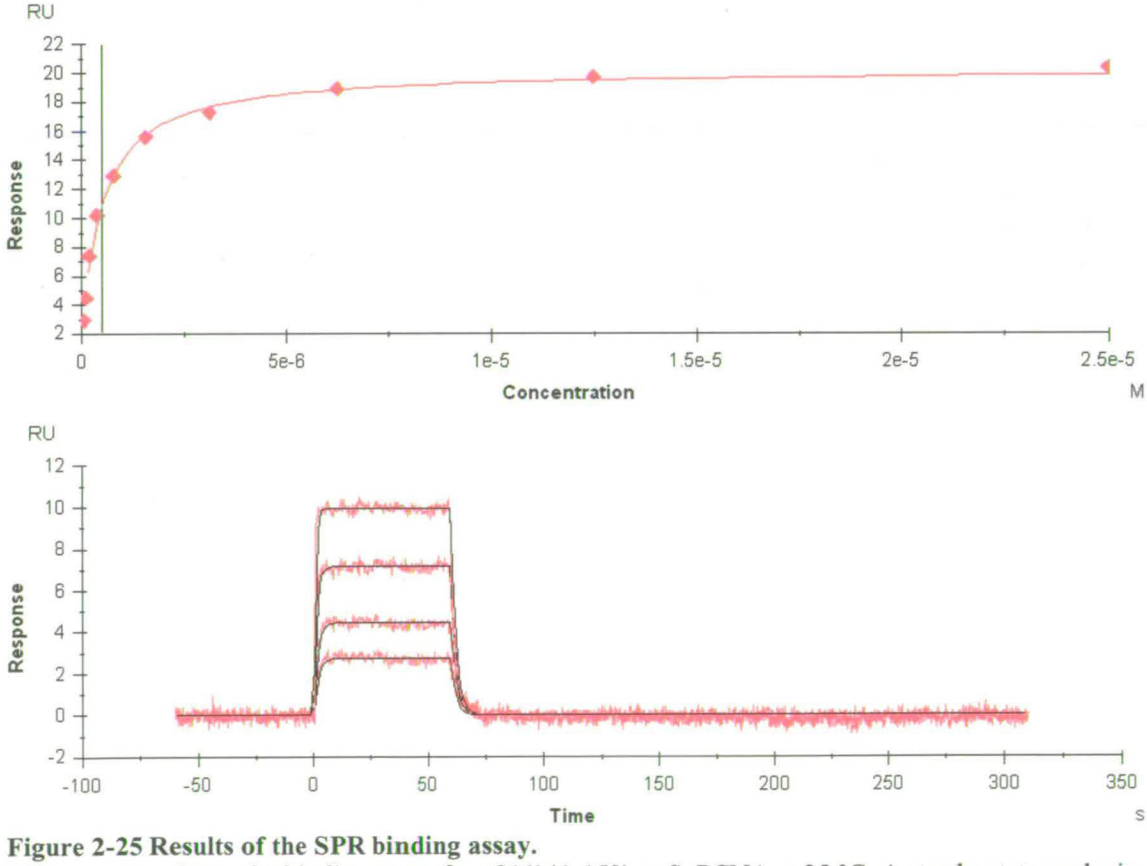


Figure 2-25 Results of the SPR binding assay.
The top curve shows the binding curve for p21(141-152) to SpPCNA at 25 °C. A steady-state analysis yielded an affinity of about 500 nM for this reaction. The bottom graph shows the quick on and off rates for this binding reaction at 25 °C.

2.4 SUMMARY, CONCLUSION AND FUTURE WORK

The crystal structures of SpPCNA in the unliganded form or in complex with p21(141-160) show four different ligands in the hydrophobic pocket. In the unliganded structure two His-tags of adjacent rings bind to the supposedly specific binding site that usually accommodates the PIP-box motif.

In the crystal complex of SpPCNA and p21(141-160) one peptide molecule fills two hydrophobic pockets that are located on two adjacent rings. One pocket contains the PIP-box motif as expected from the hPCNA/p21 co-crystal structure. The second site accommodates the C-terminus of the peptide containing hydrophobic/aromatic residues.

Histidine-rich hexapeptides mimicking the His-tag showed different binding affinities for SpPCNA and hPCNA, respectively. Furthermore, the binding mode for the p21 peptide to SpPCNA differs significantly from how it binds to hPCNA. All these results suggests that the observed changes in amino acid sequence and side chain orientation between both eukaryotic PCNA molecules have a considerable effect on specificity of the binding pocket.

The fact that both histidine-rich peptides or aromatic/hydrophobic residues can bind to the hydrophobic pocket, which is supposed to be specific to the PIP-box motif, seems to suggest that there may be a lack of specificity concerning the ligand. An exact match to the motif does not seem to be required for a low binding affinity at least.

Concerning the His-tag mimic hexapeptides, further studies are required to see if those two hexamers bind specifically to the hydrophobic pocket or if they cover the surface of SpPCNA non-specifically and help to protect it from thermal denaturation. Competition experiments (e.g. FP) using p21 peptides and the hexamers at the same time would show if binding of both at the same time is possible. Furthermore, mutant PCNA which is unable to form a trimer due to single amino acid substitution (Jonsson et al., 1995) would be small enough to fly in MALDI-TOF experiments. Analysis of the spectra of the PCNA-peptide mixture could then be used to determine the stoichiometry of the complex.

CHAPTER 3. A structure-based approach to find inhibitors for PCNA as an anti-cancer drug target

3.1 INTRODUCTION

As discussed in Chapter 1, p21 is an association control factor for PCNA in the cell. When p21 binds to the loading clamp it prevents DNA polymerase δ from attaching itself to PCNA and as result brings DNA synthesis to a stop (Waga et al., 1994). PCNA is also involved in DNA repair, and this process can continue while the cell-cycle is paused until a decision is made for either completion of DNA repair or, in case of too severe damage, apoptosis.

PCNA has been shown to be deregulated/overexpressed in several cancer forms like esophageal cancer (invasive squamous-cell carcinoma) (Kuwano et al., 1998), head and neck squamous cell carcinoma (Shin et al., 1993) and non-small cell lung cancer (Fontanini et al., 1992), while p21 has been found to be frequently downregulated in laryngeal squamous cell carcinoma (Hirvikoski et al., 1999) and primary colon carcinomas (Bukholm and Nesland, 2000).

PCNA's hydrophobic binding pocket to which p21 binds has been discussed already in this thesis. As shown in Chapter 1 it is highly conserved among the loading clamps and accommodates the conserved residues of proteins carrying the PIP-box binding motif (Q-x-x-h-x-x-a-a), where 'x' stands for any amino acid, while 'h' denotes a moderately hydrophobic and 'a' an aromatic residue. In previous work this pocket was examined and evaluated as a druggable site suited for the design of peptidomimetics (Kontopidis et al., 2005).

The idea of trying to mimic protein-protein interactions is a fairly new one. In 2004 nutlin, a small-molecule antagonist of MDM2 was found (Vassilev et al., 2004). Nutlin prevents the interaction of MDM2 with p53, a tumour suppressor protein, and keep levels of p53 high. Both compounds were developed by structure-based drug design using X-ray structures of the target proteins complexed with peptides of the binding partners.

The lack of numerous examples of drugs developed by this approach hints at possible problems inherent in the approach of structure-based drug design for protein-protein interactions. There are a number of parameters which are unique to protein-protein interactions and which need to be considered when deciding on the suitability of a site. First of all the drug molecule must achieve high affinity and selectivity for that site. Most protein-protein interaction interfaces are rather large and devoid of deep indentations making it hard to mimic by a small molecule (Fry and Vassilev, 2005). While a protein will be able to make numerous contacts with its partner a small molecule will only ever be able to cover a subset of these and hence struggle for adequate affinity. A lack of specificity will lead to toxic effect in the body due to non-specific protein binding. So it is important to maximise the energetic contribution of each of the interactions for maximum affinity. A good drug site will sequester a large portion of the surface area of the ligand away from solvent. However, if the drug target is intracellular a ligand with too much polar character will have difficulties getting across the cell membrane. The Lipinski rules that should be met by the drug candidate are 1) a molecular weight under 500g/mol, 2) less than 5 hydrogen bond donors, 3) less than 10 hydrogen bond acceptors and 4) a partition coefficient (logP) under 5 which is the $\log_{10}$ of the ratio of the concentration of the

molecule in n-octanol over the concentration in water (Lipinski, 2000; Lipinski et al., 2001).

It is difficult to design good fitting molecules from scratch, so library screens tend to be carried out. There are an estimated 10^{40} drug-like molecules (Lipinski, 2000), so it seems it should be possible to find a suitable molecule for each pocket. Unfortunately, predicting the affinity of a molecule is difficult as it requires a detailed understanding of the interactions between the compound and the protein as well as the interactions between the compound and the solvent when it is free in solution.

For this study the p21 binding pocket in PCNA was closely examined using the database-mining programme LIDAEUS (Wu et al., 2003), which searches 3D databases for potential ligands that meet the criteria linked to a specific binding pocket. This *in silico* screening (carried out by Simon Harding) provided potential ligands which were examined with two binding assays discussed in this chapter. The first assay determines the thermal denaturation/unfolding of proteins in the presence of ligands and the second assay uses labelled PCNA-binding peptides and monitors a change in Fluorescence Polarization by competitive ligand binding.

3.2 THEORY OF THE THERMAL STABILITY OF PROTEINS AND FLUORESCENCE POLARIZATION

3.2.1 Thermal Stability

A protein is functional if it exists in its native, correctly folded state. In this form the secondary and tertiary structure elements are maintained by hydrogen bonding, keeping the bulk of hydrophobic residues inside and hence hidden from the solvent. Generally charged amino acids are found on the surface of the protein where they make interactions with solvent molecules. Upon exposing the protein to increasing amounts of heat the hydrogen bonds are broken leading to partial or total unravelling of the protein. The covalent bonds between amino acids that form the backbone of a protein on the other hand remain intact. The process of heat denaturation is generally considered to be irreversible. While denaturation is also possible by exposure to extreme cold, this type of unfolding is often reversible (Privalov, 1990). From a thermodynamic point of view denaturation is achieved when sufficient energy is transferred to the native protein form to induce a change in its molecular conformation. This energy is made up of two parts: 1) the activation energy to overcome the energetic barrier and 2) the total heat of the system absorbed/released during the unfolding process (enthalpic) (Figure 3-1).

Proteins differ in their sensitivity to thermal denaturation, which can be observed by monitoring the transition or melting temperatures (T_m). This is defined as the temperature at which 50% of the protein denatures. However, the general definition for T_m is the temperature at which the rate of unfolding is at its maximum. For mammalian proteins the onset of thermal denaturation usually lies between 40 °C and 45 °C. When reaching the onset temperature first changes in the structure can be

detected. The T_m is also dependent on the rate of heating. The transition from native to denatured state is usually complete within 10 - 12°C (Lepock, 2003) from the onset.

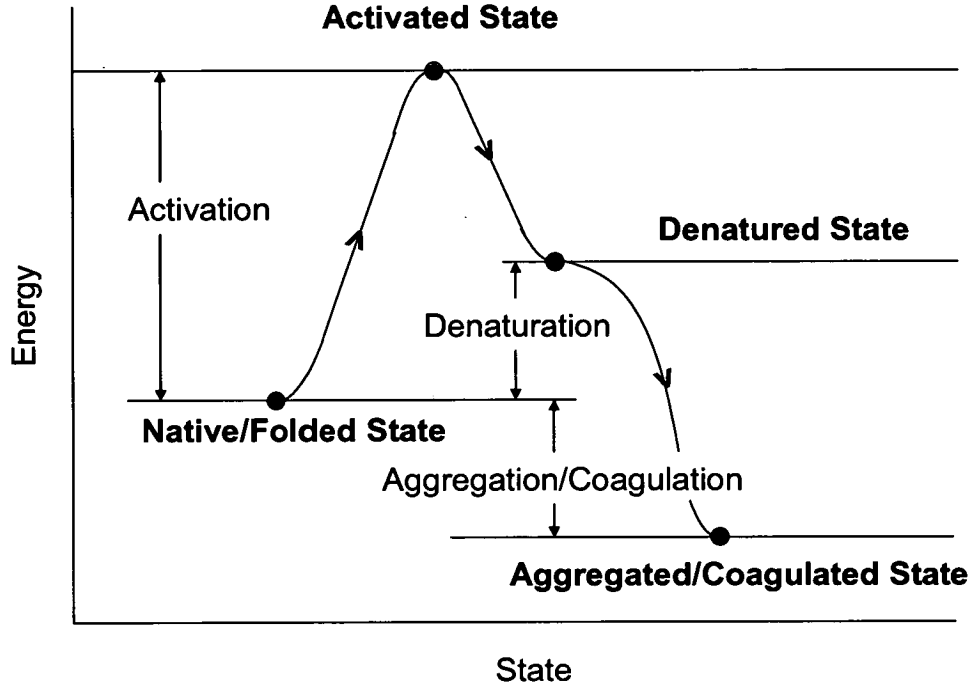


Figure 3-1 Diagram of the relationship between energy and the state of the protein.
The protein exists at the beginning in its native state. When enough energy in the form of heat is added to the system in order to overcome the energy barrier (activation energy) the protein is activated. Hydrogen bonds are broken up and the protein is denatured. At this point, due to increased entropy, the protein contains more energy than it did in its native state. Alternatively, denatured protein molecules may aggregate in a random fashion which leads to a decrease in entropy and to more favourable conformations, therefore leaving the molecule at a lower energetic state in comparison with the starting point.

Since the term denaturation is often inappropriately used for other protein states, it is worthwhile to define them here. **Aggregation** describes in general terms the formation of higher molecular weight complexes of proteins *via* protein-protein interactions. These complexes are usually not functional, in contrast to naturally occurring multimers. In more detail, we speak of **coagulation** when already

denatured protein randomly aggregates, which is usually an irreversible process. **Gelation**, however, is an orderly aggregation of native or denatured proteins and can be reversible. It is important to note that proteins tend to aggregate as part of the denaturation process and that those complexes may contain both native and denatured proteins.

A protein can be stabilised and to some degree protected from heat denaturation when a ligand in form of a peptide or a small molecule binds to the surface of the native form. This stabilization is perceived as an increase in T_m. By contrast a small molecule that binds to the even slightly unfolded form of a protein will cause a T_m shift to a lower temperature as it further destabilizes the protein (Brandts and Lin, 1990). This effect has been used in drug screening (Waldron and Murphy, 2003) and is especially useful in the absence of activity assays, as is the case for PCNA.

3.2.2 Fluorescence Polarization

Fluorescence is the process of photon emission as a result of the return of an electron in a higher energy orbital back to a lower energy orbital.

Fluorescence emitting molecules, also called fluorophores, remain excited for between 10^{-5} to 10^{-8} s, while the process of excitation, during which energy in form of a photon is absorbed, takes only about 10^{-15} s. The delay between absorption and fluorescence leaves enough time for fluorescent molecules to tumble and diffuse.

Polarization describes a preferential orientation in space. With light waves the transverse oscillation can be polarized. When light in the form of electromagnetic waves is emitted from many atoms or molecules with random orientations, those emitted waves will also have a mixture of orientations. The light in this case can be

described as unpolarized. When the emissions are selected for a particular orientation, the light is partially polarized. When the oscillating fields become aligned and oscillate in one plane only, the light is linearly polarized light.

Transition moments for excitation and emission are fixed within fluorescent molecules and are termed absorption and emission dipoles respectively. Fluorophores only absorb photons with electric oscillations parallel to the electronic transition moments of the fluorophores themselves. In solution, fluorescent molecules will assume random orientations. Hence a beam of polarized light will selectively excite only those molecules with absorption dipoles parallel to the electric oscillation of the excited photon. The result is partially polarized fluorescence emission, also known as Fluorescence Polarization (FP) (Mann and Krull, 2003).

In experiments either the change in fluorescent polarization or anisotropy (directional dependence) is measured, while the angle between the absorption and emission dipole defines the maximum measurable anisotropy.

The equation defining fluorescence anisotropy (r) is

$$r = (I_{VV} - GI_{VH}) / (I_{VV} + 2GI_{VH}) = (I_{VV} - GI_{VH}) / I_{TOT}. \qquad (1)$$

red = total fluorescence intensity

And for polarization (P) it is

$$P = (I_{VV} - GI_{VH}) / (I_{VV} + GI_{VH}). \qquad (2)$$

I_{VV} = intensity for parallel or vertical component of polarized emitted light

I_{VH} = intensity for perpendicular or horizontal component of polarized emitted light

G = correction factor to account for instrumental differences in detecting emitted components as the glass of the photo tube can introduce parallel diffraction anomalies. $G = I_{VV}/I_{VH}$

In the experiment, the sample containing the fluorescent probe is excited with linear polarized light and the vertical and horizontal components of the intensity of the emitted light are measured. These intensities are then used to calculate P or r using equations (1) and (2) (Figure 3-2).

P has possible values between -0.33 and 0.5 and r ranges from -0.25 to 0.4. The maximum anisotropy is found in solutions in which the absorption and emission dipoles are co-linear and where there are no depolarizing processes (Devlin, 1997).

In a non viscous solution molecules are free to rotate and move, so there is a larger loss of fluorescence polarization than in more viscous environments, in which the rotation of molecules is slowed down significantly. A reduction of rotational speed also occurs when the fluorophore binds to a larger molecule. By attaching a fluorophore to a small molecule or ligand the binding of that molecule to another of at least equal size can be monitored *via* its speed of rotation.

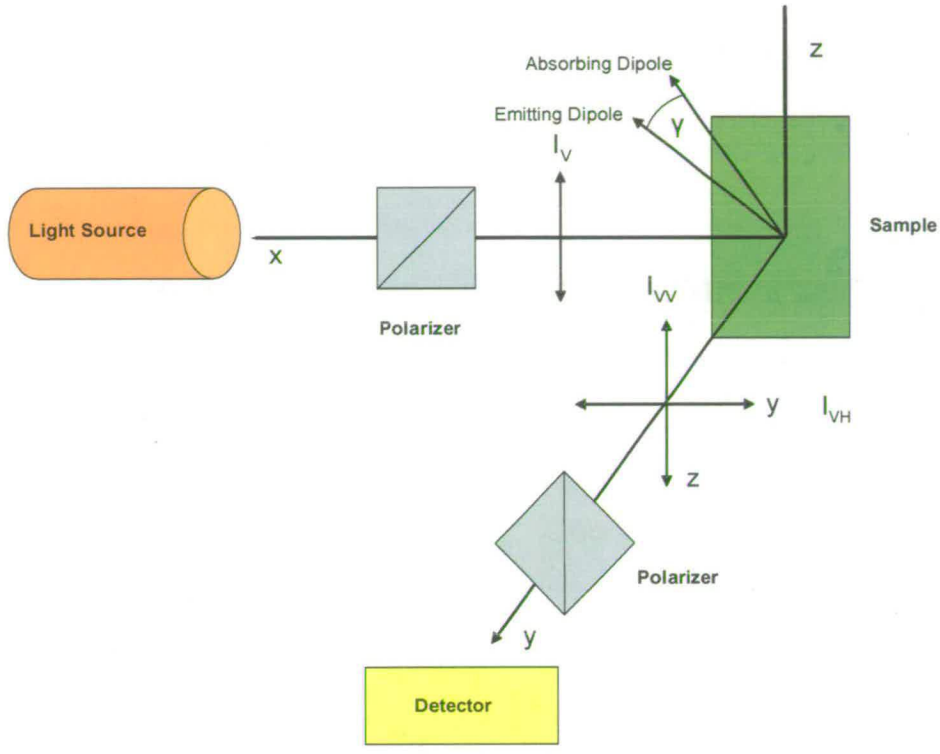


Figure 3-2 Schematic layout of a general Fluorescence Polarization experiment. [adapted from (Mann and Krull, 2003)].
A light source emits photons/waves with different polarization planes. They get filtered through a polarizer and then passed through the sample as linear polarized light. Any light that passes straight through the sample is ignored. The light that gets emitted in an angle (up to 90°) is passed through another polarizer and the intensities for vertical and horizontal emissions in relation to the original plane are monitored.

According to the original model described by Perrin (Perrin, 1926), which was derived from the Brownian motion theory, FP is dependent on a number of different factors which can be explained by equation 3:

$$(1/P-1/3) = (1/P_0-1/3)(1+\tau/\varphi) \qquad (3)$$

P = polarization

P_0 = initial polarization

τ = excited state life time of the dye

φ = rotational correlation time of the dye or dye conjugate, which can be estimated by

$$\varphi = \eta V/RT. \qquad (4)$$

V is the molecular volume of the fluorophore or dye conjugate and which can be estimated from the MW of the dye. η is the solvent viscosity, while T and R represent temperature and gas constant respectively (Haugland, 2005).

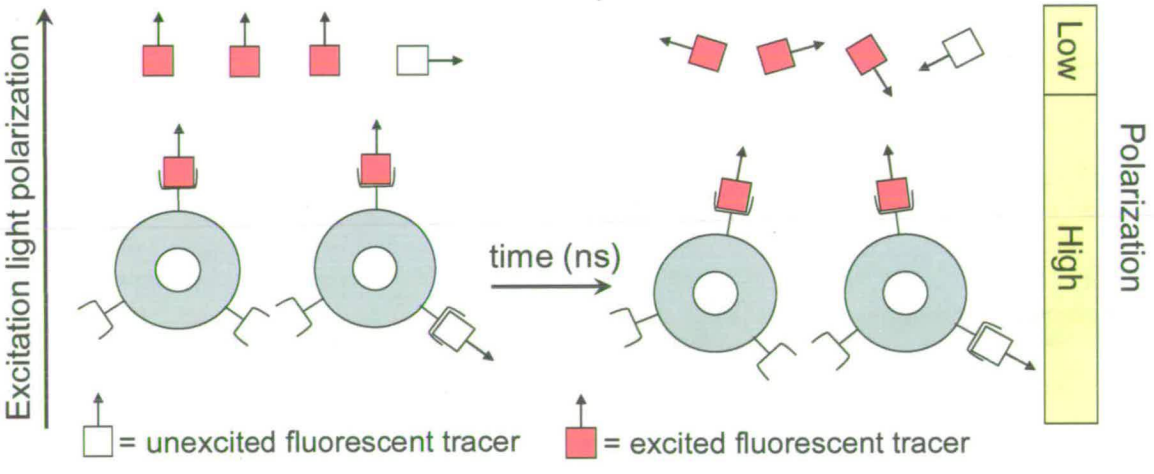


Figure 3-3 Physical basis of a fluorescence polarization assay.
(adapted from (Haugland, 2005))Only dye molecules/tracers with absorption transition vectors aligned parallel to the electric vector of linearly polarized light get excited. The labelled small molecules or peptides that remain free in solution continue to rapidly rotate and lose their photoselected orientation, which results in low FP. The labelled molecules bound to the large, slowly rotating protein molecule (here PCNA with three binding sites per trimer) change only marginally their orientation which results in high FP. The extent of signal relates directly to the extent of tracer binding to the protein.

Hence FP decreases with increasing τ, and increases with molecular weight and solvent viscosity. These assumptions, however, are based on a dye which is rigidly

attached to a spherical carrier. One of the most commonly used fluorophores is fluorescein which has an excited state life time (τ) of 4ns.

FP is also instrument dependent as implied in equations (1) and (2); hence the data quality is linked directly to the quality of the used instrument. Three points are important to note:

1) The angle at which polarized/depolarized light is detected differs between instruments.

2) Detectors measure intensities in different ways as photons can be sensed directly or with use of an enhancer. Fluorescence is measured in arbitrary units, so when applying equations (1)-(4) to the data collected with different detectors the value for r should be identical providing the sample and the angle at which polarized/depolarized light is detected remain the same.

3) The volume of the sample influences the error in the measurements: a larger volume gives smaller errors.

A number of publications report the use of FP to quantify the binding of various molecules, including peptides, to protein targets (Parker et al., 2000; Rozema and Poulter, 1999; Tota et al., 1994; Veis et al., 2000). More recently FP has also been employed for competitive assays (Figure 3-3) (Turek et al., 2001), where a known ligand is labelled and unlabeled potential ligands are added. If the unlabeled compound is able to compete with the known ligand, a decrease in FP is observed, if the concentration of the fluorophore is kept constant.

For the work described in this chapter, 5(6)-carboxyfluorescein was coupled to the PL/consensus motif peptide (Kontopidis et al., 2005; Zheleva et al., 2000) which binds *via* its PIP-box motif to PCNA with a K_d of 100 nM. The binding of this small,

labelled peptide (~1.5kDa) to the much larger PCNA molecule is observed by monitoring an increase of polarization/anisotropy. If a second molecule is able to compete with the labelled peptide for binding to PCNA a decrease of polarization is predicted.

3.2.3 LIDAEUS

LIDAEUS (Wu et al., 2003) is a molecular docking programme for screening large chemical libraries for small molecules matching a certain set of parameters based on a the physiochemical properties and structure of a target binding pocket.

Here LIDAEUS was used to construct a cubic grid of site-points representing the hydrophobic pocket into which the PIP-box residues of p21 bind. Figure 3-4 shows the p21-based template for screening runs and the interactions that need to be mimicked by a successful PCNA inhibitor.

Each site-point has associated calculated properties including van der Waals-interaction energy, hydrogen bonding capacity and buriedness. They can be separated into: hydrophobic (C), extremely hydrophobic (Csp), hydrogen bond acceptor (HBA) and hydrogen bond donor (HBD). The user can set a threshold, below which a point's potential energy must be, for it to be considered a site-point.

Initially atom types are assigned to the target molecule e.g. hydrogen donor or acceptor. The second step is docking, in which compound atoms are fitted onto site-points. This can lead to several different poses of the same compounds to the pocket. A ligand is regarded as a potential hit when a defined number of its atoms (e.g. 4 as

in this case) matches the geometry and property prerequisites of the calculated site points.

In each run a score is given to each potential ligand's pose. This is an approximation of the binding free energy of the ligand-protein complex calculated by summing the van der Waals and electrostatic interactions. An energy map, which combines information on hydrophobic and potential hydrogen bonding interactions, is allocated to each set of atoms verified in a ligand during the docking. The values for each potential atom interaction are added up after weighting factors have been applied. Optionally, as in this case, a Pose Interaction Profile (PIP) score can be added to these sums. The aspect represents the condensed description of the interactions of the posed ligand to the target protein. The score is then compared with the PIP of the model ligand (p21 in this case). The result will be a value between 0 and 1, where 1 is considered to be a perfect match of the model interaction.

Finally, LIDAEUS ranks the poses based on the above calculated values. All ligand-protein contacts are tested to ensure there are no severe steric clashes between the ligand and the protein. High ranking compounds with negative scores should have no steric clashes as any such clashes would translate into a highly positive contribution to the score. Asinex and Specs 3D libraries of drug-like small molecules were chosen on the grounds of available online pricing information, prompt delivery and cost-effectiveness.

Chemical similarity was assessed by Dr. Mike Greaney (University of Edinburgh), so that a diverse set of compounds could be ordered. For this study all work with LIDAEUS was carried out by Simon Harding.

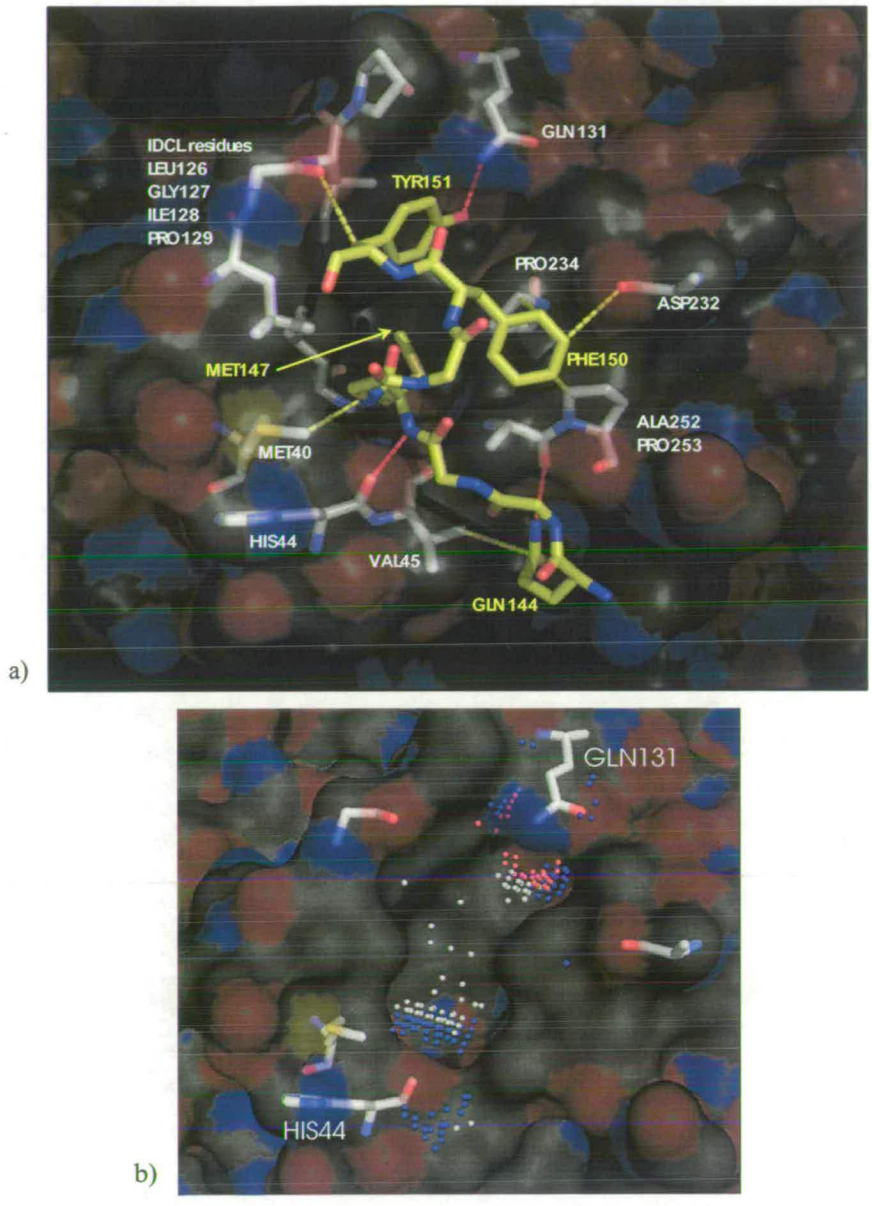


Figure 3-4 The binding of residues p21 Q144, M147, F150 and Y151 to the hydrophobic pocket of PCNA (Figure by Simon Harding).
a) Interactions between PCNA (white) and the four PIP-box residues of p21 (yellow) are displayed as hydrogen bonds (red dotted lines) and likely van der Waals contacts (yellow dotted lines). While M147 and Y151 residues of p21 are positioned directly into the pocket, F150 is lying outside, between two proline residues of PCNA, P234 and P253. The conserved PIP-box residue Q144 is not part of the hydrophobic pocket, but is vital for p21 binding to PCNA and was hence part of the LIDAEUS search template b) A typical set of site points calculated for the MFY template. The distribution of site-points depends on the template used and if Pose Interaction Profile (PIP)-scores were used to influence the site-point generation. PIP-scores will increase the likelihood of docking molecules to the pocket that are able to create specific/chosen interactions like hydrogen bonds.

3.3 MATERIALS AND METHODS

3.3.1 Analysis of PCNA Oligomerization

A 7 μM sample of purified human PCNA in 25 mM Tris pH 7.5, 2 mM DTT and 150 mM NaCl was analysed with an OmniSIZE 2.0 Dynamic Light Scattering detector (Viscotek) at 25 °C.

200 μl of a 50 μM sample of purified native (untagged) human PCNA was run on a Superdex 200 gel filtration (Amersham) column at 4 °C in 25 mM Tris pH 7.5, 2 mM DTT and 150 mM NaCl.

3.3.2 Peptides

N-terminally 5(6)-carboxyfluorescein labelled peptides p21 (141-160 = KRRQTSMTDFYSKRRLIFS-OH), PL/consensus motif 1 (SAVLQKKITDYFHPKK-OH) and p21(141-54 = KRRQTSMTDFYHSK) were synthesized by Mimotopes. 10mM stock solutions in 100% DMSO were prepared and stored at -20 °C.

An unlabelled shortened version of the p21 peptide (residues 141-152; KRRQTSMTDFYH) was synthesised by Peptide Protein Research Ltd. A 50 mM in 100% DMSO stock was prepared and stored at -20 °C.

All peptides were at least 90% pure as determined by mass spec.

3.3.3 Isothermal Titration Calorimetry (ITC).

ITC experiments were performed by Peter Brown using an AutoITC machine (Microcal Inc. Northampton, MA). A solution of purified human PCNA at 4.57 μM was titrated with 100μM of p21(141-152) at 25 °C in 10mM HEPES pH 7.4 and 100mM NaCl. The injection volume was 5μl and the final reaction volume was 2ml with a starting volume of 1.42ml. The enthalpy of binding (ΔH, $kcal.mol^{-1}$) was determined by integration of the injection peaks and the corrections for heat released due to dilution were determined in a control experiment without peptide. The resulting corrected binding isotherm was fitted by non-linear least-squares analysis to a variety of binding models using MicroCal Origin software, supplied with the instrument.

3.3.4 Compounds/Potential PCNA inhibitors candidates

The 3D libraries Specs and ASINEX which comprise about 100.000 compounds each were screened by Simon Harding for potential PCNA inhibitors using the programme LIDAEUS (Wu et al., 2003) under consideration of Lipinski rules as selection criteria. The top hits were chosen for better solubility according to logP values.

AG-205/33676017, AG-690/11632020, AG-690/08637022, AG-690/12242265, AG-205/07677038, AK-968/15252997, AP-828/41100751, AH-487/15582102, AG-205/11867190, AE-562/12222430 and AR-013/42256276 were supplied by Specs. ASN-04196551, ASN-03904373, ASN-04369399, ASN-04886713, ASN-04254797, ASN-04190395 and ASN-01516558 were ordered from Asinex.

Stock solutions of 100 mM or less were prepared in 100% DMSO and stored in the dark at room temperature.

3.3.5 Fluorescence Polarization Assay and Data analysis

Assays were carried out either using an Edinburgh Instruments FS900 T-geometry fluorescence spectrometer with a temperature-controlled sample cuvette with Glan-Thomson quartz polarisers, or a SpectraMax M5e plate reader (Molecular Devices). The experimental setup is significantly different for each instrument.

For the Edinburgh Instruments FS900 T-geometry fluorescence spectrometer a single quartz cuvette, with a 1 cm path length and 600 µl maximum volume was used. The assay was carried out in PBS (3.8 mM NaH_2PO_4, 16.2 mM Na_2HPO_4, 150 mM NaCl, pH 7.4), 5% (v/v) DMSO, 5 mM DTT using 50 nM of a labeled PL peptide only or additionally 10 µM of the short p21 peptide as model competitor. Recombinant human PCNA was titrated at volumes between 0.1 µl and 1 µl into the sample from a 17.4 µM stock solution and intensities measures were recorded as an average of 5 reads. The total sample volume was 300 µl plus any protein added. For simplicity the change of volume was considered to be insignificant (<10%).

To test the small molecules for binding, small amounts of the stocks dissolved in DMSO were diluted into the reaction buffer instead of the model competitor until maximum concentration was reached and the solutions were still clear (see Table 3-1 for solubility).

Experiments performed with the plate reader were carried out in black (both wall and bottom) 96-well plates (Greiner) with readings taken from above. The reaction buffer remained the same, but the total sample volume was reduced to 250 µl. The

concentration of labeled peptide and model competitor /small molecule remained as in the previous setup, but each well contained a fixed concentration of PCNA. Each concentration point was carried out in triplicate and intensities were recorded as an average of 100 reads. The fluorescent dye was excited at 485 nm and read at 530 nm. The anisotropy was calculated using equation (1) and the apparent affinity (K_d) for the labelled peptide was calculated using the following equation (5):

$$r = r_f + (r_b - r_f) \{(K_d + [A] + [B]) - \sqrt{((K_d + [A] + [B])^2 - (4 [A] [B])) / (2 [A])}\}. \quad (5)$$

Where [A] is the fixed concentration of labelled ligand and [B] is the concentration of titrated PCNA.

The following equation was employed to calculate the K_d of p21(141-152) (i.e. K_i) from the competition experiment:

$$K_i = K_{app} / (1+ [Lig]_L / K_d) \quad (6)$$

Here K_i is the affinity of the inhibitor, K_{app} is the apparent K_d of the labeled peptide in the presence of inhibitor, $[Lig]_L$ is the concentration of labeled peptide and K_d the affinity of labeled peptide in the absence of inhibitor as calculated using equation (5).

3.3.6 Thermal shift assay and Data analysis

The assay was carried out using thin-wall PCR 96-well plates (Bio-Rad) covered with Optical-Quality Sealing Tape (Bio-Rad) using the iCycler iQ Real Time PCR Detection System (Bio-Rad).

Each 50 µl sample contained 0.25 µM recombinant human PCNA, a 1:1000 dilution of Sypro Orange 5000X (Molecular Probes) dye and either of potential inhibitors (at maximum concentrations) or increasing concentrations of the short p21 peptide, in thermal denaturation buffer 1 (25 mM Tris pH 7.5, 100 mM NaCl, 2 mM DTT) or thermal denaturation buffer 2 (PBS, 5 mM DTT, 5% (v/v) DMSO). Each sample was carried out in triplicate and averages were calculated. The wavelengths for excitation and emission were 490 nm and 575 nm, respectively. Fluorescent readings were taken between 20 °C and 88 °C in 1 °C increments after each temperature had been maintained for 90 seconds.

The fluorescent intensities were plotted against temperature in the programme Kaleida-Graph and the T_m for each sample was calculated using the following equations:

$$F(T) = F_{post} + [(F_{pre} - F_{post}) / (1 + \exp^{(Tm - T)/C})] \qquad (7)$$

Where F(T) is the fluorescence intensity at temperature T, F_{pre} and F_{post} are the fluorescence intensities for completely folded and completely unfolded protein, respectively and C is the slope factor. Here the slope factor acts as a scaling factor for the total increase of fluorescence, which may vary for different samples.

For the model ligand p21(141-152) the K_d at different concentrations was calculated. For this purpose the simplified equation above had to be supplemented with terms for ΔH_u (the enthalpy of protein unfolding) and ΔC_{pu} (the heat capacity change of protein unfolding) (equation 8):

$$F(T) = F_{post} + (F_{pre} - F_{post}) / [1 + \exp \{(-\Delta H_u / R) (1/T - 1/T_m) + \Delta C_{pu} / R (\ln [T / T_m] + T_m / T - 1)\}]$$

(8)

When no stabilizing ligand is present $T_m = T_o$, $\Delta H_u = \Delta H_u^{To}$ and $\Delta C_{pu} = \Delta C_{pu}^{To}$.

To calculate the ligand association constant ($K_L = K_a = 1/K_d$) for the ligand at T_m following equation was used:

$$K_{L(Tm)} = \exp\{(-\Delta H_u^{To} / R) (1/T_m - 1/T_o) + \Delta C_{pu}^{To} / R (\ln [T_m / T_o] + T_o / T_m - 1)\} / [L_{Tm}]$$

(9)

where $[L_{Tm}]$ is the free ligand concentration at T_m, which is approximately the same as the total ligand concentration when the ligand is added in great excess to protein.

To calculate affinities at room temperature for comparison of data obtained previously from ITC experiments at 25 °C an extrapolation has to be carried out using the following equation (10):

$$K_{L(T)} = K_{L(Tm)} \exp \{(-\Delta H_{L(T)} / R) (1/T - 1/T_m) + \Delta C_{PL} / R (\ln [T / T_m] + 1 - T / T_m)\}$$

(10)

where $\Delta H_{L(T)}$ is the van't Hoff enthalpy of binding at temperature T. ΔC_{PL} is the heat capacity change on binding. An average value for ΔC_{PL} has been determined experimentally derived and will be considered to be -0.5 $kcal.mol^{-1}.K^{-1}$ (Wear and

Walkinshaw, 2006). $\Delta H_{L(T)}$ was equally calculated from previous ITC experiments carried out by Peter Brown.

The second exponential term in red is usually quite small compared to the first exponential part when dealing with small drug-like compounds as ligands. However, it has been shown that the heat capacity change over temperature for peptide binding can have a significant influence on affinity.

3.4 RESULTS AND DISCUSSION

3.4.1 Oligomerization state of human PCNA

It has been claimed that PCNA may occur as a double homotrimer in solution (Naryzhny et al., 2005). A single trimeric ring of PCNA has a molecular weight of about 90 kDa. Dynamic Light Scattering (DLS) showed 99% of particles were contributed by a species with a diameter of about 4.4 nm (data not shown). According to software analysis this would represent a globular, compact protein of about 200 kDa. However, the sliding clamp has a 35 Å cavity in the middle which is not picked up by DLS.

Furthermore, any rotation of hPCNA in solution would make it appear as a much larger particle, due to the overall large difference between width and thickness (flatness). It could hence be argued that the species with a diameter of 4.4 nm is probably representative of a single trimeric PCNA ring.

Analytical gel filtration analysis using a Superdex 200 column produced a single peak with a retention volume of 12.69 ml.

This is indicative of a globular protein of about 120.6 kDa (Figure 3-5). Altogether these results suggest that untagged hPCNA used in this study is probably a single trimer in solution in the stated standard buffer conditions.

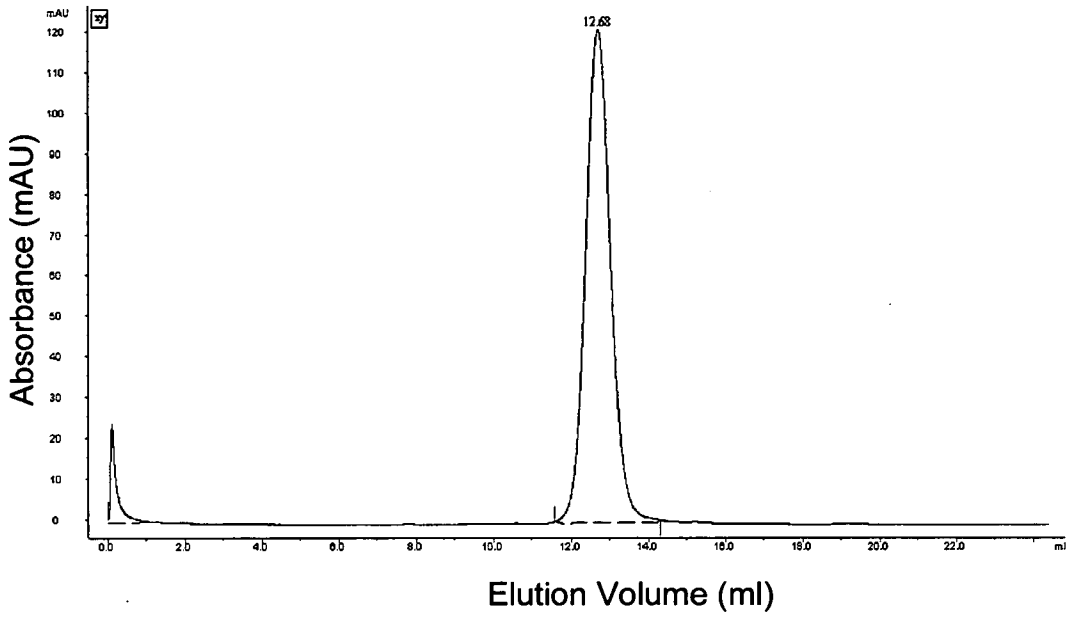


Figure 3-5 Size estimation of human PCNA by gel filtration.
The elution volume of 12.68ml on this column corresponds to an estimated size of 120 kDa by comparison with standards of known molecular weight. This agrees with the fact that PCNA appears slightly bigger than its actual molecular weight due to its ring shape.

3.4.2 Solubility of compounds from LIDAEUS runs and logP prediction

In order to find small drug-like molecules that could mimic the interactions that are made by the p21 peptide in the hydrophobic pocket two different LIDAEUS runs were set up by Simon Harding. The main interaction between p21 and PCNA is based on the four conserved residues that form the PIP-box motif and are that found in many PCNA binding proteins.

The first run investigated the interactions made by three hydrophobic residues (Met147, Phe150 and Tyr151) of p21 with PCNA. Compounds ASN-04196551, ASN-03904373, ASN-04369399, ASN-04886713, ASN-04254797, ASN-04190395, ASN-01516558, AP-828/41100751 AH-487/15582102, AG-205/11867190 and AG-

690/12242265 (Table 3-1) were selected from the best fitting hits for lower logP, chemical stability, lack of obvious toxic groups and difference in chemistry from each other. These criteria were also applied to the selection of compounds from other runs.

The second run was carried out using site points calculated for p21 homing in on Gln144 in addition to the three hydrophobic residues. From the result hits list following compounds were analysed: AG-205/33676017, AG-690/11632020, AG-690/08637022, AG-205/07677038, AK-968/15252997, AE-562/12222430 and AR-013/42256276 (Table 3-1).

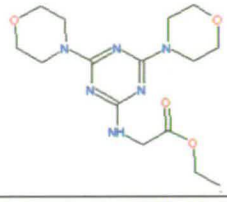
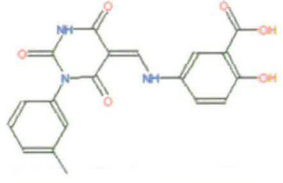
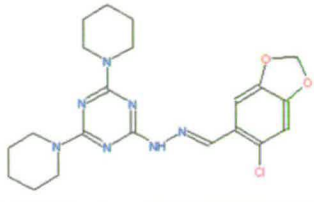
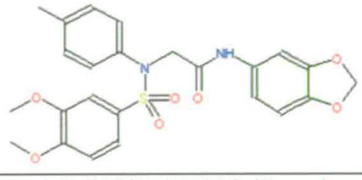
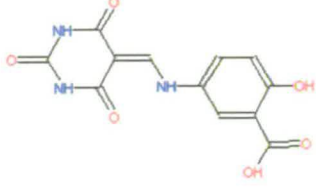
As the site-point selection for the hydrophobic pocket alone gave rise to rather hydrophobic molecules, the extension of the search pocket to include the glutamine was expected to improve the water solubility of hit compounds.

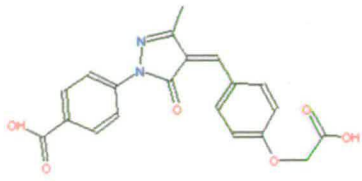
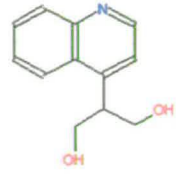
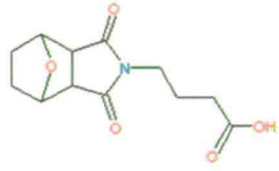
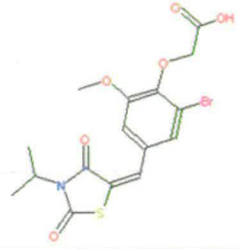
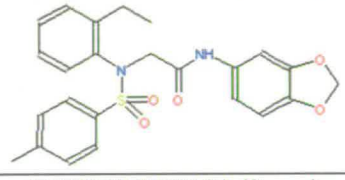
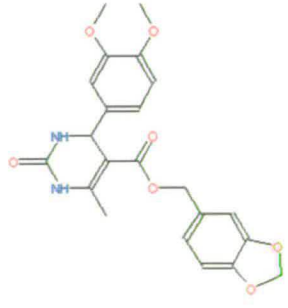
According to the Lipinski rules only compounds with a logP value below 5 were listed as a hit. In theory compounds with higher solubility in water have lower logP values. The logP values of the selected compounds were regarded to be acceptable since well-known drugs like Glivec (logP = 4.5) and Prilosec (logP = 2.6) have fairly high values, but are sufficiently soluble in the body.

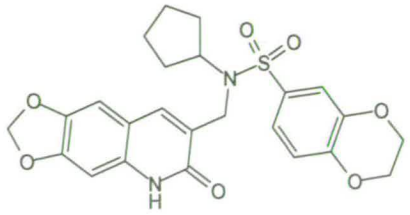
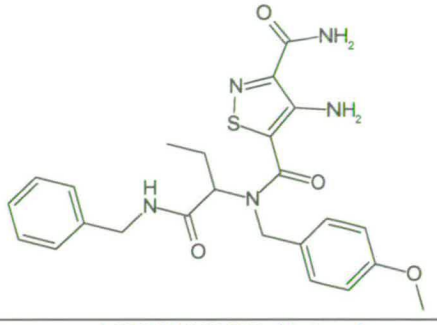
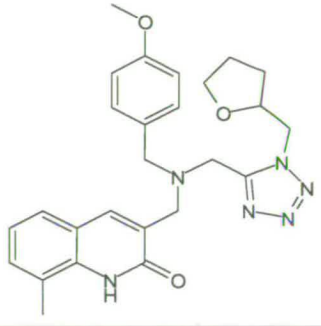
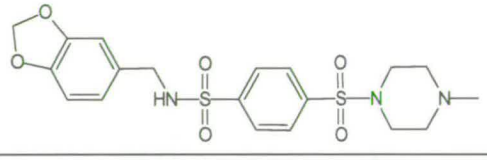
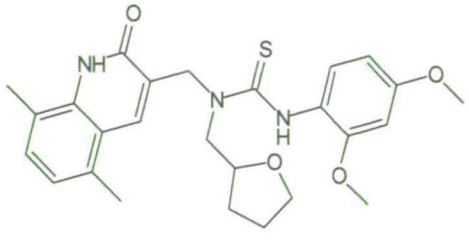
Despite the attempt to choose water soluble compounds to maximise the screening success with assays that were going to be carried out in aqueous solutions, some of the compounds turned out to be less soluble than others (Table 3-1).

There is, however, a general correlation between the solubility of the compound in an aqueous buffer plus 5% (v/v) DMSO and the logP. Nevertheless, there are a few outliers to this rule which are less soluble than predicted, especially in comparison with compounds that were allocated a similar logP.

The use of a buffer containing 5% (v/v) DMSO increased the active concentration of free ligand in solution (data not shown) and proved to be compatible with PCNA stability (see sections 3.4.3 and 3.4.4).

Compound ID (Company)	Solubility in PBS, 5% (v/v) DMSO, 5mM DTT (mM)	Stock in 100% DMSO (mM)	calculated LogP
AG-205/33676017 (Specs) 	5	100	-0.46
AG-690/11632020 (Specs) 	4	100	1.53
AG-690/08637022 (Specs) 	0.2	100	3.96
AG-690/12242265 (Specs) 	0.2	100	3.79
AG-205/07677038 (Specs) 	2.5	50	0.4

AK-968/15252997 (Specs) 	1	100	2.49
AE-562/12222430 (Specs) 	5	100	0.91
AR-013/42256276 (Specs) 	5	100	-0.9
AP-828/41100751 (Specs) 	0.2	100	3.27
AH-487/15582102 (Specs) 	0.1	100	4.38
AG-205/11867190 (Specs) 	0.3	100	3.17

ASN-04196551 (Asinex) 	0.066	33.3	2.852
ASN-03904373 (Asinex) 	0.6	100	1.645
ASN-04369399 (Asinex) 	0.6	100	0.941
ASN-04886713 (Asinex) 	0.4	100	2.3
ASN-04254797 (Asinex) 	0.133	33.3	1.806

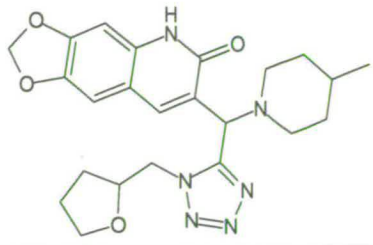
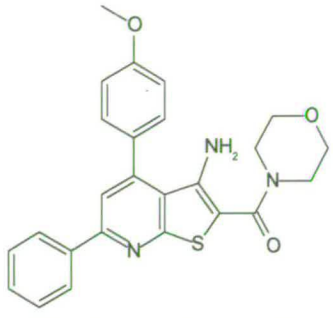
ASN-04190395 (Asinex) 	0.1	16.5	0.759
ASN-01516558 (Asinex) 	0.1	33.3	4.39

Table 3-1 List of selected potential PCNA inhibitors and their solubility.
Stocks in DMSO were prepared at a maximum concentration of 100mM. In order not to exceed 5% DMSO in the assay buffers even very soluble compounds in 100mM stocks were not added beyond 5 mM. Higher logP values predict a lower solubility in water, but without a linear correlation.

3.4.3 Thermal shift assay of human PCNA under the effect of peptide binding and changing buffer conditions

To measure the affinity of small molecule compounds to PCNA a previously described thermal shift assay for establishing conditions for protein stabilization (Ericsson et al., 2006) or for testing the structural basis of inhibitor binding (Bullock et al., 2005) was adapted for human PCNA. 0.25 μM of PCNA in combination with 1 in 1000 dilution of the commercially available Sypro Orange dye (5000X) (Molecular Probes) (originally designed as a protein gel stain) gave a sufficiently strong fluorescent signal upon protein unfolding. Similar to 1-anilinonaphthalene-8-sulphonate (Eling and DiAugustine, 1971) this probe is quenched in aqueous solution and only emits fluorescence when bound to a hydrophobic environment. This can

occur when a protein unfolds and exposes the interior hydrophobic side chains to the solvent. The increase of fluorescent intensities over a temperature range is hence directly related to the extent of protein unfolding. Thermally induced protein unfolding follows a typical two-state model with a sharp transition between native and denatured states. After reaching the completely unfolded state, the fluorescent intensity decreases again (Ericsson et al., 2006). This effect cannot be accounted for, but is assumed to be connected with protein aggregation. Therefore the data after the maximum are omitted from the calculations.

The result of the thermal shift assay of hPCNA in absence and presence of p21(141-152) are shown in Table 3-2 and Figure 3-6. The model ligand p21 (141-152) has a K_d of 307 nM at 25°C as determined by ITC. The data show that p21 is able to stabilize PCNA under heat-denaturing conditions leading to a shift in T_m towards a higher temperature. This shift is detectable under both thermal denaturation buffers.

	T_m in °C	
Concentration of peptide (µM)	**Thermal denaturation buffer 1**	**Thermal denaturation buffer 2 (incl. 5% DMSO)**
0	53.14 ± 0.24	52.71 ± 0.13
12.5	53.44 ± 0.19	52.27 ± 0.13
25	54.88 ± 0.17	53.1 ± 0.12
50	55.21 ± 0.16	53.2 ± 0.13
100	57.44 ± 0.16	54.02 ± 0.1

Table 3-2 T_m for hPCNA: dependence on p21(141-152) peptide concentration.

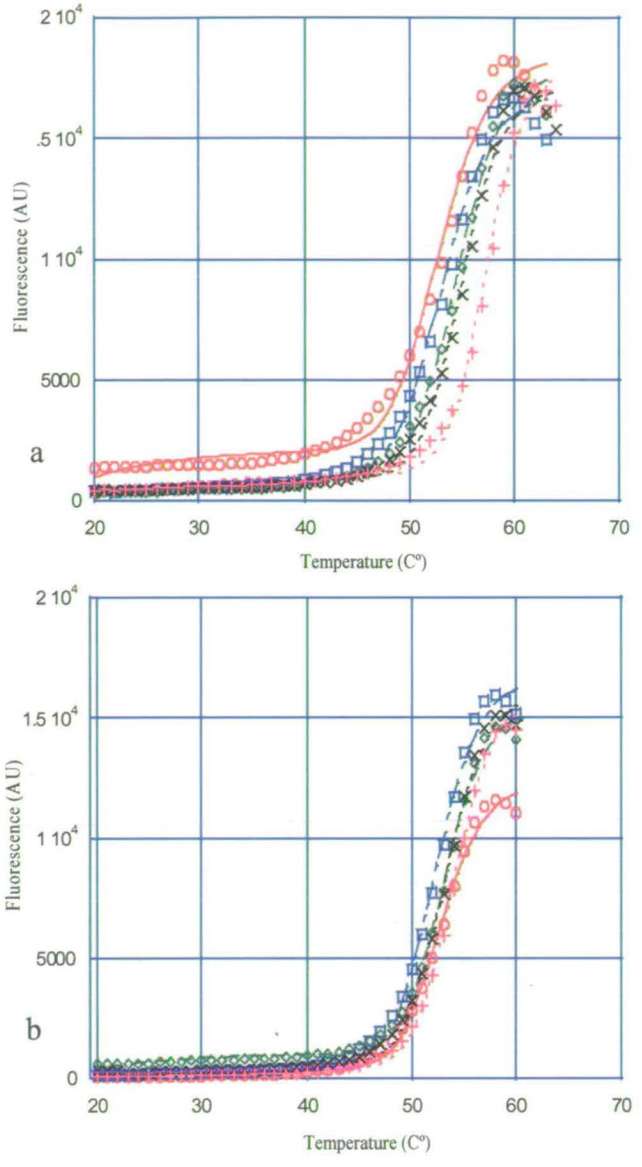


Figure 3-6 Thermal shift assay of human PCNA in the absence and presence of p21 (141-152). hPCNA at a concentration of 0.25 μM was subjected to increasing temperatures. The red trace denotes the sample without peptide, blue = 12.5 μM peptide, green = 25 μM peptide, black = 50 μM peptide and pink = 100 μM peptide. All experiments in a) were carried out in thermal denaturation buffer 1 whereas b) used thermal denaturation buffer 2. Data were fitted in Kaleida-Graph, and error-bars have been omitted for clarity.

However, the presence of 5%(v/v) DMSO has two effects: first it leads to lowered protein stability, which is marked by a small T_m shift from 53.1 °C (buffer 1) to 52.7 °C (buffer 2) in the absence of ligand; and secondly, an smaller T_m shift over the

ligand concentration range is observed when DMSO is present. The ΔT_m for 0 μM ligand and 100 μM peptide is 4.3 °C in buffer 1. For the second buffer condition this is 1.3°C. Although a significant shift outside error boundaries occurs for both buffers at 25 μM peptide, it is more obvious without DMSO.

Using equations (7)-(10) the binding data were extrapolated to calculate the apparent K_d of the peptide at 25 °C. In thermal denaturation buffer 1 peptide concentrations of 12.5 μM, 25 μM, 50 μM and 100 μM yielded affinities of 0.212 μM, 0.337 μM, 0.667 μM and 0.946 μM respectively, with the last one being the most accurate as the peptide concentration approaches saturation level (Table 3-3).

In thermal denaturation buffer 2 the same concentrations gave affinities of 0.265 μM, 0.455 μM, 0.897 μM and 1.54 μM respectively. In comparison with the ITC derived K_d of 307nM, the affinities calculated for 100 mM peptide are 3 - 5fold weaker. Nevertheless, the extrapolated affinities lie within an order of magnitude of the directly ITC derived value. A reason for observed deviations of the actual values could be differences in buffers used for the different techniques for determining K_ds (Table 3-3).

	Apparent K_d of the p21(141-152) at 25 °C in μM	
Concentration of peptide (μM)	**Thermal denaturation buffer 1**	**Thermal denaturation buffer 2**
12.5	0.212	0.265
25	0.337	0.455
50	0.667	0.897
100	0.946	1.54

Table 3-3 Apparent K_d of p21(141-152): dependence on peptide concentration.

Having established the ability of this assay to detect ligand binding, it was carried out under the same conditions with maximum compound concentrations that could be

reached in thermal denaturation buffer 2. The addition of 5% (v/v) DMSO to the buffer improved the solubility of the hydrophobic small molecules.

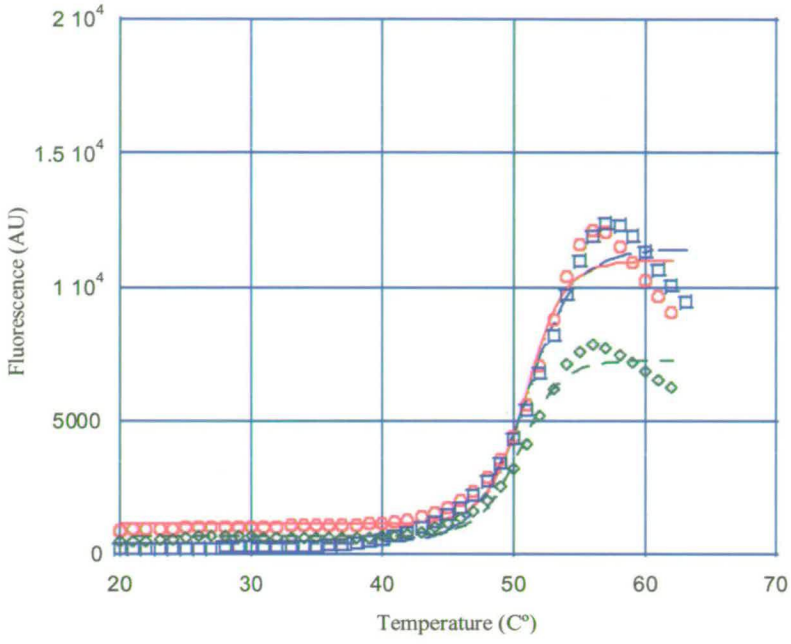


Figure 3-7 Thermal shift assay in the presence of two compounds.
Neither Asinex 04196551 (blue) nor the Specs compound AG-205/07677038 (green) showed a significant T_m shift in comparison with PCNA on its own (red). These curves are representative for the remaining compounds that were tested using this assay.

Unfortunately, none of the tested compounds showed a stabilizing effect of PCNA under stated conditions at the maximum concentration that could be reached. Figure 3-7 shows two examples curves of the set of tested compounds. Neither Asinex 04196551 nor the Specs compound AG-205/07677038 showed a significant T_m shift with 50.9 °C and 50.5 °C respectively in comparison with 50.9 C° of PCNA on its own. These are typical results for the remaining compounds. If any change was observed at all, it was rather a destabilizing effect leading to a lowered T_m (Table 3-4). This is explained by a tendency of the compound to bind to the unfolded/unfolding protein. It cannot be concluded that none of the compounds

actually bind to the surface of PCNA, however, none of them seem to have a stabilising effect.

Compound	T_m in C°	T_m control in C°
AG-205/33676017 (Specs)	50.58 ± 0.15	50.96 ± 0.22
AG-690/11632020 (Specs)	47.83 ± 0.22	48.79 ± 0.31
AG-690/08637022 (Specs)	46.58 ± 0.19	46.49 ± 0.1
AG-690/12242265 (Specs)	44.33 ± 0.23	45.85 ± 0.12
AG-205/07677038 (Specs)	50.49 ± 0.18	50.96 ± 0.22
AK-968/15252997 (Specs)	44.85 ± 0.61	48.79 ± 0.31
AE-562/12222430 (Specs)	50.068 ± 0.25	50.96 ± 0.22
AR-013/42256276 (Specs)	44.07 ±0.11	44.56 ± 0.07
AP-828/41100751 (Specs)	turbid	n/a
AH-487/15582102 (Specs)	turbid	n/a
AG-205/11867190 (Specs)	turbid	n/a
ASN-04196551 (Asinex)	50.93 ± 0.24	50.96 ± 0.22
ASN-03904373 (Asinex)	49.92 ± 0.13	50.35 ± 0.11
ASN-04369399 (Asinex)	50.43 ± 0.21	50.35 ± 0.11
ASN-04886713 (Asinex)	50.12 ± 0.2	50.35 ± 0.11
ASN-04254797 (Asinex)	49.57 ± 0.17	50.35 ± 0.11
ASN-04190395 (Asinex)	50.31 ± 0.19	50.35 ± 0.11
ASN-01516558 (Asinex)	50.22 ± 0.09	50.35 ± 0.11

Table 3-4 T_ms of assayed compounds.

3.4.4 Fluorescence Polarization (FP) Competitive Assay

The theory of this assay is based on the fact that small molecules (e.g. peptides) rotate fast in solution, but upon binding to large, slowly tumbling molecules like proteins (e.g. PCNA) their rotational speed decreases. The change in rotational speed of the small molecule can be traced as a change in fluorescent anisotropy, if a fluorescent probe is used for labelling. In this case the fluorescent signal comes from 5-(6)-carboxyfluorescein labelled peptides: p21(141-160 and PL. As determined by ITC experiments, unlabelled p21 was shown to have an affinity for PCNA of 114 nM while the PL exhibits a K_d of 100 nM (Zheleva et al., 2000).

PCNA is titrated into a fixed concentration of fluorescently labelled peptide to determine the maximum fluorescent anisotropy. For the competition assay a fixed concentration of the model ligand, p21 (141-152), or of a potential PCNA inhibitor is added prior to titration of PCNA, and anisotropies are compared.

Two different instruments were used to measure FP data for PCNA and ligands: the FS900 T-geometry fluorescence spectrometer and the M5e plate reader. While the first one is far more sensitive than the plate reader, each single point had to be titrated by hand for the FS900 T-geometry fluorescence spectrometer rendering this method very time-inefficient. The difference of sensitivity lies in the angle at which emitted light is detected (90° for the FS900 T-geometry fluorescence spectrometer and 15° for the plate reader). Figure 3-8a shows a comparison of measurements with and without p21 (141-152) for both instruments. The change of anisotropy is clearly detectable for both, but the detected absolute change in anisotropy (Δr) is considerably smaller with the plate reader.

Using equations (5) and (6) all data were analyzed in Kaleida-Graph. In order to make it easier for fairly weak binding compounds to compete against the labelled peptide a third peptide based on p21 was employed [p21(141-154]. It was shown to be the weakest binding p21-derived sequence still being able to compete against p21(141-160) (Zheleva et al., 2000). The weaker binding p21(141-154 = KRRQTSMTDFYHSK), however, gave noisy data and yielded only a small Δr. This made detection of competitive binding difficult (Figure 3-8b) and was hence rejected for this study.

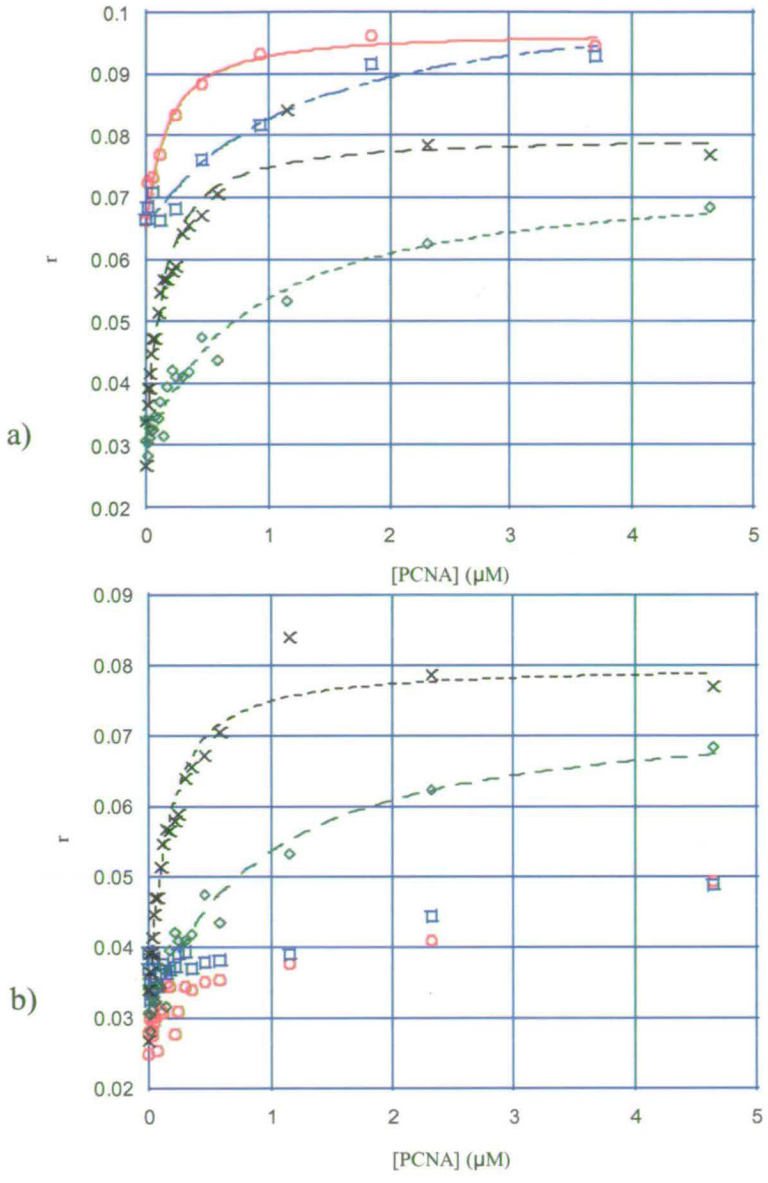


Figure 3-8 Comparison of the FP competitive assay data quality in terms of labelled probes and instruments.

a) Curves for binding of labelled PL peptide in the absence (red) and presence (blue) of 10μM p21(141-152). These data were recorded using the M5e plate reader. The affinity for PL was calculated to be 140nM ± 35nM, while it was 1190nM ± 47nM in the presence of competitor. In comparison the more sensitive FS900 T-geometry fluorescence spectrometer yielded a K_d of 106nM ± 15nM (black) for the PL binding and 847nM ± 135nM (green) in the presence of 10μM p21(141-152). Despite the difference in sensitivity both instruments measure data that lie within the same order of magnitude.

b) Binding of labelled PL (black) and weaker p21(141-154)(red) to PCNA in the absence of a competitor. The green and blue lines show the apparent affinity of PL and p21 (141-154) when 10μM of p21(141-152) are present. While the addition of unlabelled peptide has a competitive effect on PL binding, the binding of p21 (141-154) is not influenced.

The measured K_d for the model inhibitor p21(141-152) was 877 nM using the M5e plate reader and 576 nM for the FS900 T-geometry fluorescence spectrometer (Figure 3-8a). These results are close to the ITC-derived K_d for p21(141-152) and provide confidence in the ability of this assay to pick up compounds that bind specifically to PCNA *via* the hydrophobic pocket.

For testing the potential inhibitors, the M5e plate reader was chosen over the more sensitive instrument as the specific experimental setup for the latter was highly time-consuming and hence not suitable for a screening procedure. For economic reasons only four PCNA concentrations were chosen: 0 μM, 0.029 μM, 0.116 μM, 0.463 μM. The addition of model inhibitor showed a clear competitive effect despite the limited number of data points. This provides confidence in the ability of the assay to pick up competing compounds even in a restricted form (Figure 3-9).

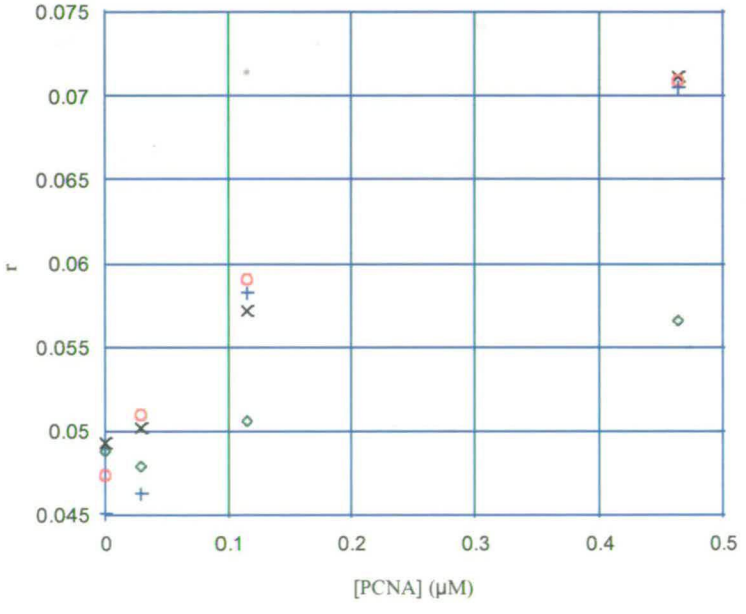


Figure 3-9 FP competitive assay of ASN 04254797 and ASN 04190395 measured using the M5e plate reader.
The black and green points show PL binding in the absence and presence of p21(141-152), respectively. Compound ASN 04254797 (blue) and ASN 04190395 (red) were added at 100μM and 133μM instead of p21 peptide and did not have a detectable affinity for PCNA.

Having established the suitability of this assay all hits were tested at their maximum concentrations in reaction buffer (PBS, 5% (v/v) DMSO, 5 mM DTT) (Table 3-1). Figure 3-9 shows the results of the FP assay for two compounds in comparison with p21(141-152). Unfortunately none of the selected compounds showed competition under the stated conditions and at maximum concentrations.

3.5 SUMMARY AND FUTURE WORK

The overall outcome of this study was that none of the potential inhibitors showed detectable binding in either of the two employed assays. There are a number of possible reasons for this result.

Despite the fact that the thermal shift assay and the FP competitive assay are able to detect peptide binding there are caveats. The binding affinity of a given ligand can decrease by 2-5 times per any 10 °C rise in temperature. A weakly binding compound at room temperature may have a significantly lower affinity at 40 °C, the onset temperature for the unfolding of PCNA. Due to a decreased number of possible contacts that a small molecule can make with the target protein this thermodynamic effect is expected to affect small molecules far more than peptide binding.

Being aware of this fact we employed the FP competitive assay as well, which is carried out at 25 °C. The drawback with this method is the use of a tight binding peptide which needs to be replaced by a molecule with an affinity in a similar order. With p21(141-152) a model competitor with a 3-fold lower affinity to PCNA than the PL peptide was used. However, a 200fold excess of this ligand had to be present in order to see a distinct competitive effect. As with the thermal shift assay, a small molecule will be at a disadvantage compared to a peptide ligand, which is able to make more contacts, and much higher concentrations in solution would be necessary to see an equivalent effect of stabilization/competition.

Furthermore, the inherent insolubility of ligands in aqueous solution was posing a problem as only low concentrations were reached in some cases.

Another explanation is that the conserved hydrophobic binding pocket on PCNA might not be a suitable drug target. Dr David Fry (Roche) carried out an evaluation of the pocket for "druggability" (private communication). A programme called cavSearch (Stahl et al., 1999) analyzes pockets on the surface of proteins by surface area and depth (Figure 3-10). According to the analysis the PCNA hydrophobic pocket lies just below the thresh-hold for a suitable drug target. Although this is the evaluation of a single programme, this might suggest that PCNA is not an easy target to be tackled by structure-based drug design.

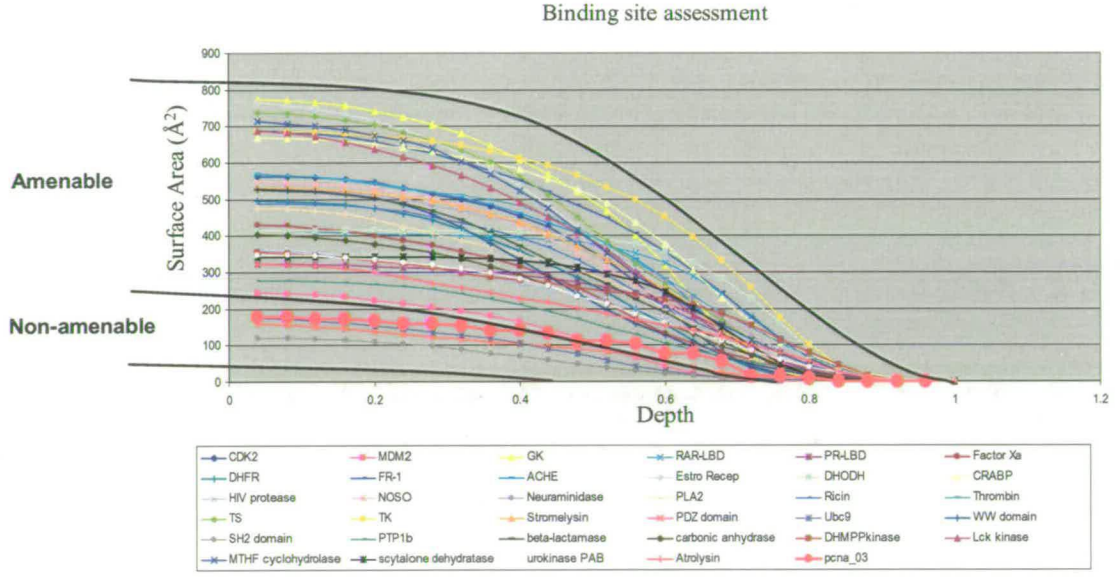


Figure 3-10 Selection of amenable and non amenable drug targets.
Pockets were evaluated according to surface area and depth. Fairly shallow pockets are classed as not druggable. The hydrophobic pocket that was targeted in this study is called pcna_03 and lies just below the threshold.

Even so, the fact that, according to cavSearch analysis, MDM2 lies only just above the threshold for an amenable site is suggesting that the PCNA pocket might still be a druggable target. This is backed up further by the fact that the MDM2/p53

interaction is in some way similar to PCNA/p21, with three hydrophobic residues binding in a hydrophobic patch.

Other researchers have adopted a different approach: they are searching for a ligand able to bridge from the hydrophobic pocket to the glutamine binding-site. Instead of searching for a whole molecule, their aim is to screen mixtures for small fragments that bind to either of those sites. When two fragments have bound to both pockets a covalent bond is formed between the two and the resulting larger compound can be identified (personal communication).

Another possibility would be to engineer a hexa His-tagged PCNA molecule to produce a stable surface for surface plasmon resonance (SPR) experiments in which potential inhibitors would be passed over the protein surface. A change in the refractive index would give information of binding affinities. However, this technique would be fairly expensive. The advantage of this method would be the possibility to carry it out at room temperature. Furthermore, there would be no requirement for the use of a tight binding peptide that any potential ligand would have to compete against.

Another approach would be to increase the molecular weight cut-off for the LIDAEUS screening to enlarge the number of possible interactions between compound and protein. The draw-back would be that the compounds would turn more into peptide-like drugs, which can be difficult to administer in an effective way. Also a lowered specificity might be the result.

CHAPTER 4. Interaction of PCNA with Gadd45

4.1 INTRODUCTION

The development of treatment resistance in malignant cancer cells originates from genetic lesions that alter the function of genes that play a role in normal cell homeostasis, especially in determining cell cycle progression and apoptosis. Hence it is of importance to understand the molecular genetic pathways that mediate growth control. The Gadd45 gene family - first described by Fornace et al (Fornace et al., 1988; Fornace et al., 1989) - contains three highly related growth-arrest and DNA damage inducible proteins; Gadd45, MyD118 and CR6 (Zhan et al., 1994; Zhang et al., 1999). Only Gadd45, a 165 amino acid, 18kDa protein, has been shown to be a direct down-stream target of p53. Upon DNA damage p53 is activated and induces expression of Gadd45 by binding to a 20 bp specific recognition site in the third intron of its target (Kastan et al., 1992). Gadd45 in turn causes arrest in G2/M phase and inhibits Cdc2. Gadd45 knockout mice show a similar phenotype to p53-deficient knockout mice, namely genomic instability as well as increased radiation carcinogenesis (Wang et al., 1999b).

Although the role of Gadd45 proteins in growth arrest and DNA repair is not well elucidated, they have been shown to interact directly with various cell cycle related proteins, such as Cdc2/cyclin B (Zhan et al., 1999), p21 (Kearsey et al., 1995) and PCNA (Smith et al., 1994). The techniques employed involved yeast-two-hybrid, immunoprecipitation, ELISA and activity assays.

The Gadd45 family shares similar, but not identical functions along different apoptotic and growth suppression pathways. Gadd45, but not MyD118/CR6 is a

target for p53 function. MyD118, but not Gadd45 is activated upon transforming growth factor β1-induced apoptosis. All three proteins appear to be induced with different expression kinetics during terminal hematopoietic differentiation, which is associated with growth arrest and apoptosis (Guillouf et al., 1995; Kastan et al., 1992; Selvakumaran et al., 1994a), and each gene is probably optimally induced by a certain subset of environmental stresses (Takekawa and Saito, 1998; Wang et al., 1999a). In addition, all three proteins also seem to be able to interact with each other in a synergetic fashion (e.g. by suppression of colony formation in several human tumour cells lines) (Vairapandi et al., 1996; Zhang et al., 1999).

The interaction between PCNA and Gadd45 was further characterized by Hall et al. (Hall et al., 1995a) using yeast two hybrid and co-immunoprecipitation experiments. In addition they analyzed the binding of recombinant human Gadd45 to a peptide array representing the whole of human PCNA and localised the domain responsible for the interaction with PCNA to the N-terminus of the protein.

Chen et al. (Chen et al., 1995) looked at the interactions between Gadd45 and PCNA and found evidence for a competitive interaction between p21 and Gadd45 for the binding of PCNA. Far-Western blots also showed that Gadd45 disrupts the ability of p21 to bind to PCNA, and conversely, p21 blocks the ability of Gadd45 to bind to the sliding clamp.

Other yeast two-hybrid experiments with Gadd45, MyD118 and CR6 (Azam et al., 2001; Vairapandi et al., 2000), in contrast to Hall et al (Hall et al., 1995a) located the PCNA-binding domain to the C-termini of all three proteins (Gadd45: residues 95-165, MyD118: residues 114-156, CR6: residues 76-159).

When recombinant Gadd45 and MyD118 were used in ELISA assays to determine the affinity of the full-length proteins to PCNA, low nanomolar K_ds were calculated for both proteins.

The interaction of MyD118 and Gadd45 with PCNA *via* their C-termini was further confirmed in mammalian cell lines by *in vivo* interaction studies using over-expression of full-length and truncated HA-Gadd45/MyD118 constructs.

Gadd45 seems to be a key protein in cell cycle regulation and possibly involved in regulation of DNA synthesis/repair *via* its interaction with PCNA and p21. Furthermore, the Gadd45 gene is down regulated in most breast cancers by methylation of its promoter region, a common mode of regulation for a cancer suppressor. Gadd45 has also been found to be frequently mutated in resectable invasive pancreatic ductal carcinomas (Yamasawa et al., 2002) and down regulation of this gene is found in ovarian cancer, although the mechanism for this is not clear.

The aim was to investigate the nature of the interaction between Gadd45 and PCNA to gain further understanding on how these two proteins might be involved in cell-cycle check point control. Localising the PCNA binding site of Gadd45, with subsequent crystal structure of a complex between the Gadd45 and PCNA would have been interesting as so far only structures exist of proteins binding to PCNA *via* the PIP-box motif and few proteins have been found to bind to PCNA that do not contain this motif. Furthermore, I was interested in the true affinity of the Gadd45/PCNA complex as an ELISA assay gives usually fairly crude results.

4.2 MATERIALS AND METHODS

4.2.1 Cloning

The full-length human Gadd45 ORF was subcloned into a pET-14b vector (Novagen) using NdeI and BamHI sites to give a hexa His-tagged protein (Carrier et al., 1994) and was a kind gift provided by Dr. Albert J. Fornace (Lombardi Comprehensive Cancer Center, Georgetown University).

4.2.1.1 Full-length untagged Gadd45

The cDNA corresponding to full-length Gadd45 was generated by polymerase chain reaction (PCR) using pET14b-Gadd45 as a template. The sequence was amplified with Elongase Enzyme Mix (Invitrogen Inc.) using forward primer 1 and reverse primer 2 (Table 4-1).The product (Figure 4-1a), digested with NdeI and XhoI (New England Biolabs) and ligated into a similarly digested pET-23a vector (Table 4-1, lane 3) (Novagen). The C-terminal hexa His-tag intrinsic to this vector was not expressed due to an inserted 5' STOP codon in the reverse product (Table 4-2).

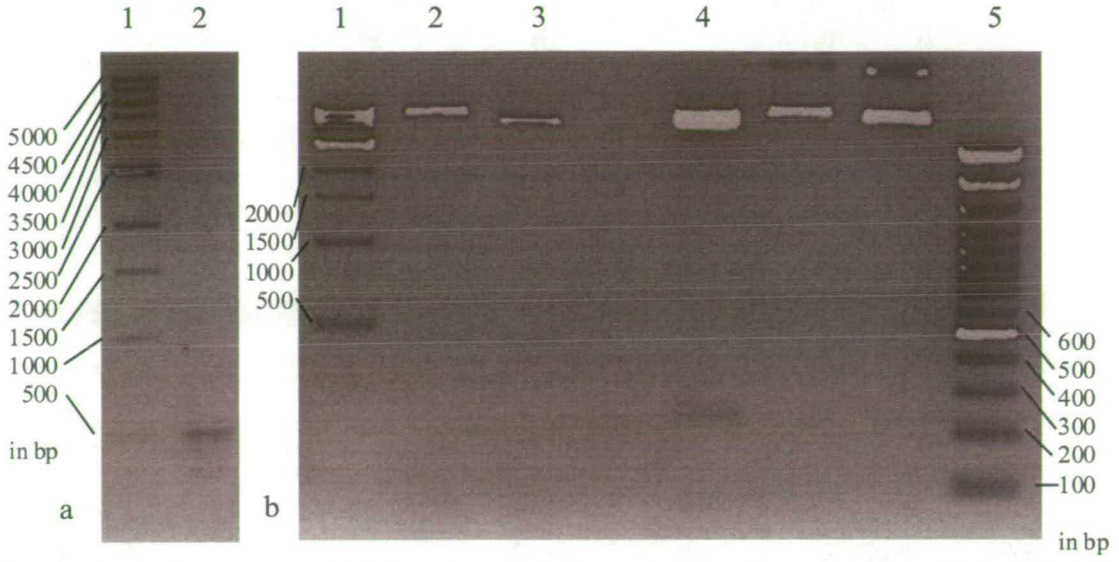


Figure 4-1 DNA agarose (1.5%) gels of PCR products and control digests for Gadd45 constructs.
a) PCR product of full length Gadd45 generated with NdeI and XhoI sites b) Restriction enzyme digest of Gadd45 constructs; gel stained with ethidium bromide. Lanes: 1= Marker, 2 = pET14b-Gadd45, 3 = pET23a-Gadd45, 4 = pGEX-KG- 95-165, 5 = Marker

4.2.1.2 GST- Gadd45 (95-165), GST- Gadd45 (93-137) and GST-Gadd45 (137-165)

Truncated GST-fusion constructs of Gadd45 were generated by PCR with forward primer 3 and reverse primer 2 (Gadd45 95-165), forward primer 3 and reverse primer 4 (Gadd45 93-137), forward primer 5 and reverse primer 2 (Gadd45 137-165) (Figure 4-2). Gadd45 (95-165) is the shortest truncation of the C-terminus that was shown to bind to PCNA in a yeast two hybrid assay (Vairapandi et al., 2000). The other two constructs divide this sequence approximately in half avoiding disruption of predicted secondary structure elements.

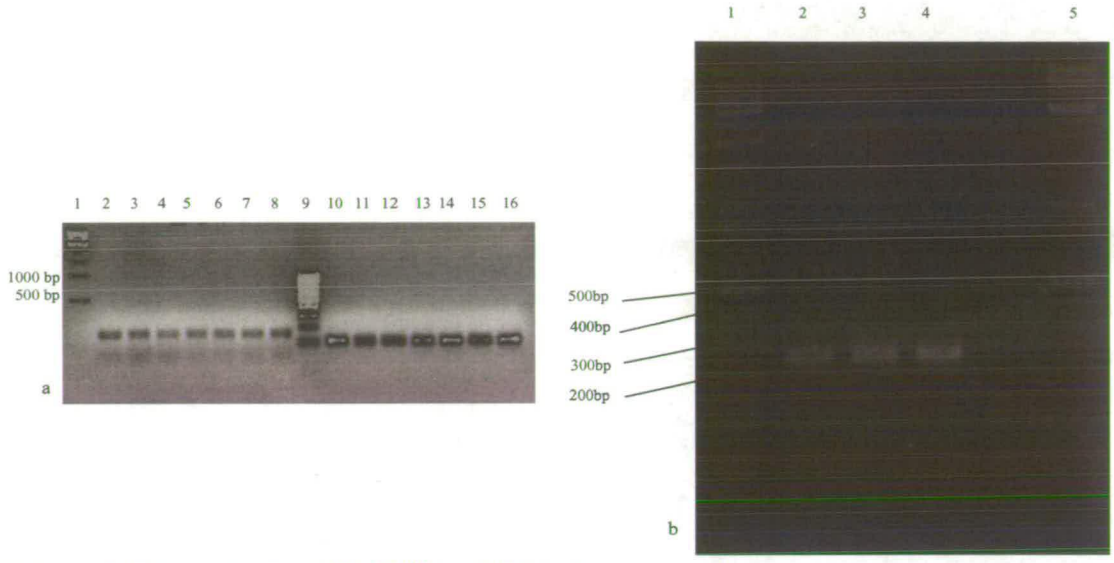


Figure 4-2 PCR product s of Gadd45 for GST-fusion constructs.
a) lane 1 = marker, lanes 2-8 = Gadd45(93-137) at different annealing temperatures (54 °C – 60 °C), lane 9 = marker, lanes 10-16 = Gadd45 (137-165) at different annealing temperatures (53 °C – 59 °C)
b) Lane 1 = marker, lanes 2-4 = Gadd45 (93-165) at different annealing temperatures (54 °C, 56 °C, 58 °C), lane 5 = marker

All three products were subcloned into pJET1.2 using a CloneJET PCR cloning kit (Fermentas) and subsequently cut out by digestion with BamHI and XhoI (Gadd45 93-137 and 137-165) or SmaI and XhoI (Gadd45 95-165) (Figure 4-1b) respectively. The resulting fragments were ligated into similarly digested pGEX-KG (Table 4-2). The sequence also contained a thrombin cleavage site for tag removal.

Primer name	Sequence	Restriction Site
Primer 1	GGGAATTCCATATGACTTTGGAGGAATTCTCGGC	NdeI
Primer 2	CCGCTCGAGTCACCGTTCAGGGAGATTAATCAC	XhoI
Primer 3	CGGGATCCAGCAACCCGGGCCGGCTG	BamHI, SmaI
Primer 4	CCGCTCGAGCTAGTCTTTCCATTGAGATGAATGTGGAT TCGTC	XhoI
Primer 5	CGGGATCCGACCCAGCTCTGTCTCAGTTGATTTGTTTTT GCCGGGAAAGTCGC	BamHI

Table 4-1 List of Primers.
The restriction sites to be incorporated into the PCR product are highlighted in green.

Plasmid	Tag	Primers used	Cloning Sites	Sequence verified
pET14b-Gadd45	6His	n/a	NdeI, BamHI	ok
pET23a-Gadd45	None	1+2	NdeI, XhoI	ok
pGEX-KG-Gadd45 (95-165)	GST	2+3	SmaI, XhoI	ok
pGEX-KG-Gadd45 (93-137)	GST	3+4	BamHI, XhoI	ok
pGEX-KG-Gadd45 (137-165)	GST	2+5	BamHI, XhoI	ok

Table 4-2 List of plasmids.
The table shows the different Gadd45 constructs obtained for this study including the presence of affinity tags and which restriction sites were engineered for cloning.

4.2.2 Expression and purification

4.2.2.1 Full-length hexa His-tagged Gadd45

Recombinant hexa His-tagged Gadd45 (20.5 kDa) was expressed in BL21 (DE3) *E. coli* (Novagen) in LB liquid media containing 50 μg/ml carbenicillin. Cultures were grown with shaking (250 rpm) at 37 °C until the A600 was ~ 0.6, then continued at 37 °C or moved to 20 °C upon which over-expression was induced by adding 0.4 mM isopropyl-β-D-thiogalactopyranoside (IPTG). Induction was continued for 2 or 5 hours, 37 °C and 20 °C, respectively and cells were harvested by centrifugation (5000 x g for 25 minutes at 4 °C. Cell pellets were flash frozen in liquid nitrogen and stored at -80 °C.

Pellets were thawed on ice and resuspended at 30 ml per 0.5 l of original cell culture in lysis buffer (Buffer A; 200 mM NaCl, 25 mM Tris pH 7.5, 2 mM β-mercaptoethanol, 10 mM imidazole, 1 mM PMSF) supplemented with 0.05% NP-40 and 400 μg/ml lysozyme. The suspension was incubated on ice for 30 minutes before sonication on ice for 6 x 15 second bursts with cooling in between. The cell lysate

was subjected to centrifugation at 15,000 x g for 30 minutes at 4 °C. The supernatant was filtered through 0.2 μm filter and applied onto a 14 ml Talon-NTA (Clontech) HR 16/10 (Pharmacia) column pre-equilibrated with Buffer A. Non-specific interactions with the resin were avoided by having 10 mM imidazole present in the equilibration buffer and bound proteins eluted with an imidazole gradient of 0% (10 mM imidazole) to 40% (206 mM imidazole) of buffer B (Buffer B; 200 mM NaCl, 25 mM Tris pH 7.5, 2 mM β-mercaptoethanol, 500 mM imidazole, 1mM PMSF) at 1 ml/minute over 10 column volumes.

Hexa His-tagged Gadd45 eluted in 4 peaks at 70 mM, 85 mM, 110 mM and 130 mM Buffer B respectively. Fractions containing Gadd45 were pooled, concentrated to 0.4 mg/ml and loaded onto a Superdex 200 HR 10/30 (Pharmacia) gel filtration column (V_t = 23.8 ml) equilibrated with Buffer C (25 mM Tris pH7.5 and 200 mM NaCl). The elution profile of recombinant hexa His-Gadd45 shows the presence of 3 distinct peaks, calculated from molecular weigh standards, corresponding to species with apparent molecular weights of 107 kDa, 66.5 kDa and 28.5 kDa.

Hexa His-Gadd45 was also purified using conventional methods. For that pellets were thawed on ice and resuspended at 30ml per 0.5 l of original cell culture in lysis buffer (Buffer D; 20 mM Tris pH 7.5, 1 mM DTT, 5 mM EDTA, 1 mM PMSF) supplemented with 400 μg/ml lysozyme. The suspension was incubated on ice for 30 minutes before sonication on ice for 6 x 10 second bursts with cooling in between. Ammonium sulphate was added slowly under stirring to give a final concentration of 1 M and the resulting solution was incubated for 1 hour on ice. The cell lysate was subjected to centrifugation at 15,000 x g for 0.5 hours at 4 °C. The supernatant was filtered through 0.2 μm filter and applied onto Phenyl Sepharose HiTrap (Pharmacia)

column (V_t = 5ml) pre-equilibrated with Buffer E (25 mM K^+MOPS pH 7.0, 1 mM EDTA, 10% glycerol, 2 mM DTT) supplemented with 1 M ammonium sulphate.

After washing the column with Buffer E supplemented with 0.4 M ammonium sulphate, recombinant hexa His-Gadd45 was eluted with a gradient of ammonium sulphate from 0.4 M to 0 M ammonium sulphate over 20 column volumes at 2 ml/minute in Buffer E. Peak fractions were analyzed by SDS-PAGE, pooled and concentrated. The pooled protein was run on a Superdex 200 HR 10/30 gel filtration column equilibrated with Buffer E supplemented with 0.2 M ammonium sulphate.

4.2.2.2 Full-length untagged Gadd45

Recombinant untagged Gadd45 was expressed in B834 (DE3) *E. coli* (Novagen) in LB liquid media containing 50 μg/ml carbenicillin. Cultures were grown with shaking (250 rpm) at 37 °C until the A600 was ~ 0.6, then moved to 28 °C upon which expression was induced by adding 0.4 mM IPTG. Induction was continued for 3 hours and cells were harvested by centrifugation at 4 °C (5000 x g for 25 minutes). Pellets were flash frozen in liquid nitrogen and stored at -80 °C.

Untagged Gadd45 was purified according to the same protocol used for conventional purification of hexa His-Gadd45 with minor modifications.

Samples of different concentration were applied to either a Superdex 200 HR10/30 or Superdex 75 HR10/30 (Pharmacia) gel filtration columns (V_t =23.8 ml) pre-equilibrated with Buffer E supplemented with 0.2 M ammonium sulphate.

Alterations to the purification protocol involved addition of 10% glycerol, DTT or NaCl to various concentrations and any effects are discussed within the results.

4.2.2.3 GST and truncated GST-Gadd45 constructs

Recombinant GST and truncated GST-Gadd45 constructs were expressed in BL21 (DE3) *E. coli* (Novagen). Cultures were grown with shaking (250 rpm) at 37 °C until an OD_{600} of ~ 0.6 was reached, and expression was induced by adding 1 mM IPTG. Induction was continued for 3 hours and cells were harvested by centrifugation at 4 °C (5000 x g for 25 minutes). Pellets were flash frozen in liquid nitrogen and stored at -80 °C.

Pellets were thawed on ice and resuspended at 30 ml per 0.5 l of original cell culture in lysis buffer (Buffer F; 50 mM Tris pH 8.0, 1 mM DTT, 1 mM EDTA, 1 mM PMSF, 2 mM benzamidine) supplemented with 0.02% Triton X-100 and 400 µg/ml lysozyme. The suspension was incubated on ice for 30 minutes before sonication on ice for 6 x 10 second bursts with cooling in between. The cell lysate was subjected to centrifugation at 15,000 x g for 30 minutes at 4 °C. The supernatant of GST, GST-Gadd45 (93-137 and 137-165) was filtered through 0.2 µm filter and applied onto a Mono Q (30µM) 10/10(Pharmacia) column (V_t ~ 5ml) pre-equilibrated with Buffer F. Recombinant proteins were eluted with a NaCl gradient from 0 M to 1 M over 20 CV at 1.5 ml/minute (Buffer F plus 1M NaCl) with the bulk of fusion protein eluting at 150 mM NaCl. Peak fractions were analyzed by SDS-PAGE and pooled. The enriched fractions were applied to a GSTPrep FF 16/10 (V_t ~ 20ml) column (Pharmacia) pre-equilibrated with Buffer F. The proteins were eluted with Buffer F supplemented with 10 mM reduced glutathione in a single step at 10 mM. Relevant fractions were pooled and buffer exchanged into storage buffer (Buffer G; 20 mM HEPES pH 7.5, 80 mM NaCl, 1 mM EDTA, 1 mM DTT, 1 mM PMSF, 2 mM benzamidine) at a ratio of 1:100 of sample to buffer and concentrated using Vivaspin

20 centrifugal concentrators (Sartorius). If necessary peak fraction were pooled, concentrated to ≤ 200 µl and applied to a Superdex 200 HR10/30 column equilibrated with buffer G. Peak fractions were analyzed by SDS-PAGE and the purest fractions were pooled and concentrated.

The supernatant of GST-Gadd45 (95-165) was applied directly to the glutathione FF column equilibrated with Buffer F and the protein was eluted with Buffer F supplemented with 10mM reduced glutathione. Peak elution fractions were pooled, applied to Mono Q 10/10(30 µm) (bed volume = 5 ml) column and eluted with a 1 M NaCl gradient from 0 M to 1 M over 20 CV at 1.5 ml/minute.

For GST-tag removal thrombin (Sigma) digest reactions were setup according to manufacturer's protocol (3 hours at 37 °C in 20 mM Tris pH 8.4, 2.5 mM $CaCl_2$, 150 mM NaCl) and quenched by addition of PMSF to give a final concentration of 2 mM. The samples were then subjected to SDS-PAGE (18 % acrylamide).

4.2.2.4 Purification of untagged human PCNA

Purified recombinant human PCNA was produced as described in section 2.3.1.

4.2.3 MALDI-TOF of GST-Gadd45 constructs

The mass spectrometry analyses were carried out on a Voyager DE-STR MALDI-TOF (Applied Biosystems) instrument using an α-cyano-4-hydroxycinnamic acid (CHCA) matrix for proteins that were digested with trypsin (Promega) prior to peptide mass fingerprinting.

4.3 PCNA/GADD45 BINDING ASSAYS

4.3.1 Gel filtration studies

10 μM of untagged Gadd45 and 10 μM or 50 μM of untagged human PCNA were mixed in Buffer H (25 mM Tris pH 7.5, 2 mM DTT and 150 mM NaCl) and incubated at 4 °C for 30 minutes before applying them to a Superdex 200 column pre-equilibrated with Buffer H. Elution volumes were compared.

4.3.2 Pull-down binding experiments

Hexa His-tagged Gadd45 was incubated at concentrations between 0.25 μM and 20 μM of with 1 μM of human PCNA in a total volume of 250 μl in Buffer I (25 mM Tris pH 7.5, 10% glycerol, 200 mM NaCl, 1 mM β-mercaptoethanol) at 4 °C for 1hr in the presence of 20 μl of Talon resin. After centrifugation for 5 minutes at 12,000 x g aliquots of the supernatants were analysed by SDS-PAGE.

4.3.3 Surface Plasmon Resonance

Pure hexa His-Gadd45 was immobilized on the surface of a NTA sensor chip essentially as described (Wear and Walkinshaw, 2006) Experiments were performed by Dr. Martin Wear.

Briefly, hexa His-Gadd45 was first captured and orientated, *via* its N-terminal His-tag, under neutral physiological conditions (pH 7.4; 150 mM NaCl) on a Ni^{2+}-NTA surface, followed by brief covalent stabilization using standard primary amine coupling chemistry. The final amount of protein immobilised on the on the surface

was 313 response units (RU). A concentration series of human PCNA from 196 nM to 50 μM, in running buffer (10 mM Hepes pH 7.4, 150 mM NaCl, 0.005% surfactant P20, and 2% ethanol), was injected over the surface at a flow rate of 50 μl/minute at 25 °C for 60 seconds. The surface was regenerated after each cycle by letting any complex dissociate in Running Buffer.

GST-tagged Gadd45 fusions were covalently immobilised on CM5 chip utilising standard amine coupling chemistries. A concentration series of human PCNA from 196 nM to 50 μM, in Running Buffer (10 mM Hepes pH 7.4, 150 mM NaCl, 0.005% surfactant P20, and 2% ethanol), was injected over the surface at a flow rate of 50 μl/minute at 25 °C for 60 seconds.

4.4 RESULTS AND DISCUSSION

4.4.1 Secondary and tertiary structure prediction of Gadd45

A sequence alignment below (Figure 4-3) of the three members of the human Gadd45 family (Gadd45, MyD118 and CR6) shows the overall high conservation of residues, with some stretches being identical.

Figure 4-4 shows the secondary structure prediction obtained May 2007 for human Gadd45 [PredictProtein (Rost et al., 2004)]. It is worth mentioning that the crystal structure of mouse Gadd45γ/CR6, which shares 53% of all residues with hGadd45, was published after the experimental work on hGadd45 for this thesis had been completed (Schrag et al., 2008).

```
hGadd45b MTLEELVACD---NAAQKMQTVTAAVEELLVAAQRQDRLTVGVYESAKLMNVDPDSVVLC 57
hGadd45a MTLEEFSAGE---QKTERMDKVGDALEEVLSKALSQRTITVGVYEAAKLLNVDPDNVVLC 57
hGadd45g MTLEEVRGQDTVPESTARMQGAGKALHELLLSAQRQGCLTAGVYESAKVLNVDPDNVTFC 60
         *****. . :   : . :*: .  *:.::*  *  *  :*.****:**::*****.*.:*

hGadd45b LLAIDEEEEDDIALQIHFTLIQSFCCDNDINIVRVSGMQRLAQLL-----GEPAETQGTT 112
hGadd45a LLAADEDDDRDVALQIHFTLIQAFCCENDINILRVSNPGRLAELLLLETDAGPAASEGAE 117
hGadd45g VLAAGEEDEGDIALQIHFTLIQAFCCENDIDIVRVGDVQRLAAIV------GAGEEAGAP 114
         :** .*::: *:*********::***:***:*:**..  *** ::       .    *:

hGadd45b EARDLHCLLVTNPHTDAWKSHGLVEVASYCEESRGNNQWVPYISLQER 160
hGadd45a QPPDLHCVLVTNPHSSQWKDPALSQLICFCRESRYMDQWVPVINLPER 165
hGadd45g G--DLHCILISNPNEDAWKDPALEKLSLFCEESRSVNDWVPSITLPE- 159
            ****:*::* : . **. .* ::  :*.***  ::*** *.* *
```

Figure 4-3 Sequence alignment of the human Gadd45 protein family using ClustalW. (Chenna et al., 2003)

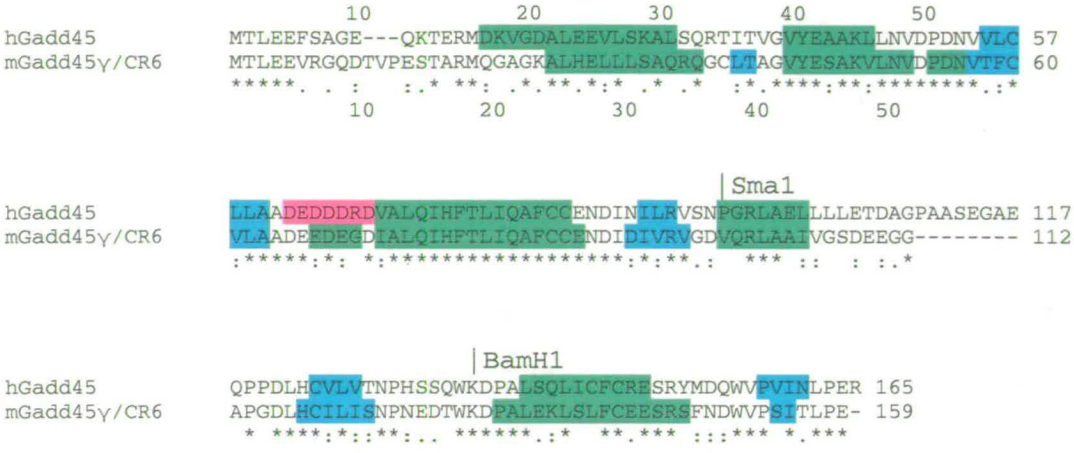


Figure 4-4 Sequence alignment of human Gadd45 and mouse Gadd45γ/CR6 with secondary structure prediction or as observed in the crystal structure (Schrag et al., 2008).
Magenta highlighting shows the sequence that has been shown to be important for Cdc2 interaction and is not present in either MyD118 or CR6. Sma1 and BamH1 mark the restriction sites used for cloning as described later. The green and blue highlighted sections mark α-helices and β-strands, respectively.

However, the regions of predicted and actual secondary structure elements are very similar in both proteins.

The C-terminal half of the Gadd45 (residues Pro95 - Arg165) that previously has been shown to interact with PCNA contains two predicted helices and two stretches of β-strand, which are also confirmed by the mouse Gadd45γ/CR6 crystal structure.

A PROSITE motif search (de Castro et al., 2006) found five potential post-translation modification sites (Table 4-3).

Post-translational Modification	Site (residue number)
Protein kinase C phosphorylation	(13) TER, (32) SQR
Casein kinase II phosphorylation	(2) TLEE, (7) SAGE
N-myristoylation	(39) GVYEAA

Table 4-3 Post-translation modification sites in Gadd45.

The latter half of the protein is rich in prolines and also contains six cysteines. The employment of DESULFIND (Ceroni et al., 2006) and DIpro (Grisendi et al., 2005)

yielded two predicted disulfide bonds; the cysteines predicted to be involved in those bonds however differ between the two results. While DIpro suggests bonds for C57/C83 and C82/C124, the most likely result for DISULFIND is C57/C82 and C124/C147.

While this information does not help with the determination of Gadd45's function in the cell, it helps to explain certain inherent problems/peculiarities of the protein which will be discussed later in this chapter.

4.4.2 Expression and Purification of Gadd45 proteins

4.4.2.1 Hexa His-Gadd45

A series of expression trials were carried out to establish the best growth and expression conditions for recombinant human hexa His-Gadd45 (20.5 kDa) (Figure 4-5). Growth to mid log phase (A600 ~ 0.6 – 0.8) at 37 °C, in LB followed by induction of over expression by addition of IPTG to 0.4 mM and further growth at either 37 °C for 3 hours or 20 °C for 5 hours gave the best results.

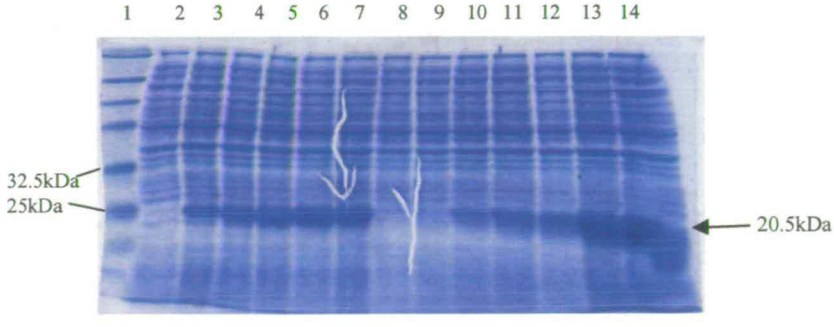


Figure 4-5 Expression trial for full-length hexa His-Gadd45.
A 12% SDS-polyacrylamide gel was run with equal amounts of culture volume and subsequently stained with coomassie. Lane 1 = marker, lanes 2-6: pre-induction, after 30min, 1 hour, 2 hours and 3 hours of induction at 37 °C, lane 8-12: after 30mins, 1 hours, 2 hours, 3 hours, 4 hours, lane 7: after 5 hours, lane 13 after 10 hours, lane 14 after 22 hours of induction at 20 °C

Hexa His-Gadd45 was expressed as described in the methods section. The first purification step used the affinity tag and a Talon-NTA column. The protein only bound well to the column when expressed at the lower temperature and could be eluted with an imidazole gradient (Figure 4-6). The pooled protein was >95% pure.

A second method of purifying Gadd45 using conventional, non-affinity chromatographic methods was adapted for hexa His-Gadd45, based on that described (Kovalsky et al., 2001) .

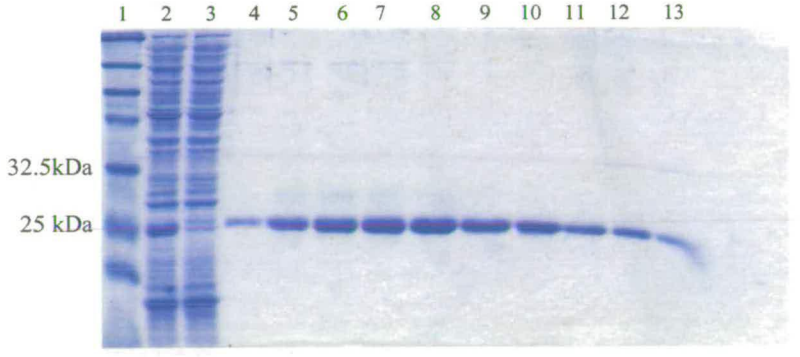


Figure 4-6 Purification of hexa His-Gadd45 on a Talon Column.
A 12% SDS-polyacrylamide gel was run with equal amounts of fraction volume ands subsequently stained with coomassie. Lane 1 = marker, lane 2 = lysate, lane 3 = flow through, lanes 4 to 13 = alternate elution fractions. An imidazole gradient from 10mM to 250mM was run over the column and Gadd45 eluted mostly between 190mM and 210mM.

A typical purification run using this protocol is shown in Figure 4-7 and Figure 4-8 using an ammonium sulfate cut to 1 M final concentration followed by hydrophobic interaction (HIC) and gel filtration chromatography. This three step protocol gave a final yield of 10 mg per liter of original cell culture and protein of 95% purity, as judged by SDS-PAGE.

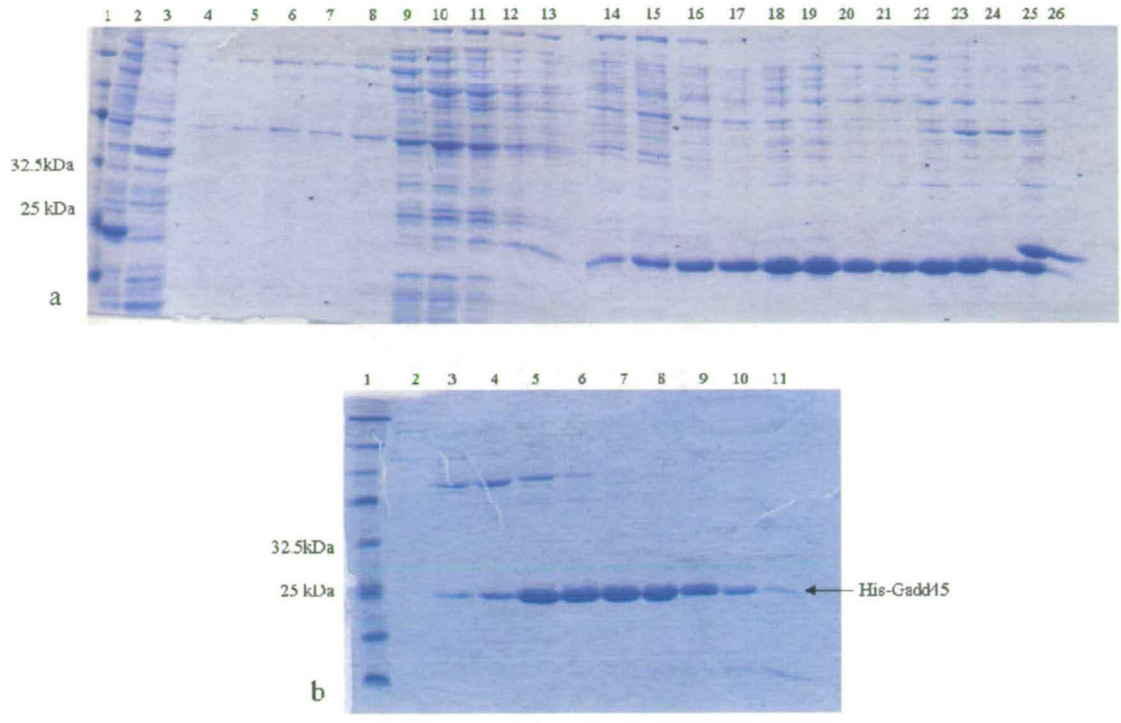


Figure 4-7 Purification of hexa His-Gadd45 using conventional methods.
a) Purification on Phenyl Sepharose column: Lane 1 = marker, lane 2 = lysate/after AS cut, lane 3 = flow through, lanes 4 – 26 = alternate elution fractions. Protein was eluted with an ammonium sulphate gradient from 0.4 M to 0 M. b) Elution fraction from Superdex 200 column: lane 1 = marker, lanes 2 – 11 = subsequent elution fractions of relevant peaks. Gels are made of 12% SDS-polyacrylamide and were stained with coomassie. Equal volumes of elution fractions were analyzed.

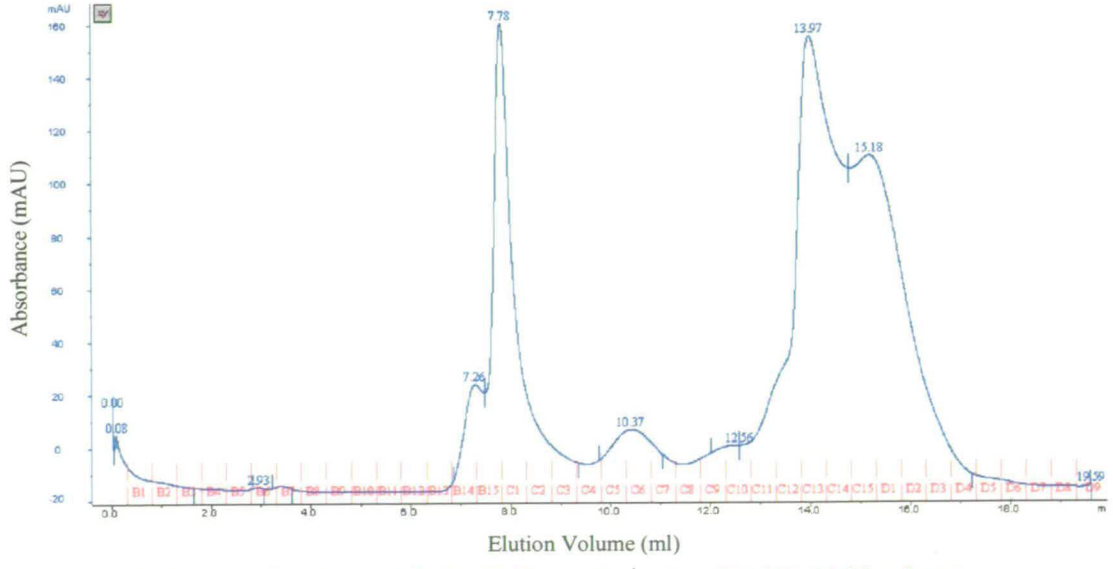


Figure 4-8 Elution profile of hexa His-Gadd45 on a Superdex 200 HR 10/30 column.
Most of the protein is found in factions C10-D4.

4.4.2.2 Untagged Gadd45

To check for expression levels of untagged Gadd45 an expression trial was carried out (Figure 4-9a). Induction of expression for 3 hours at 28 °C gave the optimal expression conditions.

Untagged Gadd45 was purified using the conventional methods protocol described above for hexa His-Gadd45. A typical purification run is illustrated (Figure 4-9b and c). Untagged Gadd45 was ≥ 95% pure after the final step, as judged by SDS-PAGE.

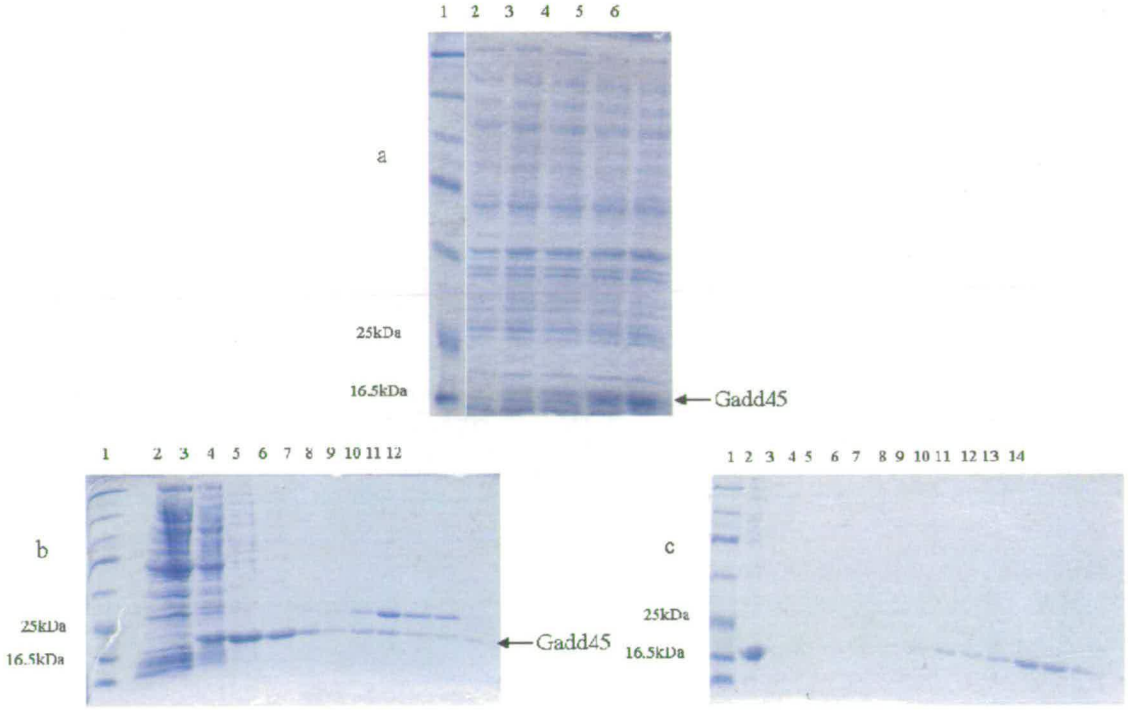


Figure 4-9 Expression and purification of untagged Gadd45.
Coomassie-stained gels in this figure are 12% SDS-polyacrylamide and equal volumes of elution fractions were analyzed. a) Expression trial. Lane 1 = marker, lanes 2 - 6: pre-induction, after 30 minutes, 1 hours, 2 hours and 3 hours of induction. b) Purification on Phenyl Sepharose. Lane 1 = marker, lane 2 = after lysate/AS cut, lane 3 = flow through, lane 4 = wash at 0.2 M AS, lanes 5 – 12 = alternate fractions of elution. Elution of Phenyl Sepharose started at 0.2 M AS instead of 0.4 M AS, so Gadd45 eluted in 0.2 M AS wash. c) Purification on Superdex 200 column. Lane 1 = marker, lane 2 = pool of fractions of lane 4 and 5 from b) (2.5 mg/ml), lanes 3 – 14 = subsequent elution fractions.

4.4.2.3 GST and GST-Gadd45 constructs

Expression trials were carried out for all three GST-Gadd45 fusion proteins. Good levels of over-expression were achieved, for all three fusion constructs, by induction for 3 hours at 37 °C with 1 mM IPTG (Figure 4-10).

The calculated molecular weights of the constructs were 31 kDa for GST-Gadd45 (93-137), 29.7 kDa for GST-Gadd45 (137-165) and 34.2 kDa for GST-Gadd45 (95-165). GST-Gadd45 (93-137) ran at about 32.5 kDa which is close to its calculated weight, GST-Gadd45 (137-165) was visible as a band running slightly lower than its calculated molecular weight.

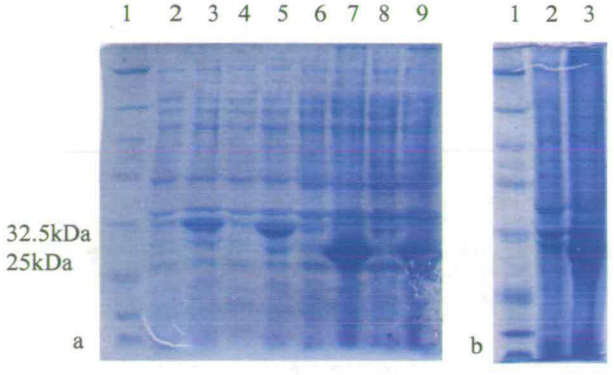


Figure 4-10 Expression trial of GST- Gadd45 constructs.
12% SDS-polyacrylamide gels were stained with coomassie. Equal amounts of culture volume were analyzed. a) Fragments 93-137 and 137-165, lane 1 = marker, lanes 2 + 4 = pre-induction 93-137, lane 3 + 5 = 3 hours after induction, lanes 6 + 8 = pre-induction, lanes 7+ 9 = 3 hours after induction b) fragment 95-165. Lane 1 = marker, lane 2 = pre-induction, lane 3 = 3 hours after induction

Interestingly, the largest construct GST-Gadd45 (95-165) ran between those first two constructs, with an apparent molecular weight of approximately 30 kDa. It is worth noting that residues 95-165 contain 8 proline residues in total. The constrained conformation of the peptidyl-prolyl bond, and the resultant 'kinks" it creates in the polypeptide chain, can result in proteins that contain a higher than normal prevalence

of proline residues running with higher or lower apparent molecular weights on SDS polyacrylamide gels depending on the shape of the kink. This might explain a higher than expected apparent molecular weight of this construct.

After expression of the three constructs they were purified utilizing a protocol with ion-exchange and affinity chromatography steps (Figure 4-11, Figure 4-12), as described above. The proteins were either judged to be sufficiently pure after the glutathione column or required a further 'polishing' gel filtration step, as in the case of GST-Gadd45 (137-165) (Figure 4-13).

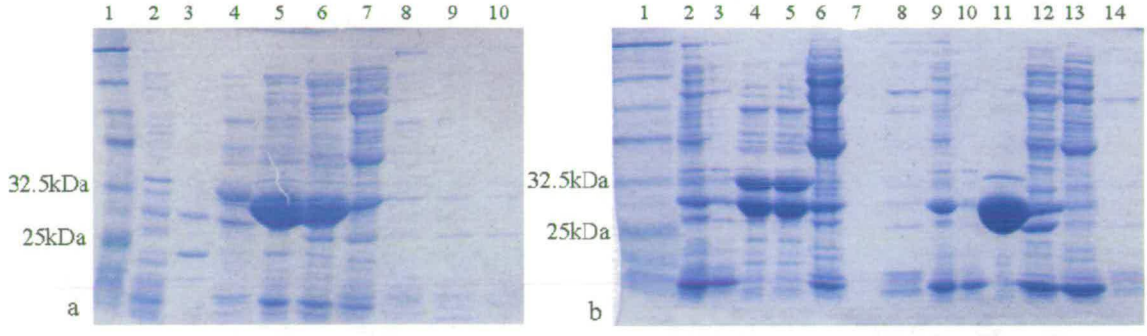


Figure 4-11 Purification of GST constructs on a MonoQ column.
a) GST-Gadd45 (93-137): lane 1 = marker, lane 2 = flow through, lanes 3-10 = selected elution peaks fractions b) GST-Gadd45 (137-165) and GST only: lane 1 = marker, lane 2 = GST-Gadd45 (137-165) lysate, lane 3 = flow through, lanes 4 to 7 = selected elution peak fractions, lane 8 = GST only lysate, lane 9 = flow through, lanes 10-13 = selected elution peak fractions. 12% SDS-polyacrylamide gels were stained with coomassie. Equal amounts of fraction volume were analyzed.

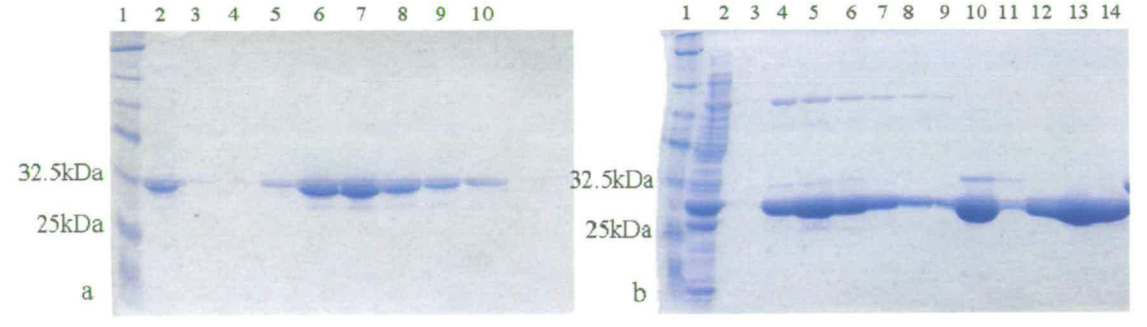


Figure 4-12 Purification of GST constructs on a GST FF column.
a) GST-Gadd45(93-137): lane 1 = marker, lane 2 = pool from MonoQ step, lane 3-10 = subsequent elution fractions b) GST-Gadd45(137-165) and GST: lanel = marker, lane 2 = GST-Gadd45(137-165) pool from MonoQ step, lane 3-9 = subsequent elution fractions, lane 10 = GST pool from MonoQ step, lanes 11-14 = subsequent elution fractions. 12% SDS-polyacrylamide gels were stained with coomassie. Equal amounts of fraction volume were analyzed.

The supernatant from cells expressing GST-Gadd45 (95-165) only contained small amounts of the tagged fusion-protein. Most of the protein was found insoluble in the cell debris pellet (Figure 4-14a). For enrichment the lysate was applied directly to the GST FF column and following elution of the fusion protein with reduced glutathione, subjected to ion exchange chromatography on a MonoQ column. The apparent lack of solubility of most of the expressed protein was not observed for the other two constructs. It is suggestive of an inherent instability/solubility problem for the largest Gadd45 fusion protein.

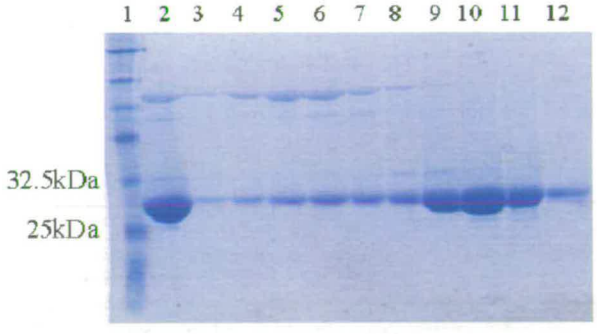


Figure 4-13 Purification of GST-Gadd45 (137-165) on Superdex 200.
Lane 1 = marker, lane 2 = pool from GST FF step, lanes 3-12 = subsequent elution fractions. A 12% SDS-polyacrylamide gel was run with equal amounts of fraction volume and stained with coomassie.

All three GST fusion constructs showed signs of instability in terms of the fusion partner losing their GST tag (Figure 4-12, Figure 4-14). A lower molecular weight band runs below the GST fusion band and could correspond to GST alone, which has a molecular weight of ~ 26 kDa, or an intermediate break-down product(s).

This breakdown instability was most prevalent for the GST-Gadd45 (95-165) fusion protein (Figure 4-14) and manifested itself most obviously following the glutathione (Fast Flow) step.

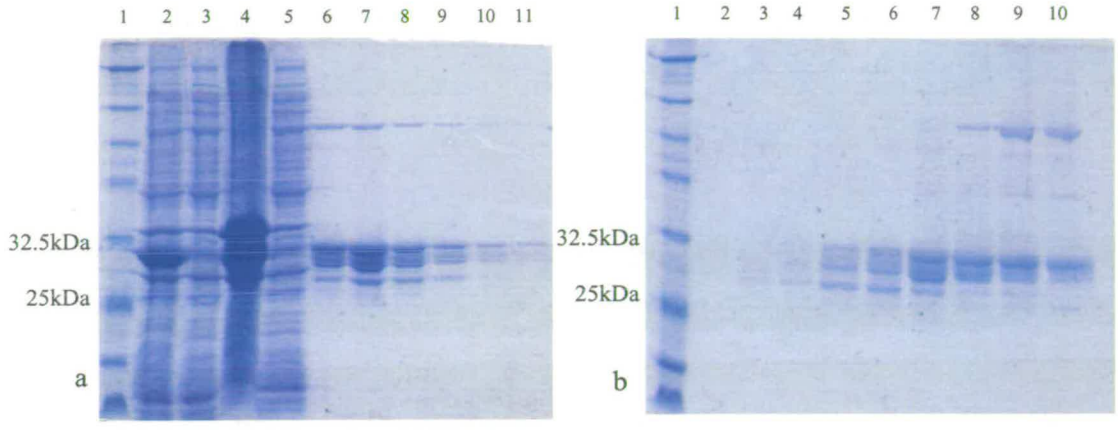


Figure 4-14 Purification of GST-Gadd45(95-165) on GST FF and MonoQ.
a) GST-FF step: lane 1 = marker, lane 2 = lysate before centrifugation, lane 3 = supernatant after centrifugation, lane 4 = pellet, lane 5 = flow through after GST FF, lanes 6-11 = subsequent elution fractions b) MonoQ step: lane 1 = marker, lanes 2-10 = subsequent elution fractions. 12% SDS-polyacrylamide gels were run with equal amounts of fraction volume and stained with coomassie

As the Gadd45 constructs were going to be used in *in vitro* solution binding assays, they were tested to see whether the shorter Gadd45 fragments were accessible to the solvent and available for interaction with other proteins or if they had become occluded/included by being proximal to GST.

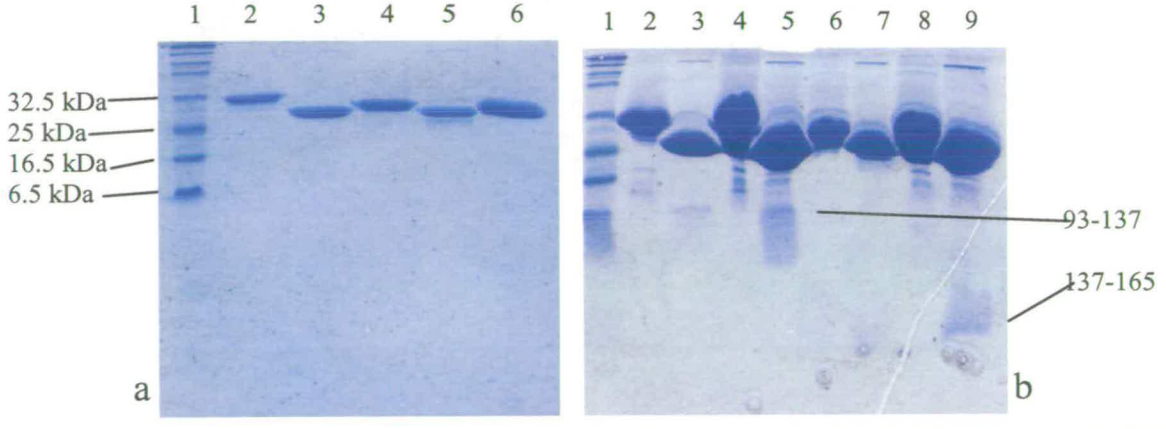


Figure 4-15 Thrombin digest of GST-Gadd45(93-137) and GST-Gadd45(137-165) on 18% SDS PAGE with various protein amounts.
a) Lane 1 = marker, lane 2 + 3 = undigested and digested GST-Gadd45(93-137) respectively, lane 4 + 5 = undigested and digested GST-Gadd45(137-165), lane 6 =purified GST only. 0.35 µl of each sample were loaded b) lane 1 = marker, lane 2 + 3 = undigested and digested GST-Gadd45(93-137) respectively, amount loaded = 1µl of each sample. Lane 4+5, same as 2+3, but with 3µl of each sample loaded, lanes 6-9 like 2-5, but with GST-Gadd45(137-165)

Therefore the peptides were cleaved from the GST tag by thrombin (only for Gadd45 93-137 = 4909 Da and Gadd45 137-165 = 3624 Da).

The cleavage of peptides from the GST tag was successful (Figure 4-15) as cleaved peptides and tag-free protein could clearly be distinguished from the uncleaved construct. Unfortunately, the incubation time for the digestion might have been too long as only little peptide was present in the samples in comparison to cleaved protein which suggests that some of the peptide might have been destroyed by unspecific thrombin activity. The resolving power of the 18% SDS polyacrylamide gel was not high enough to give realistic values of the molecular weight of the two peptides. The large differences of apparent molecular weight might be due to the same reason that leads to the variations of apparent molecular weight in the GST-Gadd45 constructs prior digestion.

4.4.3 Oligomeric state of full-length Gadd45

As part of the purification protocol both hexa His-tagged and untagged Gadd45 were subjected to analytical gel filtration using a Superdex 200 HR 10/30 column.

Hexa His-Gadd45 eluted from a Talon-NTA column between ~ 50 mM and ~160 mM imidazole. When samples of selected fractions across the elution gradient were applied to a Superdex 200 column, it was apparent that the protein eluting at higher imidazole concentrations contained higher order multimers compared to fractions from earlier in the imidazole gradient (Figure 4-16).

At all concentrations tested, Gadd45 eluted as a mixed oligomeric population, including large aggregates that eluted in the void volume (≥1.3 MDa). However, the

occurrence of higher order oligomers was directly related to higher protein concentrations. A sample of the protein eluted in the earlier part of the imidazole gradient was concentrated from 0.3 mg/ml (~14 μM) (Figure 4-16) to 2.5 mg/ml (~115 μM) and run on the Superdex 200 gel filtration which yielded just protein in the void volume. This suggests that aggregation is concentration dependent. Addition of 10% glycerol as well as 2 mM DTT resulted in a minimal stabilizing effect on the protein and the levels of aggregation were slightly reduced (Figure 4-17).

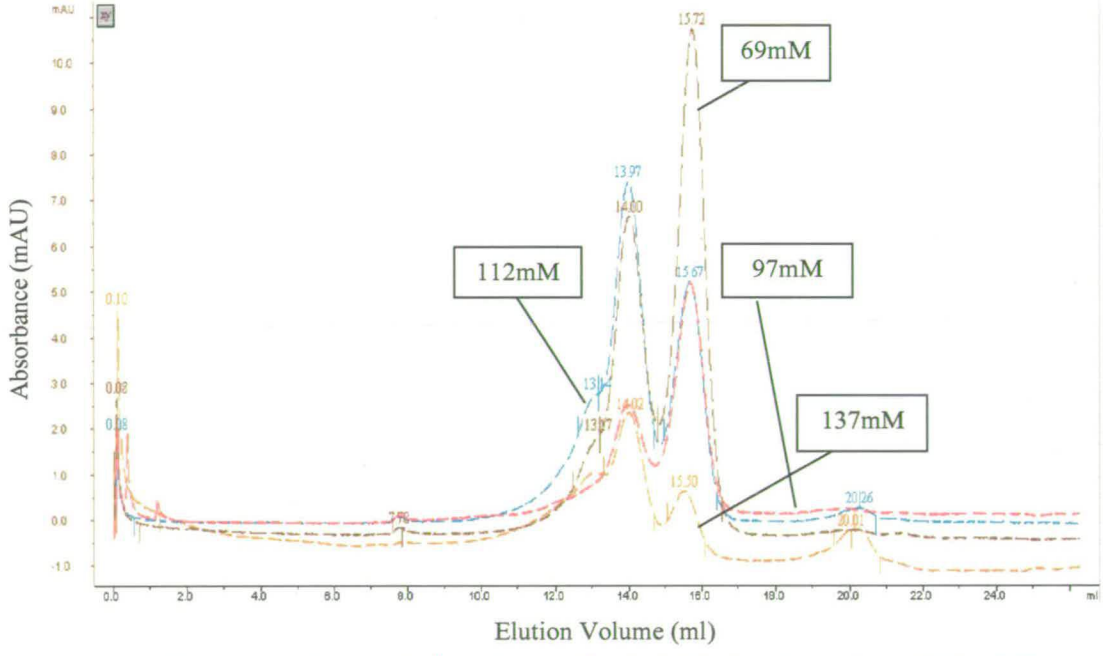


Figure 4-16 Elution profile of various hexa His-Gadd45 fractions from Talon-NTA on a Superdex 200 column.
The boxes show the imidazole concentration at which the protein was eluted from a Talon-NTA column. Protein which eluted at a low imidazole concentration contains more of a species of about 29.7 kDa, as well as 64.6 kDa and about 97 kDa. With increasing imidazole concentration protein was eluted that contained more of higher molecular weight species. Only small amounts appeared in the void volume, which is due to the very low total protein concentration. The calculated molecular weight of hexa His-Gadd45 is 20.5 kDa, so it seems that the smallest species present runs half way between a monomer and a dimer. (15.63 ml = 29716 Da, 14 ml = 64662 Da, 13.15 ml = 96896 Da)

Furthermore, untagged Gadd45 was purified using an AS cut to 1 M final concentration and a phenyl sepharose HIC step. Even at 10 times higher concentration hexa His-Gadd45 did not aggregate as badly compared to when it was

passed over a metal affinity column using Talon-NTA (all void, data not shown) (Figure 4-18).

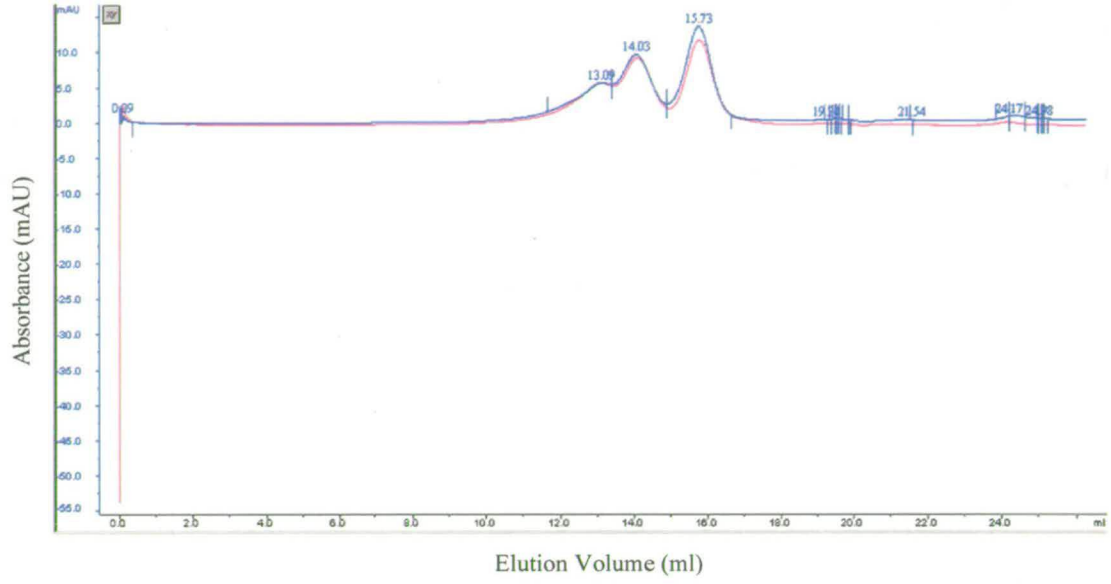


Figure 4-17 Elution profile of hexa His-Gadd45 on Superdex 200.
The same protein sample was applied to the column in 25 mM Tris pH 7.5, 200 mM NaCl plus just 10% glycerol (magenta) or both 10% glycerol and 2 mM DTT(blue). There seems to be a slight difference in the latter run, which contains proportionally more of the 29.7 kDa species than the other run, which suggest a small stabilizing effect of glycerol and DTT together.

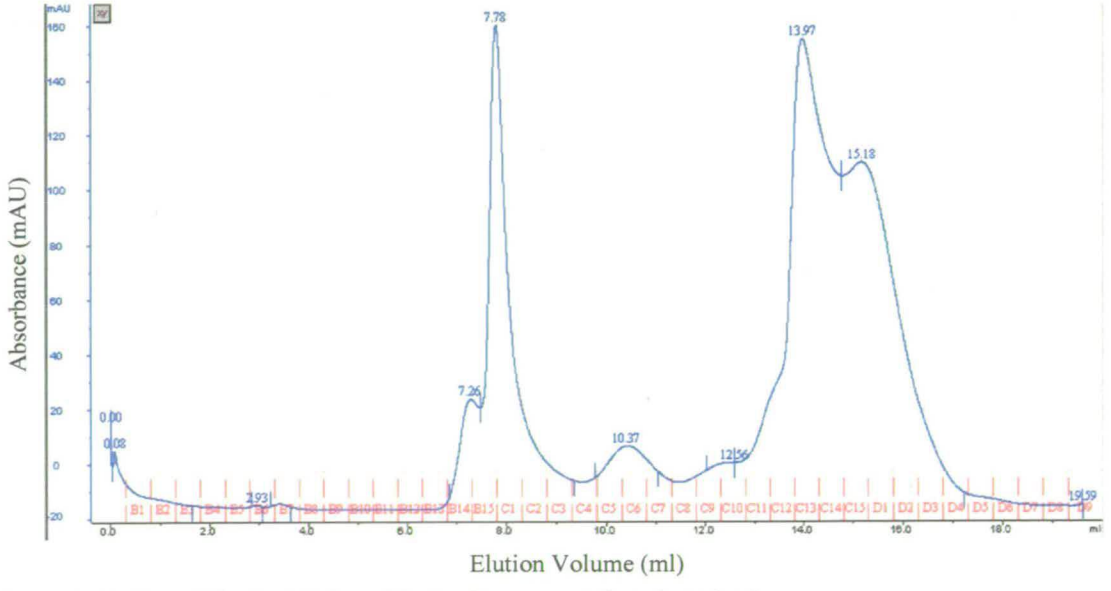


Figure 4-18 Hexa His-Gadd45 purified using conventional methods.
The protein coming off the Phenyl Sepharose was pooled, concentrated and applied to a Superdex 200 column. Despite a large amount of the protein ending up in the void volume, a lot of the protein is still present as the 29.7kDa and 64.6kDa species. A similar concentration of NTA purified protein aggregated completely (data not shown).

The previously published gel filtration studies on untagged full-length Gadd45 carried out on a S200 column and cross-linking experiments suggested Gadd45 had a natural tendency to oligomerize, however it appeared to be predominantly dimers (Kovalsky et al., 2001). The apparent affinity for self-association was calculated to be 2.5 μM.

Untagged Gadd45 purified using the non-metal affinity/conventional chromatographic methods approach ran as a single species (Figure 4-19), with an apparent molecular weight of 29.7 kDa. There was, however, protein in the void volume.

Furthermore, Kovalsky et al. (Kovalsky et al., 2001) realized that Ni-purified protein was not able to protect DNA in a nucleosome assay and had an increased level of oligomerization which is consistent with our findings. As no protection in general was observed with Gadd45 on naked DNA, it is assumed that the interaction between Gadd45 and nucleosomes is probably due to protein/protein complex formation with histones.

The authors predicted hence that the mode of interaction is dependent on the oligomeric status of Gadd45, which could also be a regulating step for Gadd45, assuming that a certain level of oligomerization in the cell is natural and not an artefact of protein purification.

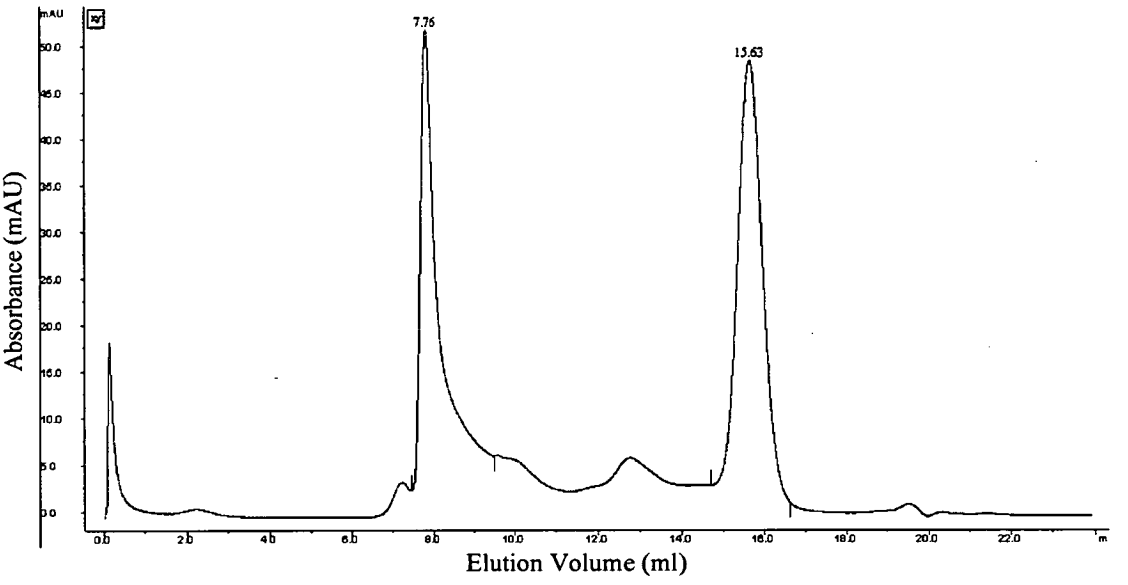


Figure 4-19 Untagged Gadd45 appears to run as a dimer.
Despite the high amount of protein in the void volume untagged Gadd45 purified using conventional methods seems to be less prone to aggregation into various multimeric states than hexa His-Gadd45

4.4.4 Interaction of human PCNA and full-length Gadd45

4.4.4.1 Pull-down (depletion) assay

Pull-down experiments were carried out to analyze the ability of hexa His-tagged Gadd45 to bind human PCNA. Human PCNA does not bind to the NTA-Talon resin on its own (Figure 4-20, Lane 3). The band of 0.25 µM Gadd45 without resin (lane 4) and the one with 1 µM Gadd45 plus resin and PCNA (lane 8) show about the same amount of free protein in the supernatant. As the resin has a maximum binding capacity of about 5–10 mg/ml even most of the His-tagged protein of the highest concentration (20 µM, lane 11) should be bound to the resin and hence not appear on

the gel. This further points towards misfolding/aggregation of Gadd45 leading to occlusion of the stretch of histidine residues necessary for resin binding.

Unsurprisingly, the amount of human PCNA in the supernatant does not change which shows a lack of affinity of hPCNA for His-tagged Gadd45 purified in the stated way.

As mentioned above the way of purification, the concentration and the presence of a tag influences the degree of aggregation.

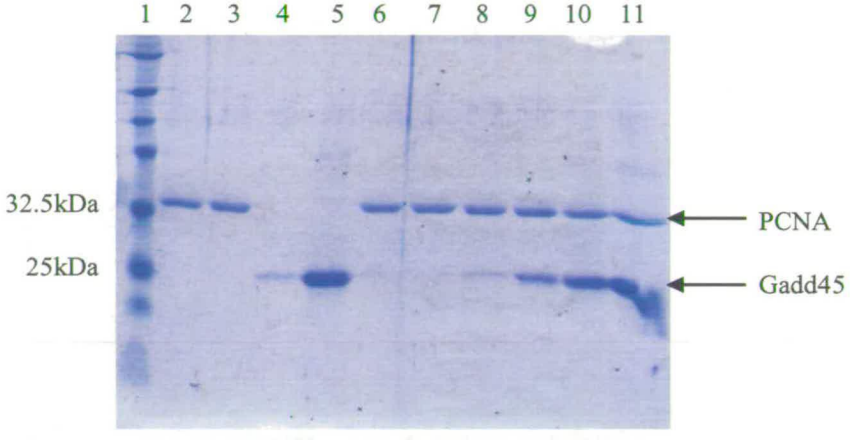


Figure 4-20 Pull-down binding assay of hexa His-Gadd45 and PCNA.
A 12% SDS-polyacrylamide gel was run with equal amounts of supernatant and stained with coomassie.
It appears that hexa His-Gadd45 is not able to bind to PCNA as the latter protein remains in the supernatant despite addition of Gadd45 in molar excess. Also most of Gadd45 does not get bound to the resin, which suggests a folding problem with this protein. However, a positive binding result in this assay would require a fairly tight and stable association between both proteins. A transient interaction would probably be missed. Key: lane 1 = marker, lane 2 = 1 μM PCNA, lane 3 = 1 μM PCNA + resin, lane 4 = 0.25 μM Gadd45, lane 5 = 20 μM Gadd45 + resin, lanes 6-11 = 0.25, 0.5, 1, 5, 10, 20 μM Gadd45 + resin + 1μM PCNA.

4.4.4.2 Gel filtration studies

200 μl samples were run on a Superdex 200 column with 50 μM untagged Gadd45 and 50 μM or 250 μM hPCNA. hPCNA eluted at about 12.69 ml and Gadd45 at 15.63 ml (Figure 4-21). There is a tiny shift towards a higher retention volume for

hPCNA (12.81 ml) at the higher concentration. According to a calibration curve (Dr. Martin Wear) Gadd45 eluted as 29.7 kDa and hPCNA as 120.6 kDa at 12.69 ml and 113.9 kDa at 12.81 ml. Although the over all mass of trimeric hPCNA is about 90 kDa, it has been shown previously that it elutes as a protein with a higher mass due to the ring-shape with cavity. Gadd45 has been shown to occur in a mainly dimeric form (Kovalsky et al., 2001), but 29.7 kDa lies in between a monomer and a dimer of Gadd45 (18.3 kDa). The structure of the closely related mouse CR6 is now known (Schrag et al., 2008) and its shape is of a globular nature. The authors of the paper suggest fairly globular dimer possibilities, which could explain a retention volume below the one expected for the Gadd45 dimer mass.

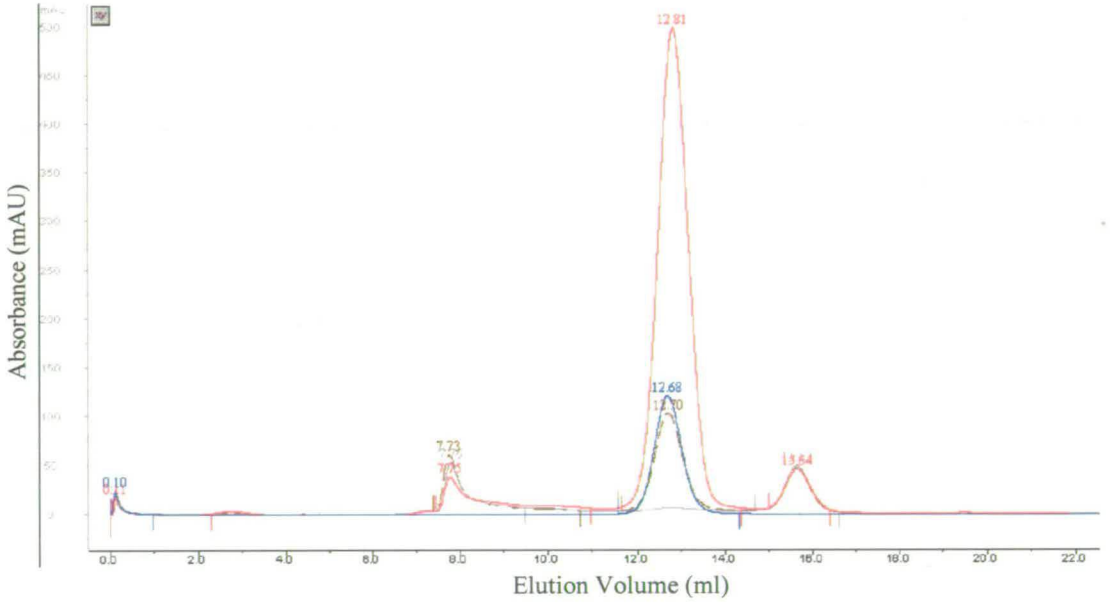


Figure 4-21 Testing untagged Gadd45/PCNA complex formation using a Superdex 200 gel filtration column.
Incubation of proteins together did not yield any significant differences in elution volume in comparison with running the proteins separately, which suggests a lack of interaction between these two proteins under given conditions. It is worth mentioning that this assay is also more suitable to pick up tight and stable interactions as in the case of the pull-down assay in 3.4.1. Key: Blue trace = 50μM hPCNA, red trace = 250μM hPCNA and 50μM Gadd45, grey trace = 50μM Gadd45, brown trace = 50μM Gadd45 and 50μM PCNA.

The total amount of free Gadd45 does not seem to differ at all which would suggest that no interaction between hPCNA and Gadd45 takes places. Neither is there any sign of the formation of new peak or increased trailing of existing ones which would be expected for complex formation. The only other peak that is present in the chromatogram is signifying the void volume (Mre = 1.3 MDa), presumably from Gadd45 aggregates.

Overall the verdict is that Gadd45 and hPCNA show no sign for direct interaction, that Gadd45 is inherently unstable and shows a tendency for aggregation. On the other hand, Gadd45 might not be able to interact in this form due to misfolding. Also this method is best suited to pick up tight interactions ($\sim \leq 10$ μM) and ones with rapid kinetics, so a possible weak interaction might have been missed.

4.4.4.3 SPR

Further analysis of the putative hexa His-Gadd45: human PCNA interaction was performed using a surface plasmon resonance (SPR). Representative sensorgrams from a PCNA concentration series generated on a surface of covalently stabilized hexa His-Gadd45, are shown in Figure 4-22. PCNA solutions of increasing concentrations (195 nM, 390 nM, 781 nM, 3125 nM, 6.25 μM, 12.5 μM 25 μM and 50 μM) were passed over a Gadd45 surface. Figure 4-22 shows the response of the surface to human PCNA. The only model that was able to fit the data at all well was a two-state reaction which includes a reaction that allows for an internal conformational change upon binding. The sensorgrams have clear bi-phasic character, both for the association and dissociation phases. A rapid binding

interaction followed by a slower internal conformational rearrangement that stabilizes the complex could very likely explain the shape of these sensorgrams. However, an interaction with mass-transport limitations can also generate curves with a similar bi-phasic character, where interaction kinetics are faster than the flow rate of the cycle is able to deliver or remove material to and from the surface. A further complicating factor is the presence of varying amounts of glycerol in the samples which may affect the relative viscosities, and thus the refractive index of the sample, exacerbating any mass transport problems in the system. However, it should be noted that the mass transport coefficient (which relates rate of transfer of mass to the surface, through the transition zone and the dextran matrix, in the non-turbulent laminar flow-cell to the molecular weight, the diffusion coefficient and flow rate) derived from the fitting procedure is significantly higher then the normally considered threshold for mass transport problems for protein:ligand interactions of $>10^9$ $RU.M^{-1}.s^{-1}$, so the data were not affected by mass transport through the presence of glycerol.

Counter to the model used in fitting, there is no evidence in the literature of a conformational change in PCNA upon protein or peptide binding, which is supported by numerous crystal structures of unliganded and complexed PCNA. Furthermore, gel-filtration analysis has already established the general instability of hexa His- and untagged Gadd45, so a non-specific interaction between human PCNA and a non-natively folded/misfolded/denatured and thus non-specifically active Gadd45 surface, that is just generally 'sticky' is a more likely explanation for the observed binding responses.

The on-rate is 0.00022 $M^{-1}.s^{-1}$, which is very slow considering that the diffusion limited on-rate for average proteins is somewhere in the range of 5 - 10 $M^{-1}.s^{-1}$ and one would have expected somewhere between 0.1 – 5 $M^{-1}.s^{-1}$ for two protein of this size, with little conformational complexity involved in their binding, from SPR analysis. This is likely a further reflection of the stability/heterogeneity problems observed with the Gadd45 surfaces.

The calculated equilibrium affinity (K_d) ~ 80 μM (response was measured 5 seconds before the end of the injection phase for 5 seconds) suggests a rather low affinity for this protein/protein interaction. This value seems at odds with the low nM affinities reported for the PCNA:Gadd45 interaction by other researchers (Vairapandi et al., 2000) and it is hence unlikely to bare any biological relevance in the cell. Even considering that a higher localized concentration is possible (e.g. nucleus) in order to achieve useful, functional amounts of complex with such a weak affinity, the kinetics of the interaction are very slow and would appear not to be chemically sensible for check point/DNA regulation functions as this is likely to require rapid kinetics.

The reverse experiment where human PCNA was immobilized on the surface of a sensor chip, using standard amine chemistries on a CM5 chip was attempted, but unfortunately proved to be unsuccessful. However, the PCNA was intolerant to the drop in pH (below ~ pH 6) and salt concentration, required to for the weak electrostatic interaction with the dextran surface to facilitate covalent coupling with the activated surface. Dropping then pH below ~ 6 resulted in PCNA precipitating out of solution.

The K_ds we calculated might be comparable with affinities calculated from ELISA binding assays carried out by others (Vairapandi et al., 2000), but they do not seem

to be due to any specific interactions. To sum up, we have reasons to believe that their data could be accredited to non-specific binding due to misfolded protein.

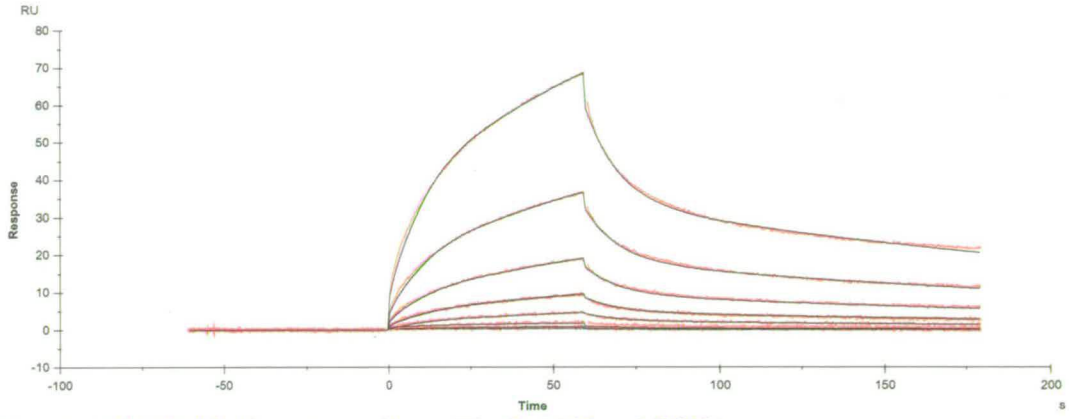


Figure 4-22 SPR binding assay of hexa His-Gadd45 and PCNA.
Increasing concentrations of PCNA (195 nM, 390 nM, 781 nM, 3125 nM, 6.25 μM, 12.5 μM 2 5μM and 50 μM) were passed over the Gadd45 surface. The response to increased PCNA concentrations can be interpreted as a 2-state binding process with conformational change or, more likely, as non-specific binding due to misfolded/sticky protein.

4.4.5 GST-Gadd45 constructs

4.4.5.1 Interaction of GST-Gadd45 fragments with human PCNA using SPR

Covalent coupling of GST-Gadd45 fragments provided a stable surface on CM5 (CM4 for GST-Gadd45 (95-165)). Increasing concentrations of human PCNA were applied to the surface, but neither specific nor non-specific binding was observed. The lack of stickiness of the GST-Gadd45 surfaces suggest correctly folded and stable GST, while the C-terminal parts of Gadd45 do not appear to have an affinity for PCNA as even very high concentrations of human PCNA did not seem to have an effect.

The recently published crystal structure of mouse Gadd45γ/CR6 shows that the tertiary structure of these GST-constructs is likely to be compromised due to the location of those regions in folded protein (Figure 4-23).

This could explain the lack of binding in the SPR experiments, but would also lead to questioning of previous results (Vairapandi et al., 2000) on which this work had been based.

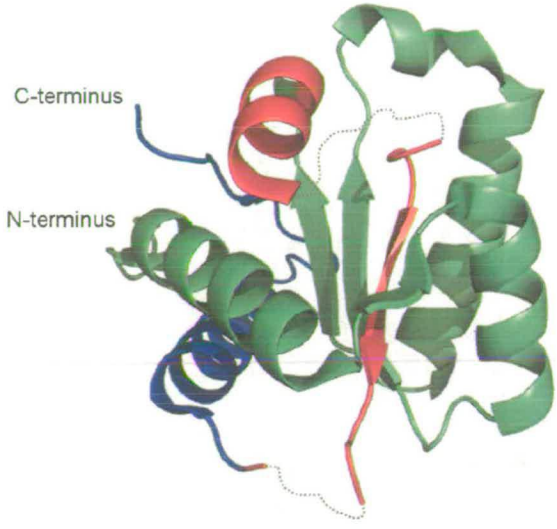


Figure 4-23 The crystal structure of mouse Gadd45γ/CR6 (Schrag et al., 2008).
The regions equivalent to residues 95-137 and residues 137-165 in hGadd45 are marked in red and blue, respectively. Grey dotted lines show loops that are not included in the model. The secondary structure elements contained in either of the fragments seem to interact with helices or strands located in other parts of the structure. It is hence likely that the Gadd45 fragments in our GST constructs might not be able to assume the secondary structure elements possibly needed for binding to PCNA.

4.5 SUMMARY, CONCLUSION AND FUTURE WORK

To sum up, despite literature results claiming a direct *in vitro* interaction between recombinantly expressed PCNA and full-length Gadd45 (Vairapandi et al., 2000), our data fail to confirm this. While results of both gel filtration and pull-down depletion assays suggest no interaction, SPR data point towards a non-specific interaction, which is likely to be due to misfolded Gadd45. We believe, in concord with previous work, that Gadd45 is unstable with a tendency to aggregate.

Private communication with Dr. Fornace implied that crystallization of full-length Gadd45 has been attempted and failed, which could be explained on the basis of the oligomeric state of hexa His-tagged Gadd45 purified using Ni-NTA resin.

Interaction of PCNA with the C-terminal parts of Gadd45 *in vitro* could not be confirmed by our work either.

Previously it had been found that the N-terminal residues 1-94 of Gadd45 are sufficient for interacting with PCNA (Hall et al., 1995a), which is at odds with the findings of Vaiparandi et al. Confirmative experiments could be designed in form of making tagged N-terminal Gadd45 constructs to use with SPR.

Furthermore, it cannot be ruled out that the PCNA:Gadd45 interaction needs some sort of co-factor (e.g. DNA) for the formation of a stable complex, which could be investigated in the future.

CHAPTER 5. Final Summary and Conclusion

This thesis is based on structural and biochemical work carried out with SpPCNA and hPCNA. I have determined the structure of SpPCNA and so added to the information available about this model organism.

As pointed out in Chapter 1 the hydrophobic pocket in PCNA and other sliding clamps is highly conserved. This new structure further confirms this notion.

Having obtained a complex structure of SpPCNA with p21 (141-160) allowed us to compare the binding mode of the peptide to the yeast protein with the complex of hPCNA and p21 (141-160). This structure has hence yielded information about protein-protein interactions. Despite the high degree of conservation between hPCNA and SpPCNA and the presence of most hydrogen donors/acceptors observed in the human complex structure, the peptide was unable to bind to the IDCL in the yeast protein and the C-terminal end of the peptide bound to the hydrophobic pocket of a nearby second PCNA molecule instead.

In order to investigate if this is merely an artefact of crystal packing we used various techniques to test the affinity of peptides p21 (141-160) and p21 (141-152) for both SpPCNA and hPCNA. Since p21 (141-152) is missing the portion that is making the contacts to the IDCL in hPCNA it is expected to bind with the same affinity to SpPCNA as the full length peptide. Our affinity data suggest that this is the case, but for confirmation further experiments using both ITC and SPR should be set up with the two different peptides.

The significance of the p21 (141-160) binding mode in SpPCNA *in vivo* is difficult to predict as fission yeast does not have a p21 homologue. Since the IDCL feature is conserved between all areas of life and since other proteins like DNA polymerase δ

seem to make interactions with this loop, it can be assumed that it also plays an important role in yeast protein-protein interactions.

The second interesting feature of this structure is that the C-terminus of p21 (141-160) is found in the hydrophobic pocket of an adjacent ring. That an untypical ligand can bind to the apparently specific pocket beneath the IDCL is also observed in the unliganded structure of SpPCNA, where histidine residues from the hexa His-tags are located in the pockets of adjacent rings. The unusual feature of His-tags bound to the hydrophobic pocket challenges somewhat the concept that this pocket is specific for accommodating the PIP-box binding residues found in various PCNA partner proteins. This finding, however, may be useful for future work on finding potential inhibitors for PCNA, which leads to the second part of this thesis.

The work for this thesis also involved a search for novel drug-leads that would bind specifically to the hydrophobic pocket. A small molecule like that could be used as an anti-cancer drug agent by inhibiting DNA synthesis and other PCNA-related functions. A number of small drug-like molecules from *in silico* screening were tested using an FP competitive assay and a thermal denaturation assay. None of the compounds, however, showed any binding in either assay.

Designing drugs based on structural information is inherently difficult and there are few successful examples. To find a small molecule that mimics in essence the interactions formed at a protein:protein interface is tricky as this usually involves a large number of both polar and non-polar interactions. A drug molecule, however, is supposed to make only a defined maximum number of interactions in order to fulfil the Lipinski rules. These rules ensure that the compound remains non-peptide-like, which is important for oral activity in humans.

So one reason why our selected compounds have not shown any binding may be that the interactions predicted by LIDEAUS are actually not enough to mimic the p21 interaction and to bind tightly enough to be picked up in the assays.

Furthermore, while LIDEAUS showed good fits for all selected compounds, it does not take into consideration rotational bonds. In solution the small molecules may assume a preferred conformation different to the one display in the library file. This could mean that there are enthalpic and entropic advantages for the molecules to stay in solution and to form hydrogen bonds with bulk water molecules, rather than replace water from the protein surface. It has been known that water molecules near a hydrophobic pocket are forming tighter bonds which would have to be broken up for compound binding. In future the use of programmes that can screen compounds based on their preferred conformation in solution will give rise to hits that have the best chance of binding to a given pocket.

Another limiting factor is the poor solubility of many compounds in aqueous solution despite good logP ratings. The data suggest that this predicted value can not be relied on when making decisions on solubility. Hence it is possible that some compounds would have shown binding at a higher concentration.

Failure in finding effective molecules may also have been down to the choice of libraries that were screened. ASINEX and Specs were chosen on grounds of speed of delivery and price. Other compounds that would have been harder to obtain might have yielded a positive hit in the assays.

There are a few points to consider when using the assays mentioned above. The FP competitive assay relies on a molecule to compete with the labelled PL peptide, which has a high affinity for PCNA. While a shift in polarization could be seen with

p21 (141-152) none of the compounds showed an effect. One reason may be that the compounds are not binding to that pocket for reasons discussed above. It is also possible, however, that some of the small molecules can bind to the pocket, but with a much lower affinity than that of the model inhibitor peptide. At maximum concentrations that could be obtained in the experimental buffer the compounds may not have been able to compete successfully with the labelled probe as their affinity may be fairly weak.

The thermal denaturation assay looks at the stabilizing effect of compounds when the protein is exposed to increasing temperatures. The onset of unfolding starts only at about 40°C for most proteins and only from that point any effects of a compound on the protein stability can be observed. While a molecule might have a reasonable affinity at 25°C, at 40°C this will have potentially reduced significantly. Another point to consider is that even if a compound bound specifically to the pocket, the number of interactions with PCNA might be too few to have a stabilizing effect under the given conditions.

Since PCNA does not have a specific enzymatic activity it is fairly difficult to develop binding assays that are both specific and sensitive. A functional assay that monitors DNA replication would give the opportunity to check for binding at a lower temperature and avoid the need for competition against high concentrations of tightly binding peptides.

Finally, in an attempt to learn more about how proteins lacking the PIP-box motif interact with PCNA, we tried to further characterize the direct *in vitro* interaction between PCNA and Gadd45, which had been previously observed by Vaiparandi et

al. However, our data fail to confirm what was shown before which leads to questioning these previous results.

One possible reason for failing to observe a direct interaction between PCNA and Gadd45 is the inherent instability of recombinantly expressed Gadd45. Although the method of purification has a slight effect on oligomerization states, it is possible that all of the protein that was used in this study was mis-folded or partially aggregated which would lead to a loss of function and impair its ability to bind to partner proteins i.e. PCNA. The instability of Gadd45 has been reported by other groups. Previously published Gadd45/PCNA binding results were based amongst others on an ELISA assay that used a method of immobilizing PCNA on the plates that may have deteriorating effects on it. The affinity data that this group derived from that assay are may be based on non-specific binding due to suboptimal quality of proteins used.

Most interaction studies between Gadd45 and PCNA carried out by other groups involved the use of yeast-two-hybrid assays. Gadd45 is likely to be more stable in the cell and hence able to bind to PCNA. It is also possible that this interaction might need another factor freely present in yeast cells for stabilization.

While p21 and Gadd45 compete for association with PCNA their mode of binding is different as Gadd45 does not contain a PIP-box motif. Gadd45's suppressive checkpoint effect is mainly due to disrupting the Cdc2-cyclin B1 complex important for G2/M transition, whereas p21 can act as a Cdk-cyclin inhibitor. By developing a drug that mimics the interaction of Gadd45 with PCNA cell division could possibly be stopped by disrupting the Cdc2-cyclin B1 complex which also contains PCNA during G2/M transition.

Finally, it may be necessary to turn to another example of the non-PIP PCNA partners to discover the basis of these interactions as working with Gadd45 has proven to be problematic.

References

Amin, N. S., and Holm, C. (1996). In vivo analysis reveals that the interdomain region of the yeast proliferating cell nuclear antigen is important for DNA replication and DNA repair. Genetics *144*, 479-493.

Amson, R., Lassalle, J. M., Halley, H., Prieur, S., Lethrosne, F., Roperch, J. P., Israeli, D., Gendron, M. C., Duyckaerts, C., and Checler, F. (2000). Behavioral alterations associated with apoptosis and down-regulation of presenilin 1 in the brains of p53-deficient mice. Proc Natl Acad Sci US A *97*, 5346-5350.

Arakawa, H., Moldovan, G. L., Saribasak, H., Saribasak, N. N., Jentsch, S., and Buerstedde, J. M. (2006). A Role for PCNA Ubiquitination in Immunoglobulin Hypermutation. PLoS Biol *4*, e366.

Arias, E. E., and Walter, J. C. (2006). PCNA functions as a molecular platform to trigger Cdt1 destruction and prevent re-replication. Nat Cell Biol *8*, 84-90.

Arroyo, M. P., Downey, K. M., So, A. G., and Wang, T. S. F. (1996). Schizosaccharomyces pombe proliferating cell nuclear antigen mutations affect DNA polymerase delta processivity. Journal of Biological Chemistry *271*, 15971-15980.

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

Azam, N., Vairapandi, M., Zhang, W., Hoffman, B., and Liebermann, D. A. (2001). Interaction of CR6 (GADD45g) with Proliferating Cell Nuclear Antigen Impedes Negative Growth Control. Journal of Biological Chemistry *276*, 2766-2774.

Bienko, M., Green, C. M., Crosetto, N., Rudolf, F., Zapart, G., Coull, B., Kannouche, P., Wider, G., Peter, M., and Lehmann, A. R. (2005). Ubiquitin-Binding Domains in Y-Family Polymerases Regulate Translesion Synthesis (American Association for the Advancement of Science), pp. 1821-1824.

Bird, A. (2002). DNA methylation patterns and epigenetic memory. Genes & Development *16*, 6-21.

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

Brand, S. R., Bernstein, R. M., and Mathews, M. B. (1994). Autoreactive Epitope Profiles of the Proliferating Cell Nuclear Antigen Define 2 Classes of Autoantibodies. Journal of Immunology *152*, 4120-4128.

Brandts, J. F., and Lin, L. N. (1990). Study of strong to ultratight protein interactions using differential scanning calorimetry. Biochemistry *29*, 6927-6940.

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

Bukholm, I. K., and Nesland, J. M. (2000). Protein expression of p53, p21 (WAF1/CIP1), bcl-2, Bax, cyclin D1 and pRb in human colon carcinomas. Virchows Archiv *436*, 224-228.

Bullock, A. N., Debreczeni, J. E., Fedorov, O. Y., Nelson, A., Marsden, B. D., and Knapp, S. (2005). Structural basis of inhibitor specificity of the human protooncogene proviral insertion site in Moloney murine leukemia virus (PIM1) kinase. J Med Chem *48*, 7604-7614.

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

Burnouf, D. Y., Olieric, V., Wagner, J., Fujii, S., Reinbolt, J., Fuchs, R. P. P., and Dumas, P. (2004). Structural and biochemical analysis of sliding clamp/ligand interactions suggest a competition between replicative and translesion DNA polymerases. Journal of Molecular Biology *335*, 1187-1197.

Carrier, F., Smith, M. L., Bae, I., Kilpatrick, K. E., Lansing, T. J., Chen, C. Y., Engelstein, M., Friend, S. H., Henner, W. D., and Gilmer, T. M. (1994). Characterization of human Gadd45, a p53-regulated protein. Journal of Biological Chemistry *269*, 32672-32677.

Carson, M., Johnson, D. H., McDonald, H., Brouillette, C., and DeLucas, L. J. (2007). His-tag impact on structure. Acta Crystallographica D *63*, 295-301.

Ceroni, A., Passerini, A., Vullo, A., and Frasconi, P. (2006). DISULFIND: a disulfide bonding state and cysteine connectivity prediction server. Nucleic Acids Research *34*, W177.

Chang, D. Y., and Lu, A. L. (2002). Functional interaction of MutY homolog (MYH) with proliferating cell nuclear antigen (PCNA) in fission yeast, Schizosaccharomyces pombe. J Biol Chem *277*, 11853-11858.

Chant, A., Kraemer-Pecore, C. M., Watkin, R., and Kneale, G. G. (2005). Attachment of a histidine tag to the minimal zinc finger protein of the Aspergillus nidulans gene regulatory protein AreA causes a conformational change at the DNA-binding site. Protein Expr Purif *39*, 152–159.

Chapados, B. R., Hosfield, D. J., Han, S., Qiu, J. Z., Yelent, B., Shen, B. H., and Tainer, J. A. (2004). Structural basis for FEN-1 substrate specificity and PCNA-mediated activation in DNA replication and repair. Cell *116*, 39-50.

Chen, I. T., Smith, M. L., O'Connor, P. M., and Fornace Jr, A. J. (1995). Direct interaction of Gadd45 with PCNA and evidence for competitive interaction of Gadd45 and p21Waf1/Cip1 with PCNA. Oncogene *11*, 1931-1937.

Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T. J., Higgins, D. G., and Thompson, J. D. (2003). Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Research *31*, 3497-3500.

Chuang, L. S., Ian, H. I., Koh, T. W., Ng, H. H., Xu, G., and Li, B. F. (1997). Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1.

Clark, A. B., Valle, F., Drotschmann, K., Gary, R. K., and Kunkel, T. A. (2000). Functional Interaction of Proliferating Cell Nuclear Antigen with MSH2-MSH6 and MSH2-MSH3 Complexes (ASBMB), pp. 36498-36501.

Copani, A., Copani, A., Angela Sortino, M., Nicoletti, F., Bruno, V., Nicoletti, F., Ubertia, D., and Memo, M. (2001). Activation of cell-cycle-associated proteins in neuronal death: a mandatory or dispensable path? Trends in Neurosciences *24*, 25-31.

Daimon, K., Kawarabayasi, Y., Kikuchi, H., Sako, Y., and Ishino, Y. (2002). Three proliferating cell nuclear antigen-like proteins found in the hyperthermophilic archaeon Aeropyrum pernix: Interactions with the two DNA polymerases. Journal of Bacteriology *184*, 687-694.

de Castro, E., Sigrist, C. J. A., Gattiker, A., Bulliard, V., Langendijk-Genevaux, P. S., Gasteiger, E., Bairoch, A., and Hulo, N. (2006). ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. Nucleic Acids Research *34*, W362.

Del Carratore, R., Della Croce, C., Simili, M., Taccini, E., Scavuzzo, M., and Sbrana, S. (2002). Cell cycle and morphological alterations as indicative of apoptosis promoted by UV irradiation in S. cerevisiae. Mut Res-Genetic Toxicology and Environmental Mutagenesis *513*, 183-191.

Devlin, J. P. (1997). High Throughput Screening: the discovery of bioactive substances: Marcel Dekker).

Dianova, II, Bohr, V. A., and Dianov, G. L. (2001). Interaction of human AP endonuclease 1 with flap endonuclease 1 and proliferating cell nuclear antigen involved in long-patch base excision repair. Biochemistry *40*, 12639-12644.

Dionne, I., Nookala, R. K., Jackson, S. P., Doherty, A. J., and Bell, S. D. (2003). A heterotrimeric PCNA in the hyperthermophilic archaeon Sulfolobus solfataricus. Molecular Cell *11*, 275-282.

Ducoux, M., Urbach, S., Baldacci, G., Hubscher, U., Koundrioukoff, S., Christensen, J., and Hughes, P. (2001). Mediation of Proliferating Cell Nuclear Antigen (PCNA)-dependent DNA Replication through a Conserved p21Cip1-like PCNA-binding Motif Present in the Third Subunit of Human DNA Polymerase d. Journal of Biological Chemistry *276*, 49258-49266.

Eissenberg, J. C., Ayyagari, R., Gomes, X. V., and Burgers, P. M. J. (1997). Mutations in yeast proliferating cell nuclear antigen define distinct sites for interaction with DNA polymerase delta and DNA polymerase epsilon. Molecular and Cellular Biology *17*, 6367-6378.

Eling, T. E., and DiAugustine, R. P. (1971). A role for phospholipids in the binding and metabolism of drugs by hepatic microsomes. Use of the fluorescent hydrophobic probe 1-anilinonaphthalene-8-sulphonate. Biochem J *123*, 539-549.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta crystallographica Section D, Biological crystallography *60*, 2126-2132.

Ericsson, U. B., Hallberg, B. M., DeTitta, G. T., Dekker, N., and Nordlund, P. (2006). Thermofluor-based high-throughput stability optimization of proteins for structural studies. Analytical Biochemistry *357*, 289-298.

Evans, P. (1997). SCALA. Medical Research Council, CCP 4 Newsletter on Protein Crystallography, 22-24.

Fien, K., and Stillman, B. (1992). Identification of replication factor C from Saccharomyces cerevisiae: a component of the leading-strand DNA replication complex. Molecular and Cellular Biology *12*, 155-163.

Fontanini, G., Pingitore, R., Bigini, D., Vignati, S., Pepe, S., Ruggiero, A., and Macchiarini, P. (1992). Growth fraction in non-small cell lung cancer estimated by proliferating cell nuclear antigen and comparison with Ki-67 labeling and DNA flow cytometry data. American Journal of Pathology *141*, 1285-1290.

Fornace, A. J., Alamo, I., and Hollander, M. C. (1988). DNA Damage-Inducible Transcripts in Mammalian Cells. Proceedings of the National Academy of Sciences *85*, 8800-8804.

Fornace, A. J., Nebert, D. W., Hollander, M. C., Luethy, J. D., Papathanasiou, M., Fargnoli, J., and Holbrook, N. J. (1989). Mammalian genes coordinately regulated by

growth arrest signals and DNA-damaging agents. Molecular and Cellular Biology *9*, 4196-4203.

Fröhlich, K. U., and Madeo, F. (2000). Apoptosis in yeast–a monocellular organism exhibits altruistic behaviour. FEBS Letters *473*, 6-9.

Fry, D. C., and Vassilev, L. T. (2005). Targeting protein–protein interactions for cancer therapy. Journal of Molecular Medicine *83*, 955-963.

Fukuda, K., Morioka, H., Imajou, S., Ikeda, S., Ohtsuka, E., and Tsurimoto, T. (1995). Structure-Function Relationship of the Eukaryotic DNA-Replication Factor, Proliferating Cell Nuclear Antigen. Journal of Biological Chemistry *270*, 22527-22534.

Gaillard, P. H., Martini, E. M., Kaufman, P. D., Stillman, B., Moustacchi, E., and Almouzni, G. (1996). Chromatin assembly coupled to DNA repair: a new role for chromatin assembly factor I. Cell *86*, 887-896.

Gogol, E. P., Young, M. C., Kubasek, W. L., Jarvis, T. C., and Hipple, P. H. v. (1992). Cryoelectron microscopic visualization of functional subassemblies of the bacteriophage T4 DNA replication complex. Journal of Molecular Biology *224*, 395-412.

Grisendi, S., Bernardi, R., Rossi, M., Cheng, K., Khandker, L., Manova, K., and Pandolfi, P. P. (2005). Role of nucleophosmin in embryonic development and tumorigenesis. Nature *437*, 147-153.

Guillouf, C., Grana, X., Selvakumaran, M., De Luca, A., Giordano, A., Hoffman, B., and Liebermann, D. A. (1995). Dissection of the genetic programs of p53-mediated G1 growth arrest and apoptosis: blocking p53-induced apoptosis unmasks G1 arrest. Blood *85*, 2691.

Gulbis, J. M., Kelman, Z., Hurwitz, J., Odonnell, M., and Kuriyan, J. (1996). Structure of the C-terminal region of p21(WAF1/CIP1) complexed with human PCNA. Cell *87*, 297-306.

Hall, P. A., Kearsey, J. M., Coates, P. J., Norman, D. G., Warbrick, E., and Cox, L. S. (1995a). Characterization of the interaction between PCNA and Gadd45. Oncogene *10*, 2427-2433.

Hall, P. A., Kearsey, J. M., Coates, P. J., Norman, D. G., Warbrick, E., and Cox, L. S. (1995b). Characterization of the Interaction between Pcna and Gadd45. Oncogene *10*, 2427-2433.

Hata, S., Kouchi, H., Tanaka, Y., Minami, E., Matsumoto, T., Suzuka, I., and Hashimoto, J. (1992). Identification of Carrot Cdna Clones Encoding a 2nd Putative Proliferating Cell-Nuclear Antigen, DNA Polymerase-Omicron Auxiliary Protein. European Journal of Biochemistry *203*, 367-371.

Haugland, R. P. (2005). The Handbook — A Guide to Fluorescent Probes and Labeling Technologies, 10 edn: Invitrogen Corp.).

Henderson, D. S., Wiegand, U. K., Norman, D. G., and Glover, D. M. (2000). Mutual Correction of Faulty PCNA Subunits in Temperature-Sensitive Lethal mus209 Mutants of Drosophila melanogaster Genetics *154*, 1721-1733.

Henneke, G., Koundrioukoff, S., and Hübscher, U. (2003). Phosphorylation of human Fen1 by cyclin-dependent kinase modulates its role in replication fork regulation. Oncogene *22*, 4301-4313.

Hermann, A., Gowher, H., and Jeltsch, A. (2004). Biochemistry and biology of mammalian DNA methyltransferases. Cellular and Molecular Life Sciences (CMLS) *61*, 2571-2587.

Hirvikoski, P., Kellokoski, J. K., Kumpulainen, E. J., Virtaniemi, J. A., Johansson, R. T., and Kosma, V. M. (1999). Downregulation of p21/WAF1 is related to advanced and dedifferentiated laryngeal squamous cell carcinoma. British Medical Journal *52*, 440.

Hoege, C., Pfander, B., Moldovan, G. L., Pyrowolakis, G., and Jentsch, S. (2002). RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. Nature *419*, 135-141.

Iida, T., Suetake, I., Tajima, S., Morioka, H., Ohta, S., Obuse, C., and Tsurimoto, T. (2002). PCNA clamp facilitates action of DNA cytosine methyltransferase 1 on hemimethylated DNA. Genes to Cells *7*, 997-1007.

Ink, B., Zornig, M., Baum, B., Hajibagheri, N., James, C., Chittenden, T., and Evan, G. (1997). Human Bak induces cell death in Schizosaccharomyces pombe with morphological changes similar to those with apoptosis in mammalian cells. Mol Cell Biol *17*, 2468-2474.

Jaenicke, R., and Bohm, G. (1998). The stability of proteins in extreme environments Current opinion in structural biology *8*, 738-748.

Jiricny, J. (2006). The multifaceted mismatch-repair system. Nat Rev Mol Cell Biol.

Jones, S., and Thornton, J. M. (1996). Principles of protein-protein interactions. Proceedings of the National Academy of Sciences *93*, 13-20.

Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991). Improved Methods for Building Protein Models in Electron Density Maps and the Location of Errors in these Models. Acta Cryst *47*, 110-119.

Jonsson, Z. O., Hindges, R., and Hubscher, U. (1998). Regulation of DNA replication and repair proteins through interaction with the front side of proliferating cell nuclear antigen. Embo Journal *17*, 2412-2425.

Jonsson, Z. O., and Hubscher, U. (1997). Proliferating cell nuclear antigen: more than a clamp for DNA polymerases. Bioessays *19*, 967-975.

Jonsson, Z. O., Podust, V. N., Podust, L. M., and Hubscher, U. (1995). Tyrosine-114 Is Essential for the Trimeric Structure and the Functional Activities of Human Proliferating Cell Nuclear Antigen. Embo Journal *14*, 5745-5751.

Jurgensmeier, J. M., Krajewski, S., Armstrong, R. C., Wilson, G. M., Oltersdorf, T., Fritz, L. C., Reed, J. C., and Ottilie, S. (1997). Bax-and Bak-induced cell death in the fission yeast Schizosaccharomyces pombe. Molecular Biology of the Cell *8*, 325-339.

Kantardjieff, K. A., and Rupp, B. (2003). Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. Protein Science *12*, 1865.

Kastan, M. B., Zhan, Q., El-Deiry, W. S., Carrier, F., Jacks, T., Walsh, W. V., Plunkett, B. S., Vogelstein, B., and Fornace Jr, A. J. (1992). A mammalian cell cycle checkpoint pathway utilizing p53 and GADD45 is defective in ataxia-telangiectasia. Cell *71*, 587-597.

Kazmirski, S. L., Zhao, Y. X., Bowman, G. D., O'Donnell, M., and Kuriyan, J. (2005). Out-of-plane motions in open sliding clamps: Molecular dynamics simulations of eukaryotic and archaeal proliferating cell nuclear antigen. Proceedings of the National Academy of Sciences of the United States of America *102*, 13801-13806.

Kearsey, J. M., Coates, P. J., Prescott, A. R., Warbrick, E., and Hall, P. A. (1995). Gadd45 is a nuclear cell cycle regulated protein which interacts with p21Cip1. Oncogene *11*, 1675-1683.

Kelley, L. A., MacCallum, R. M., and Sternberg, M. J. E. (2000). Enhanced genome annotation using structural profiles in the program 3D-PSSM. Journal of Molecular Biology *299*, 501-522.

Kelman, Z., and Odonnell, M. (1995a). Structural and Functional Similarities of Prokaryotic and Eukaryotic DNA-Polymerase Sliding Clamps. Nucleic Acids Research *23*, 3613-3620. Corrigendum: 3623, 4938.

Kelman, Z., and Odonnell, M. (1995b). Embryonic Pcna - a Missing Link. Current Biology *5*, 814-814.

Ko, R., and Bennett, S. E. (2005). Physical and functional interaction of human nuclear uracil-DNA glycosylase with proliferating cell nuclear antigen. DNA Repair *4*, 1421-1431.

Kong, X. P., Onrust, R., Odonnell, M., and Kuriyan, J. (1992). 3-Dimensional Structure of the Beta-Subunit of Escherichia-Coli DNA Polymerase-Iii Holoenzyme - a Sliding DNA Clamp. Cell *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. (2005). 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 *102*, 1871-1876.

Kovacevic, M., Diligenti, M., Gori, M., and Milutinovic, V. (2002). Recognition of Common Areas in a Web Page Using Visual Information: a possible application in a page classification. Proceedings of the 2002 IEEE International Conference on Data Mining (ICDM'02), 250.

Kovalsky, O., Lung, F. D. T., Roller, P. P., and Fornace, A. J. (2001). Oligomerization of Human Gadd45a Protein. Journal of Biological Chemistry *276*, 39330-39339.

Krawitz, D. C., Kama, T., and Kaufman, P. D. (2002). Chromatin Assembly Factor I Mutants Defective for PCNA Binding Require Asf1/Hir Proteins for Silencing. Molecular and Cellular Biology *22*, 614.

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

Kuriyan, J., and Odonnell, M. (1993). Sliding Clamps of DNA-Polymerases. Journal of Molecular Biology *234*, 915-925.

Kuwano, H., Saeki, H., Kawaguchi, H., Sonoda, K., Kitamura, K., Nakashima, H., Toh, Y., and Sugimachi, K. (1998). Proliferative activity of cancer cells in front and center areas of carcinoma in situ and invasive sites of esophageal squamous-cell carcinoma. Int J Cancer *78*, 149-152.

Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. Journal of Applied Crystallography *26*, 283-291.

Leibovici, M., Gusse, M., Bravo, R., and Mechali, M. (1990). Characterization and Developmental Expression of Xenopus Proliferating Cell Nuclear Antigen (Pcna). Developmental Biology *141*, 183-192.

Lepock, J. R. (2003). Cellular effects of hyperthermia: relevance to the minimum dose for thermal damage. International Journal of Hyperthermia *19*, 252-266.

Leslie, A. G. W. (1992). Recent changes to the MOSFLM package for processing film and image plate data. Joint CCP4+ ESF-EAMCB Newsletter on Protein Crystallography *26*, 11-20.

Lipinski, C. A. (2000). Drug-like properties and the causes of poor solubility and poor permeability. Journal of Pharmacological and Toxicological Methods *44*, 235-249.

Lipinski, C. A., Lombardo, F., Dominy, B. W., and Feeney, P. J. (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Advanced Drug Delivery Reviews *46*, 3-26.

Lo, M. C., Aulabaugh, A., Jin, G., Cowling, R., Bard, J., Malamas, M., and Ellestad, G. (2004). Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. Analytical Biochemistry *332*, 153-159.

Ludovico, P., Sousa, M. J., Silva, M. T., Lea~ o, C., Co, and rte-Real, M. (2001). Saccharomyces cerevisiae commits to a programmed cell death process in response to acetic acid (Soc General Microbiol), pp. 2409-2415.

Lyall, M. S., Dundas, S. R., Curran, S., and Murray, G. I. (2006). Profiling Markers of Prognosis in Colorectal Cancer. Clinical Cancer Research *12*, 1184.

Madeo, F., Frohlich, E., and Frohlich, K. U. (1997). A Yeast Mutant Showing Diagnostic Markers of Early and Late Apoptosis. The Journal of Cell Biology *139*, 729-734.

Maga, G., and Hubscher, U. (2003). Proliferating cell nuclear antigen (PCNA): a dancer with many partners. Journal of Cell Science *116*, 3051-3060.

Mann, T. L., and Krull, U. J. (2003). Fluorescence polarization spectroscopy in protein analysis. Analyst *128*, 313-317.

Martin, I. V., MacNeill, S. A., and Journals, O. (2004). Functional analysis of subcellular localization and protein–protein interaction sequences in the essential DNA ligase I protein of fission yeast. Nucleic Acids Research *32*, 632-642.

Matsumiya, S., Ishino, S., Ishino, Y., and Morikawa, K. (2002). Physical interaction between proliferating cell nuclear antigen and replication factor C from Pyrococcus furiosus. Genes to Cells *7*, 911-922.

Matsumiya, S., Ishino, Y., and Morikawa, K. (2001). Crystal structure of an archaeal DNA sliding clamp: Proliferating cell nuclear antigen from Pyrococcus furiosus. Protein Science *10*, 17-23.

Millar, D., Trakselis, M. A., and Benkovic, S. J. (2004). On the solution structure of the T4 sliding clamp (gp45). Biochemistry *43*, 12723-12727.

Miyata, T., Suzuki, H., Oyama, T., Mayanagi, K., Ishino, Y., and Morikawa, K. (2005). Open clamp structure in the clamp-loading complex visualized by electron microscopic image analysis. Proceedings of the National Academy of Sciences of the United States of America *102*, 13795-13800.

Moarefi, I., Jeruzalmi, D., Turner, J., O'Donnell, M., and Kuriyan, J. (2000). Crystal structure of the DNA polymerase processivity factor of T4 bacteriophage. Journal of Molecular Biology *296*, 1215-1223.

Modrich, P. (2006). Mechanisms in Eukaryotic Mismatch Repair. Journal of Biological Chemistry *281*, 30305.

Moggs, J. G., Grandi, P., Quivy, J. P., Jonsson, Z. O., Hubscher, U., Becker, P. B., and Almouzni, G. (2000). A CAF-1-PCNA-mediated chromatin assembly pathway triggered by sensing DNA damage. Molecular and Cellular Biology *20*, 1206-1218.

Moldovan, G. L., Pfander, B., and Jentsch, S. (2006). PCNA Controls Establishment of Sister Chromatid Cohesion during S Phase. Molecular Cell *23*, 723-732.

Mortusewicz, O., Schermelleh, L., Walter, J., Cardoso, M. C., and Leonhardt, H. (2005). Recruitment of DNA methyltransferase I to DNA repair sites. Proceedings of the National Academy of Sciences of the United States of America *102*, 8905.

Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997). Refinement of protein structures by the maximum likelihood method. Acta Crystallogr D *53*, 240-255.

Naktinis, V., Turner, J., and Odonnell, M. (1996). Molecular switch in a replication machine defined by an internal competition for protein rings. Cell *84*, 137-145.

Naryzhny, S. N., Zhao, H., and Lee, H. (2005). Proliferating cell nuclear antigen (PCNA) may function as a double homotrimer complex in the mammalian cell. Journal of Biological Chemistry *280*, 13888-13894.

Odonnell, M., Onrust, R., Dean, F. B., Chen, M., and Hurwitz, J. (1993). Homology in Accessory Proteins of Replicative Polymerases - Escherichia-Coli to Humans. Nucleic Acids Research *21*, 1-3.

Oku, T., Ikeda, S., Sasaki, H., Fukuda, K., Morioka, H., Ohtsuka, E., Yoshikawa, H., and Tsurimoto, T. (1998). Functional sites of human PCNA which interact with p21 (Cip1/Waf1), DNA polymerase delta and replication factor C. Genes to Cells *3*, 357-369.

Oyama, M., Wakasugi, M., Hama, T., Hashidume, H., Iwakami, Y., Imai, R., Hoshino, S., Morioka, H., Ishigaki, Y., and Nikaido, O. (2004). Human NTH1 physically interacts with p53 and proliferating cell nuclear antigen. Biochemical and Biophysical Research Communications *321*, 183-191.

Painter, J., and Merritt, E. A. (2006). TLSMD web server for the generation of multi-group TLS models. J Appl Crystallogr *39*, 109-111.

Papouli, E., Chen, S., Davies, A. A., Huttner, D., Krejci, L., Sung, P., and Ulrich, H. D. (2005). Crosstalk between SUMO and Ubiquitin on PCNA Is Mediated by Recruitment of the Helicase Srs2p. Molecular Cell *19*, 123-133.

Parker, G. J., Law, T. L., Lenoch, F. J., and Bolger, R. E. (2000). Development of High Throughput Screening Assays Using Fluorescence Polarization: Nuclear Receptor-Ligand-Binding and Kinase/Phosphatase Assays. Journal of Biomolecular Screening *5*, 77.

Parrilla-Castellar, E. R., Arlander, S. J. H., and Kamitz, L. (2004). Dial 9-1-1 for DNA damage: the RaO-Hus1-Rad1 (9-1-1) clamp complex. DNA Repair *3*, 1009-1014.

Pascal, J. M., O'Brien, P. J., Tomkinson, A. E., and Ellenberger, T. (2004). Human DNA ligase I completely encircles and partially unwinds nicked DNA. Nature *432*, 473-478.

Paunesku, T., Mittal, S., Protic, M., Oryhon, J., Korolev, S. V., Joachimiak, A., and Woloschak, G. E. (2001). Proliferating cell nuclear antigen (PCNA): ringmaster of the genome. International Journal of Radiation Biology *77*, 1007-1021.

Perrin, M. F. (1926). Polarisation de la lumiére de fluorescence. Vie moyenne des molécules dans l'état excité. Journal de Physique et le Radium *7*, 390-401.

Pfander, B., Moldovan, G. L., Sacher, M., Hoege, C., and Jentsch, S. (2005). SUMO-modified PCNA recruits Srs2 to prevent recombination during S phase. Nature *436*, 428-433.

Podust, V. N., Podust, L. M., Muller, F., and Hubscher, U. (1995). DNA Polymerase-Delta Holoenzyme - Action on Single-Stranded-DNA and on Double-Stranded DNA in the Presence of Replicative DNA Helicases. Biochemistry *34*, 5003-5010.

Privalov, P. L. (1990). Cold denaturation of proteins. Critical Reviews in Biochemistry and Molecular Biology *25*, 281-305.

Qin, J., and Li, L. (2003). Molecular Anatomy of the DNA Damage and Replication Checkpoints. Radiation Research *159*, 139-148.

Reynolds, N., Warbrick, E., Fantes, P. A., and MacNeill, S. A. (2000). Essential interaction between the fission yeast DNA polymerase delta subunit Cdc27 and Pcn1 (PCNA) mediated through a C-terminal p21(Cip1)-like PCNA binding motif. Embo Journal *19*, 1108-1118.

Rolig, R. L., and McKinnon, P. J. (2000). Linking DNA damage and neurodegeneration. Trends in Neurosciences *23*, 417-424.

Roos, G., Landberg, G., Huff, J. P., Houghten, R., Takasaki, Y., and Tan, E. M. (1993). Analysis of the Epitopes of Proliferating Cell Nuclear Antigen Recognized by Monoclonal-Antibodies. Laboratory Investigation *68*, 204-210.

Rost, B., Yachdav, G., and Liu, J. (2004). The PredictProtein server. Nucleic Acids Research *32*, W321-W326.

Rozema, D. B., and Poulter, C. D. (1999). Yeast protein farnesyltransferase. pK a s of peptide substrates bound as zinc thiolates. Biochemistry *38*, 13138-13146.

Ryser, S., Vial, E., Magnenat, E., Schlegel, W., and Maundrell, K. (1999). Reconstitution of caspase-mediated cell-death signalling in Schizosaccharomyces pombe. Current Genetics *36*, 21-28.

Sakakura, C., Hagiwara, A., Tsujimoto, H., Ozaki, K., Sakakibara, T., Oyama, T., Ogaki, M., and Takahashi, T. (1994). Inhibition of Gastric-Cancer Cell-Proliferation by Antisense Oligonucleotides Targeting the Messenger-Rna Encoding Proliferating Cell Nuclear Antigen. British Journal of Cancer *70*, 1060-1066.

Sancar, A., Lindsey-Boltz, L. A., Unsal-Kacmaz, K., and Linn, S. (2004). Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. Annual Review of Biochemistry *73*, 39-85.

Schrag, J. D., Jiralerspong, S., Banville, M., Jaramillo, M. L., and O'Connor-McCourt, M. D. (2008). The crystal structure and dimerization interface of GADD45 {gamma}. Proceedings of the National Academy of Sciences *105*, 6566.

Scott, M., Bonnefin, P., Vieyra, D., Boisvert, F. M., Young, D., Bazett-Jones, D. P., and Riabowol, K. (2001). UV-induced binding of ING1 to PCNA regulates the induction of apoptosis. Journal of Cell Science *114*, 3455-3462.

Selvakumaran, M., Lin, H. K., Sjin, R. T., Reed, J. C., Liebermann, D. A., and Hoffman, B. (1994a). The novel primary response gene MyD118 and the proto-oncogenes myb, myc, and bcl-2 modulate transforming growth factor beta 1-induced apoptosis of myeloid leukemia cells. Molecular and Cellular Biology *14*, 2352-2360.

Shamoo, Y., and Steitz, T. A. (1999). Building a replisome from interacting pieces: Sliding clamp complexed to a peptide from DNA polymerase and a polymerase editing complex. Cell *99*, 155-166.

Shibahara, K., and Stillman, B. (1999). Replication-dependent marking of DNA by PCNA facilitates CAF-1-coupled inheritance of chromatin. Cell *96*, 575-585.

Shin, D. M., Voravud, N., Ro, J. Y., Lee, J. S., Hong, W. K., and Hittelman, W. N. (1993). Sequential Increases in Proliferating Cell Nuclear Antigen Expression in Head and Neck Tumorigenesis: A Potential Biomarker. Journal of the National Cancer Institute *85*, 971-978.

Skaliter, R., Bergstein, M., and Livneh, Z. (1996). beta*, a UV-inducible shorter form of the beta subunit of DNA polymerase III of Escherichia coli .2. Overproduction, purification, and activity as a polymerase processivity clamp. Journal of Biological Chemistry *271*, 2491-2496.

Smith, M. L., Chen, I. T., Zhan, Q., Bae, I., Chen, C. Y., Gilmer, T. M., Kastan, M. B., O'Connor, P. M., and Fornace Jr, A. J. (1994). Interaction of the p53-regulated protein Gadd45 with proliferating cell nuclear antigen. Science *266*, 1376.

Smith, S., and Stillman, B. (1989). Purification and Characterization of Caf-I, a Human Cell Factor Required for Chromatin Assembly During DNA-Replication Invitro. Cell *58*, 15-25.

Stahl, M., Bur, D., and Schneider, G. (1999). Mapping of proteinase active sites by projection of surface-derived correlation vectors. Journal of Computational Chemistry *20*, 336-347.

Stukenberg, P. T., Studwellvaughan, P. S., and Odonnell, M. (1991). Mechanism of the Sliding Beta-Clamp of DNA Polymerase-Iii Holoenzyme. Journal of Biological Chemistry *266*, 11328-11334.

Takekawa, M., and Saito, H. (1998). A family of stress-inducible GADD45-like proteins mediate activation of the stress-responsive MTK1/MEKK4 MAPKKK. Cell *95*, 521-530.

Tanaka, H., Ryu, G. H., Seo, Y. S., and MacNeill, S. A. (2004). Genetics of lagging strand DNA synthesis and maturation in fission yeast: suppression analysis links the Dna2-Cdc24 complex to DNA polymerase {delta}. Nucleic Acids Research *32*, 6367.

Tao, W., Kurschner, C., and Morgan, J. I. (1997). Modulation of Cell Death in Yeast by the Bcl-2 Family of Proteins. Journal of Biological Chemistry *272*, 15547-15552.

Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res *22*, 4673-4680.

Tinker, R. L., Kassavetis, G. A., and Geiduschek, E. P. (1994). Detecting the ability of viral, bacterial and eukaryotic replication proteins to track along DNA. EMBO Journal *13*, 5330-5337.

Tota, M. R., Daniel, S., Sirotina, A., Mazina, K. E., Fong, T. M., Longmore, J., and Strader, C. D. (1994). Characterization of a Fluorescent Substance P Analog. Biochemistry *33*, 13079-13086.

Tsuchimoto, D., Sakai, Y., Sakumi, K., Nishioka, K., Sasaki, M., Fujiwara, T., Nakabeppu, Y., and Journals, O. (2001). Human APE2 protein is mostly localized in the nuclei and to some extent in the mitochondria, while nuclear APE2 is partly associated with proliferating cell nuclear antigen. Nucleic Acids Research *29*, 2349-2360.

Tsurimoto, T. (1998). PCNA, a multifunctional ring on DNA. Biochimica et biophysica acta *1443*, 23-39.

Turek, T. C., Small, E. C., Bryant, R. W., Adam, W., and Hill, G. (2001). Development and Validation of a Competitive AKT SerineThreonine Kinase Fluorescence Polarization Assay Using a Product-Specific Anti-phospho-serine Antibody. Analytical Biochemistry *299*, 45-53.

Uberti, D., Ferrari Toninelli, G., and Memo, M. (2003). Involvement of DNA damage and repair systems in neurodegenerative process. Toxicology Letters *139*, 99-105.

Umar, A., Buermeyer, A. B., Simon, J. A., Thomas, D. C., Clark, A. B., Liskay, R. M., and Kunkel, T. A. (1996). Requirement for PCNA in DNA mismatch repair at a step preceding DNA resynthesis. Cell *87*, 65-73.

Unk, I., Haracska, L., Gomes, X. V., Burgers, P. M. J., Prakash, L., and Prakash, S. (2002). Stimulation of 3' to 5' Exonuclease and 3'-Phosphodiesterase Activities of Yeast Apn2 by Proliferating Cell Nuclear Antigen. Molecular and Cellular Biology *22*, 6480.

Vagin, A., and Teplyakov, A. (1997). MOLREP: an Automated Program for Molecular Replacement. Journal of Applied Crystallography *30*, 1022-1025.

Vairapandi, M., Azam, N., Balliet, A. G., Hoffman, B., and Liebermann, D. A. (2000). Characterization of MyD118, Gadd45, and Proliferating Cell Nuclear Antigen (PCNA) Interacting Domains PCNA IMPEDES MyD118 AND Gadd45-MEDIATED NEGATIVE GROWTH CONTROL. Journal of Biological Chemistry *275*, 16810-16819.

Vairapandi, M., Balliet, A. G., Fornace Jr, A. J., Hoffman, B., and Liebermann, D. A. (1996). The differentiation primary response gene MyD118, related to GADD45, encodes for a nuclear protein which interacts with PCNA and p21WAF1/CIP1. Oncogene *12*, 2579-2594.

Vassilev, L. T., Vu, B. T., Graves, B., Carvajal, D., Podlaski, F., Filipovic, Z., Kong, N., Kammlott, U., Lukacs, C., and Klein, C. (2004). In Vivo Activation of the p53 Pathway by Small-Molecule Antagonists of MDM2. Science *303*, 844.

Veis, A., Tompkins, K., Alvares, K., Wei, K., Wang, L., Wang, X. S., Brownell, A. G., Jengh, S. M., and Healy, K. E. (2000). Specific Amelogenin Gene Splice Products Have Signaling Effects on Cells in Culture and in Implants in Vivo. Journal of Biological Chemistry *275*, 41263-41272.

Vogt, G., and Argos, P. (1997). Protein thermal stability: hydrogen bonds or internal packing? Folding & Design *2*, S40-S46.

Waga, S., Hannon, G. J., Beach, D., and Stillman, B. (1994). The p 21 inhibitor of cyclin-dependent kinases controls DNA replication by interaction with PCNA. Nature *369*, 574-578.

Waldron, T. T., and Murphy, K. P. (2003). Stabilization of proteins by ligand binding: application to drug screening and determination of unfolding energetics. Biochemistry *42*, 5058-5064.

Wang, S. C., Nakajima, Y., Yu, Y. L., Xia, W., Chen, C. T., Yang, C. C., McIntush, E. W., Li, L. Y., Hawke, D. H., and Kobayashi, R. (2006). Tyrosine phosphorylation controls PCNA function through protein stability. Nature Cell Biology *8*, 1359-1368.

Wang, X., Gorospe, M., and Holbrook, N. J. (1999a). gadd45 Is Not Required for Activation of c-Jun N-terminal Kinase or p38 during Acute Stress (ASBMB).

Wang, X. W., Zhan, Q., Coursen, J. D., Khan, M. A., Kontny, H. U., Yu, L., Hollander, M. C., O'Connor, P. M., Fornace, A. J., and Harris, C. C. (1999b). GADD45 induction of a G2/M cell cycle checkpoint (National Acad Sciences), pp. 3706-3711.

Warbrick, E. (2000). The puzzle of PCNA's many partners. Bioessays *22*, 997-1006.

Warbrick, E., Lane, D. P., Glover, D. M., and Cox, L. S. (1995). A Small Peptide Inhibitor of DNA-Replication Defines the Site of Interaction between the Cyclin-Dependent Kinase Inhibitor P21(Waf1) and Proliferating Cell Nuclear Antigen. Current Biology *5*, 275-282.

Warbrick, E., Lane, D. P., Glover, D. M., and Cox, L. S. (1997). Homologous regions of Fen 1 and p 21 Cip 1 compete for binding to the same site on PCNA: a potential mechanism to co-ordinate DNA replication and repair. Oncogene *14*, 2313-2321.

Warren, A. J. (2002). Eukaryotic transcription factors. Current Opinion in Structural Biology *12*, 107-114.

Watanabe, K., Morishita, J., Umezu, K., Shirahige, K., and Maki, H. (2002). Involvement of RAD9-Dependent Damage Checkpoint Control in Arrest of Cell Cycle, Induction of Cell Death, and Chromosome Instability Caused by Defects in Origin Recognition Complex in Saccharomyces cerevisiae. Eukaryotic Cell *1*, 200.

Wear, M. A., and Walkinshaw, M. D. (2006). Thermodynamics of the cyclophilin-A/cyclosporin-A interaction: A direct comparison of parameters determined by surface plasmon resonance using Biacore T100 and isothermal titration calorimetry. Analytical Biochemistry *359*, 285-287.

Wood, V., Gwilliam, R., Rajandream, M. A., Lyne, M., Lyne, R., Stewart, A., Sgouros, J., Peat, N., Hayles, J., and Baker, S. (2002). The genome sequence of Schizosaccharomycespombe. Nature *415*, 871-880.

Wu, S. Y., McNae, I., Kontopidis, G., McClue, S. J., McInnes, C., Stewart, K. J., Wang, S., Zheleva, D. I., Marriage, H., and Lane, D. P. (2003). Discovery of a novel family of CDK inhibitors with the program LIDAEUS: structural basis for ligand-induced disordering of the activation loop. Structure *11*, 1537-1546.

Yamasawa, K., Nio, Y., Dong, M., Yamaguchi, K., and Itakura, M. (2002). Clinicopathological Significance of Abnormalities in Gadd45 Expression and Its Relationship to p53 in Human Pancreatic Cancer (AACR), pp. 2563-2569.

Yao, N., Turner, J., Kelman, Z., Stukenberg, P. T., Dean, F., Shechter, D., Pan, Z. Q., Hurwitz, J., and Odonnell, M. (1996). Clamp loading, unloading and intrinsic stability of the PCNA, beta and gp45 sliding clamps of human, E-coli and T4 replicases. Genes to Cells *1*, 101-113.

Zhan, Q., Antinore, M. J., Wang, X. W., Carrier, F., Smith, M. L., Harris, C. C., and Fornace Jr, A. J. (1999). Association with Cdc2 and inhibition of Cdc2/Cyclin B1 kinase activity by the p53-regulated protein Gadd45. Oncogene *18*, 2892-2900.

Zhan, Q., Lord, K. A., Alamo, I., Hollander, M. C., Carrier, F., Ron, D., Kohn, K. W., Hoffman, B., Liebermann, D. A., and Fornace, A. J. (1994). The gadd and MyD genes define a novel set of mammalian genes encoding acidic proteins that synergistically suppress cell growth. Molecular and Cellular Biology *14*, 2361-2371.

Zhang, P., Sun, Y. B., Hsu, H., Zhang, L. F., Zhang, Y. N., and Lee, M. (1998). The interdomain connector loop of human PCNA is involved in a direct interaction with human polymerase delta. Journal of Biological Chemistry *273*, 713-719.

Zhang, W., Bae, I., Krishnaraju, K., Azam, N., Fan, W., Smith, K., Hoffman, B., and Liebermann, D. A. (1999). CR6: A third member in the MyD118 and Gadd45 gene family which functions in negative growth control. Oncogene *18*, 4899-4907.

Zhang, Z. G., Shibahara, K., and Stillman, B. (2000). PCNA connects DNA replication to epigenetic inheritance in yeast. Nature *408*, 221-225.

Zheleva, D. I., Zhelev, N. Z., Fischer, P. M., Duff, S. V., Warbrick, E., Blake, D. G., and Lane, D. P. (2000). A quantitative study of the in vitro binding of the C-terminal domain of p21 to PCNA: Affinity, stoichiometry, and thermodynamics. Biochemistry *39*, 7388-7397.