eman ta zabal zazu

Universidad
del País Vasco

Euskal Herriko
Unibertsitatea

Molecular recognition between PCNA and proteins involved in DNA replication and repair: p15 and p12

Amaia González Magaña

2019

Supervisor: Dr. Francisco J. Blanco

Resumen

El antígeno nuclear de células en proliferación (PCNA) es un factor esencial en replicación y reparación de ADN en eucariotas. PCNA forma un anillo homotrimérico que rodea y se desliza a lo largo del ADN, actuando como una plataforma de anclaje de enzimas que modifican el ADN y de proteínas reguladoras a la horquilla de replicación. Estas interacciones ocurren frecuentemente a través de una secuencia consenso de aminoácidos denominada motivo PIP que se une a un bolsillo hidrofóbico de PCNA. En esta tesis doctoral se estudian proteínas de unión a PCNA implicadas en la replicación y reparación del ADN, con el fin último de entender sus interacciones y función utilizando una aproximación estructural.

Se ha realizado un exhaustivo análisis conformacional de la proteína p15^{PAF} doblemente monoubiquitinada y el impacto de esta modificación postraduccional en su unión a PCNA y a ADN. p15^{PAF} es una proteína intrínsecamente desordenada que regula la reparación de ADN durante la replicación celular. El gen de p15^{PAF} está sobreexpresado en varios tipos de tumores y su función está regulada por la monoubiquitinación simultánea en sus lisinas 15 y 24. La caracterización estructural de p15^{PAF} di-monoubiquitinada muestra que se mantiene monomérica y desordenada en solución y es capaz de unir PCNA del mismo modo que su forma no ubiquitinada, aunque interacciona con el ADN con menor afinidad. Así mismo, los datos de SAXS indican la formación transitoria de dímeros de ubiquitina, lo que sugiere que p15^{PAF} en su forma doblemente ubiquitinada es una diana de unión de la DNA metil-transferasa Dnmt1. Los datos de ITC confirman está hipótesis, sugiriendo que p15^{PAF} tiene una función en el mantenimiento de la metilación del DNA durante la replicación cromosómica a través del reclutamiento de Dnmt1.

También ha caracterizado la interacción de la subunidad p12 de la polimerasa delta (pol δ) humana con PCNA mediante RMN, calorimetría y cristalografía. La polimerasa delta es esencial para la replicación de la hebra retardada y la maduración de los fragmentos de Okazaki en eucariotas. Está formada por cuatro subunidades: la subunidad catalítica p125, con actividad polimerasa y exonucleasa, y tres subunidades reguladoras: p50, p66 y p12. En eucariotas superiores la cuarta subunidad juega un papel fundamental en el mantenimiento de la estabilidad genómica. A pesar de la relevancia central de pol δ en la replicación del ADN, se conoce muy poco sobre su estructura y el modo de interacción las diferentes subunidades entre ellas, así como con PCNA. La estructura cristalográfica de PCNA en complejo con un

fragmento de p12 revela que se une a través de un motivo PIP muy divergente de la secuencia consenso. Los datos indican que las subunidades p66 y p12 contribuyen al reconocimiento de PCNA con diferentes especificidades y afinidades. El motivo PIP divergente de p12 une PCNA con mayor afinidad que el motivo PIP canónico de la helicasa RecQ5 y se propone que la presencia de residuos cargados positivamente dentro de la secuencia PIP es desfavorable para la unión a PCNA.

Summary

Proliferating cell nuclear antigen (PCNA) is an essential factor for DNA replication and repair in eukaryotes. PCNA forms a homotrimeric ring that encircles and slides on the DNA, acting as an anchoring platform for DNA-editing enzymes and regulatory proteins to the replication fork. These interactions occur mainly through a canonical amino acid sequence known as PIP motif that binds to a hydrophobic pocket of PCNA. In this thesis, the interactions between PCNA and proteins involved in DNA replication and repair are structurally studied with the ultimate aim of understanding their functional role.

A conformational analysis of double monoubiquitinated p15^{PAF} is performed and the impact of this postranslational modification in its binding to PCNA and DNA is assessed . p15^{PAF} is an intrinsically disordered protein that regulates DNA repair during DNA replication. The p15^{PAF} gene is overexpressed in several types of cancer and its function is regulated by monoubiquitination at two lysines 15 and 24 simultaneously. The structural characterization of di-monoubiquitinated p15^{PAF} shows that it is monomeric, disordered and binds PCNA as the non-ubiquitinated form, but interacts with the DNA with lower affinity. SAXS analysis indicates that the ubiquitin moieties forms a transient dimer, suggesting that ubiquitinated p15^{PAF} is a binding partner of DNA methyl transferase Dnmt1. This hypothesis is confirmed by ITC data, suggesting that p15^{PAF} is involved in the maintenance of DNA methylation during chromosomal replication by recruitment of Dnmt1.

The interaction between p12 subunit of human polymerase delta (pol δ) and PCNA has been characterized by NMR, ITC and crystallography. Human pol δ is an essential enzyme for the replication of DNA lagging strand and the maturation of Okazaki fragments in eukaryotes. It consists of 4 subunits: a catalytic subunit p125, with polymerase and exonuclease activities, and three regulatory subunits: p50, p66, and p12. In high eukaryotes the fourth subunit of the polymerase plays an important role in the maintenance of genome stability. Despite the central relevance of pol δ in DNA replication, little is known about its structure and how it interacts with PCNA. The crystallographic structure of the PCNA-p12 peptide complex reveals that p12 binds through a highly divergent PIP motif. These data show that p12 and p66 subunits of pol δ contribute to the recognition of PCNA with different structural specificities and affinities. We also show that the p12 divergent PIP motif binds PCNA with higher affinity

than the canonical PIP box of RecQ5 helicase. It is proposed that the presence of positive charged residues inside the PIP sequence is unfavourable for PCNA binding.

INDEX

ABBREVIATIONS

ATP: Adenosine triphosphate

BAH: Bromo-adjacent homology

BSA: Bovine Serum Albumin

CD: Circular Dichroism

CS: Chemical Shift

CSP:Chemical Shifts Perturbations

Dmax: Maximum particle diameter

dmUbp15: Doubly monouiquitinated p15

DNA: Deoxyribonucleic acid.

DNMT: DNA methyl transferase

DSB: Double strand break

dsDNA: Double strand DNA

DSS: 4,4-dimethyl-4-silapentane-1-sulfonic acid

DTT: Dithiothreitol

EDTA: Ethylenediaminetetraacetic acid

FID: Free induction decay

HEPES: 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

HR: Homologous recombination

HSQC: Heteronuclear Single Quantum Coherence

IDCL: Interdomain connenting loop

IDP: Intrinsically disordered protein

IDR:Intrinsically disordered regions

IPTG: Isopropyl β-D-1-thiogalactopyranoside

IPTG: Isopropyl β-D-1-thiogalactopyranoside

Kd: Dissociation constant

LB: Luria-Bertani broth

MALDI-TOF: Matrix-Assisted Laser Desorption/Ionization- Time-Of-Flight

MALS: Multi Angle static Light Scattering

MD: Molecular Dynamics

NLS: N-terminal nuclear localization signal

NMR: Nuclear Magnetic Resonance

PAGE: Polyacrylamide gel electrophoresis

PBS: Phosphate-Buffered Saline

PCNA: Proliferating nuclear antigen

PDB: Protein Data Bank

PEG: Polyethylene glycol

PIP: PCNA interacting Protein.

RC: Random coil

RCI: Random Coil Index

RFC: Replication factor C

RFTS: Replication foci targeting sequence

Rg: Radius of gyration

RPA: Replication protein A

SAXS: Small angle X-ray Scattering

SDS: Sodium dodecyl sulphate

SEC: Size Exclusion Chromatography

ssDNA: Single strand DNA

TCEP: Tris(2-carboxyethyl)phosphine

TEV: Tobacco Etch Virus cysteine protease

TLS: Translesion synthesis

TMS: etramethylsilane

TRD: target recognition domain

TROSY: Transverse relaxation optimized spectroscopy

Ub: Ubiquitin

UV:Ultra-violet

NOE: Nuclear Overhauser effect

bp: base pair

SLiMs: small linear motifs

APC/C: Anaphase promoting complex/cyclosome

GST: Glutathione S-transferases

1.Introduction

1.1. DNA sliding clamps

DNA polymerases are the enzymes that replicate chromosomal DNA during the S-phase of the cell cycle. They can quickly polymerize thousands of nucleotides without detaching from the genomic template [1]. This fast and processive activity is conferred to polymerases by their association with multimeric ring shaped proteins known as DNA sliding clamps. They encircle and slide freely on the DNA, tethering replicative polymerases and other factors to the DNA duplex [2]. First evidence of sliding clamp structure was obtained from the polymerase III β subunit of *E. coli* Pol-III complex. Biochemical assays demonstrated that the β subunit interacts tightly with nicked circular plasmid while it dissociates when the DNA is linearized [3]. DNA Sliding clamps are loaded onto DNA by clamp loaders, evolutionarily conserved proteins of the AAA+ family of ATPases [4, 5]. They form pentameric complexes that bind and open the sliding clamp, placing it onto DNA 3' end called a primer/template-junction in an ATP dependent process [6].

Sliding clamps are functionally and structurally similar across all living organisms, including some viruses, and can assembly in homodimeric, homotrimeric or heterotrimeric rings, with the protomers binding each other in a head-to-tail way [7]. Despite the low sequence similarity between sliding clamps of different domains of life (Figure 1.1), they all adopt a three-dimensional pseudo-six fold symmetry structure consisting of an outer layer of 6 β-sheets and an inner layer of 12 α-helices facing the central channel [8, 9].

The bacterial β clamp is a homodimeric ring that comprises two protomers [10], each one with three topologically similar domains (Figure 1.2). In contrast, the functional equivalents of β clamp in T4 bacteriophage (gene 45 protein, gp45) and PCNA (Proliferating cell nuclear antigen) in eukaryotes and archaea assemble in a homotrimeric ring, in which each protomer contains two similar domains connected by an interdomain-connecting loop (IDCL) [11, 12].

In the T4 bacteriophage ring, the two domains are less similar between each other and consequently gp45 clamp has a triangular appearance instead of the hexagonal shape of the others.
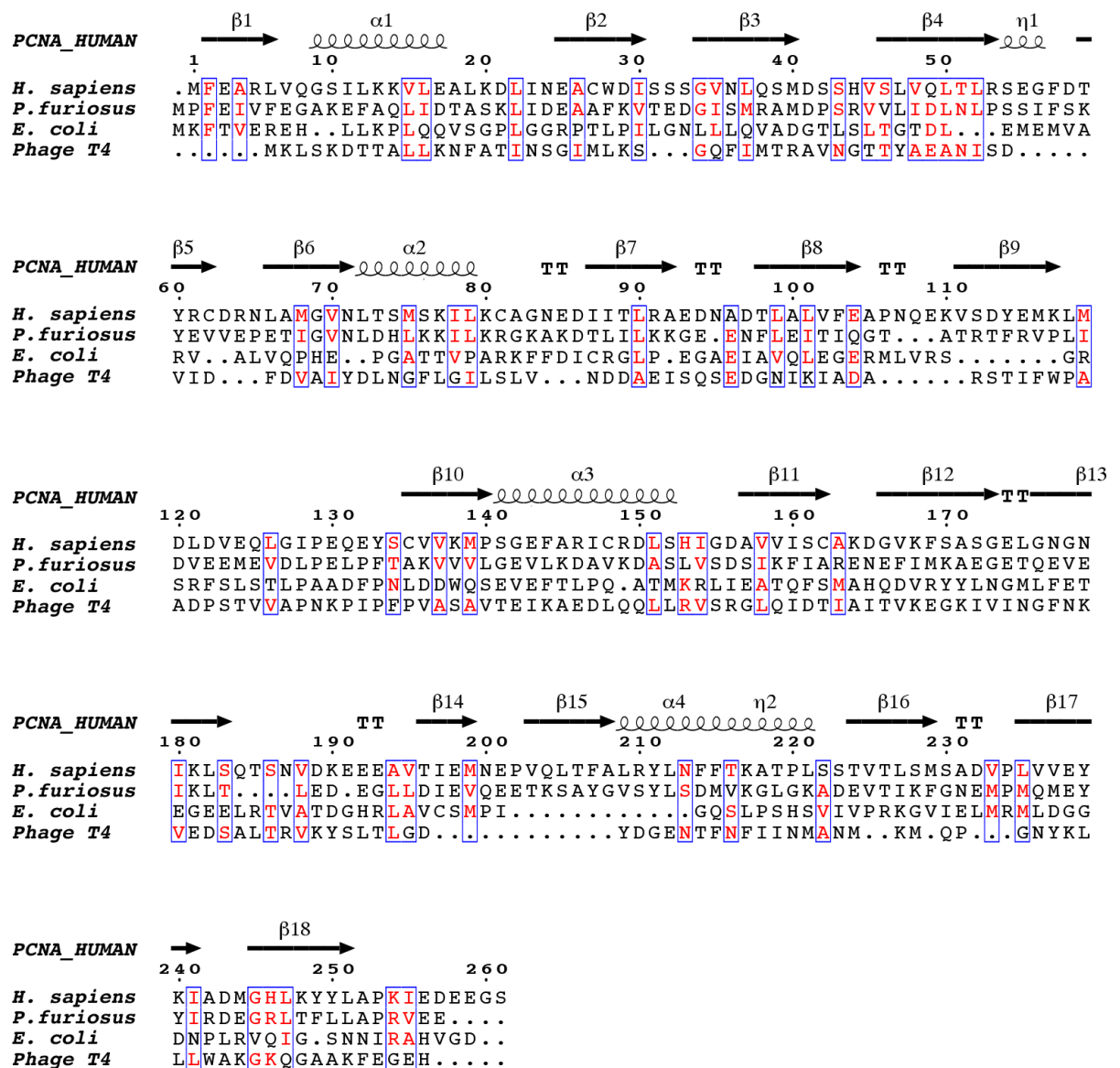
```
                β1           α1                    β2          β3          β4      η1
PCNA_HUMAN  ────────▶  ℓℓℓℓℓℓℓℓℓℓℓ            ───────▶    ───────▶    ───────▶  ℓℓℓ   ─
              1          10        20          30          40          50

H. sapiens  .MFEARLVQGSILKKVLEALKDLINEACWDISSSGVNLQSMDSSHVSLVQLTLRSEGFDT
P.furiosus  MPFEIVFEGAKEFAQLIDTASKLIDEAAFKVTEDGISMRAMDPSRVVLIDLNLPSSIFSK
E. coli     MKFTVEREH..LLKPLQQVSGPLGGRPTLPILGNLLLQVADGTLSLTGTDL...EMEMVA
Phage T4    ...............MKLSKDTTALLKNFATINSGIMLKS...GQFIMTRAVNGTTYAEANISD.....


                β5          β6          α2              β7          β8          β9
PCNA_HUMAN  ────▶   ──────────▶  ℓℓℓℓℓℓℓℓ      TT  ───────▶  TT  ──────▶  TT  ──────▶
              60        70         80            90         100         110

H. sapiens  YRCDRNLAMGVNLTSMSKILKCAGNEDIITLRAEDNADTLALVFEAPNQEKVSDYEMKLM
P.furiosus  YEVVEPETIGVNLDHLKKILKRGKAKDTLILKKGE.ENFLEITIQGT...ATRTFRVPLI
E. coli     RV..ALVQPHE..PGATTVPARKFFDICRGLP.EGAEIAVQLEGERMLVRS.......GR
Phage T4    VID...FDVAIYDLNGFLGILSLV...NDDAEISQSEDGNIKIADA......RSTIFWPA


                         β10          α3              β11         β12          β13
PCNA_HUMAN        ──────────▶  ℓℓℓℓℓℓℓℓℓℓℓℓℓℓ      ──────▶    ──────▶ TT─        ▶
              120        130         140         150         160         170

H. sapiens  DLDVEQLGIPEQEYSCVVKMPSGEFARICRDLSHIGDAVVISCAKDGVKFSASGELGNGN
P.furiosus  DVEEMEVDLPELPFTAKVVVLGEVLKDAVKDASLVSDSIKFIARENEFIMKAEGETQEVE
E. coli     SRFSLSTLPAADFPNLDDWQSEVEFTLPQ.ATMKRLIEATQFSMAHQDVRYYLNGMLFET
Phage T4    ADPSTVVAPNKPIPFPVASAVTEIKAEDLQQLLRVSRGLQIDTIAITVKEGKIVINGFNK


                          β14         β15         α4      η2      β16         β17
PCNA_HUMAN  ──────▶  TT  ─────▶    ─────▶  ℓℓℓℓℓℓℓℓℓℓℓℓℓℓ    ──────▶  TT        ▶
              180        190         200         210         220         230

H. sapiens  IKLSQTSNVDKEEEAVTIEMNEPVQLTFALRYLNFFTKATPLSSTVTLSMSADVPLVVEY
P.furiosus  IKLT....LED.EGLLDIEVQEETKSAYGVSYLSDMVKGLGKADEVTIKFGNEMPMQMEY
E. coli     EGEELRTVATDGHRLAVCSMPI............GQSLPSHSVIVPRKGVIELMRMLDGG
Phage T4    VEDSALTRVKYSLTLGD............YDGENTFNFIINMANM..KM.QP...GNYKL


                β18
PCNA_HUMAN  ▶  ──────▶
              240        250         260

H. sapiens  KIADMGHLKYYLAPKIEDEEGS
P.furiosus  YIRDEGRLTFLLAPRVEE....
E. coli     DNPLRVQIG.SNNIRAHVGD..
Phage T4    LLWAKGKQGAAKFEGEH.....
```

Figure 1.1. Structure-based sequence alignment of DNA sliding clamps from *H. sapiens*, *P. furiosus, E. coli* and gp45 gene of bacteriophage T4. The structure alignment was performed with chain A from each of the PDB files, which correspond to one of the three protomers (except that of the *E. coli*, which has two protomers, where only the N-terminal two thirds of the sequence is shown).   Similar residues are coloured red. Secondary structure elements corresponding to human PCNA are shown above the alignment and were generated with ESPript. β-strands are indicated as arrows, α-helices as spirals, and β-turns as TT. Since *E.coli* β-clamp assembles in a dimer instead of a trimer, the sequence of chain A is longer and it is not fully plotted.
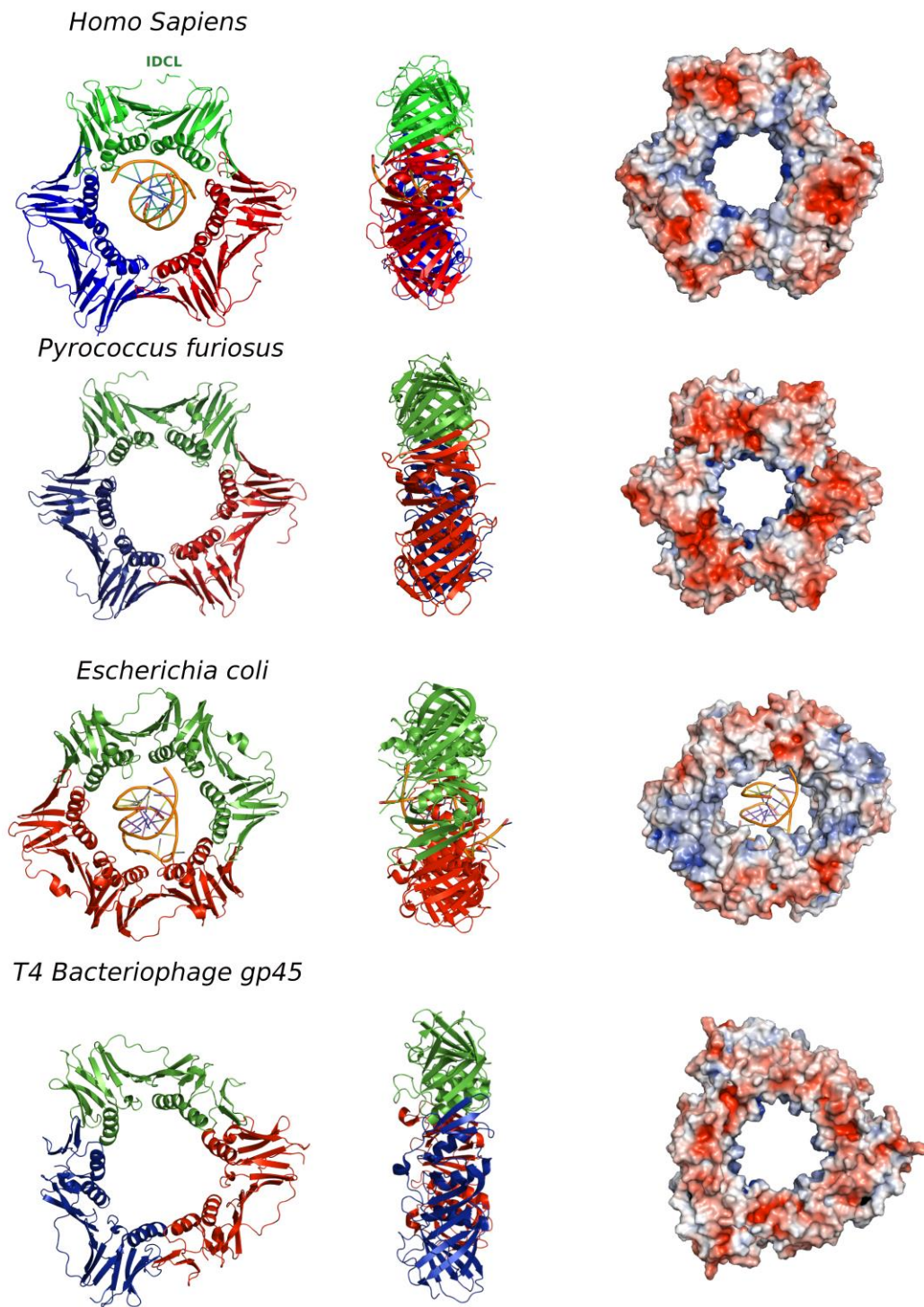
Figure 1.2. Sliding clamps from different organisms. Crystal structures of gene 45 antigen of *Bacteriophage T4* (Protein Data Bank [PDB] code 1CZD), the *Escherichia coli* β clamp (PDB code 3BEP), (D) *Pyrococcus furiosus* PCNA (PDB code 1GE8), and (D) *Homo sapiens* PCNA (PDB code 1VYM). For each organism, the top and side views of the three-dimensional crystal structure are shown. Each protomer is represented in one colour. The surface electrostatic potential is represented on the right with positive charges depicted in blue and negative charges in red.

The central pore has an internal diameter of approximately 35 Å, larger than the 24 Å of the double stranded DNA helix (dsDNA) in a canonical B-form [13]. Overall, clamps are acidic proteins with a net negative charge. The outer surface possesses a strong negative potential, but the α-helices facing the central cavity are rich in positively charged amino acids (primarily Lys and Arg) and generate a positive electrostatic potential that facilitates the DNA to pass through [14]. The negatively charged surface might contribute to prevent unspecific interactions, facilitating the correct disposition of the DNA inside the ring. There are two distinct faces of the ring: the front face, also known as the C-face (since the carboxyl terminal tail of the protomers is located there), contains a hydrophobic pocket where polymerases and other proteins bind, while the back face serves as target for post-translational modifications that alter the clamp's properties [6].

For several years considerable efforts have been made to structurally assess the association between sliding clamps and DNA, as well as to understand the molecular mechanism by which these ring-shape, multimeric proteins slide on DNA. However, the weak and non-sequence specificity of PCNA-DNA interplay made this task elusive. The first available high resolution crystallographic structure of a sliding clamp in complex with DNA , was the *E. coli* DNA polymerase III β subunit bound to a designed 10 bp dsDNA [15] (Fig.1A). In order to mimic the primed DNA strand while being copied by the polymerase, the designed dsDNA had a four-base long 5′-overhang of ssDNA on one end. The dsDNA portion appeared inside the PCNA ring, tilted with an angle of 22°, which could be explained by contacts of the front side of the ring with the DNA, but also by the interaction of the ssDNA with another symmetry related ring in the crystal lattice. The authors proposed that the β-clamp-ssDNA contact acts as "placeholder", attaching the clamp at the 3' end of primed site and preventing it from sliding off the DNA before initiation of replication [15]. Surprisingly, cryo-EM work performed on the catalytic core of the bacterial replisome (this is, β-clamp bound to DNA, polymerase IIIα, exonuclease, Θ or τ) showed that interactions between the clamp and DNA are different from the ones described in the crystal structure of the clamp-DNA complex. In particular, the cryo-EM derived structure describes an extensive rearrangement of the DNA between synthesis and editing modes that moves toward the inner wall of the clamp acting as a "lock"[16].

The crystal structure of a single-chain chimera of *S. cerevisiae* PCNA bound to primed DNA was also solved, but little information could be obtained from DNA structure due to low

occupancy and the presence of disordered regions. Nevertheless, the model showed dsDNA within the central cavity of PCNA ring, facilitating contacts between the negatively charged phosphate backbone of DNA with positively charged residues of the inner side of PCNA. The ssDNA overhang in this case was not observed [17].

1.2. Human PCNA

1.2.1. *Structure and function*

Proliferating cell nuclear antigen (PCNA) was concurrently discovered by two different groups. On one hand, Miyachi et al. [18] detected an auto-antigen in the sera of some patients with systemic lupus erythematosus and because the protein was detected in nuclei of dividing cells they named it PCNA [19]. On the other hand, Bravo and Celis [20] identified a protein which was synthetized during the S-phase cell cycle and decided to call it cyclin. Further experiments showed that both were the same protein of 29 kDa, which behaves as a trimer in solution and is stable over 25-45 °C. PCNA 86 kDa ring is opened by the clamp loader Replication Factor C and it is placed encircling DNA duplex in a process where ATP is hydrolysed [21]. Extensive studies demonstrate that PCNA is an auxiliary factor for the replication polymerases δ and ε, increasing their processivity by tethering them and sliding along the double-stranded DNA helix. Especially important is its role in the synthesis of the lagging strand, acting as a platform where polymerase δ, FEN1 (flap endonuclease 1) and LIG1 (DNA ligase I) bind to synthesize, process, and join Okazaki fragments [22]. PCNA recruits other factors to the replication fork participating in DNA repair (translesion synthesis, homologous recombination, mismatch repair, and nucleotide excision repair), chromatin remodelling and cell cycle control [22]. Not all these partners bind simultaneously to PCNA, and switching may be triggered by different pathways: affinity-competition, proteolysis or post-translational modifications such as acetylation, phosphorylation, sumoylation and ubiquitination [2].

The first crystal structure of human PCNA was solved in 1996 bound to the C-terminal region of the cell-cycle checkpoint protein p21[WAF1/CIP1] (PDB entry: 1AXC) [23]. Yet, it was not until 2004 when two structures of the human PCNA trimer alone were elucidated (PDB entries 1VYM and 1W60) corresponding to two distinct spatial groups (monoclinic C2 and trigonal P3, respectively) [24]. Both structures are essentially the same but a bucking trimer is seen in the monoclinic crystal compared with the planar ring in the trigonal one, indicating a certain degree

of plasticity. The front face of the ring contains the three IDCLs linking the two domains of each protomer, and the back face has prominent loops emerging into the solvent. Each protomer presents a hydrophobic pocket next to the IDCL loop on the front face of the ring where PCNA binding proteins accommodate (Figure 1.1.D).
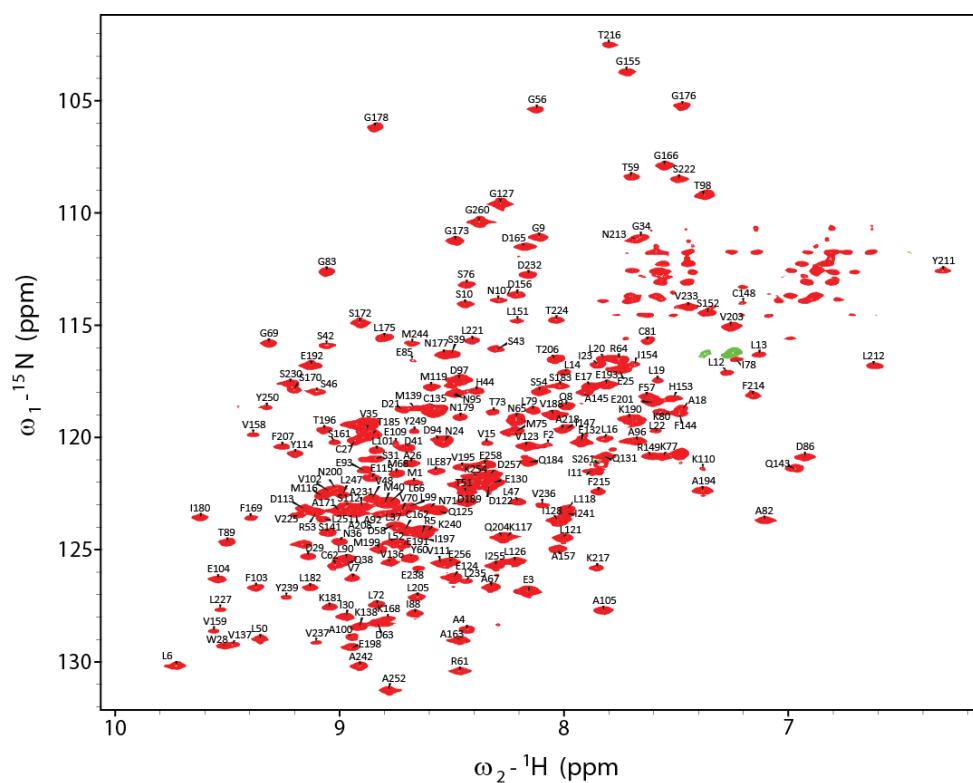

Figure 1.3 ^{2}D 1H–^{15}N TROSY fingerprint spectrum of human PCNA (recorded at 35°C and 800 MHz). Out of the 251 backbone amides (excluding the first signal and the 8 prolines) only 12 remain unassigned [25]. The unassigned signals correspond to the side chains of Asn, Gln and Arg residues. Green signals are folded in the indirect dimension (arginine side chains).

Subsequently, the solution NMR spectrum of PCNA was assigned and the backbone chemical shifts and long range NOEs were measured [25] (Figure 1.3). These data were consistent with the crystallographic results, indicating that the structure in solution is the same. Moreover, size exclusion chromatography coupled to multiangle light scattering (SEC-MALS) measurements showed that the trimeric form is predominant even at high concentrations, confirming that recombinant PCNA behaves as one single trimer in solution [26]. Experiments in intact cells and cell extracts showed that PCNA can form dimers of trimers [27], but even if PCNA may exist transiently as a loosely bound dimer of trimers inside cells, their functional role is unclear.

Thermal and chemical denaturation studies indicate that although human PCNA has the same three-dimensional structure as its *S. cerevisiae* homolog it is less stable. Human PCNA also displays an increased backbone dynamics compared with yeast PCNA. Altogether, the biophysical and structural data of PCNA suggest that this highly dynamic and plastic behaviour could be an evolutionary advantage to bind to a large number of ligands [26].
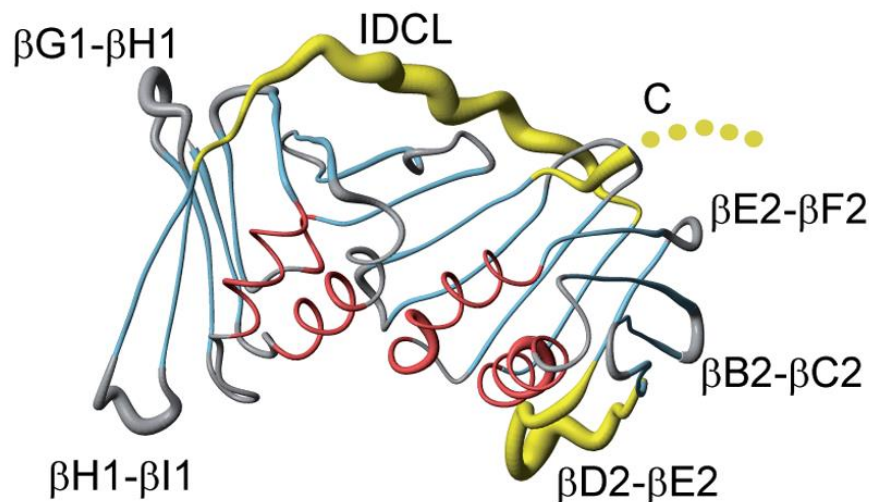

Figure 1.4. Backbone dynamics of human PCNA derived from ^{15}N NMR relaxation measurements. Representation of the backbone structure of one protomer of the PCNA as coil with the secondary structure elements coloured blue (β-strands) or red (α-helices). The thickness illustrates the relative dynamics of the corresponding backbone NH bond. The flexible regions are shaded in yellow: IDCL (residues 117–134), the βD2-βE2 loop (184–195), and the C-terminus (252–261). (Figure adapted from De Biasio et al., 2012).

1.2.2. *PCNA sliding on DNA*

Single molecule studies showed two modes of PCNA sliding on DNA. In one mode, the clamp tracks the helical pitch of the DNA duplex, resulting in a rotation of the protein around the DNA. In the second mode, less, frequent, the protein translationally diffuse at higher rates [28]. The crystal structure of human PCNA bound to 10 bp DNA duplex inside the channel shows the DNA tilted by an angle of 15° relative to the three-fold rotation axis of the ring [29]. The PCNA-DNA interface comprises six conserved basic residues spread along five α-helices of two subunits that establish short-living polar contacts with five consecutive phosphates of one of

the DNA strands. NMR studies confirmed that primed DNA does not bind the hydrophobic pocket of PCNA. In order to figure out the sliding clamp motions, structural data were combined with molecular dynamics simulations [30]. These simulations presented a more pronounced tilting of the DNA of approximately 30° and a larger portion of DNA interacting simultaneously with two protomers of the trimer. Still, the basic residues within the PCNA channel establish short-live interactions with the phosphates of DNA and switch between adjacent phosphates in a non-coordinated manner. The data suggest a mechanism for PCNA sliding on the PCNA: when a sufficient number of polar contacts are established with adjacent phosphate groups of DNA, the ring rotates on the DNA, resulting in the advancement of one base pair. Therefore, the conserved patch of basis residues of the inner wall enables PCNA to rotationally track the helical pitch by spiral diffusion, named "cogwheel diffusion" (Figure 1.5 A). This process keeps DNA-protein contacts and maintains the clamp orientation invariant relative to the DNA.
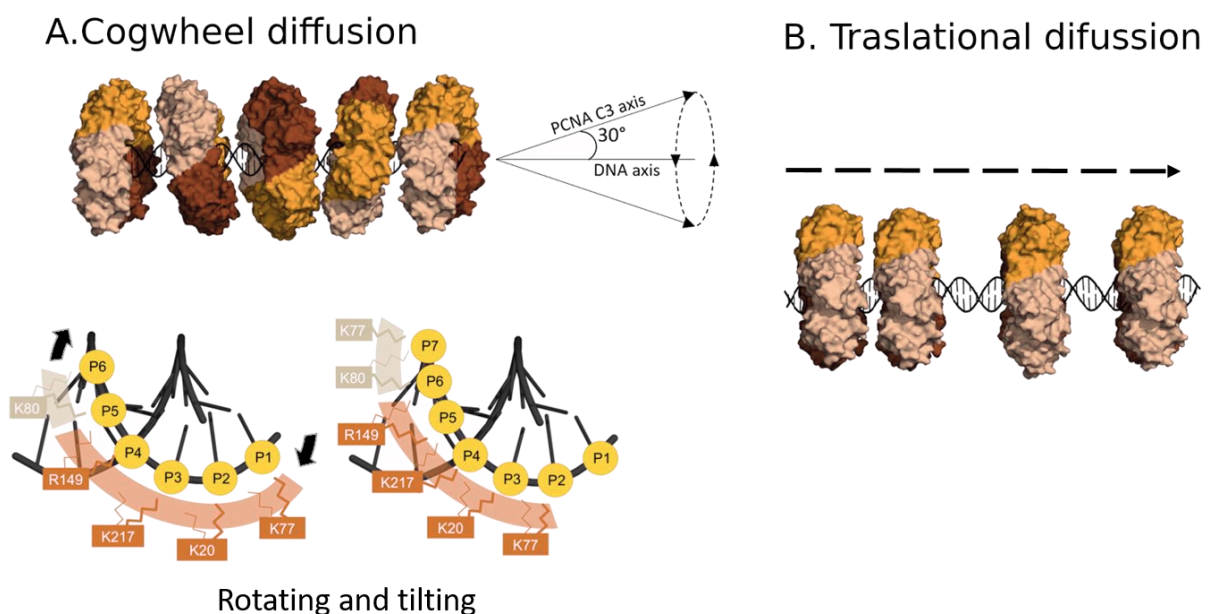

Figure 1.5. Human PCNA sliding on DNA. A) Cogwheel diffusion mode: PCNA rotation tracks the DNA through a spiral motion, establishing transient interactions with DNA, which keeps the orientation of the clamp relative to DNA invariant. The threefold rotation axis of PCNA processes around the DNA helical axis (at a fixed 30° angle). In the lower panel the evolution of PCNA–DNA contacts during cogwheel sliding are illustrated. Interacting side chains can rapidly switch between adjacent phosphates randomly (indicated by the thin and thick lines), enabling PCNA to advance along the phosphate backbone by rotating and tilting motions. B) Translational diffusion mode: PCNA travels along the DNA, without contacting the DNA. (Figure adapted from de March et al. 2017).

Due to the clamp symmetry, DNA can interact with three equivalent sites of PCNA. Site switching is generated by 120° rotations about the threefold axis of the ring and requires the interruption and subsequent rearrangement of the clamp–DNA interface (Figure 1.5B). Thus, the occasional exchange of DNA among the three positions may explain the translational component of PCNA sliding observed in single-molecule studies [28].

The basic residues at the sliding interface are essential to stablish a defined orientation of the clamp relative to DNA. This specific orientation may be required by polymerase δ holoenzyme to assemble and initiate elongation of primed Okazaki fragments [29]. Mutations of two conserved lysines of yeast PCNA-DNA interface impaired completely yeast polymerase δ function in replicating circular DNA templates. Moreover, the sliding surface of DNA can be modulated by both, specific lysine 240 acetylation to control DNA damage response [31] [32] and binding of protein factor p15 [33]. Therefore, although less studied that the outer surface, the inner part of PCNA is also evolutionary conserved, highly regulated and crucial for the PCNA role in the replication fork [30] [34].

1.2.3. *PCNA binding proteins: Revisiting PIP box definition*

A large network of proteins is responsible for replicating DNA with high fidelity and for repairing DNA damage through different pathways. PCNA is a global hub in DNA metabolism that interacts with a large number of proteins involved in a variety of DNA related processes [35]. As PCNA is a symmetric homotrimeric in solution, it has three identical hydrophobic pockets to bind simultaneously different partners and coordinate a variety of functions in space and time. Most of the proteins that bind PCNA either are IDPs (intrinsically disordered proteins) or have IDRs (intrinsically disordered regions). These disordered regions have accessible surface areas which increase their ability to interact with diverse binding partners through short linear motifs (SLiMs) [36]. Many of the proteins that interact with PCNA display a characteristic SLiM known as PIP box (PCNA interacting Protein-box) or an extended version called PIP degron. The canonical PIP box motif is *QXXhXXaa*, where *h* is an aliphatic hydrophobic residue (frequently I, L or M), *a* is an aromatic hydrophobic one (F, W or Y) whereas X can be any amino acid [9]. The PIP degron motif targets PCNA for degradation and harbors also a basic residue (K,R) four amino acids ahead from the last aromatic as well as a TD motif (threonine and aspartic) just before the aromatic residues [37].

Structural studies of PIP box containing proteins or derived peptides in complex with human PCNA have unveiled the molecular details of the PIP motif-PCNA interface. The crystal structure of a p21-derived peptide in complex with human PCNA was the first structural characterization of this interface (PDB code: 1AXC) [23]. Since then, several crystal structures of PCNA bound to different ligands have been solved. Overall, all PCNA interacting proteins adopt the same conformation, which consists of: an extended N-terminal region; a 3_{10} helical turn of four residues enclosed by the hydrophobic residues of the PIP box, and a C-terminal region of variable length that sometimes adopts a β strand secondary structure and interacts with the IDCL. The conserved helix inserts like a plug into the hydrophobic pocket of PCNA whereas the initial glutamine sticks into the so-called Q-pocket establishing hydrogen bonds with the backbone of PCNA [23 38]. In addition, the assignment of human PCNA spectrum (Figure 1.3) [25] allowed to analyse the ^1H-^{15}N correlation spectra of the sliding clamp in the presence and absence of different ligands and calculate perturbations in the backbone amide chemical shifts (CSP) and/or intensities. Thereby, NMR data provided complementary or new information of these interactions in solution at the residue level [39].

The available structural information of PCNA in complex with different partners reveals that three ligands are able to bind simultaneously the three identical protomers of PCNA ring. Isothermal titration calorimetry (ITC) data with different peptides are all well fitted with a model assuming one set of equivalent sites, with no evidence of binding cooperativity. Therefore, it seems that ligands compete to bind PCNA based on their affinity, which can be modulated by posttranslational modifications [9]. PCNA itself suffers different posttranslational modifications that regulate its binding to different partners or RFC unloading [40]. Monoubiquitination at Lys 164 has been described to increase the affinity of PCNA for TLS polymerase η [41]. However, this result is in contradiction with new evidence showing that both pol η biding to PCNA and DNA synthesis by this holoenzyme are independent of PCNA monoubiquitination [42].

Up to know, the tumour suppressor p21 is the PCNA binding partner showing the highest affinity (Table1), probably it needs to displace other proteins to block replication in response to DNA damage. This high affinity lies on an efficient hydrophobic packing as well as more

number of electrostatic interactions with the C-terminal region of PCNA [38]. Studies with variable length peptides of p21 suggest that basic residues at N and C-termini encompassing the PIP box protein also contribute to increase the binding affinity. Interestingly, a recent study combining experimental data with a computational approach confirmed this correlation between positive patches of residues flanking the PIP box motif and strong affinity for PCNA [36]. Moreover, FEN-1 endonuclease and p68 (also known as p66, the third subunit of polymerase δ) exhibit less basic residues than p21 and accordingly present lower affinity (Table1). However, the comparison of the structure of PCNA with both full length FEN-1 and a 20-residue long PIP fragment reveals additional contacts involving the core domain and the C-terminal part of FEN-1, which increase the affinity by three orders of magnitude.

Emerging evidence shows that the canonical PIP box definition is becoming too narrow because many other sequences diverging from the canonical PIP box have been described to bind PCNA. In particular, several non-canonical sequences have been reported in Y-family of DNA polymerases (Polη, Polι and Polκ), where the first glutamine at position 4 and the aromatic residue at position 8 is not conserved [43]. An alternative PCNA binding motif have been found: the AlkB homologue PCNA interacting motif (APIM), which comprises only 5 amino acids with the sequence [K/R]–[F/Y/ W]–[L/I/V/A]–[L/I/V/A]–[K/R], present in several cytosolic proteins [44,45]. The crystal structured of PCNA in complex with APIM peptides showed a similar binding mode than PIP motif despite holding a highly divergent sequence from the canonical PIP box Furthermore, a recent work points out that the PIP sequence has overlapping regions with RIR (Rev1 interacting region) and MIP (Mlh1 interacting proteins) motives. All these findings suggest that a more general class of PIP-like motifs should be considered. PCNA might be able of bind a broader class of partners and these ligands may be capable of recognizing more than one hub. However, the exact features that determine the selectivity and the affinity of the ligands are not well understood.

So far, almost 80 PIP-like proteins have been described to interact with PCNA in the literature using different techniques. However, not all these experimental techniques are equally reliable, and some of them are known to yield false positives (as yeast two hybrid, ELISA, or co-immunoprecipitation). Only those quantitative methods that use pure protein such as ITC, crystallography, NMR or Electron Microscopy provide unambiguous proof of a direct

interaction[9]. The available information on the interaction of PCNA with several proteins and peptides is summarized in Table1.1. Only those interactions that have been characterized by one or more of the techniques mentioned above have been included.

Table 1.1. PCNA interacting proteins summary. The selected proteins positions with residues specific to the PIP box motif are marked in red.

Protein	Activity	Sequence	T(°C)	Kd (µM)	Method	PDB code	Ref
CANONICAL							
p21	CDK1 inhibitor	[139]GRKRRQTSMTDFYHSKRRLIFS[160]	30	0.080	ITC, XR, NMR	1AXC	(Gulbis et al. 1996)
Pol δ p68	Replicative polymerase	[451]GKANRQVSITGFFQRK[466]	30	1.5	ITC, XR, NMR	1U76	(Bruning and Shamoo 2004)
Fen-1	Endonuclease	[331]SRQGSTQGRLDDFFKVTG[350]	30	59.9	ITC, NMR	1U7B	(Bruning and Shamoo 2004)
p15[PAF]	PCNA ass. factor	[50]GNPVCVRPTPKWQKGIGEFFRLSPKDSE[77]	25	5.56	ITC, XR, NMR	4D2G / 6GWS	(De Biasio et al. 2015) / (De March et al. 2018)
ZRANB3 (PIP)	Helicase/Endonuclease	[511]FTHFEKEKQHDIRSFFVPQPKK[532]	25	4.8	ITC,XR	5MLO	(Sebesta et al. 2017)
DVC1	Adaptor protein	[321]SNSHQNVLSNYFPRV[336]	25	15.55	ITC, XR	5IY4	(Wang, Xu, and Jiang 2016)
DNMT1	Methyltransferase	[161]STRQTTITSHFAKGPAKRKP[180]	25	1	ITC, XR	6K3A	(Jimenji et al. 2019)
UHRF2	E3 ubiquitin ligase	[784]NEILQTLLDLFFPGYSK[800]	20	25.7	ITC, XR	5ICO	(Chen et al. 2017)
Cdt1	Replication factor	[1]MEQRRVTDFFARRR[14]	ND	7.2	FP, XR	6QCG	(Hayashi et al. 2018)
RecQ5	Helicase	[952]KTSPGRSVKEEAQNLIRHFF HGRARCESE[980]	35	210	NMR	-	(Gonzalez-Magaña et al. 2019)
NON-CANONICAL							
pol ι	TLS polymearase	[419]CAKKGLIDYYLMPSLST[435]	25	0.39	SPR	2ZVM	(Hishiki et al. 2009)
pol η	TLS polymerase	[694]CKRPRREGMQTLESFFKPLTH[713]	25	0.4	SPR	2ZVK	(Hishiki et al. 2009)
Pol κ	TLS polymerase	[856]CIKPNNPKHTLDIFFK[870]	25	ND	SPR	2ZVL	(Hishiki et al. 2009)
ZRANB3 (APIM)	Helicase/Endonuclease	[1058]QVRRQSLASKHGSDITRFLVKK[1079]	25	9.24	ITC, XR	5MLW 5YD8	(Sebesta et al. 2017)(Hara et al. 2018)
PARG	Glycosylase	[402]QHGKKDSKITDHFMRLPKA[420]	25	3.3	ITC, XR	5MAV	(Kaufmann et al. 2017)
p12	Polymerase δ regulator subunit	[1]MGRKRLITDSYPVVKRREG[19]	25	38	ITC,XR, NMR	6HVO	(Gonzalez-Magaña et al. 2019b)
Cdt2	E3 ubiquitin ligase	[704]SSMRKICTYFHRKS[717]	ND	0.057	FP, XR	6QC0	(Hayashi et al. 2018)
TRAIP	E3 ubiquitin ligase	[447]KQRVRVKTVPSLFQAKLDTFLWS[469]	25	30.7	ITC, XR	4ZTD	(Hoffmann et al. 2016)
RNH2B	RNase	[290]DKSGMKSIDTFFGVKNKKKIGKV[312]	-	-	XR	3P87	(Bubeck et al. 2011)

1.2.4. *The PCNA-p15 interaction*

1.2.4.1. Structure and function of p15

Intrinsically disordered proteins (IDPs) are those that lack secondary and tertiary structure under physiological conditions. Although they fulfil important biological functions across all life domains, it is well known that they are more abundant in eukaryotes [46]. In particular, most transcription factors as well as proteins involved in signal transduction in eukaryotic organisms are predicted to be disordered or contain disordered regions (IDRs) [47]. This reveals a correlation between complex cell regulation and the greater presence of IDPs /IDRs [48]. Moreover, the majority of proteins associated with human cancer have been identified as IDPs or IDR containing proteins (79%), which underlines the crucial roles they play in several cellular events that are altered in cancer such as cell proliferation, DNA repair, and apoptosis [49]. The evolutionary advantage of IDPs likely lies in their plasticity to interact with plenty of ligands because of their large available binding surface and their vast regulation through post-translational modifications [50].

In order to exert their function, IDPs bind folded partners. The interaction occurs primarily through SLIMs, which adopt a relative stable conformation upon binding. SLIMS can partially fold whereas the rest of the IDP remains disordered. This may happen by different mechanisms: (I) induced fit, where the binding partner forces the IDP to adopt its folded conformation, (II) conformational selection, when the binding protein selects the most efficient conformation for the bound state, or (III) a combination of induced fit and conformational selection [51,52]. The structural characterization of IDPs has always been challenging since they exist as an ensemble of multiple conformers in fast exchange. Thereby, NMR is the main method used to analyse the structure and dynamics of partially or completely unfolded proteins [50] . SAXS has also emerged as powerful tool to analyse the structure and dynamics of biomolecular complexes and highly flexible proteins that can complement NMR studies [48].

PCNA associated factor p15, PAF15, p15PAF, PCLAF or KIAA0101 (herafter p15) was first identified as PCNA binding factor by co-immunoprecipitation and yeast two hybrid assays. It

is an intrinsically disordered protein of 111 residues which is overexpressed in several type of human cancers and correlates with poor prognosis [53, 54]. It is primarily present in the nucleus and mitochondria of the cells [55]. The expression levels of p15^{PAF} vary during the cell cycle with the major peak in the S phase where the DNA replication occurs [56]. p15 is involved in DNA replication as well as in DNA damage bypass and cell cycle progression by interacting with PCNA trough a canonical PIP box sequence. The degradation of p15 is mediated by the ubiquitin ligase anaphase-promoting complex/cyclosome (APC/C) that targets it for degradation. Both PCNA and p15 co-localize in the nucleus during the S phase of the cell cycle and they are associated with chromatin. The absence of p15 reduces DNA replication importantly, suggesting that it regulates PCNA processive activity [33, 57]. UV irradiation up-regulates p15 expression and p15-PCNA interaction [55]. Co-immunoprecipitation experiments from pancreatic cancer cell lines indicate that p15 appears as part of DNA-replication foci together with PCNA, pol δ and the endonuclease FEN-1 [58].

The structural characterization of p15 showed that is monomeric in solution and exhibits typical features of a disordered protein. Although CD data as well as the lack of dispersion in the NMR backbone amide signals are compatible with a largely disordered protein (Figure 1.6 A), RDC measurements give evidence of transient non-random structural elements. Specific regions of helical structure are detected in the central part of p15 encompassing the PIP box sequence. These partially structured regions are proposed as potential molecular recognition elements for p15 [50].

1.2.4.2. PCNA-p15-DNA interaction

The molecular recognition of PCNA/p15 has been characterized by an integrative approach and shows a unique mode of binding to PCNA that extends outside the canonical PIP box. The central segment of p15, where the PIP box is located, binds tightly to the front and inner side of the ring, while the N and C tail remains disordered at opposites sides of the ring (Figure 1.6 and 1.7). ITC experiments revealed that three molecules of p15 are able to bind the three promoters of PCNA ring in a non-cooperative manner with a K$_d$ of 1.1 μM [59]. However, in the cell it is unlikely to have three identical molecules occupying the three hydrophobic pockets of the ring. PCNA ligands compete for binding the three protomers of the ring in the cell and

the preferential binding of some proteins over the others presumably depends on the affinities regulated by post-translational modifications, steric hindrance and relative protein concentrations at specific space and time. The IDP p21 is the only protein that might be able to displace efficiently other competitors from PCNA binding (Kd=80 nM, table 1.1) which underlines its predominant role in the inhibition of DNA replication and cell cycle progression [23].
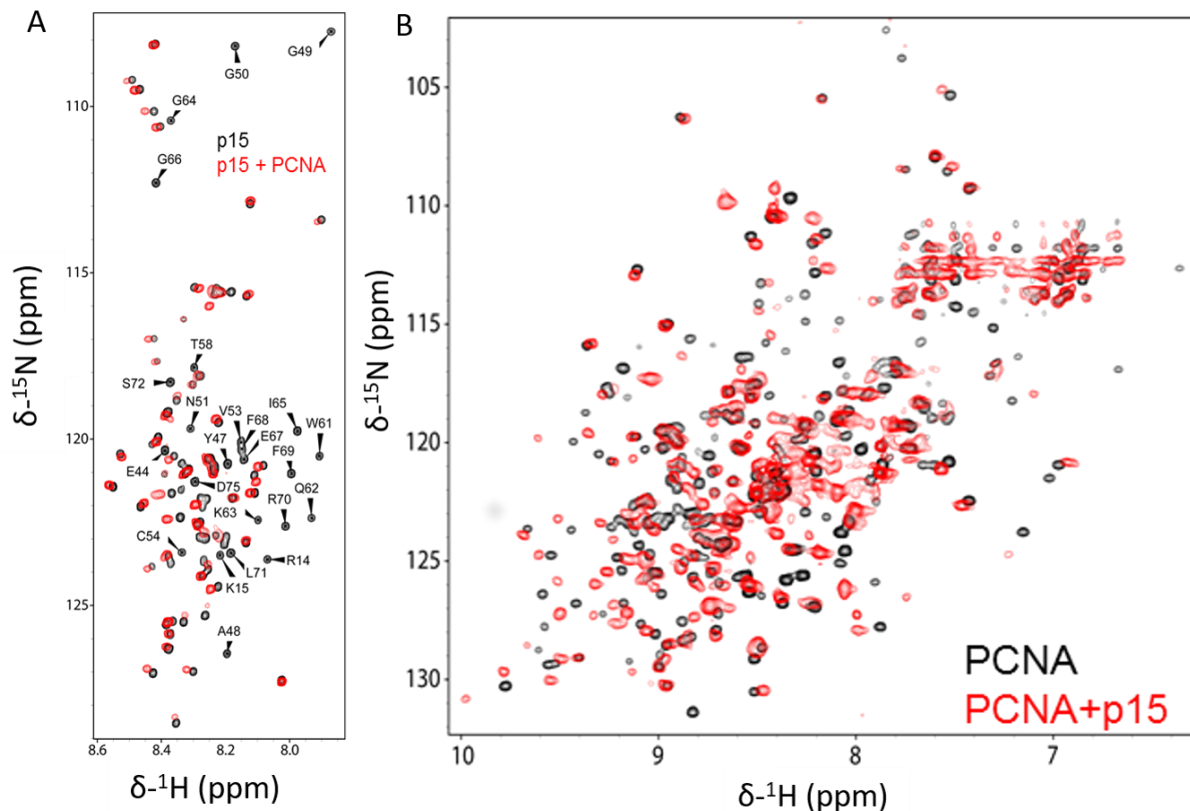

Figure 1.6. PCNA-p15 interaction by NMR. A) Overlay of ^{1}H-^{15}N HSQC spectra of p15 in the presence or absence of PCNA. The residues that present larger CSPs or intensity decrease are indicated. B) Overlay of ^{1}H-^{15}N TROSY spectra of perdeuterated PCNA in the presence (red) or absence of p15 (black) (Figure adapted from de Biasio et al 2015).

Similar to other PCNA binding partners, the canonical PIP box residues of p15 (^{62}QKGIGEFF^{69}) bind the groove on the front face of PCNA protomer under the IDCL. The side chain of Q62 interact with residues A252 and A208 of PCNA (Q pocket); residues I65-F67 adopt the characteristic 3_{10} helix; and the hydrophobic trident formed by I65, F68 and F69 insert into the hydrophobic pocket of PCNA. In contrast to other PCNA binding partners, the PIP box of p15 does not form a stable intermolecular anti-parallel β-strand with the IDCL of PCNA protomer. Moreover, crystallographic results and NMR data with full-length p15 strongly

suggest that the flexible N-terminus passes through the PCNA ring and non-specifically contact its back-face (Figure 1.7A). Molecular modelling simulations [50] and SAXS data [33] support this premise (Figure 1.7 B).
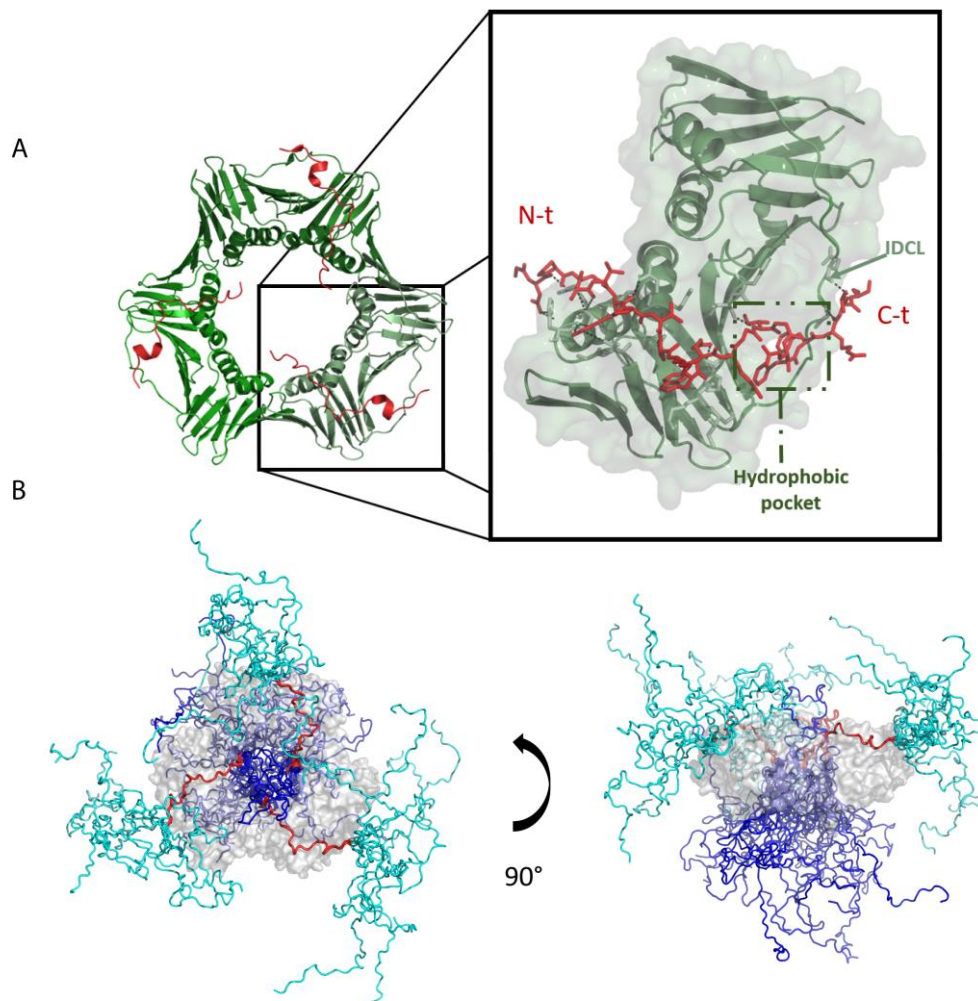

Figure 1.7. The central region of p15 binds PCNA while the N- and C-terminal tails remain disordered. A) Crystal structure of PCNA trimer bound to three p15[41-72] fragments (PDB code: 6GWS) in cartoon representation. PCNA is coloured in a different tone of green while p15 fragments are shown in red. On the right, a zoom view of the binding interface is shown, where polar contacts between PCNA and p15 are marked with discontinuous black lines. p15 peptide and PCNA are represented with sticks and cartoon respectively. Only those residues of PCNA that interact with p15 are drawn with sticks. The surface representation of the monomer of PCNA is shown with transparency. B) Model of the full-length p15–PCNA complex. Front and side views of 10 modelled structures of p15–PCNA complex based on experimental data. PCNA is shown as grey surface and p15 as ribbons, with the central region coloured in red and the added disordered N- and C-terminal extensions in cyan and blue, respectively. In one of the 10 selected complex models, one p15 N-terminus folds back towards the front-face of PCNA instead of passing through the hole (Figure adapted from de Biasio et al. 2015).

Despite p15 can bind PCNA alone, the functionally relevant form appear to be the ternary complex constituted by p15 bound to PCNA loaded on DNA. p15 binds DNA and PCNA

through two independent sites and similar affinities. p15 binding to DNA occurs mainly via electrostatic interactions with the positively charged N-terminal tail [59]

Recently, the crystallographic structure of the ternary complex of p15-PCNA-DNA was solved, where two molecules of p15 bind to the ring of PCNA while DNA occupies the other PIP site. This evidence was further confirmed by MD simulations showing that when p15 binds two subunits of the PCNA ring, the DNA is able to pass through and interact with the remaining subunit [33] . Solution NMR reveals that the presence or absence of DNA does not alter the PCNA-p15 interaction and thereby it is proposed that p15 reduces the available sliding surface of PCNA. In particular, p15 may act as a belt that attaches loosely the DNA to the clamp during synthesis by replicative polymerase δ. Yet, this tether may be released after replication blockage to allow the entrance of translesion synthesis polymerases (TLS) to bypass the DNA lesion.

1.2.4.3. p15 ubiquitination

Genomic DNA can suffer damage, which leads to replication blockage, resulting in DNA strand breaks, chromosomal rearrangements and cell death. Replicative polymerases δ and ε, are unable to efficiently incorporate nucleotides opposite to DNA damage templates [60]. Translesion synthesis is one of the paths responsible of overcoming stalled replications forks and involves specialized polymerases that bypass DNA lesions, called translesion synthesis polymerases [61].

During normal DNA replication, p15 in complex with PCNA is mono-ubiquitinated at two lysines located in the N-terminal tail (K14 and K25) [62]. Upon UV induced damaged, this doubly monoubiquitinated p15 (hereafter dmUbp15) is targeted for proteasomal degradation and is likely replaced by its non-ubiquitinated form at the stalled replication fork. Simultaneously, PCNA is ubiquitinated at Lys 164, which facilitates the recruitment of TLS polymerase η to the damaged site to bypass the lesion. [63, 62]. In view of these data, it is speculated that p15 might play a role regulating the sliding velocity of PCNA on DNA, which allows the switch from replicative to TLS polymerases (Figure 1.8).
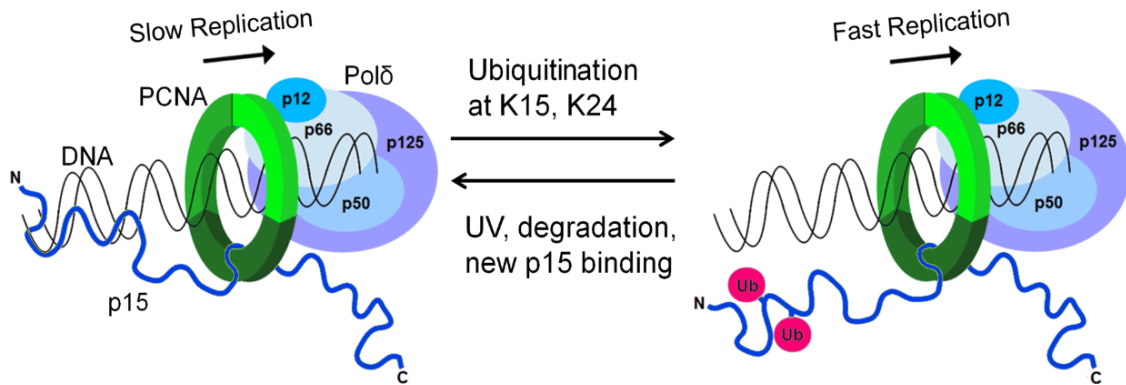
Figure 1.8. Possible mechanism by which p15 might regulate the sliding velocity of PCNA on DNA. During normal replication p15 is ubiquitinated and slides fast on DNA. Upon DNA damage ubiquitinated p15 is degraded and replaced by non-ubiquitinated p15, slowing down the sliding velocity and thus, facilitating the switch for replicative polymerases. (Figure modified from De Biasio et al, 2015).

p15 ubiquitination is accomplished by UHRF1 ubiquitin ligase, which is a pivotal enzyme for DNA methylation maintenance [64]. Previous studies demonstrate that UHRF1 is able to recognize hemimethylated DNA and enhance the recruitment of DNMT1 methyl transferase to fully methylate the newly synthetized DNA strand [65, 66]. Moreover, UHRF1 also ubiquitinates two close lysines of the N-terminal tail of Histone H3 that are specifically recognized by the RFTS domain of DNMT1. The family of DNA methyl transferases (DNMTs) consists of three different classes of enzymes responsible of the DNA methylation in the cell: DNMT1, DNMT2 and DNMT3A/3B/3L. Overall, DNMT3A and DNMT3B are related with de novo methylation while DNMT2 function is poorly understood. DNMT1 is reported to be involved primarily in the methylation inheritance by recognizing methylation patterns during DNA replication [65].

Eukaryotic DNMT1 is a large protein composed of different domains: N-terminal nuclear localization signal (NLS), a replication foci targeting sequence (RFTS) that recruits DNMT1 to the DNA replication fork, a zinc finger CXXC domain that specifically recognizes unmethylated CpG dinucleotides[67], bromo-adjacent homology (BAH) domains, and a C-terminal methyltransferase domain including the catalytic core and the target recognition domain (TRD)[68].

Under normal conditions, DNMT1 enzyme is inactive, because the RFTS domain precludes the active site of the enzyme. It is suggested that the binding of doubly monoubiquitinated histone H3 to RFTS triggers the re-arrangement of DNMT1 which activates its methyl

transferase activity [69]. The activation of DNMT1 methyl transferase at specific time and position is of utmost importance to maintain the DNA methylation pattern in the cell. Remarkably, p15 exhibits a very similar N-terminal tail as the one of histone H3 that is also ubiquitinated at two close lysines by UHRF1. These evidences suggest that RFTS domain might also recognize specifically the two ubiquitins of p15 and recruit it to replication fork and play a role in epigenetic DNA methylation [70](Figure 1.9).

Although during the last decade a considerable progress has been made in the knowledge regarding p15 structure and interactions, the molecular mechanism underlying p15 function in DNA replication and lesion bypass remains controversial. Moreover, the monoubiquitination of lysines 15 and 24 in the N-terminal tail of p15 appears to be extremely important and further studies should be done in this topic.
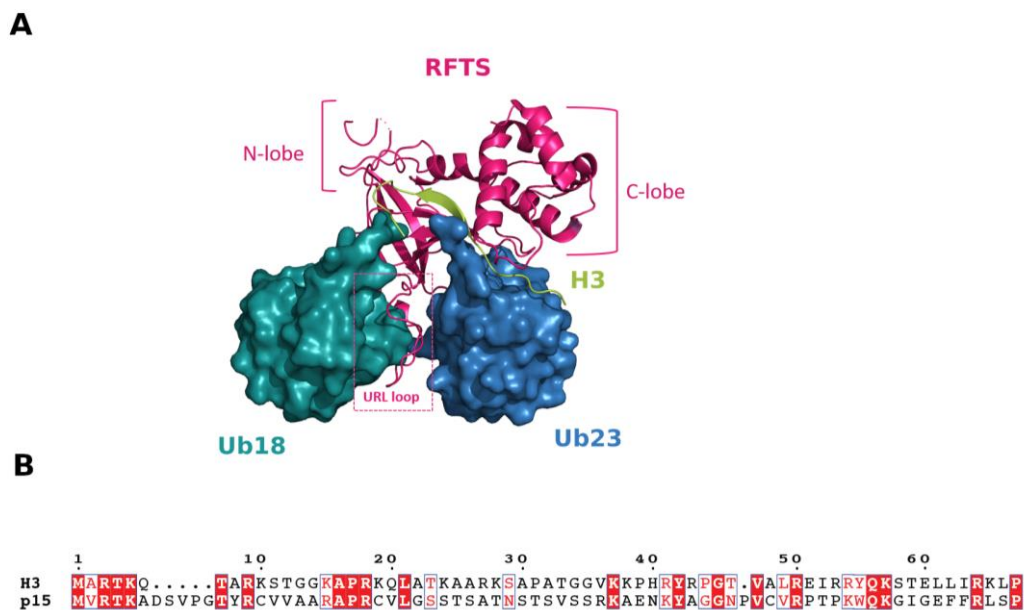

Figure 1.9. Doubly monoubiquitinated p15 might be recognized by the reader module RFTS of DNMT1 methyl transferase. A) Crystal structure of DNMT1 RFTS domain in complex with K18/K23 mono-ubiquitinated histone H3 (PDB code: 5WV). The two ubiquitins are coloured in red and blue, the histone H3 N-terminal tail is depicted in yellow and the RFTS domain in pink. B) Sequence alignment of Histone H3 and p15 N-terminal tail. The structure alignment was performed with the Structure Alignment software. Similar residues are indicated with red letters while strictly conserved residues are in white surrounded by red colour.

1.2.5. *The PCNA- p12 subunit of polymerase delta interaction*

1.2.5.1. Eukaryotic polymerase delta

DNA replication is an essential event for every living organism and it needs to be performed with exquisite fidelity in order to ensure species survival [71]. Therefore, an efficient machinery is pivotal to maintain genetic information during cell division, DNA repair, DNA recombination, and DNA damage bypass [72].

Replication machinery is in charge of replicating DNA in a semiconservative way in every living entity. Moreover, it is optimized to sort out the problem of antiparallel dsDNA (the two strands of the DNA are not identical but complementary to each other and they run in opposite direction). The replicative polymerases always read the DNA in 3-'5' sense and incorporate the new nucleotides in 5'-3' direction [73]. Since the dsDNA strands are antiparallel, the leading strand is synthesized in the same direction that the replication machinery moves, whereas the lagging strand needs to be replicated discontinuously in the opposite direction [74]. The lagging strand is synthetized in fragments, termed Okazaki fragments, which are subsequently joined by a ligase. Thus, replication of both strands of the DNA duplex is accomplished by many different proteins and enzymes that function in a highly coordinated manner [71]. Three DNA polymerases are responsible for chromosomal DNA replication: Pol α, Pol δ, and Pol ε. Polα acts as a DNA primase for both the leading and the lagging strand, Polε synthesizes primarily the leading strand while Polδ predominantly elongates the lagging strand [75]. Historically, only Pol α, Pol β and Pol γ were identified in mammals. Polymerase α was considered the replicative polymerase although it did not possess an intrinsic 3' 5' exonuclease activity such as the one observed in *E.coli* or T4 Bacteriophage polymerases. In the late 70s however, the discovery of DNA polymerase δ with intrinsic 3'-5' exonuclease proofreading activity entailed a great advance [76, 77]. Shortly after, a second polymerase with intrinsic 3'-5' exonuclease activity and proofreading function was identified, named polymerase ε [78]. Moreover, the discovery of polymerase δ is associated to the identification of PCNA as its processivity/auxiliary factor [79]. Polymerase δ belongs to the family of B polymerases and it is primarily involved in the replication of the lagging strand as well as the maturation of Okazaki fragments. It elongates primers synthesized on the lagging strand in 5'-3' direction until it reaches the 5' terminus of the previous Okazaki fragment. In eukaryotes, each Okazaki fragment is initiated by the DNA pol α/primase complex that place cRNA/DNA

hybrid primers every 100–250 nucleotides (nt). The replication protein A (RPA) binds and protects intermittent single-stranded DNA (ssDNA) from nucleases degradation and avoids the formation of undesired DNA structures. The clamp loader RFC, recognize these hybrids primers and loads PCNA onto the DNA, so that the front side of the ring face the 3' end of the nascent primer template junction. Polymerase δ then enters to constitute the holoenzyme necessary to initiate DNA synthesis [80] (Figure 1.10).
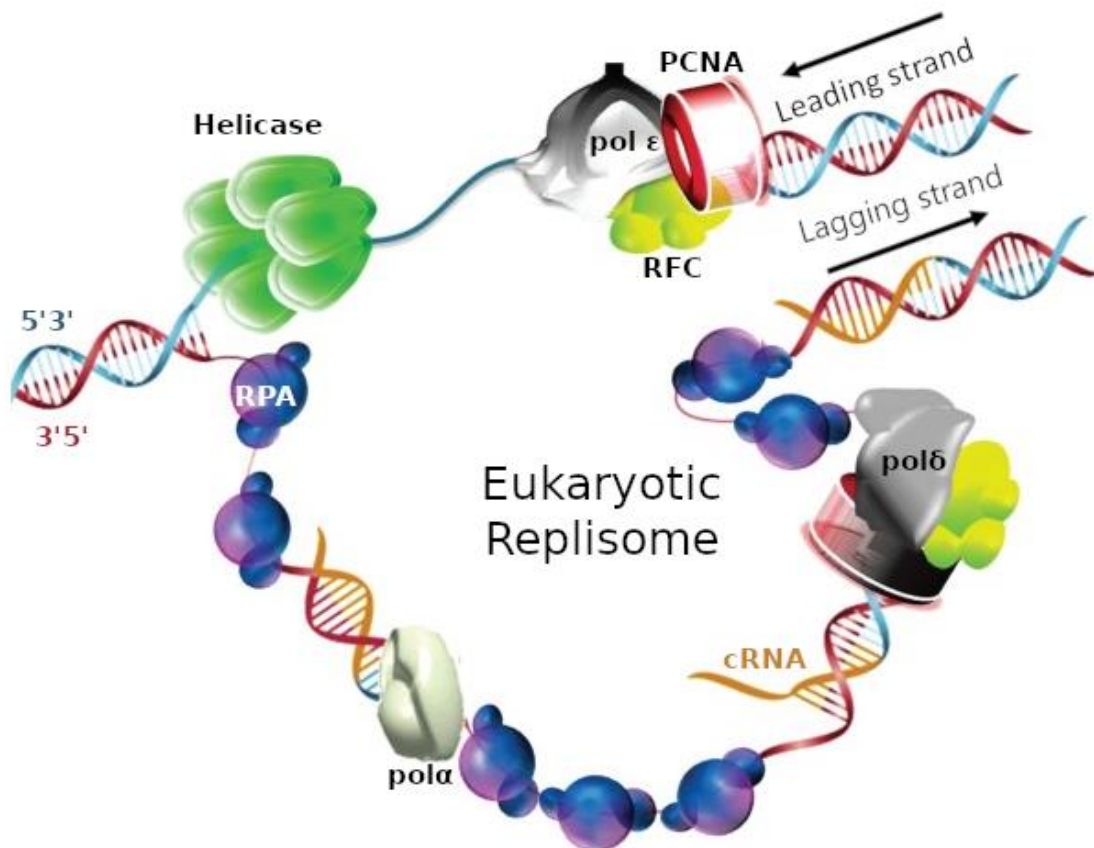

Figure 1.10. Eukaryotic replication fork. DNA helicase unwinds the duplex. On the leading strand, replication factor C (RFC) loads PCNA and pol ε to synthesize the leading strand.. The lagging strand is primed by polα and then pol δ enters to elongate discontinuously. The single-stranded DNA (ssDNA) is coated by replication protein A (RPA) (figure modified from Balakrishnan et al., 2013). The orange strand represents the primed DNA synthetized by the polymerase α.

According to biochemical and genetic assays, the initiator cRNA of the 5' is removed by a coordinated action of polymerase δ and FEN-1 before fragments ligation [81, 82]. In addition, pol δ possess an 3'-5' exonuclease activity, which is important not only to detect and eliminate

miss-incorporated nucleotides but also for the maturation of Okazaki fragments [75]. It also participates in DNA repair in gap filling processes and it is recruited to DNA lesion sites.

Initially, only two subunits of mammalian polymerase delta were defined: p125, the catalytic subunit that harbours the polymerase and exonuclease activity and p50, which tightly associate with p125. Later, two other accessory subunits were discovered: p68 (also known as p66) [83] and p12 [84]. Saccharomyces cerevisiae (*S.cerevisiae*) Pol δ has only three subunits, the equivalents of p125, p50 and p68 whereas Schizosaccharomyces pombe (*S. pombe*) Pol δ has an additional fourth subunit, Cdm1 which presents only 25% homology to p12 [84,85] (Table1.2) . In contrast, the catalytic domain p125 of human shares over 60% similarity with those of *S. cerevisiae* and *S. pombe* Pol δ, suggesting a strong evolutionary conservation. These evidences indicate that the fourth subunit p12 exists as an evolutionary adaptation of higher eukaryotes to a more complex chromatin organization due to a larger genome [86].

Table 1.2. Equivalent subunits of polymerase δ in *Homo Sapiens, Saccharomyces pombe* and *Saccharomyces cerevisiae.*

H. Sapiens	p125	p50	p68	p12
S. Pombe	Pol3	Cdc1	Cdc27	Cdm1
S. Cerevisiae	Pol3	Pol31	Pol32	-

1.2.5.2. Functional roles of human p12

The fourth and smallest subunit of human polymerase δ is a 12 KDa protein which is degraded in vivo in response to DNA damage during the S-phase [87, 88]. Degradation of p12 is mediated by ubiquitination and subsequent proteasomal degradation, leading to the conversion of a three subunit polymerase δ3 [89]. Two ubiquitin E3 ligases have been described to ubiquitinate p12: RNF8 and CRL4[Cdt2], although the number of E3 ligases involved in p12 regulation has to be further investigated .

Recent biochemical and cellular studies demonstrated that the loss of the small subunit of the polymerase does not affect cell viability and leads to a completely functional form of the enzyme with an increased proofreading capacity [90]. Pol δ3 is less prone to introduce incorrect

nucleotides during DNA synthesis and therefore, it is the preferred form under DNA stress or damage conditions with higher rate of mutations [91] .

Moreover, kinetic assays allow measuring the ratio of excision/polymerization, which provides an index of the intrinsic proofreading capacity of the polymerase. When pol δ encounters a wrong nucleotide leads to a conformational change in the polymerase that produces the decrease of polymerization rate and facilitates the dissociation of the incorrect base [92, 93]. The ratio of excision/polymerization is significantly higher (40 times) in polymerase δ3 than in its tetrameric form and consequently is predominant in DNA repair [76]. Cellular and biochemical results show that polymerase δ3 is the major form during the S phase of the cell cycle and therefore, is responsible for the replication of the lagging strand, since it synthesizes the DNA with greater fidelity. It is suggested a possible role of CRL4^{Ctd2} as a regulator of p12 degradation during the S phase, which also ubiquitinates p21 or Ctd1. CRL4^{Ctd2} recognizes its substrates when PCNA is loaded onto DNA by the RFC, which occurs in primer template junction, before the initiation of replication [94, 95, 96].

Although the exact function of Pol δ4 holoenzyme is not clear, recent evidence shows that it is essential for HR [91]. Homologous recombination is a mechanism that repairs a variety of DNA lesions, including double-strand DNA breaks (DSBs), single-strand DNA gaps and interstrand crosslinks [97]. p12 knock out cells exhibit defects in homologous recombination and consequently are more sensitive to genotoxic agents that can lead to DNA strand breaks and subsequently the development of tumours. Future work has to be done regarding the potential impact of the regulation of pol δ3/pol δ4 balance and the function of the tetrameric form of pol δ [91, 98] (Figure 1.11).
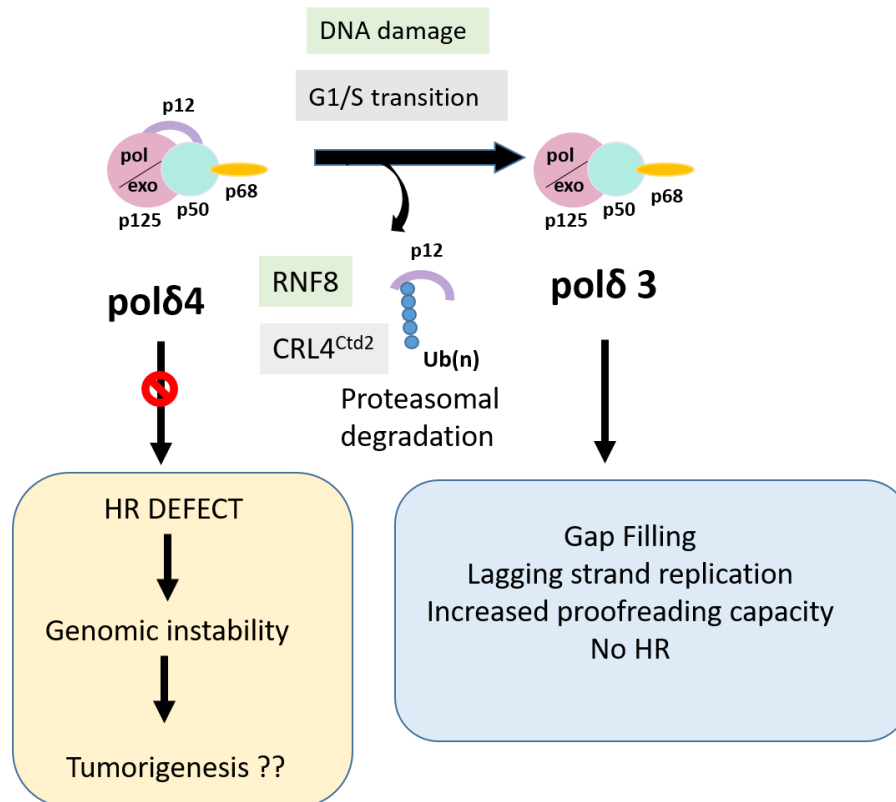
Fig 1.11. Pol δ4 conversion to pol δ3 in response to DNA damage and during the G1/S transition. The functions of polymerase δ3 are indicated in the blue square. The deficiency of pol δ4 leads to a defects in homologous recombination, that might produce alterations that eventually produce cancer (yellow square) .

1.2.5.3. Structure and interaction with PCNA

The structural characterization of all the subunits of human polymerase δ and its interaction with PCNA is a goal that remains elusive, and would provide very valuable information to understand its mechanism of action. As mentioned above, polymerase δ belongs to the family of B polymerases that are present in all kingdoms of life and even in viruses. Several of these structures have been solved, except in eukaryotes. Due to the high conservation of the catalytic subunit, it is likely to exhibit very similar structure. However in order to function properly Pol δ needs the accessory subunits and until recently there was no structural information about any of these regulatory domains [99].

The catalytic core of pol δ is formed by a tight heterodimer of p125 and p50. The third subunit, p68, binds p50 through is N-terminal domain while p12 is believed to bridge p125 and p50 subunits [100]. The only subunit with catalytic activity is p125 whereas the other three

perform a regulatory role by stimulating p125 and stabilizing the whole complex. p50 acts as a scaffold that interacts simultaneously with the other three subunits and facilitates the recruitment of diverse proteins involved in DNA metabolism such as p21 [101].

Yeast two hybrid assays and co-immunoprecipitation experiments reported that three of the four subunits interact with PCNA: p125, p68 and p12 subunits. Nonetheless, only the interaction of p68 has been structurally characterized. In particular, p68 binds through a canonical PIP box to PCNA with an affinity of 1.5 μM (Table 1.1). While the interaction of p125-PCNA seems to be very weak [87], mutations of p68 and p12 PIP boxes affect the activity of pol δ [88, 102].

Pulldown experiments with different GST-fused peptides spanning the sequence of p12 indicate that the p12 PCNA binding site is located the N-terminus and is highly divergent from the canonical PIP sequences [87].

1.3. References

1. Kuriyan, J. Sliding Clamps of DNA Polymerases. *J. Mol. Biol.* (1993).

2. Moldovan, G. L., Pfander, B. & Jentsch, S. PCNA, the Maestro of the Replication Fork. *Cell* **129**, 665–679 (2007).

3. Stukenberg, P. T., Studwell-Vaughan, P. S. & O'Donnell, M. Mechanism of the sliding β-clamp of DNA polymerase III holoenzyme. *J. Biol. Chem.* **266**, 11328–11334 (1991).

4. Kelch, B. A., Makino, D. L., Donnell, M. O. & Kuriyan, J. Clamp loader ATPases and the evolution of DNA replication machinery. 1–14 (2012).

5. Jeruzalmi, D., O'Donnell, M. & Kuriyan, J. Clamp loaders and sliding clamps. *Curr. Opin. Struct. Biol.* **12**, 217–224 (2002).

6. Kelch, B. A. Review The Lord of the Rings: Structure and Mechanism of the Sliding Clamp Loader. **105**, 532–546 (2016).

7. Indiani, C. & Donnell, M. O. The replication clamp-loading machine at work in the three domains of life. **7**, 751–761 (2006).

8. Krishna, T. S. R., Kong, X. P., Gary, S., Burgers, P. M. & Kuriyan, J. Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA. *Cell* **79**, 1233–1243 (1994).

9. De Biasio, A. & Blanco, F. J. *Proliferating cell nuclear antigen structure and interactions: Too many partners for one dancer? Advances in Protein Chemistry and Structural Biology* **91**, (Elsevier Inc., 2013).

10. Kong, X. P., Onrust, R., O'Donnell, M. & Kuriyan, J. Three-dimensional structure of the β subunit of E. coli DNA polymerase III holoenzyme: A sliding DNA clamp. *Cell* **69**, 425–437 (1992).

11. Hedglin, M., Kumar, R. & Benkovic, S. J. Replication Clamps and Clamp Loaders. 1–19 (2013).

12. Matsumiya, S. & Ishino, Y. Crystal structure of an archaeal DNA sliding clamp : Proliferating cell nuclear antigen from Pyrococcus furiosus. 17–23 (2001). doi:10.1110/ps.36401.system.

13. Wing, R. *et al.* Crystal structure analysis of a complete turn of B-DNA. *Nature* **287**, 755–758 (1980).

14. Kelman, Z. & Donnell, M. O. Structural and functional similarities of prokaryotic and eukaryotic DNA polymerase sliding clamps. **23**, 3613–3620 (1995).

15. Georgescu, R. E. *et al.* Structure of a Sliding Clamp on DNA. *Cell* **132**, 43–54 (2008).

16. Fernandez-Leiro, R., Conrad, J., Yang, J., Freund, S. M. V & Sjors, H. W. Self-correcting mismatches during high-fidelity DNA replication. **24**, 140–143 (2017).

17. McNally, R., Bowman, G. D., Goedken, E. R., O'Donnell, M. & Kuriyan, J. Analysis of the role of PCNA-DNA contacts during clamp loading. *BMC Struct. Biol.* **10**, (2010).

18. Miyachi, K., Fritzler, M. J. & Tan, E. M. Autoantibody to a Nuclear Antigen in Proliferating Cells.

J. Immunol. **121**, 2228 LP – 2234 (1978).

19. Kelman Z. PCNA: structure, functions and interactions. *Oncogene* **14**, 629-640. (1997).

20. Bravo, R. & Celis, J. E. A search for differential polypeptide synthesis throughout the cell cycle of hela cells. *J. Cell Biol.* **84**, 795–802 (1980).

21. Tsurimoto, T. & Stillman, B. Replication factors required for SV40 DNA replication in vitro. I. DNA structure-specific recognition of a primer-template junction by eukaryotic DNA polymerases and their accessory proteins. *J. Biol. Chem.* **266**, 1950–1960 (1991).

22. Slade, D. Maneuvers on PCNA Rings during DNA Replication and Repair. *Genes (Basel).* **9**, 416 (2018).

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

24. Kontopidis, G. *et al.* Structural and biochemical studies of human proliferating cell nuclear antigen complexes provide a rationale for cyclin association and inhibitor design. (2004).

25. Sánchez, R., Torres, D., Prieto, J., Blanco, F. J. & Campos-Olivas, R. Backbone assignment of human proliferating cell nuclear antigen. *Biomol. NMR Assign.* **1**, 245–247 (2007).

26. de Biasio, A. *et al.* Reduced stability and increased dynamics in the human Proliferating Cell Nuclear Antigen (PCNA) relative to the yeast homolog. *PLoS One* **6**, (2011).

27. Naryzhny, S. N., Zhao, H. & Lee, H. Proliferating cell nuclear antigen (PCNA) may function as a double homotrimer complex in the mammalian cell. *J. Biol. Chem.* **280**, 13888–13894 (2005).

28. Kochaniak, A. B. *et al.* Proliferating cell nuclear antigen uses two distinct modes to move along DNA. *J. Biol. Chem.* **284**, 17700–17710 (2009).

29. De March, M. *et al.* Structural basis of human PCNA sliding on DNA. *Nat. Commun.* **8**, 13935 (2017).

30. De March, M. & De Biasio, A. The dark side of the ring : role of the DNA sliding surface of PCNA. *Crit. Rev. Biochem. Mol. Biol.* **0**, 1–11 (2017).

31. Cazzalini, O. *et al.* CBP and p300 acetylate PCNA to link its degradation with nucleotide excision repair synthesis. *Nucleic Acids Res.* **42**, 8433–8448 (2014).

32. Billon, P., Li, J., Lambert, J., Sugiyama, T. & Brunzelle, J. S. Acetylation of PCNA Sliding Surface by Eco1 Promotes Genome Stability through Homologous Recombination. 78–90 (2017). doi:10.1016/j.molcel.2016.10.033

33. De March, M. *et al.* p15 PAF binding to PCNA modulates the DNA sliding surface. 1–13 (2018). doi:10.1093/nar/gky723

34. Yao, N. Y. & O'Donnell, M. DNA Replication: How Does a Sliding Clamp Slide? *Curr. Biol.* **27**, R174–R176 (2017).

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

36.   Prestel, A. *et al.* The PCNA interaction motifs revisited: thinking outside the PIP-box. *Cell. Mol. Life Sci.* (2019). doi:10.1007/s00018-019-03150-0

37.   Havens, C. G. & Walter, J. C. Docking of a specialized PIP box onto chromatin-bound PCNA creates a degron for the ubiquitin ligase CRL4Cdt2. *Mol Cell* **23**, 1–7 (2009).

38.   Bruning, J. B. & Shamoo, Y. Structural and Thermodynamic Analysis of Human PCNA with Peptides Derived from DNA Polymerase- ▯ p66 Subunit and Flap Endonuclease-1. **12**, 2209–2219 (2004).

39.   De Biasio, A. *et al.* Proliferating cell nuclear antigen (PCNA) interactions in solution studied by NMR. *PLoS One* **7**, e48390 (2012).

40.   Wang, S.-C. *et al.* Tyrosine phosphorylation controls PCNA function through protein stability. *Nat. Cell Biol.* **8**, 1359–1368 (2006).

41.   Kannouche, P. L., Wing, J. & Lehmann, A. R. Interaction of human DNA polymerase η with monoubiquitinated PCNA: A possible mechanism for the polymerase switch in response to DNA damage. *Mol. Cell* **14**, 491–500 (2004).

42.   Hedglin, M., Pandey, B. & Benkovic, S. J. Characterization of human translesion DNA synthesis across a UV-induced DNA lesion. 1–18 (2016). doi:10.7554/eLife.19788

43.   Hishiki, A. *et al.* Structural basis for novel interactions between human translesion synthesis polymerases and proliferating cell nuclear antigen. *J. Biol. Chem.* **284**, 10552–10560 (2009).

44.   Gilljam, K. M. *et al.* Identification of a novel, widespread, and functionally important PCNA-binding motif. *J. Cell Biol.* **186**, 645–654 (2009).

45.   Olaisen, C., Müller, R., Nedal, A. & Otterlei, M. PCNA-interacting peptides reduce Akt phosphorylation and TLR-mediated cytokine secretion suggesting a role of PCNA in cellular signaling. *Cell. Signal.* **27**, 1478–1487 (2015).

46.   Click, T. H., Ganguly, D. & Chen, J. Intrinsically disordered proteins in a physics-based world. *Int. J. Mol. Sci.* **11**, 5292–5309 (2010).

47.   Wright, P. E. & Dyson, H. J. Intrinsically disordered proteins in cellular signalling and regulation. *Nat. Rev. Mol. Cell Biol.* **16**, 18–29 (2015).

48.   Cordeiro, T. N., Chen, P., Biasio, A. De & Sibille, N. Disentangling polydispersity in the PCNA – p15 PAF complex , a disordered , transient and multivalent macromolecular assembly. **45**, 1501–1515 (2017).

49.   Iakoucheva, L. M., Brown, C. J., Lawson, J. D., Obradović, Z. & Dunker, A. K. Intrinsic disorder in cell-signaling and cancer-associated proteins. *J. Mol. Biol.* **323**, 573–84 (2002).

50.   De Biasio, A. *et al.* P15PAF Is an intrinsically disordered protein with nonrandom structural

preferences at sites of interaction with other proteins. *Biophys. J.* **106**, 865–874 (2014).

51. Ganguly, D., Zhang, W. & Chen, J. Synergistic folding of two intrinsically disordered proteins: Searching for conformational selection. *Mol. Biosyst.* **8**, 198–209 (2012).

52. Knott, M. & Best, R. B. A preformed binding interface in the unbound ensemble of an intrinsically disordered protein: Evidence from molecular simulations. *PLoS Comput. Biol.* **8**, (2012).

53. Yu, P. *et al.* p15(PAF), a novel PCNA associated factor with increased expression in tumor tissues. *Oncogene* **20**, 484–489 (2001).

54. Xie, C., Yao, M. & Dong, Q. Proliferating cell unclear antigen-associated factor (PAF15): A novel oncogene. *International Journal of Biochemistry and Cell Biology* **50**, 127–131 (2014).

55. Simpson, F. *et al.* The PCNA-associated factor KIAA0101/p15PAF binds the potential tumor suppressor product p33ING1b. *Exp. Cell Res.* **312**, 73–85 (2006).

56. Mailand, N., Gibbs-Seymour, I. & Bekker-Jensen, S. Regulation of PCNA-protein interactions for genome stability. *Nat. Rev. Mol. Cell Biol.* **14**, 269–82 (2013).

57. Emanuele, M. J., Ciccia, A., Elia, A. E. H. & Elledge, S. J. Proliferating cell nuclear antigen (PCNA)-associated KIAA0101/PAF15 protein is a cell cycle-regulated anaphase-promoting complex/cyclosome substrate. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 9845–9850 (2011).

58. Hosokawa, M. *et al.* Oncogenic role of KIAA0101 interacting with proliferating cell nuclear antigen in pancreatic cancer. *Cancer Res.* **67**, 2568–76 (2007).

59. De Biasio, A. *et al.* Structure of p15(PAF)-PCNA complex and implications for clamp sliding during DNA replication and repair. *Nat. Commun.* **6**, 6439 (2015).

60. E. Sale, J., R.Lehmann, A. and & Woodgate, R. Y-family DNA polymerases and their role in tolerance of cellular DNA damage. *Nat Rev Mol Cell Biol.* **13**, 141–152 (2012).

61. Boehm, E. M., Spies, M. & Washington, M. T. PCNA tool belts and polymerase bridges form during translesion synthesis. *Nucleic Acids Res.* **44**, 8250–8260 (2016).

62. Povlsen, L. K. *et al.* Systems-wide analysis of ubiquitylation dynamics reveals a key role for PAF15 ubiquitylation in DNA-damage bypass. *Nat. Cell Biol.* **14**, 1089–1098 (2012).

63. M. Boehm, E., Gildenberg, M. S. and & M. Washington, T. The many roles of PCNA in eukaryotic DNA replication. *Enzymes* **25**, 289–313 (2016).

64. Karg, E. *et al.* Ubiquitome analysis reveals PCNA-associated factor 15 (PAF15) as a specific ubiquitination target ofUHRF1 in embryonic stem cells. *J. Mol. Biol.* **15**, (2017).

65. Hervouet, E., Peixoto, P., Delage-Mourroux, R., Boyer-Guittaut, M. & Cartron, P. F. Specific or not specific recruitment of DNMTs for DNA methylation, an epigenetic dilemma. *Clin. Epigenetics* **10**, 1–18 (2018).

66. Bostick, M. *et al.* UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science (80-.).* **317**, 1760–1764 (2007).

67. Pradhan, M. *et al.* CXXC domain of human DNMT1 is essential for enzymatic activity. *Biochemistry* **47**, 10000–10009 (2008).

68. Xue, B. *et al.* Epigenetic mechanism and target therapy of uhrf1 protein complex in malignancies. *Onco. Targets. Ther.* **12**, 549–559 (2019).

69. Ishiyama, S. *et al.* Structure of the Dnmt1 Reader Module Complexed with a Unique Two-Mono-Ubiquitin Mark on Histone H3 Reveals the Basis for DNA Methylation Maintenance. *Mol. Cell* **68**, 350-360.e7 (2017).

70. Garvilles, R. G. *et al.* Dual functions of the RFTS domain of dnmt1 in replication-coupled DNA methylation and in protection of the genome from aberrant methylation. *PLoS One* **10**, 1–19 (2015).

71. Zhang, D. & O'Donnell, M. *The Eukaryotic Replication Machine. Enzymes* **39**, (Elsevier Inc., 2016).

72. Hübscher, U., Maga, Giovanni & Spadari, S. Eukaryotic dna polymerases. *Annu. Rev. Biochem.* 39–92 (2002). doi:10.1201/9781351071857

73. Balakrishnan, L. & Bambara, R. A. Okazaki Fragment Metabolism. *Cold Spring Harb Perspect Biol 2013;5a010173* (2013). doi:10.1016/B978-0-12-374984-0.01087-1

74. Burgers, P. M. J. & Kunkel, T. A. Eucariotic DNA Replication Fork. *Annu. Rev. Biochem.* **86**, 417–438 (2017).

75. Stodola, J. L. & Burgers, P. M. Mechanism of lagging-strand DNA replication in eukaryotes. *Adv. Exp. Med. Biol.* **1042**, 117–133 (2017).

76. Lee, M. Y. W. T., Xiaoxiao, W., Zhang, S., Zhang, Z. & Y.C. Lee, E. Regulation and Modulation of Human DNA Polymerase δ Activity and Function. *Genes (Basel).* **8**, 190 (2017).

77. Byrnes, J. J., Downey, K. M., Black, V. L. & So, A. G. A New Mammalian DNA Polymerase with 3'to 5'Exonuclease Activity: DNA Polymerase δ. *Biochemistry* **15**, 2817–2823 (1976).

78. Pospiech, H. & Syväoja, J. E. DNA Polymerase e - More Than a Polymerase. *Sci. World J.* **3**, 87–104 (2003).

79. Prelich, G. *et al.* Functional identity of proliferating cell nuclear antigen and a DNA polymerase-δ auxiliary protein. *Nature* **326**, 517–520 (1987).

80. Hedglin, M., Pandey, B. & Benkovic, S. J. Stability of the human polymerase δ holoenzyme and its implications in lagging strand DNA synthesis. (2016). doi:10.1073/pnas.1523653113

81. Grasby, J. A., Finger, L. D., Tsutakawa, S. E., Atack, J. M. & Tainer, J. A. Unpairing and gating: Sequence-independent substrate recognition by FEN superfamily nucleases. *Trends Biochem.*

Sci. **37**, 74–84 (2012).

82. Balakrishnan, L. & Bambara, R. A. Flap Endonuclease 1. *Annu. Rev. Biochem.* **82**, 119–138 (2013).

83. Hughes, P., Tratner, I., Ducoux, M., Piard, K. & Baldacci, G. Isolation and identification of the third subunit of mammalian DNA polymerase δ by PCNA-affinity chromatography of mouse FM3A cell extracts. *Nucleic Acids Res.* **27**, 2108–2114 (1999).

84. Liu, L., Mo, J. Y., Rodriguez-Belmonte, E. M. & Lee, M. Y. W. T. Identification of a fourth subunit of mammalian DNA polymerase δ. *J. Biol. Chem.* **275**, 18739–18744 (2000).

85. Kesti, T., Frantti, H. & Syvaoja, J. E. Molecular cloning of the cDNA for the catalytic subunit of human DNA polymerase δ *J. Biol. Chem.* **268**, 10238–10245 (1993).

86. Lee, M. Y. W. T. *et al.* The tail that wags the dog. 23–31 (2014).

87. Li, H. *et al.* Functional Roles of p12 , the Fourth Subunit of Human DNA Polymerase δ . **281**, 14748–14755 (2006).

88. Zhang, S. *et al.* A novel DNA damage response: Rapid degradation of the p12 subunit of DNA polymerase δ *J. Biol. Chem.* **282**, 15330–15340 (2007).

89. Meng, X. *et al.* DNA damage alters DNA polymerase δ to a form that exhibits increased discrimination against modified template bases and mismatched primers. *Nucleic Acids Res.* **37**, 647–657 (2009).

90. Barton Laws, M. *et al.* The p12 subunit of human polymerase delta modulates the rate and fidelity of DNA synthesis. *NIH Public access* **17**, 148–159 (2013).

91. Lee, M. Y. W. T. *et al.* Two forms of human DNA polymerase δ: Who does what and why? *DNA Repair (Amst).* 102656 (2019). doi:10.1016/j.dnarep.2019.102656

92. Kunkel, T. A. DNA Replication Fidelity Downloaded from. *J. Biol. Chem.* **279**, 16895–16898 (2004).

93. A. Kunkel, T. and & Bebenek, K. DNA REPLICATION FIDELITY. *Annu. Rev. Biochem.* **69**, 497–529 (2000).

94. Terai, K., Shibata, E., Abbas, T. & Dutta, A. Degradation of p12 subunit by CRL4Cdt2 E3 ligase inhibits fork progression after DNA damage. *J. Biol. Chem.* **288**, 30509–30514 (2013).

95. Darzynkiewicz, Z. *et al.* Initiation and termination of DNA replication during S phase in relation to cyclins D1, E and A, p21WAF1, Cdt1 and the p12 subunit of DNA polymerase d revealed in individual cells by cytometry. *Oncotarget* **6**, 11735–11750 (2015).

96. Lin, S. H. S. *et al.* Dynamics of enzymatic interactions during short flap human Okazaki fragment processing by two forms of human DNA polymerase δ. *DNA Repair (Amst).* **12**, 922–935 (2013).

97. Krejci, L., Altmannova, V., Spirek, M. & Zhao, X. Homologous recombination and its regulation.

33

Nucleic Acids Res. **40**, 5795–5818 (2012).

98.   Zhang, S. *et al.* Loss of the p12 subunit of DNA polymerase delta leads to a defect in HR and sensitization to PARP inhibitors. *DNA Repair (Amst).* **73**, 64–70 (2019).

99.   Baranovskiy, A. G. *et al.* X-ray structure of the complex of regulatory subunits of human DNA polymerase delta. *Cell cycle* **7**, 3026–3036 (2008).

100.   Xie, B. *et al.* Reconstitution and characterization of the human DNA polymerase delta four-subunit holoenzyme. *Biochemistry* **41**, 13133–13142 (2002).

101.   Li, H., Xie, B., Rahmeh, A., Zhou, Y. & Lee, M. Y. W. T. Direct Interaction of p21 with p50, the Small Subunit of Human DNA Polymerase Delta Hao. *Cell Cycle* **5**, 585–588 (2006).

102.   Rahmeh, A. A. *et al.* Phosphorylation of the p68 subunit of pol δ acts as a molecular switch to regulate its interaction with PCNA. *Biochemistry* **51**, 416–424 (2012).

2.Objetives

This doctoral thesis aims at the structural study of the interactions between PCNA with two proteins involved in DNA replication and repair: p15 and p12.

The first part of the thesis focuses on the study of the effect of p15 ubiquitination on the molecular recognition properties of this oncogenic protein. The specific objectives are:

- Conformational characterization of doubly monoubiquitinated p15 (dmUbp15).
- Structural characterization of dmUbqp15 binding to PCNA.
- Structural characterization of dmUbqp15 binding to DNA.
- Characterization of the possible interaction between dmUbp15 and the RFTS domain of Dnmt1.

The fourth objective was defined after it was reported (during the course of this thesis) that doubly monoubiquitinated histone H3 binds the RFTS domain of Dnmt1 methyl transferase.

The second part of the thesis focuses on the study of the molecular recognition of polymerase δ p12 subunit by PCNA. The specific objectives of this part are:

- Structural characterization of the interaction between a fragment of polymerase δ and PCNA.
- Comparative structural study of canonical and non-canonical PCNA binding motives.

3.Methods

3.1. Circular Dichroism

Circular dichroism (CD) spectroscopy is a quick method to identify the secondary structure content of a protein in solution. It is based in the differential absorption of left- and right-handed circularly polarized light by certain molecules [1]. This differential absorption may be expressed in units of molar ellipticity ($[\theta]$, deg.cm^2.dmol^{-1}). Since CD is an absorption event, the chromophores that contribute to the CD spectrum are the same that contribute to the protein absorption spectrum, but the chromophores should be either chiral or be in an asymmetric environment in order to produce a CD signal. The protein chromophores that contribute the most to the far UV (below 250 nm) CD signal are the peptide bonds (with absorbance transitions between 220 and 190 nm), although some side chains may also contribute to this region, like Tyr, Trp, Phe, His and Met.

The secondary structure motifs of proteins present characteristic far-UV CD spectra. For instance, α-helical proteins have negative bands at 222 nm and 208 nm and a positive band at 193 nm whereas antiparallel β-sheets have negative bands at 218 nm and a positive band at 195 nm. In the case of disordered proteins, they exhibit the typical random coil CD spectrum with low ellipticity above 210 and a strong negative band at around 195 nm [2] (Figure 3.1).
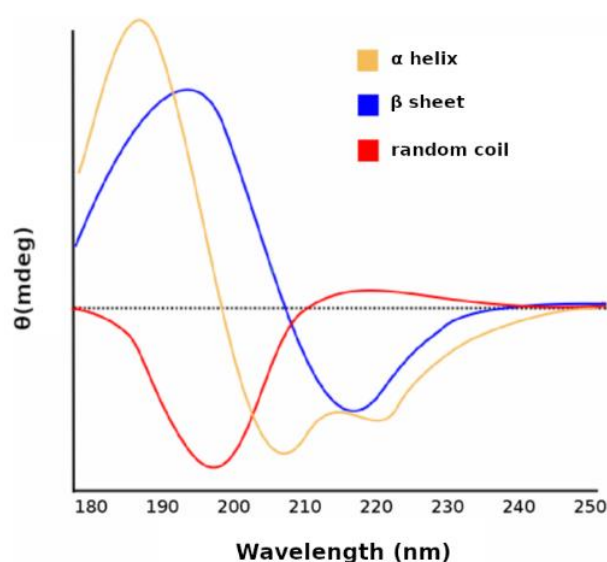

Figure 3.1. Circular dichroism reference spectra for 100% of each type of secondary structure.

3.2. Size Exclusion Chromatography-Multi angle static Light Scattering (SEC-MALS)

Like all light scattering methods, including X-ray diffraction SAXS and, MALS is based on the fact that electromagnetic radiation is scattered by inducing polarization in electrons comprising molecular matter [3]. MALS measures the scattering of a red laser by a protein in solution at different angles. The measured intensities depend on the molar mass and concentration of the protein [4]. Thus, MALS allows to calculate the molar mass of the protein and, therefore, to determine its oligomeric state. It is often coupled with size exclusion chromatography (SEC), which serves to separate the protein from possible impurities and large aggregates. Size Exclusion Chromatography by itself is frequently used to determine the size of a protein, when the column elution volume is calibrated with proteins of known size, but this procedure is only reliable for globular proteins and that do not interact with the column matrix. It is particularly unsuited for disordered proteins, or proteins with long disordered regions, as these macromolecules typically elute at smaller volumes than a globular protein of the same mass, overestimating its size. SEC coupled with a MALS detector allows to unambiguously determine the molar mass independently of the protein shape and possible interaction with the column matrix [5].

3.3. Small Angle X Ray Scattering (SAXS)

The SAXS experiments presented in this thesis were recorded and analyzed in collaboration with Dr. Tiago Cordeiro from Instituto de Tecnología Química e Biológica António Xabier, ITQB NOVA (Portugal) and Dr. Pau Bernadó from Centre de Biochimie Structurale, INSERM, CNRS, and Université Montpellier (France).

In SAXS measurements the elastic scattering of X-rays by a protein solution at small angles (typically 0.1 - 10°) is registered, providing low-resolution (1-100 nm) structural information of the protein [6]. SAXS has emerged as a powerful tool to characterize highly dynamic proteins like intrinsically disordered proteins [7]. Together with fast methods for conformational ensemble generation, the SAXS data can be used to define an ensemble of structures that

represent the conformational state of a protein without a defined three-dimensional structure.

The solvent scattering must be subtracted from the scattering of the protein sample, and small differences in solvent composition can significantly affect SAXS data, this is the reason why SAXS is often perform in-line with SEC, which also serves to separate possible large aggregates [8] [9].

In a typical SAXS experiment, the scattering data are collected in a two-dimensional detector and are represented as a scattering curve in which the logarithm of the intensity is plotted against the transfer moment. The Fourier transform of the scattering curve makes it possible to calculate the distance distribution (Pr), from which the molar mass, the radius of gyration (Rg), and the maximum interatomic distance (Dmax) can be determined, as well as the overall three-dimensional shape of the protein.

3.4. X Ray Crystallography

X-Ray Crystallography, is the most powerful experimental technique to study macromolecular structures at atomic level. It allows obtaining the electron density in a given crystal from the intensity of the x rays diffracted by it. The scattered waves can either interact cooperatively or cancel out between them, which is known as constructive or destructive interference, respectively. A protein crystal functions as a three-dimensional diffraction grid that causes destructive and constructive interferences, creating a diffraction pattern that contains information about the distribution and type of atoms that build up the crystal. The crystal contains millions of repetitions of the same molecule that forms a three-dimensional ordered system, magnifying the detected signal. Due to their periodicity, crystals can decompose into simpler units related by translations in the three dimensions. The asymmetric unit is the minimum portion of the crystal that generates the unit cell by application of its symmetry elements. The repetition of the unit cell by translational operators form the entire crystal. For chiral molecules, like proteins, there are only 65 possible arrangements (space groups), that form a three dimensional periodic system through combinations of translation and rotation symmetry operations. Packing imperfections disrupt crystal symmetry, which results in poor

quality diffraction. Therefore, obtaining monocrystals of sufficient quality is one of the bottlenecks of X-ray crystallography.

3.4.6. *Crystallization*

In order to form crystals, it is necessary to prepare a highly pure protein solution that has to become supersaturated. A supersaturated solution is thermodynamically metastable, which helps nucleation and facilitates crystal growth while it reaches equilibrium. The most extended method for protein crystallization is the vapor diffusion technique, in which the protein solution is mixed with the precipitant solution and is placed in a closed container with a siliconized cover. The mixture drop (from 0.1 to 1 μl) would equilibrate against the reservoir containing the precipitant solution. Vapor diffusion crystallization can be performed in either hanging-drop or sitting-drop way. In general, the initial screening for suitable crystallization conditions is carried out in two different steps. First, a wide selection of chemical conditions, available from commercial kits, is analyzed. Secondly, conditions where crystals have grown are selected and refined in larger plates, varying different parameters such as the concentration of protein and precipitant, the pH as well as the presence of additives [10].

Once the crystals reach a sufficient size to be fished from the crystallization drops, they can be tested by X-ray diffraction (Figure 3.2). The objective is to collect diffraction patterns of the crystals in different angular positions with respect to the detector. With the help of a goniometer, the crystal is rotated in small angle intervals, typically 0.1°, in order to collect all the reflections without overlapping each other. The spots observed in a diffraction pattern are called reflections.
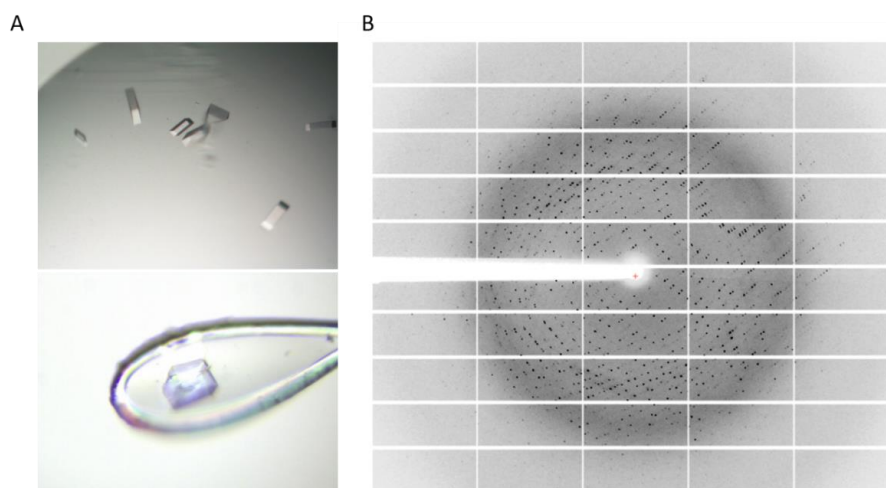
Figure 3.2. Crystallization and data collection. A) Crystals of the complex of a p12 fragment with PCNA in 15 well optimizing plates corresponding to the best crystallization conditions (above) and the best crystal inside a nylon loop. **B)** Diffraction pattern a 2.1 Å of the crystal in A obtained from synchrotron radiation at the XALOC beamline (ALBA, Barcelona).

3.4.7. Data Processing

Once the images have been collected, at different angles, the set of reflections have to be processed. First, the spatial group and the size of the unit cell are determined from the distances between the reflections. The space group can be established with the help of merging statistics, systematic absences, and space group preferences, although there may be ambiguities due to enantiomorphic space group pairs. In those cases, all possible space groups must be tested in the phasing stage. Then, in the indexing process, Miller indices are assigned to each reflection for a given spatial group. The next step is the integration process in which the intensity of each reflection is defined. Subsequently, the list of reflections of each diffraction pattern is combined with the total set of images and they are normalized on an absolute intensity scale. At this first stage, a descriptive statistical table is obtained with information on the space group, asymmetric unit, the resolution limit, and some other parameters that indicate the quality of the data.

The next step consists of extracting the structural information from the diffraction data and this needs to calculate both the amplitude and phase of the dispersed waves. The amplitude, which is related to the intensities, is obtained directly from the diffraction pattern, while the

phase, must be calculated by indirect methods. The "Phase problem" can be solved using different approaches such as molecular replacement, isomorphic replacement, or the introduction of anomalous dispersers.

Molecular Replacement can be used to calculate the phases when the structure of the same or a similar protein is available. Briefly, the Patterson vectors of the known structure are superimposed with those of the problem structure by rotation and translation operations. Once the orientation and position in the unit cell have been correctly determined, the structure factors of the protein under study can be defined using the calculated phases and the experimental amplitudes [11]. In this way, a preliminary electronic density map is obtained. However, the initial phases must be improved in order to obtain the final structure. This improvement process is termed the phase refinement and consist of a series of cyclic steps that add information to the electron density map. The final refined model must be mathematically validated to ensure its accuracy relative to the experimental data. One of the main indicators of the correct spatial distribution of the atoms is the Rfactor. This parameter is accompanied by a complementary value called R_{free}, in which a set of reflections is excluded from the refinement process. In addition, there are different stereochemical quality indicators such as the deviations of the bond distances and angles from ideal values (expressed as root mean square deviations, rmsd), and the Ramachandran plot, which defines the energetically allowed regions for backbone dihedral angles φ and Ψ of the amino acid residues. An important parameter in crystallography is the atomic B factor, also known as temperature factor, which reflects the corresponding dynamics of the atom [3].

3.5. Isothermal titration calorimetry (ITC)

Isothermal titration calorimetry is a biophysical technique that quantifies the reaction energy of an interaction between molecules and provides the thermodynamic parameters of the binding [12]. It is based on an adiabatic process, in which the released or absorbed heat of the reaction is balanced by a thermic core to keep the reference and the sample cells at the same temperature (Figure 3.3). One of the reagents (ligand) is loaded in an injection device and it is added in small volumes inside the cell containing the other reagent (protein). Between each

injection, the temperature of the two cells (reference and sample cell) is equalised. The increment or reduction of the applied energy by the thermic core in each injection is plotted against the molar ratio between the ligand and protein in the cell (binding isotherm). The stoichiometry of the reaction (n), the dissociation constant (K_d) as well as the enthalpy (ΔH) and the entropy change (ΔS) can be directly calculated from the isotherm. The molar ratio in the middle of the isotherm indicates the stoichiometry of the reaction while the curve adjustment to a binding model provides the K_d. The amount of the released energy allows to calculate the enthalpy and entropy variations. Finally, with all these data it is possible to calculate the change in the Gibbs free energy (ΔG) using the ecuation $\Delta G = \Delta H - T\Delta S$ [13].
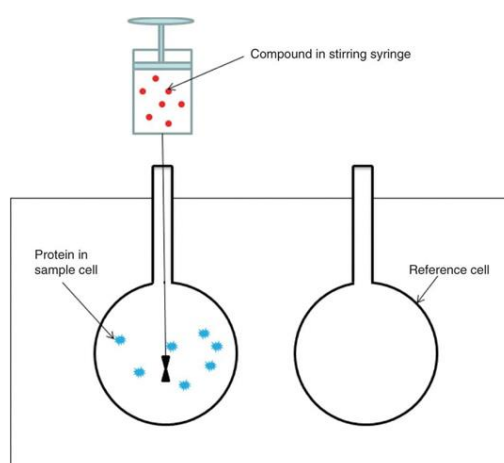
Figure 3.3. Scheme of an ITC instrument, consisting of a stirring syringe, a sample cell and a reference cell (figure obtained from Creative Biostructures web page)

3.6. Protein Nuclear Magnetic Resonance

Nuclear Magnetic Resonance Spectroscopy (NMR) is a versatile tool for studying the structure and dynamics of macromolecules. Macromolecules are made up by atoms, which are formed by nuclei and their surrounding electrons. The nuclei have four fundamental physical properties: mass, electric charge, magnetism and spin. Nuclear magnetism and spin make NMR possible. They have a negligible effect on the normal chemical and physical behaviour of substances at the macroscopic level, allowing NMR data acquisition without altering the chemical properties of the sample [14].

The spin of a nucleus is characterized by the quantum spin number I. Those nuclei with $I \neq 0$ are active in NMR. Moreover, those with $I = ½$ are especially favourable for NMR observation. They can be in either of two states which, in the presence of a magnetic field, differ in their

energy. The energy difference is in the frequency range of MHz, therefore radio waves are used to induce transitions between both states.

At equilibrium, the population of the two states produces a bulk magnetization aligned with the direction of the applied magnetic field (the Z axis). When the equilibrium magnetization is perturbed by a radiofrequency pulse, a transversal magnetization in the XY plane is generated. In a typical NMR experiment, the sample is irradiated by a radiofrequency pulse with the frequency necessary to excite a certain type of nuclei. The excited nuclei will return to the equilibrium state and the evolution of the transversal magnetization in the XY plane induces an oscillatory voltage (Free Induction Decay of FID) in the coil that surrounds the sample. The FID is a time-domain signal that has to be deconvoluted by the Fourier Transform, yielding a frequency-domain NMR spectrum. When a spin returns to equilibrium two relaxation events occur: longitudinal and transversal relaxation. The longitudinal relaxation is important to determine the velocity of the repetitive acquisitions in NMR experiments, and transverse relaxation is important for the total time of signal acquisition [15].

The nuclei within a molecule will resonate at different frequencies depending on their local chemical environment, the so called Chemical Shift (δ). The chemical shifts will depend on the intensity of the magnetic field applied, which means that depends on the spectrometer used. Besides, the difference in chemical shifts are very small to be reported in MHz. To overcome these problems chemical shifts are expressed relative to a reference compound, typically TMS (tetramethylsilane). In this way, the chemical shift is independent of the magnetic field and is expressed as ppm (parts per million of the field). Since TMS is not soluble in water, similar compounds are used instead, such as DSS (sodium trimethylsilylpropanesulfonate).

In protein NMR the resonance frequency of each kind of nuclei will depend on the chemical environment as well as on the amino acid type. Chemical environment is directly related with the 3D structure of the protein (secondary, tertiary and quaternary), ligand binding, or even by the solvent conditions, including pH, salt, or temperature. 1D NMR spectra are of limited utility to analyse proteins, because of the strong signal overlap due to the large number of atoms. Hence, in protein NMR it is necessary to perform multidimensional experiments to obtain intensity data at two or more frequencies (corresponding to two or more nuclei). In these experiments a series of separated pulses are applied during variable time lapses. The

subsequent Fourier transform of the direct and indirect dimensions results in a multidimensional spectrum which increases substantially the spectral resolution, facilitating the analysis.

The preferred nuclei in protein NMR are ^{1}H, ^{13}C and ^{15}N due to their favourable spin quantum number I=1/2 and their ubiquity in the polypeptide chain. Nevertheless, while ^{1}H is naturally abundant (99%), ^{15}N and ^{13}C are present only in the 0.4% and 1% of the nuclei at natural abundancy, respectively. Consequently, proteins are often produced isotopically enriched in ^{13}C and ^{15}N to carry out NMR experiments. The NMR study of large proteins (MW > 20KDa) presents two main challenges: low sensitivity caused by large transverse relaxation rates (due to the slow molecular tumbling in solution), and low resolution because of the large number of overlapping signals. Deuterium enrichment improves signal-to-noise ratio by reducing the transverse relaxation rate of the remaining protons while reducing the spectral complexity [16]. Therefore, deuterium labelling is frequently used with large proteins to improve the quality of the spectra.

3.6.1. *NMR spectral assignment*

One of the advantages of NMR over other forms of spectroscopy is that the excited states are long lived and we can manipulate them to pass excitation from one nucleus to another. Therefore, it is possible to obtain signals that correlate the frequencies of two or more nuclei. In the correlation spectra, each resonance is represented as an independent nuclear frequency axis, and those signals that appear at the intersection of two or more frequencies indicate a correlation between the corresponding nuclei. The information obtained from multidimensional spectra allow to determine unambiguously which signal of the spectrum correspond to which nuclei of the molecule [17].

Heteonuclear Single Quantum Coherence (HSQC) spectra are the most commonly recorded experiment in protein NMR, where the signals correspond to the correlation of two directly bound nuclei. Particularly useful is the ^{1}H ^{15}N-HSQC, because it correlates NH groups and provides information about each peptide bond. This spectrum is considered the "fingerprint" of a protein, since each signal belongs to one amino acid residue (apart from certain signals from the side chains of Arg, Trp, Gln, Asn and Lys) and it is very useful tool for protein-ligand

binding studies. Heteronuclear correlations can also be measured by the Transverse Relaxation Optimized Spectroscopy (TROSY) experiment, which is particularly useful for large deuterated proteins with slow tumbling rates, although it can also be used for non-deuterated proteins (increasing the resolution with a certain loss of sensitivity). TROSY experiment allows selecting the signal with the smaller transverse relaxation rate among the four components that produces an ^{15}N-^{1}H in a non-decoupled heteronuclear correlation spectrum [18].

Protein backbone assignment requires recording triple resonance experiments that enable frequency separation of signals overlapping in 2D spectra and frequency correlations between nuclei of the same or sequential amino acids. The most frequently recorded triple resonance experiments for the assignment of the backbone resonances are HNCO, HN(CA)CO, HNCA, HN(CO)CA, HNCACB, HN(CO)CACB, HN(CA)HA, and HN(COCA)HA which are explained in figure 3.4. HN(CO)CACB and HNCACB experiments are also used but they are not suitable for large proteins since their sensitivity is low [19].
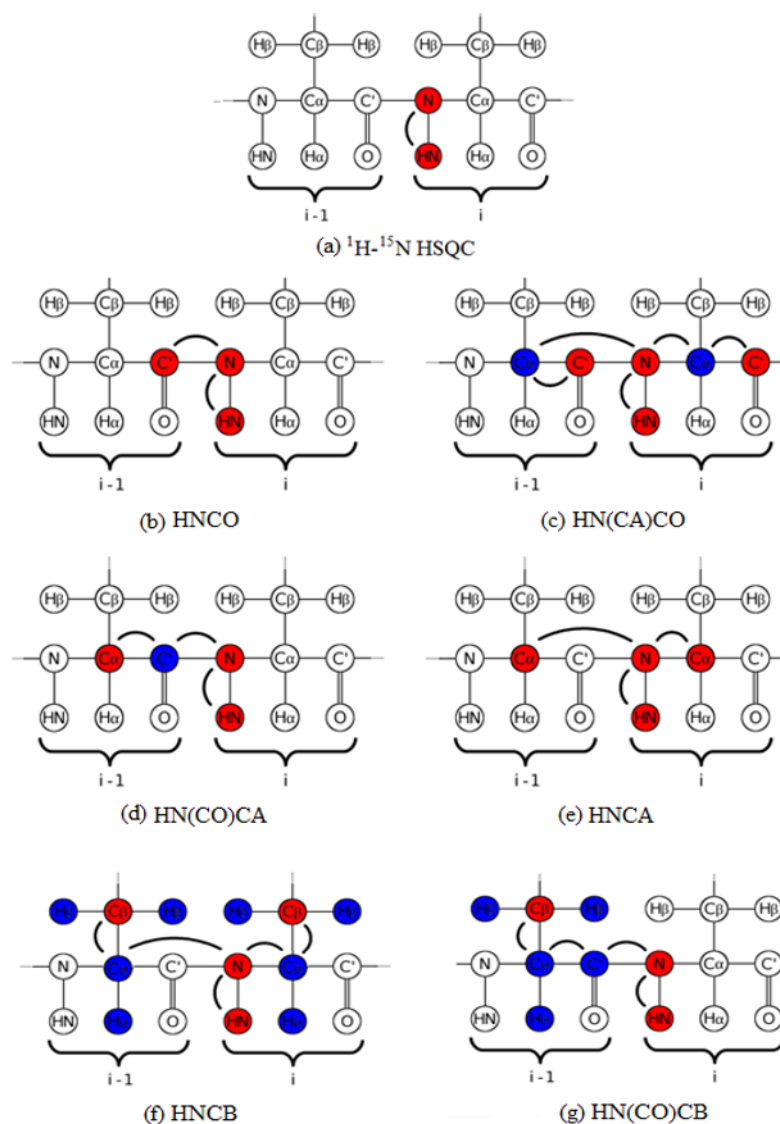
Figure 3.4. Magnetisation transfer diagrams for backbone assignment. Nuclei that participate in the transfer processes but are not observed are depicted in blue and nuclei whose frequency is measured are indicated in red.

3.6.2. Protein-ligand binding studies

Once the backbone resonances are assigned, it is very useful to map ligand binding by calculating Chemical Shift Perturbations (CSPs). In a typical experiment, a ligand (which can be a compound, DNA or another protein) is titrated into the labelled protein and each addition is monitored by subsequent 2D HSQC or TROSY spectra acquisitions. Chemical shifts are very sensitive to even slight environmental changes and therefore even a low affinity binding can be detected. The analysis consists of monitoring the peak shifts at each point of the titration, which allows calculating the CSPs (a number that indicates how much the frequency of a certain resonance in the protein has shifted). The peaks that present large CSPs

are most likely to be located at the binding site or nearby. Moreover, the shape of the titration curve (perturbations vs ligand/protein ratio) can be fitted to a binding model (typically assuming a single set of equivalent binding sites), yielding a dissociation (K_d) constant per residue. The most frequent method of calculating CSPs is measuring ^{1}H-^{15}N HSQC spectra, since it is sensitive, well resolved (better than ^{1}H-^{13}C HSQC spectra) and it is easy and cheap to obtain the protein sample. The standard way to calculate the CSPs is combining the ^{1}H and ^{15}N CSP values (in ppm), multiplying the ^{15}N shift changes by a scaling factor to compensate their different spectral widths. In this work, the formula used to calculate CSP from the ^{1}H and ^{15}N resonances is:

$$CSP = \sqrt{\frac{1}{2}\left(\Delta\delta_H^2 + \frac{\Delta\delta_N^2}{5}\right)}$$

The equation to calculate the dissociation constant from the CSP values is (assuming a simple 1:1 binding stoichiometry):

$$CSP = K_d + [P] + x[P] - \sqrt{\frac{(K_d + [P] + x[P])^2 - (4\,x[P]^2)}{2[P]CSPmax}}$$

Where K_d is the dissociation constant, [P] is the total protein concentration (considered constant), x is the molar ratio ([ligand]/[protein]), CSP is the measured chemical shift perturbation, and CSP_{max} is the maximum CSP at saturation. The adjustable parameters are CSP_{max} and K_d.

Despite chemical shift perturbation is a very sensitive and frequently used method to study protein-ligand interactions by NMR, it provides reliable values of the dissociation constant only when the protein-ligand equilibrium is in the fast exchange regime. This occurs when the exchange rate (the sum of the on and off kinetic rates) is much larger than the difference between the chemical shifts (in Herzts) of the signals of the protein in the free and bound states. Typically, this condition holds when the K_d is larger than 3 µM, which makes NMR extremely powerful to study low affinity interactions that ITC cannot reliably measure [20].

3.7. References

1. Woody, W., R. Circular Dichroism. *Methods Enzymol.* **246**, 163 (1995).

2. Martin, S. R. & Schilstra, M. J. Circular Dichroism and Its Application to the Study of Biomolecules. *Methods Cell Biol.* **84**, 263–293 (2008).

3. Rupp, B. *Biomolecular Crystallography: Principles, Practice, and Application to Structural Biology.* (2010).

4. Wyatt, P. J. Light scattering and the absolute characterization of macromolecules. *Anal. Chim. Acta* **272**, 1–40 (1993).

5. Folta-Stogniew, E. & Williams, K. R. Determination of molecular masses of proteins in solution: Implementation of an HPLC size exclusion chromatography and laser light scattering service in a core laboratory. *J. Biomol. Tech.* **10**, 51–63 (1999).

6. Mertens, H. D. T. & Svergun, D. I. Structural characterization of proteins and complexes using small-angle X-ray solution scattering. *J. Struct. Biol.* **172**, 128–141 (2010).

7. Bernadó, P. & Svergun, D. I. Analysis of Intrinsically Disordered Proteins by Small-Angle X-ray Scattering. *Methods Mol. Biol.* **896**, 71–87 (2012).

8. Lee, M. Y. W. T. *et al.* Two forms of human DNA polymerase δ: Who does what and why? *DNA Repair (Amst).* 102656 (2019). doi:10.1016/j.dnarep.2019.102656

9. Jacques, D. A. & Trewhella, J. Small-angle scattering for structural biology - Expanding the frontier while avoiding the pitfalls. *Protein Sci.* **19**, 642–657 (2010).

10. Weber, P. C. Overview of protein crystallization methods. *Methods Enzymol.* **276**, 13–22 (1997).

11. Evans, P. & McCoy, A. An introduction to molecular replacement. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **64**, 1–10 (2007).

12. Pierce, M. M., Raman, C. S. & Nall, B. T. Isothermal titration calorimetry of protein-protein interactions. *Methods* **19**, 213–221 (1999).

13. Falconer, R. J. Applications of isothermal titration calorimetry - the research and technical developments from 2011 to 2015. *J. Mol. Recognit.* **29**, 504–515 (2016).

14. Levitt, M. H. *Spin Dynamics-Basics of NMR*. *Wiley* **53**, (2013).

15. DW Claridge, T. *High-Resolution NMR Techniques in Organic Chemistry. Tetrahedron Organic Chemistry Series Volume 19 By Timothy D. W. Claridge (Dyson Perrins Laborotory, Oxford). Pergamon: Oxford. 1999. ix + 384 pp. Hardbound $134.50, ISBN 0-08-042799-5. Paperback $49.50,. Journal of the American Chemical Society* **122**, (2000).

16. Sattler, M. & Fesik, S. W. Use of deuterium labeling in NMR: Overcoming a sizeable problem. *Structure* **4**, 1245–1249 (1996).

17. Kwan, A. H., Mobli, M., Gooley, P. R., King, G. F. & MacKay, J. P. Macromolecular NMR spectroscopy for the non-spectroscopist. *FEBS J.* **278**, 687–703 (2011).

18. Pervushin, K. Impact of transverse relaxation optimized spectroscopy (TROSY) on NMR as a technique in structural biology. *Q. Rev. Biophys.* **33**, 161–197 (2000).

19. Ferentz, A. E. & Wagner, G. NMR spectroscopy: A multifaceted approach to macromolecular structure. *Q. Rev. Biophys.* **33**, 29–65 (2000).

20. Williamson, M. P. Using chemical shift perturbation to characterise ligand binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **73**, 1–16 (2013).

4. Double monoubiquitination modifies the molecular recognition properties of p15^{PAF} promoting binding to the reader module of Dnmt1

Amaia González-Magaña[1], Alain Ibáñez de Opakua[1], Nekane Merino[1], Hugo Monteiro[2],
Tammo Diercks[1], Javier Murciano-Calles[3], Irene Luque[3], Pau Bernadó[4], Tiago N. Cordeiro[2],
Alfredo De Biasio[5] and Francisco J. Blanco[1,6,*]

4.1. Abstract

The Proliferating Cell Nuclear Antigen-associated factor p15[PAF] is a nuclear protein that acts as a regulator of DNA repair during DNA replication. The p15[PAF] gene is overexpressed in several types of human cancer and its function is regulated by monoubiquitination of two lysines (K15 and K24) at the protein N-terminal region. We have previously shown that p15[PAF] is an intrinsically disordered protein which partially folds upon binding to PCNA and independently contacts DNA through its N-terminal tail. Here we present an NMR conformational characterization of p15[PAF] monoubiquitinated at both K15 and K24 via a disulfide bridge mimicking the isopeptide bond. We show that doubly monoubiquitinated p15[PAF] is monomeric, intrinsically disordered, binds to PCNA as non-ubiquitinated p15[PAF] does, but interacts with DNA with reduced affinity. Our SAXS-derived conformational ensemble of doubly monoubiquitinated p15[PAF] shows that the ubiquitin moieties, separated by 8 disordered residues, form transient dimers due to the high local effective ubiquitin concentration. This observation and the sequence similarity with histone H3 N-terminal tail suggest that doubly monoubiquitinated p15[PAF] is a binding target of DNA methyl transferase Dnmt1, as confirmed by calorimetry. Therefore, doubly monoubiquitinated p15[PAF] directly interacts with PCNA and recruits Dnmt1 for maintenance of DNA methylation during replication.

4.2. Introduction

Ubiquitination is a multifunctional post-translational modification that may not only drive the protein through the proteasome degradation pathway, but also change the affinity for binding to other biomolecules and, thus, affect protein function in a variety of cellular processes. A few ubiquitinated proteins have been structurally studied and shown to preserve their three-dimensional structure, however, the impact of this modification may be particularly strong in the case of intrinsically disordered proteins (IDPs), whose conformational equilibrium might be significantly altered[1]. A conformational and molecular recognition characterization of an ubiquitinated IDP is still missing, likely because of the

difficulties in obtaining high amounts of the pure modified protein. Here we present a thorough structural analysis of the p15[PAF] oncogenic protein monoubiquitinated at two lysine residues via a disulfide bridge that mimics the isopeptide bond.

The p15[PAF] protein (hereafter named p15) is a nuclear 12 kDa polypeptide initially identified as a proliferating cell nuclear antigen (PCNA) binding protein[2]. The DNA sliding clamp PCNA has the shape of a ring and acts as a docking platform for many enzymes that edit DNA[3] . p15 binds to the front face of the PCNA ring as it encircles the DNA and slides along it[4]. Regulatory mono-ubiquitination at residues K15 and K24 (Figure 4.1A) selectively occurs on PCNA-bound p15 during the S phase of the cell cycle[5], when replicative DNA Polymerases copy the DNA. The ubiquitination of p15 is mediated by the E3 ligase UHFR1 (ubiquitin-like PHD and RING finger containing domain 1)[6], which also monoubiquitinates the histone H3 N-terminal tail at two sites[7,8] Following UV stress, the interaction of doubly monoubiquitinated p15 (hereafter named dmUbp15) with PCNA is disrupted, inducing the recruitment of the translesion synthesis (TLS) polymerase η to PCNA at stalled replisomes and, thus, facilitating the bypass of replication-fork blocking lesions [5]. It is thought that, upon DNA-replication blockage, the monoubiquitinated sites are polyubiquitinated and p15 is removed from PCNA in a proteasome-dependent manner[5]. p15 is an emergent cancer target as it is overexpressed in multiple types of human cancer and is associated with poor prognosis[9-11]. Recently it has been described as a factor required for the maintenance of breast cancer and glioma cell stemness[12,13].

The p15 protein is intrinsically disordered in solution[14] but adopts a folded structure for the central region when bound to PCNA[4]. This central region contains the PCNA interacting protein motif (PIP), and tightly binds on the front face of the PCNA ring. In contrast to other PCNA interacting proteins, however, p15 contacts the inside of, and passes through, the PCNA ring, with the disordered p15 termini emerging at opposite faces of the ring[4]. Both free and PCNA-bound p15 contact DNA mainly with its N-terminal tail, rich in positively charged amino acids, suggesting that p15 acts as a flexible drag to regulate PCNA sliding along the DNA[4]. The dustace between the two ubiquitin moieties at the N-terminal region of the p15 sequence and its similarity to the N-terminal tail of histone H3, suggest that dmUbp15 binds the replication foci targeting sequence (RFTS) of DNA methytransferase Dnmt1 as doubly monoubiquitinated histone H3 does[15].

Here, we present a detailed characterization of the conformational and molecular recognition properties of doubly monoubiquitinated p15. To prepare dmUbqp15 we followed a non-enzymatic approach consisting of a direct disulfide linkage that mimics the isopeptide bond. The doubly monoubiquitinated protein remains disordered based on circular dichroism (CD), Small-Angle X-ray Scattering (SAXS), and NMR chemical shift data analysis. We show that dmUbp15 binds PCNA and DNA, yet with lower affinity than unmodified p15 in the case of DNA, suggesting a role of p15 in regulating the PCNA sliding velocity on the DNA. Furthermore, dmUbp15 binds the RFTS of Dnmt1 as seen by calorimetry, suggesting that it recruits Dnmt1 to replication forks for maintenance of DNA methylation.

4.3. Results

4.3.1. *Doubly monoubiquitinated p15 remains intrinsically disordered*

In order to monoubiquitinate p15 at two residues, we used a non-enzymatic approach comprising the introduction of reactive thiol groups both in the ubiquitin and in the target p15 protein[16,17]. The disulfide bond directed ubiquitination yields a sterically similar linkage to the native isopeptide bond[18]. We expressed and purified a p15 mutant with the ubiquitination target lysine residues mutated to cysteine whereas two native cysteines in the sequence were changed to serine. A reactive ubiquitin added to the p15CCSS mutant spontaneously produced dmUbp15. With this method we were able to produce mg amounts of homogeneous dmUbp15 (Figure S4.1).

The SEC-MALS analysis of dmUbp15 demonstrated that the protein is a monomer (Figure 4.1B). The far-UV CD spectrum shows one minimum at 205 nm and a shoulder at 225 nm, consistent with a mixture of random coil and regular secondary structure (Figure 4.1C). The CD spectrum of the quadruple p15 mutant p15CCSS (used for ubiquitination at positions 15 and 24; see the methods section) is typical for a random coil. The difference spectrum shows the pattern of minima corresponding to the ubiquitin moieties[19]. These results suggest that isolated dmUbp15 lacks any defined structure apart from in the ubiquitin moiety.
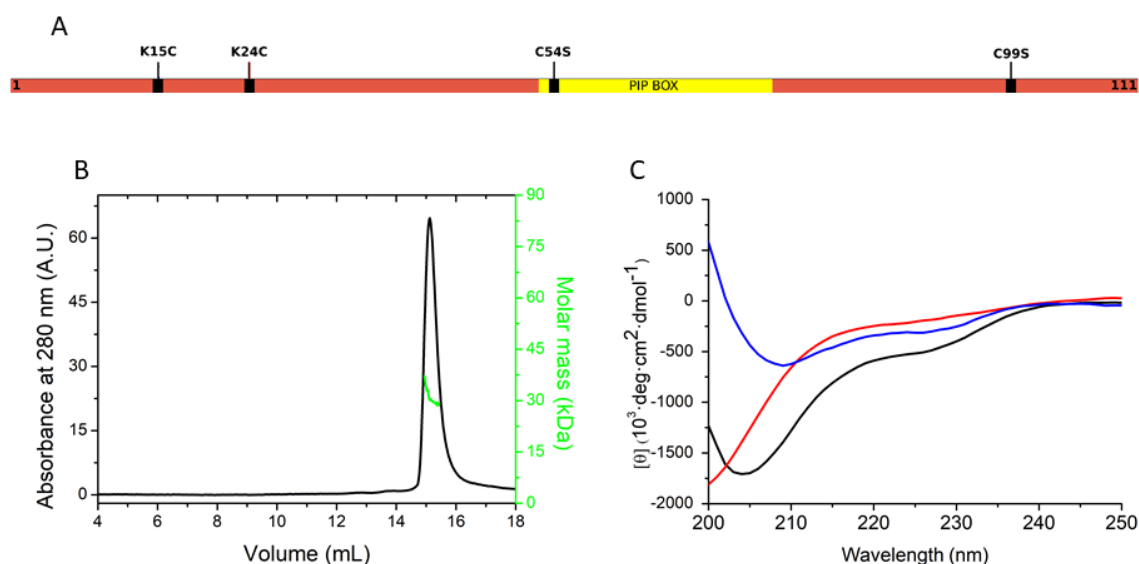
Figure 4.1. dmUbp15 is monomeric and unstructured in solution. (A) Scheme of p15 showing with position of the mutations introduced in this study (p15CCSS) as well as the PIP-box region (yellow), encompassing all the residues observed in the crystal structure with PCNA[20]. (B) Size exclusion chromatogram of dmUbp15 (black, axis on the left), and molar mass derived from MALS (green, axis on the right) in PBS pH 7.0 at 25 °C. The theoretical molar mass of dmUbp15 is 29,019 Da (Figure S4.1). (C) CD spectra of dmUbp15 (black), p15CCSS (red), and their difference (blue).

In contrast to CD and other analytical methods, our studies of dmUbp15 constructs by heteronuclear NMR correlation spectroscopy are selective for the p15 moiety due to its labelling with NMR active ^{13}C and ^{15}N isotopes, whereas the unlabeled fused ubiquitin moieties remain invisible. The assignment of the p15 backbone and some C^β side chain NMR signals was then obtained from a set of conventional triple resonance experiments with out-and-back transfer and detection on the amide protons. The 2D ^{1}H,^{15}N HSQC fingerprint spectrum of the p15 moiety in dmUbp15 shows 99 assigned amide signals (Figure 4.2A), where only the two N-terminal residues and the nine prolines (in the absence of an observable amide proton) are missing. All prolines have *trans* peptide bonds, as indicated by small [C^α - C^β] chemical shift differences[21]. A few minor signals with ca. 5 % intensity of the major ones are visible in the ^{1}H,^{15}N HSQC spectrum (data not shown). At least one of these signals corresponds to a residue following a proline with a [C^α - C^β] chemical shift difference typical for a *cis*-configuration of the peptide bond. Therefore, we interpret the minor signals as arising from small populations of dmUbp15 proteins with peptidyl-prolyl *cis*-peptide bonds, as occurs with non-ubiquitinated p15. The dispersion of backbone amide ^{1}H chemical shifts is similarly low for unmodified and for doubly monoubiquitinated p15 (Figure 4.2B), indicating

62

that the latter remains disordered and flexible. The major changes occur at sites of ubiquitination and/or mutation (Figure 4.3 and 4.4).

The distribution of p15 ^{15}N transverse relaxation rates, R_2, along the sequence (Figure S4.3) is likewise similar in both states except for the region between residues 11 and 27, which contains the ubiquitination sites (15 and 24). The approximately twofold increase in ^{15}N R_2 rates can, thus, be explained by a locally restricted mobility due to the attached ubiquitin moieties in dmUbp15. As in isolated p15, relatively large R_2 values are also found for residues 60 to 72, which could be explained by steric restrictions between bulky side chains and the backbone[22], or by the presence of minor populations of non-random conformations[14].
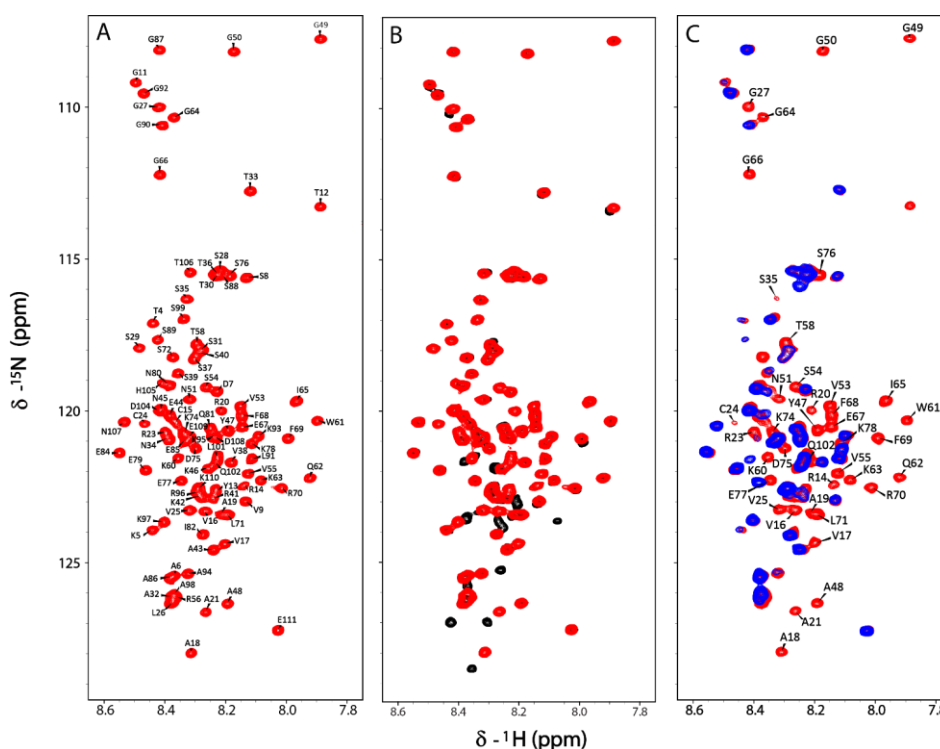
Figure 4.2. NMR fingerprint spectra of dmUbp15 with only p15 enriched in ^{15}N. (A) 2D ^{1}H,^{15}N HSQC NMR spectrum of dmUbp15 with assignment of the backbone amide signals (in PBS pH 6.2 at 25 °C and 800 MHz). (B) Overlay of dmUbp15 (red) and non-ubiquitinated p15 (black) spectra. (C) Overlay of dmUbp15 spectra at pH 6.8in the absence (red) and in the presence (blue) of PCNA at 1:4 molar ratio of PCNA protomer. The residues that suffer the highest intensity drop in the complex are labeled.

The chemical shift deviations for H^{α}, C$^{′}$, C^{α}, and C^{β} from random coil values (Figure 4.3), taken from intrinsically disordered proteins under native conditions[23], do not indicate preferential local conformations in dmUbp15. Overall, our NMR data indicates the same level of intrinsic

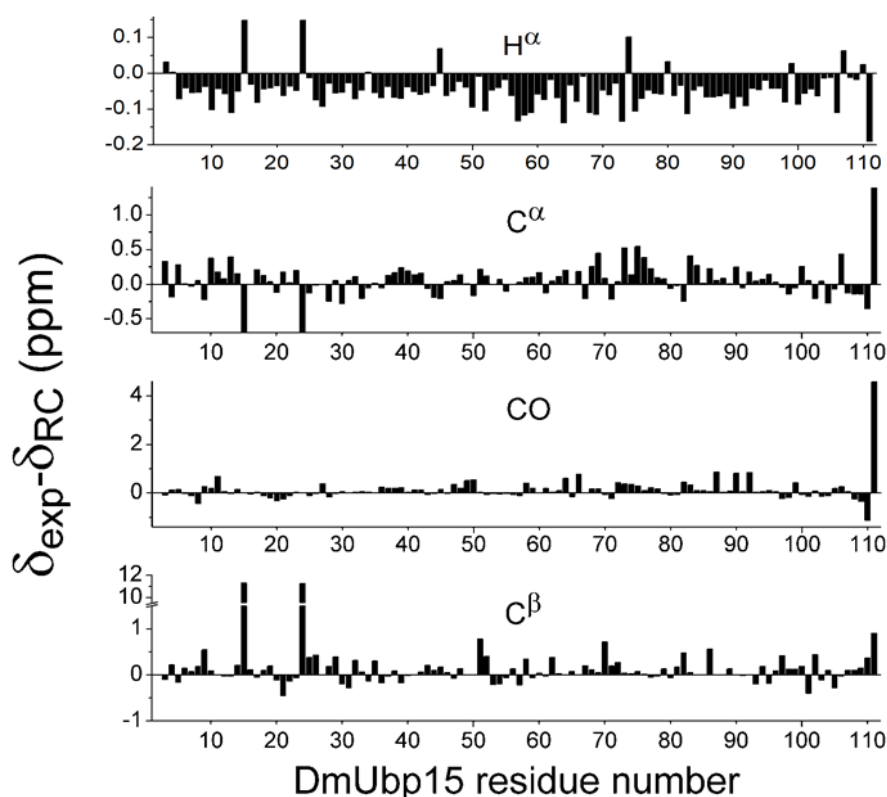disorder and flexibility for p15 in the dmUbp15 fusion protein as in the previously analyzed isolated state [14].

Figure 4.3. Experimental and random coil chemical shift[23] differences (δ_{exp} - δ_{RC}) *versus* dmUbp15 residue number, derived for the H^α, C^α, CO, and C^β nuclei in PBS buffer pH 6.2, 25ºC.

To further investigate the structural effects of ubiquitination in p15, we applied Small-angle X-ray scattering (SAXS) as the best suited technique to probe the overall structure of disordered proteins[24]. Our synchrotron SEC-SAXS data revealed that dmUbp15 is a non-globular particle in solution (Figure 4.4A) with a radius of gyration (R_g) of 33.7±0.4 Å and a maximum intermolecular distance (D_{max}) of 139.0±10.0 Å (Figure 4.4B). dmUbp15 is slightly larger than non-ubiquitinated p15CCSS, which has an R_g of 29.4±0.5 Å and a D_{max} of 124.0±5.0 Å (Figure 4.4A and 4.4B). The SAXS-derived structural parameters of p15CCSS are in the range of those previously measured for p15 wild-type p15, R_g of 28.1±0.3 Å and a D_{max} 120.0±10.0 Å[14]. The Kratky representation of dmUbp15 SAXS data (Figure 4.4A) displays dual features. On the one hand it shows a broad peak, indicating an important level of compaction. On the other hand, this peak does not decrease completely suggesting that p15 remains disordered upon

di-monoubiquitination. Note that the Kratky representation of dmUbp15 is entirely different from that of p15CCSS, which displays a monotonous increase of $(sR_g)^2 \cdot I(s)$ at wide angles characteristic of fully disordered proteins (Figure 4.4A). The pair-wise distance distribution, $P(r)$, of dmUbp15 (Figure 4.4B) also suggests this order-disorder duality. It is asymmetric with a smooth decrease towards a large D_{max}, as observed for highly flexible proteins like p15CCSS, but also displays two main peaks that encapsulate the intra- and inter-domain pair-wise distances of ubiquitin tags[25]. The SAXS data is, therefore, compatible with a protein containing folded domains tethered to a disordered region, and is in clear contrast with those measured for ubiquitin, which is a globular protein.

To get more detailed structural information from the SAXS data, we applied the Ensemble Optimization Method (EOM)[26]. The NMR analysis indicates that the presence of the two grafted ubiquitins does not perturb the conformational features of p15, which remains disordered in solution. Prompted by this observation, we have modelled the p15 protein as an IDP in both scenarios, with or without randomly oriented ubiquitin-tags at residues 15 and 24. An atomistic ensemble model of the doubly monoubiquitinated p15 was built combining *Flexible-Meccano* (FM), Molecular Dynamics (MD) simulations, and in-house scripts (see methods for details). A second ensemble for p15CCSS was built by removing both ubiquitin moieties and the linkers from the dmUbp15 ensemble (Figure 4.4C). Using these structural pools, the EOM derived ensembles provided an excellent fit to the experimental SAXS curves of p15CCSS and dmUbp15, with χ_i^2 values of 0.94 and 0.99, respectively (Figure 4.4D). The EOM derived R_g distributions, compared with those obtained from the initial pools (before EOM), are displayed in Figures 4.4E and S4.4 for dmUbp15 and p15CCSS, respectively. These distributions sample a broad range of R_g values, indicating high flexibility of the protein regardless of the ubiquitination. Interestingly, both ensembles display an enrichment in extended conformations (large R_g values) when compared with the initial pools. This observation suggests that the transient secondary elements previously observed in wild-type p15 are also present in both cases[14], and substantiates the absence of major changes in the conformational sampling of p15 in the presence of the two covalently bound ubiquitin proteins, as also observed by NMR.

4.3.2. *dmUbp15 adopts inter-ubiquitin "closed" and "open" conformations*

SAXS is highly sensitive to the spatial distribution of globular domains within flexible proteins[25]. To exploit this feature, we analyzed the relative position of both ubiquitin moieties in the EOM-optimized ensemble of dmUbp15. By computing the distance between their centers of mass (d_{Ub-Ub}) for each conformer, we observed two families of structures that we classified as "*closed*" and "*open*". To better visualize these two families, we used a joint distribution of d_{Ub-Ub} *vs* end-to-end distances (d_{ee}) of the structures from the initial pool (in grey) and the refined ensemble (in red) (Figure 4.4F). The pool, representing a random distribution of ubiquitin positions when grafted to dmUbp15, presents a single peak with a maximum at d_{Ub-Ub} around 53 Å. Upon ensemble optimization against SAXS data, the single peak was resolved into two main clusters centered on d_{Ub-Ub} values of 30.0±4.0 Å and 52.4±8.9 Å, hereafter named *closed* and *open* states, respectively. The inspection of the conformations belonging to the *closed* state indicates that both ubiquitin moieties are in contact forming dimers in different three-dimensional arrangements. It is worth noting that both peaks in the optimized distribution appear isolated, and there is not a continuum of conformations between the closed and the open states. These observations indicate that, for doubly monoubiquitinated p15 the two ubiquitin moieties form a transient dimer. The presence of inter-ubiquitin contacts does not perturb the overall size of the molecule, which remains highly disordered upon ubiquitination (Figures 4.4B-F).

To further substantiate these observations, SAXS curves for the two states were computed by averaging the profiles of conformations belonging to each peak. Then, we performed a grid search by varying the relative population of the ensemble-based SAXS pattern representing the *closed* and *open* states. Although the minimum is relatively broad, the optimal level of agreement ($\chi^2 = 0.99$) with the SAXS curve of dmUbp15 was obtained at a 69/31 ratio for the *open/closed* models (Figure 4.4G and S4.5).
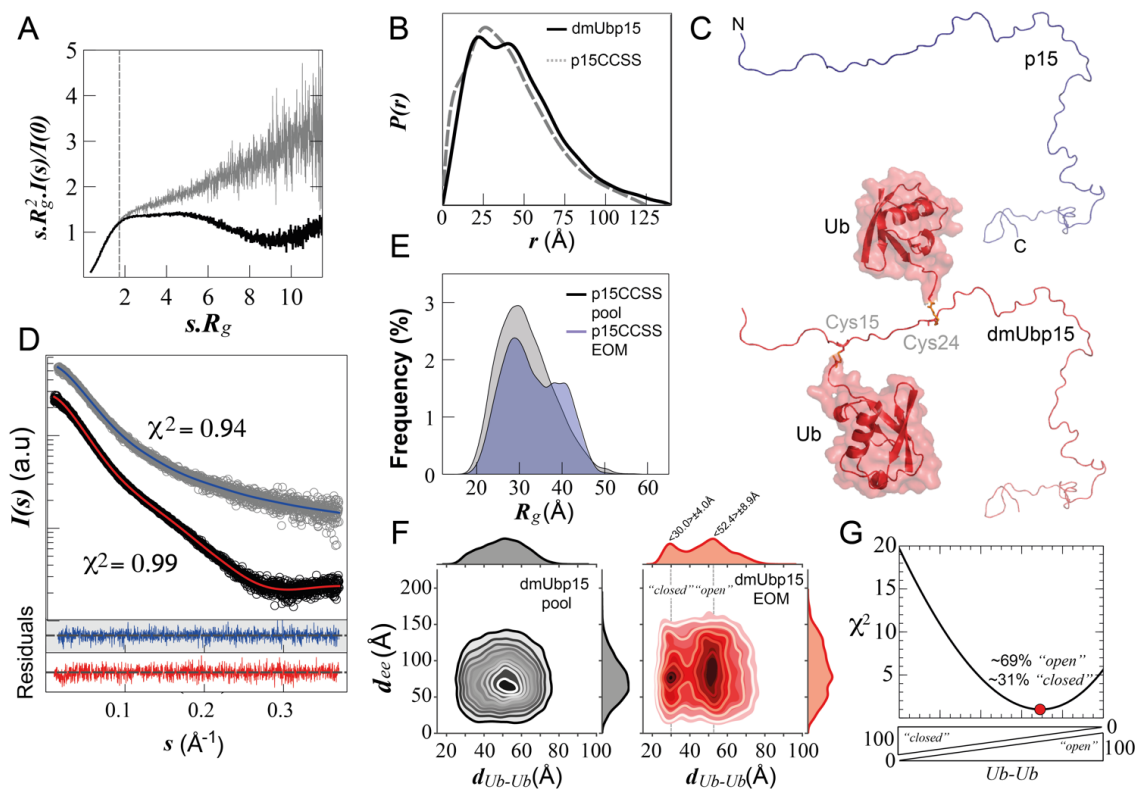
Figure 4.4. SAXS data analysis of dmUbp15 and p15CCSS in PBS pH 7.0 at 25 °C. (A) Dimensionless Kratky representation of the SEC-SAXS data measured for p15CCSS (grey) and dmUbp15 (black). The absence of a clear peak maximum at sR_g=√3 (dashed line) indicates that dmUbp15 remains conformationally heterogeneous upon ubiquitination. (B) Normalized pair-wise distance distribution, $P(r)$, computed from experimental SAXS curves of p15CCSS (dashed grey line) and dmUbp15 (solid black line). As p15CCSS, dmUb1p15 has a large D_{max} and smooth ending to $P(D_{max})$=0. (C) Structural cartoon representation of p15CCSS (blue) and dmUbp15 (red). The linker connecting the ubiquitin tags to each cysteine residue is represented in yellow sticks. (D) Logarithmic-scale representation of scattering intensity, $I(s)$, as a function of the momentum transfer, s, measured for p15CCSS (empty grey circles) and dmUbp15 (empty black circles). Solid lines are averaged back-calculated curves derived from sub-ensembles of p15CCSS (blue) and dmUbp15 (red) conformations derived from the EOM fit of the SAXS data. Residuals using absolute values are displayed at the bottom with the same color code, with a scale ranging from –4.0 to 4.0 (white and grey bands). (E) Comparison of the R_g distributions of the EOM-selected structures (blue) and the initial pool of 12,000 random-coil structures of p15CCSS (grey) built with *Flexible-Meccano*. The SAXS-based ensemble of p15CCSS is slightly more extended than the statistically random pool (~10% of enrichment). (F) Bivariate kernel density of the initial pool (grey-palette contours) and EOM-refined ensembles of dmUbp15 (red-palette contours) plotted in the d_{ee} and d_{Ub-Ub} conformational space. SAXS data analysis reveals that dmUbp15 exists as a dynamic ensemble of conformations displaying a large population of states with short inter-ubiquitin distances (*closed* state). (G) Evolution of χ^2 as a function of the relative population of *closed* and *open* states. A well-defined minimum is observed for an ensemble with ~31% of the *closed* state.

Ubiquitin has been reported to form highly dynamic symmetric dimers with very low affinity ($K_d \sim$ 5 mM)[27]. Using this dissociation constant and the relative population of the *open/closed* conformations derived from SAXS data, we estimated an effective ubiquitin concentration, C_{eff}, of 2.24 mM in the context of dmUbp15[28]. To further characterize the interaction between the two ubiquitin moieties within dmUbp15 and determine the dimerization surface, we recorded NMR spectra of singly monoubiquitinated p15 with only one ^{15}N-labelled ubiquitin molecule at position 15 or at position 24 (mUbp15_15 or mUbp15_24), and compared them with the spectrum of dmUbp15 (with two ^{15}N labeled ubiquitin chains). The differences in chemical shifts were very small (Figure S4.6) and do not match the interacting residues in ubiquitin dimers[27], suggesting that the interaction detected by SAXS is too transient to be observed by NMR chemical shit differences.

4.3.3. *dmUbp15 interacts with PCNA as non-ubiquitinated p15*

The interaction of dmUbp15 with PCNA was studied by ITC and NMR. Calorimetry shows that one molecule of dmUbp15 binds to each protomer of PCNA with a K_d of 1.6 ± 0.1 µM at 25°C (Figure S4.7 and Supplementary Table S4.1), which is similar to the isolated p15, with a K_d of 1.1 ± 0.1 µM [4]. The mutation C54S, within the PCNA binding region of p15, does not affect the affinity (K_d= 1.3 ± 0.1 µM, Figure S4.7 and Table S4.1). This result appears in contradiction with the previously reported promotion of PCNA binding by ubiquitination of p15 observed by nuclear co-localization[6].

Our quantitative calorimetry measurements demonstrate that ubiquitination does not affect the direct interaction between p15 and PCNA, and our NMR data shows that the PCNA binding mode of dmUbp15 is equivalent to that of p15 (Figure 4.2C). On the side of p15, this is indicated by the same general decrease in ^{1}H,[15]N signal intensities and the same group of disappearing signals in the central region of the dmUbp15 chain (Figure S4.8)[4]. On the side of PCNA, however, the ^{1}H,[15]N TROSY spectrum only shows a few signals from its most flexible regions (Figure S4.9), most likely due to the slower tumbling rate of the complex (with an increased molar mass of ~180 kDa)

4.3.4. *dmUbp15 interacts with DNA with reduced affinity as compared with non-ubiquitinated p15*

p15 interacts with DNA in a sequence independent manner mainly through its N-terminal region [4], which is rich in basic amino acids and contains both ubiquitination sites. Our working hypothesis was that double monoubiquitination of p15 might reduce its affinity for DNA via the neutralization of two positive charges (at the two lysine side chains) and via the increased steric hindrance exerted by the two attached ubiquitin moieties at the center of the DNA binding region at the p15 N-terminus. We examined this hypothesis by measuring the DNA affinity by NMR. This yielded a fivefold reduced DNA affinity of dmUbp15 (K_d= 24 ± 3 μM; Figure 4.5) as compared with isolated p15 (K_d = 5 ± 1 μM).
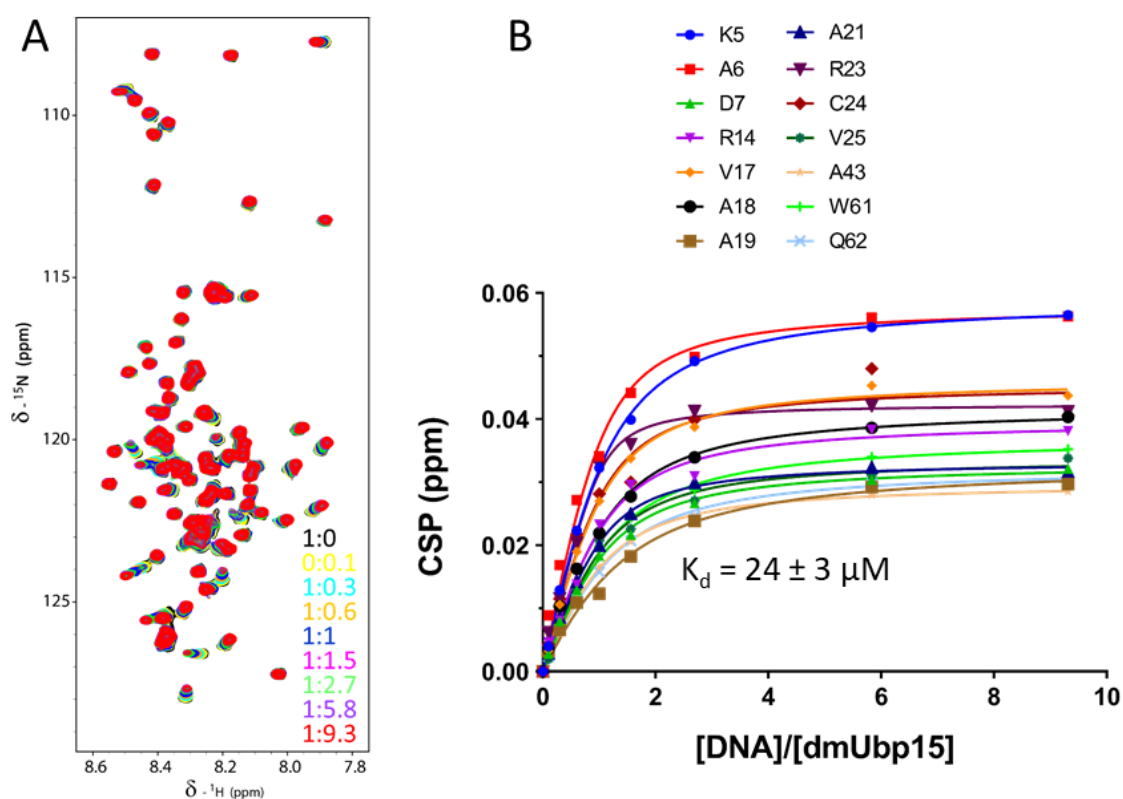
Figure 4.5. NMR analysis of dmUbp15 binding to a 24 bp DNA duplex (in PBS pH 6.2 at 25 ºC). (A) Overlay of ^{1}H,^{15}N HSQC spectra at increasing DNA:dmUbp15 molar ratios with colors as indicated. (B) Chemical shift perturbation, CSP, versus DNA:dmUbp15 molar ratio for residues with CSP larger than the average plus one standard deviation. The indicated curves show fitting to a single-site binding model and yield *per residue* K_d values. The mean value of the dissociation constant and the standard error are indicated.

To further analyze whether the main inhibiting effect of p15 ubiquitination upon DNA binding is the entailed removal of two positive charges, rather than the increased steric hindrance,

we also measured the DNA affinity of p15CCSS with double lysine-to-cysteine mutation, and found a K_d = 91 ± 5 μM (Figure S4.10). This result clearly confirms the positive charge neutralization as the dominant effect of p15 ubiquitination upon DNA binding.

NMR analysis of ^{15}N labelled free ubiquitin (not conjugated with p15) shows that ubiquitin does not directly bind DNA, p15 or PCNA (Figure S4.11).

However, by isotopically labeling only the ubiquitin moieties of dmUbp15, we measured small chemical shift perturbations in the presence of an excess of DNA (Figure S4.12 top). These perturbations occur mainly at the C-terminal end, the site of conjugation with p15, and are likely due to DNA binding to the positively charged arginine residues in this region, which enhances the overall DNA affinity of dmUbp15 as compared to p15CCSS. Even smaller perturbations were found in the chemical shifts of ubiquitin for dmUbp15 in the presence of PCNA (Figure S4.12 bottom).

4.3.5. *The RFTS domain of Dnmt1 specifically binds dmUbp15*

The double monoubiquitination of p15, separated by 8 disordered residues, makes it a favorable target for binding the reader module of DNA methyl transferase Dnmt1. This enzyme converts hemi-methylated DNA (CpG methylated only in one of the strands) into its fully methylated form (in both strands), a crucial process in embryonic development [29]. Association of Dnmt1 with the replication machinery enhances methylation efficiency and occurs through a direct interaction of Dnmt1 with PCNA [30,31]. More recently it was shown that post-replicative DNA methylation is controlled by Dnmt1 recruitment to doubly monoubiquitinated histone H3[15]. The RFTS of Dnmt1 binds with nanomolar affinity (K_d = 18 ± 4 nM at 20 ºC) the histone H3 N-terminal tail doubly monoubiquitinated at two lysine residues separated by 3 to 8 residues[15]. We found that dmUbp15 is recognized by the RFTS domain of Dnmt1 with a similar affinity (K_d = 29.3 ± 4.5 nM at 25 ºC; Figure 4.6).
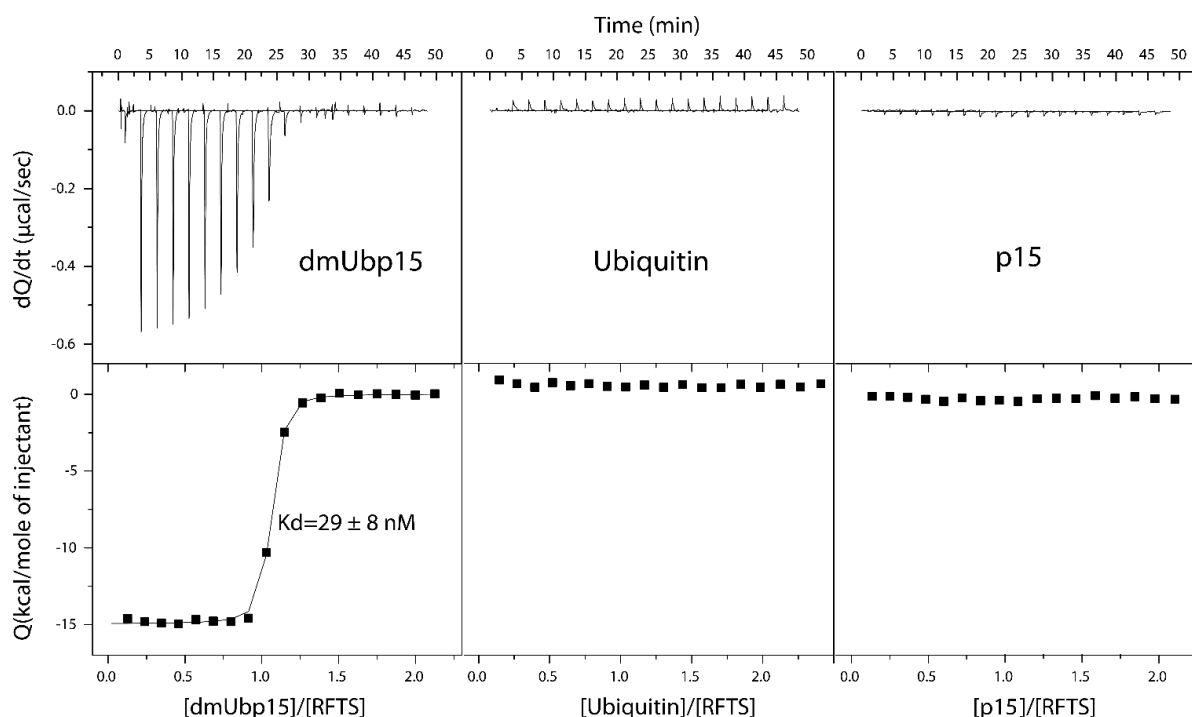
Figure 4.6. Calorimetric titration of dmUbp15, p15 and Ub with the RFTS domain of Dnmt1. The upper panel represents the heat effect associated with the peptide injection (differential power) and the lower panel represents the ligand concentration dependence of the heat released upon binding at 25 ºC. Squares in the lower panels correspond to the experimental data and the continuous line to the best fit to a model of one set of identical binding sites.

However, this binding is completely abolished when p15 is not ubiquitinated, in contrast to histone H3, which still binds with micromolar affinity[15]. To confirm that Dnmt1 specifically binds dmUbp15 we performed ITC experiments titrating monoubiquinated p15 at only one position (mUbp15_15 or mUbp15_24) with RFTS (Figure S4.12). Both monoubiquinated p15 molecules bind RFTS, but with a 20-fold lower affinity and an impaired stoichiometry (Table S4.2). These results indicate that double monoubiquitination of p15 is required for specific interaction with the RFTS domain of Dnmt1.

The structure of the complex is not known, but it is likely to be similar to that of the doubly monoubiquitinated histone H3 one, with the ubiquitin recognition loop of RFTS binding at the interface of both ubiquitin moieties (Fig 1.9). The preferential location of the two ubiquitin moieties close to each other in the conformational ensemble of dmUbp15, as observed by SAXS and the sequence similarity between the histone H3 and p15[4] support this hypothesis.

4.4. Discussion

IDPs play diverse biological roles and structural disorder is important for the function of proteins that regulate processes often altered in cancer[32]. One of the functional advantages of IDPs presumably derives from their conformational adaptability, allowing for interactions with different biomolecular partners and for extensive regulation through post-translational modifications at many accessible sites[33]. Despite the lack of persistent structure they can form fuzzy complexes[34] with low to very high affinity[35], often driven by electrostatic forces. Conformational analysis of IDPs is challenging since their polypeptide backbone exhibits a high degree of flexibility with rapid interconversion of multiple conformers. NMR is then the method of choice for conformational analysis, together with SAXS, which has the capacity to report on the three-dimensional space sampled by disordered states and, thus, complements the local information provided by NMR[26,36].

The interest in IDPs has recently increased due to their involvement in the formation of membraneless organelles via protein phase separation[37]. Solutions of p15 or dmUbp15, however, do not show phase separation (at least up to 20 g/L at 25 ºC), and our SEC-MALS data demonstrate that both are monomeric. Both proteins are intrinsically disordered, as confirmed by NMR. In the case of isolated p15, residual dipolar couplings showed the existence of low populations of nonrandom conformations in certain regions of the chain that could not be identified by local chemical shift deviation[14]. Possibly, dmUbp15 also shows some preferential local conformations, but these have too low populations for detection by chemical shift analysis (Figure 4.3). Ubiquitination causes only local perturbations in p15 chemical shifts and backbone dynamics, as sampled via ^{15}N NMR relaxation. Therefore, double ubiquitination does not change the intrinsically disordered nature of p15. SEC-SAXS data is also consistent with a disordered nature of the p15 chain in dmUbp15, however these data also indicate that the ubiquitin-ubiquitin distance distribution is far from random, with an ensemble of conformations enriched in those with the ubiquitin moieties close to each other. This result is consistent with the observation of low affinity ubiquitin dimers in ubiquitin solutions[27], and with a high effective local concentration of ubiquitin in dmUbp15. The interaction of p15 with PCNA is unchanged by ubiquitination. Both p15 and dmUbp15

bind PCNA with the same low micromolar affinity through the central region of the p15 polypeptide chain. This result is in disagreement with the proposed promotion of PCNA binding by p15 ubiquitination[6]. Upon PCNA binding, p15 traverses the central channel of the PCNA ring, with its disordered N- and C-terminal regions emerging on opposite sides of the ring[4,38]. We assume that dmUbp15 adopts the same binding mode as it can still thread through the PCNA ring with its unmodified C-terminal end (Figure 4.7). Yet, dmUbp15 binds DNA with a fivefold reduced affinity, and this effect could be interpreted in two ways: i) removal of positive charge on the targeted lysines and its substitution by a globally neutral polypeptide chain (ubiquitin has a pI of 6.6), and ii) steric impediment from attaching two ubiquitin globules (7 kDa each) at the N-terminal region of p15 that is mainly involved in DNA binding. The importance of the former charge effect is confirmed by an analysis of the p15CCSS mutant, whose affinity for DNA is reduced by a factor of 19. The steric effect seems to be much less important and is likely overridden by a favorable effect of the local positive charge at the C-terminal region of the ubiquitin molecule.
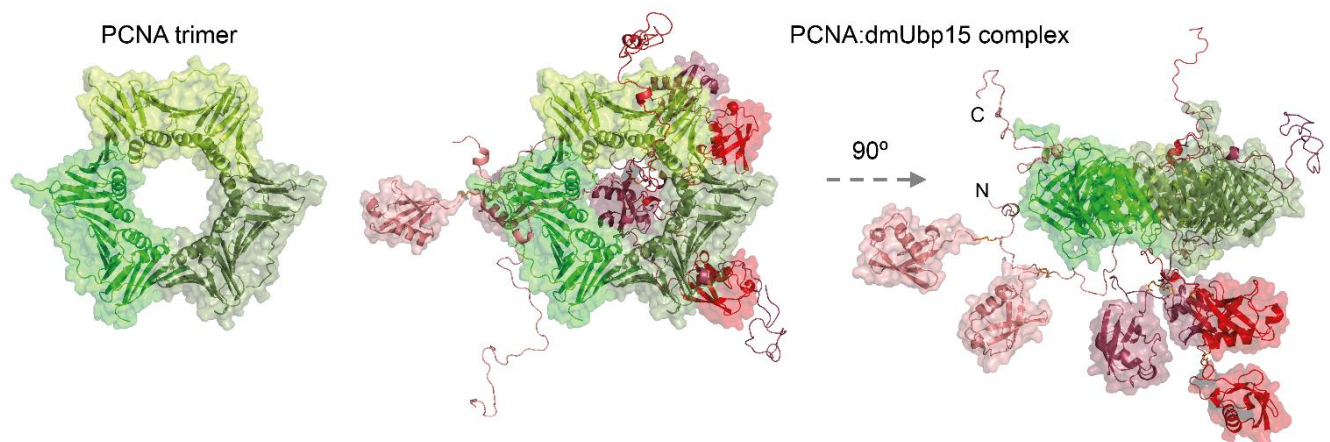

Figure 4.7. Model of the PCNA homotrimer (green gradation) with the maximum number of three bound dmUbp15 proteins. The largely unstructured p15 chains funnel through the PCNA ring and contact it via their PIP-box region[4]. dmUbp15 molecules, with two ubiquitin moieties at p15 residues 15 and 24, are colored in a red palette. The N- and C-terminal of one of the p15 chains are indicated.

The reduced affinity of ubiquitinated p15 for DNA might have functional implications for the regulatory role of p15 on DNA replication and repair. Unperturbed DNA replication occurs with dmUbp15. In response to DNA damage PCNA-bound dmUbp15 is degraded and replaced

by non-ubiquitinated p15, with higher DNA affinity, which may slow down the sliding of PCNA thus facilitating the switch from replicative to translesion synthesis (TLS) polymerases.

At the same time, the double monoubiquitination of p15 separated by 8 disordered residues makes it a target for binding the DNA methyl transferase Dnmt1, which also binds PCNA [39], during non-perturbed DNA replication p15 is doubly ubiquitinated and lies at the back face of the PCNA ring, because the N-terminal tail of p15 emerges from the back face. Therefore, dmUbp15 may act synergistically with PCNA to recruit and activate the Dnmt1 enzyme to the back of the replication fork (where the hemimethylated DNA localises) to complete the methylation on the newly synthesized strand.

It has been shown that Dnmt1 is inhibited by its RFTS domain, which binds the catalytic domain to the exclusion of DNA[40,41]. Activation of Dnmt1 at specific times and locations is crucial for the maintenance of proper DNA methylation during cell proliferation, which ensures tissue or cell type specific functions. Thus, constitutively active Dnmt1 leads to abnormal methylation patterns that affect this cell type specific functional differentiation. Previous studies proposed that the binding of H3 doubly monoubiquitinated at K18 and K23 to RFTS leads to Dnmt1 activation through dissociation of inserted RFTS from the catalytic domain[15]. Therefore, the two ubiquitin moieties act as a unique module recognized by the RFTS domain in an essential process to activate Dnmt1 at the proper time and location. Our results suggest that dmUbp15 binds RFTS and leads to the activation of Dnmt1 in the same way as doubly monoubiquitinated histone H3. Further studies should confirm if dmUbp15 activates Dnmt1 and whether this activation leads to a proper or abnormal function of Dnmt1. Future work will clarify the proposed role of dmUbp15 in the context of maintenance of DNA methylation.

4.5. Methods

4.5.1. *Protein expression and purification*

The disulfide directed ubiquitination is a non-enzymatic method that requires the introduction of thiol groups both in ubiquitin and in the target protein. Thus, we designed a quadruple p15 mutant with the ubiquitination target lysines mutated to cysteine, whereas two native cysteines were mutated to serine. The amino acid sequence of human p15[PAF] used

in this work corresponds to isoform 1, the canonical sequence of the full-length protein (Uniprot Q15004), with the following four mutations: K15C, K24C, C54S, and C99S (p15CCSS). We used a synthetic gene with a codon usage optimized for bacterial expression and cloned into a pET11d vector. The protein was expressed and purified as previously described for wild type p15 [14]. Briefly, after BL21(DE3) bacterial growth at 37 ºC in LB or M9 minimal medium with the appropriate isotopic enrichment (1 g/L 98% $^{15}NH_4Cl$ and 2 g/L 99% $^{13}C_6$-glucose) and after cell sonication, the protein was isolated from the supernatant by three chromatography steps at room temperature: anion exchange, reverse phase and cation exchange. Then, a final reverse phase was done to be able to lyophilize the pure protein without buffer. After preparation for reverse phase chromatography (by adding trifluoroacetic acid, TFA, to a final concentration of 0.1 %), the solution was loaded on a Phenomenex Jupiter C_4 250x10 mm column with 10 μm particles and 300 Å pore diameter previously equilibrated with aqueous 0.1% TFA. The protein was eluted with a buffer gradient (to 90% aqueous acetonitrile, 0.1% TFA), eluting at approximately 40% acetonitrile. Fractions containing the protein (as seen by SDS-PAGE) were pooled, freeze-dried. In some instances, the protein was purified from the insoluble fraction by solubilization in 6 M urea and tenfold dilution in water at pH 3.0, followed by reverse phase and cation exchange chromatography (plus a final reverse phase step before lyophilization). Mass spectrometry measurements indicated that the initial methionine was partially processed by bacterial enzymes.

To prepare p15 mutants singly monoubiquitinated at positions 15 or 24 we reversed the mutation of one of the cysteines to lysine. Therefore, the new mutants could only be ubiquitinated at the remaining cysteine. The expression and purification strategy for these mutants was the same as for p15CCSS.Human PCNA was produced and purified with or without isotope enrichment as described [4].

The RFTS domain of human Dnmt1 (Uniprot P26358; residues 351-600) was produced from a synthetic gene that includes and N-terminal MGSSH$_6$GSSG tag sequence followed by the human rhinovirus type 14 3C protease cleavage site. The protein was expressed in BL21(DE3) cells grown overnight at 20- ºC in autoinducible ZYP5052 medium supplemented with 50 μM of $ZnCl_2$. After sonication the protein was found in the pellet, solubilized with 8 M urea and refolded by dropwise 50-fold dilution in 20 mM Tris pH 8, 300 mM NaCl, 1 mM DTT, 50 μM $ZnCl_2$, at 4 ºC. The soluble protein was applied at room temperature to a His-Trap 5 mL Ni^{2+}

column equilibrated in 20 mM Tris pH 8, 300 mM NaCl, 1 mM DTT and was eluted with 250 mM imidazole. The fractions containing the protein were pooled and cleaved with the protease o/n at 4 °C while dialyzed against 20 mM Tris pH 8.0, 1 mM DTT, to eliminate the imidazole. The dialyzed solution was applied to a His Trap column equilibrated in the same buffer and the flowthrough, containing the RFTS protein, was polished by gel filtration on a Superdex 200 26/60 column in PBS pH 7.0.

Ubiquitin fused to intein was prepared with natural isotope abundance or enriched in ^{15}N using the Ub-AvaDnaE-AAFN-H$_6$ clone[16] in BL21(DE3) cells, and was purified from the soluble fraction by Co^{2+} loaded HisTrap FF column in 20 mM Tris pH 8, 300 mM NaCl and 1 mM DTT (with or without 0.5 M imidazol).The final protein concentration was determined by UV absorbance at 280 nm. After integrity checks by SDS-PAGE and MALDI-TOF the proteins were stored at -80ºC.

4.5.2. *Synthesis of disulfide-linked doubly monoubiquitinated p15*

This synthesis was done as described before[17], with some modifications. Ub-AvaDnaE-AAFN-H$_6$ protein[16] was dialyzed in 100 mM sodium phosphate pH 7.8, 150 mM NaCl, 1 mM EDTA, 1 mM TCEP. The ubiquitin moiety was cleaved by adding 100 mM cysteamine and 50 mM TCEP, and incubating for 12 hours at 25 ºC. After adding 0.1% TFA, ubiquitin with C-terminal aminoethane thiol linker was purified by reverse phase chromatography on a C$_4$ column with a gradient of 90% aqueous acetonitrile 0.1% TFA, eluting at approximately 50% acetonitrile. After lyophilization, a yield of 40 mg Ub-SH per liter of culture was obtained. The Ub-SH protein was activated for ubiquitin linkage using 20.0 equivalents of 2,2´-dithiobis(5-nitropyridine) (DTNP) in 2.0 ml of a 3:1 (v/v) acetic acid:water mixture. This solution was added to the Ub-SH powder and the reaction proceeded for 72 hours at 25 ºC with agitation. Ubiquitin-S-nitro-2-pyridinesulfenyl disulfide adduct (Ub-DTNP) was purified by reverse phase chromatography on a C$_4$ column, as described above.

1.0 equivalent of p15CCSS mutant (0.5 μmol) and 4.0 equivalents of Ub-DTNP (2 μmol) were dissolved in 1 M HEPES (10 ml) pH 6.8 and shaken at 25 ºC for 1 hour. Ubiquitin-conjugated p15 was purified by C$_4$ reverse phase chromatography to yield ca. 0.15 μmol (30%) of dmUbp15. The sample identity and purity were confirmed by MALDI-TOF and SDS-PAGE

(Figure S4.1). An equivalent procedure was used to obtain the singly monoubiquitinated mUbp15_15 and mUbp15_24 molecules. Of note, both ubiquitin and p15 mutants used in this coupling reaction could have distinct isotopic enrichment for subsequent NMR studies.

4.5.3. *Size exclusion chromatography-multi angle light scattering (SEC-MALS).*

This experiment was performed at 25 ºC using a Superdex 200 10/300 GL column (GE HealthCare) connected to a DAWN-HELEOS light scattering detector and an Optilab rEX differential refractive index detector (Wyatt Technology). The column was equilibrated with 0.1 µm filtered PBS (10 mM phosphate, 140 mM chloride, 153 mM sodium, 4.5 mM potassium) at pH 7.0, and the SEC-MALS system was calibrated with a sample of Bovine Serum Albumin (BSA) at 1 g/L in the same buffer. 0.1 mL of dmUbp15 protein at 1 g/L were injected at 0.5 mL/min. The data was acquired and analysed by ASTRA (Wyatt). From repeated measurements of BSA samples at 1 g/L under identical or similar conditions we estimate an experimental error of ca. 5% in the molar mass.

4.5.4. *CD spectroscopy*

Spectra of 3 µM dmUbp15 or p15CCSS samples in PBS pH 7.0 were recorded on a Jasco-815 spectropolarimeter at 25 ºC in a quartz cuvette (path length 2 mm).

4.5.5. *NMR Spectroscopy*

All NMR experiments were recorded at 35 ºC (for observation of PCNA) and 25 °C (for observation of dmUbp15 or free ubiquitin) on a Bruker Avance III spectrometer operating at 18.8 T (800 MHz ^{1}H Larmor frequency) and equipped with a TCI cryo-probe and z gradients. Spectra were processed with TopSpin (Bruker) and analyzed using Sparky[42]. ^{1}H chemical shifts were referenced directly, ^{13}C and ^{15}N chemical shifts indirectly[43], to added 2,2-dimethyl-2-silapentane-5-sulfonate (DSS, methyl ^{1}H signal at 0.00 ppm). Spectra for p15 backbone assignment were recorded using a 150 µM dmUbp15 [U-^{13}C,^{15}N] sample (with unlabeled ubiquitin) in PBS, pH 6.2, 5% (v/v) ^{2}H$_2$O, 0.01% NaN$_3$. The p15 resonance assignment was

derived from a set of 3D HNCO, HN(CA)CO, HNCA, HN(CO)CA, HNCACB, HN(CO)CACB, HN(CA)HA, and HN(COCA)HA experiments, all implemented as BEST–TROSY version[44]. Automatic assignment of backbone and $^{13}C^\beta$ resonances was obtained with MARS[45] and completed manually. The assignment for p15CCSS mutant was obtained in a similar way except that the sample was 50 μM and that no $^{13}C^\beta$ or $^1H^\alpha$ resonances were assigned. The NMR resonance assignment for p15CCSS and dmUbp15 have been deposited in the BioMagResBank with the entries 27698 and 27696, respectively. Secondary chemical shift values were calculated from the difference between measured $^1H^\alpha$, $^{13}C'$, $^{13}C^\alpha$, $^{13}C^\beta$ chemical shifts and their amino acid specific random-coil values[23]. Backbone amide ^{15}N transverse (T_2) relaxation times were measured from a series of 2D ^1H-^{15}N HSQC spectra modulated by ^{15}N CPMG periods[46] of different duration (16.02, 48.05, 96.09, 160.16, 256.26, and 400.4 ms). These spectra were acquired in an interleaved mode and signal intensities were fitted to a single exponential decay. Signal overlap prevented reliable intensity measurement for some residues that were not included in the further analysis.

For titrations of dmUbp15 with dsDNA we used a 65 μM protein sample in PBS pH 6.2 and a stock solution of concentrated dsDNA (1.72 mM) in the same buffer[4]. The dsDNA was formed by a 5'-TCAACATGATGTTCATAATCCCAA-3'oligonucleotide and its reverse complementary sequence. This was the same dsDNA that was used to measure the p15 interaction with DNA[4]. The dissociation constants (K_d) were derived from those residues that at the last point of the titration showed a weighted[47] amide ^1H and ^{15}N chemical shift perturbation (CSP) larger than the average plus one standard deviation. The CSP values along the titration were fitted to a 1:1 binding model for each residue, and the reported K_d values are the mean ± the standard error of the mean.

4.5.6. *Molecular Modeling and Small-Angle X-ray Scattering*

A random coil ensemble of p15CCSS containing *ca* 15,000 conformations was calculated by Flexible-Meccano (FM)[48,49] followed by side-chain modelling with SCCOMP[50], and energy-refinement in explicit solvent using GROMACS 5.1.1[51]. In parallel, we created a structural library for the aminoethane thiol linker (NH$_2$-CH$_2$-CH$_2$-SH) by running a Molecular Dynamics (MD) simulation of a neutral Ac-GlyGlyCysGlyGly-NH$_2$ peptide (PET5, with acetylated and

amidated N and C termini) in water at 25 ºC, where the linker was attached to the central cysteine residue and prepended by Ac-Gly to mimic a minimal protein context (i.e., Ac-Gly-CO-NH-CH$_2$-CH$_2$-**S**-S-CH$_2$-PET5). We used the force field AMBER99sb-ILDN[52] with no ions added, and TIP3P[53] as the water model. The system was equilibrated for 0.2 ns in the NVT ensemble and then for another 0.2 ns in the NPT ensemble with a 2 fs time-step. The production MD run spanned 14 ns, and from the last 10 ns a frame was extracted every 10 ps to create a library of 1000 different linker positions. From this library, the two engineered cysteine residues of each p15CCSS structure were *in-silico* labelled with one randomly selected and sterically allowed linker conformation, followed by grafting an ubiquitin-moiety (PDB: 1UBQ), with flexible C-terminal residues 74-76, onto the amino group of the attached linker (i.e., Ub-Gly76-CO-NH-CH$_2$-CH$_2$-S-S-CH$_2$-Cys15/24-p15), resulting in 12,000 doubly ubiquitinated p15 (dmUbp15) conformers with no steric clashes. To obtain the equivalent pool of p15CCSS conformers, we removed the linkers and the Ub-tags, so that each dmUbp15 conformer had its untagged version. We also created an ensemble model of PCNA-dmUbp15 complex with three dmUbp15 molecules bound to PCNA trimer via the PIP-Box region, as observed in the crystal structure of the PCNA-p15 complex[4].

Synchrotron SEC-SAXS data were collected at 25 ºC on the B21 beamline at the Diamond Light Source (Didcot, UK) using an in-line Agilent HPLC system. Samples of 45 µL with 5.05 and 13.05 g/L of p15CCSS and dmUbp15, respectively, were injected into a 4.6 ml Shodex KW402.5-4F size exclusion column at a flow rate of 0.16 mL min^{-1}. The SEC mobile phase consisted of PBS pH 7.0 (plus 1mM DTT in the case of p15CCSS). Two-second frames were acquired using a Pilatus 2M pixel detector (DECTRIS) at a sample-detector distance of 4 m and a wavelength of $\lambda = 1.0$ Å, covering a moment of transfer range of $0.0032 < s < 0.37$ Å^{-1}. Scattering intensities from p15CCSS and dmUbp15 SEC-peak regions were integrated and buffer subtracted to produce the SAXS-profile from each sample using the ScÅtter software [54]. From the profiles the pairwise distance distribution functions, *P(r)*, were obtained by indirect Fourier Transform with GNOM[55] using a momentum transfer range of $0.010 < s < 0.37$ Å^{-1}. The R_g values were estimated by applying the Guinier approximation in the range $s < 0.8/R_g$ for p15CCSS and $s < 1.3/R_g$ for dmUbp15. We used CRYSOL[56] to compute the theoretical SAXS profiles from conformational ensembles of untagged p15CCSS and doubly ubiquitinated dmUbp15. All theoretical curves were obtained with 101 points and a maximum scattering vector of 0.37 Å$^-$

[1] using 25 harmonics. Assuming that p15 remains disordered upon ubiquitination, we employed the ensemble optimization method (EOM)[26] to select from the p15CCSS and dmUbp15 structure pools, a set of 50 conformers whose theoretical SAXS profiles fit the SEC-SAXS profiles obtained for p15CCSS and dmUbp15. EOM was performed by minimizing the χ^2 between experimental (I_{exp}^j) and average theoretical (I_{theo}^j) SAXS intensities:

$$\chi_j^2 = \frac{1}{K-1} \sum_{i=1}^{K} \left[\frac{I_{exp}^j(s_i) - \mu I_{theo}^j(s_i)}{\sigma(s_i)} \right]^2$$

where K is the number of data points of each SEC-SAXS profile (I_{exp}^j), $\sigma(s)$ the standard deviations of the scattering intensity, and μ a scaling factor. I_{theo}^j was obtained by averaging the scattering of 50 explicit models per species (i.e., j = p15CCSS or dmUbp15):

$$I_{theo}^j(s) = \frac{1}{N} \sum_{1}^{N=50} j(s)$$

By running EOM multiple times (i.e., 220 runs) we were able to compare the Rg distribution and joint distribution of distances between the centers of mass of the two ubiquitin moieties (d_{Ub-Ub}) vs end-to-end distances (d_{ee}) of the selected ensemble and the initial pool.

4.5.7. *Isothermal Titration Calorimetry*

For dmUbp15 binding to PCNA we employed a Microcal iTC200 calorimeter with 25 µM PCNA in the cell at 25 ºC. The PCNA protein solution (dialyzed against PBS, pH 7.0) was titrated with a 0.4 mM stock solution of dmUbp15 prepared by dissolving the lyophilized material in the dialysis buffer and adjusting the pH to 7.0 with NaOH. For binding to p15CCSS and p15C54S, PCNA was dialyzed against PBS pH 7.0, 2 mM TCEP. PCNA 20 µM solutions in the cell were titrated with 400 µM of the corresponding p15 variant. The binding isotherms were analyzed by non-linear least-squares fitting of the experimental data to a model assuming a single set of equivalent sites[57], using Microcal Origin (OriginLab).

For experiments with the RFTS domain of Dnmt1 we used a Microcal PEAQ-ITC (Malvern) calorimetry system. dmUbp15, mUbp15_15, mUbp15_24, p15 or ubiquitin were placed in the cell and RFTS in the syringe, in PBS, pH 7.0. The binding isotherms were analyzed by non-linear least-squares fitting of the experimental data to a model assuming a single set of equivalent

sites, using the Malvern software and plotted with Origin. All the experiments were performed at 25 °C.

4.6. Supplementary material

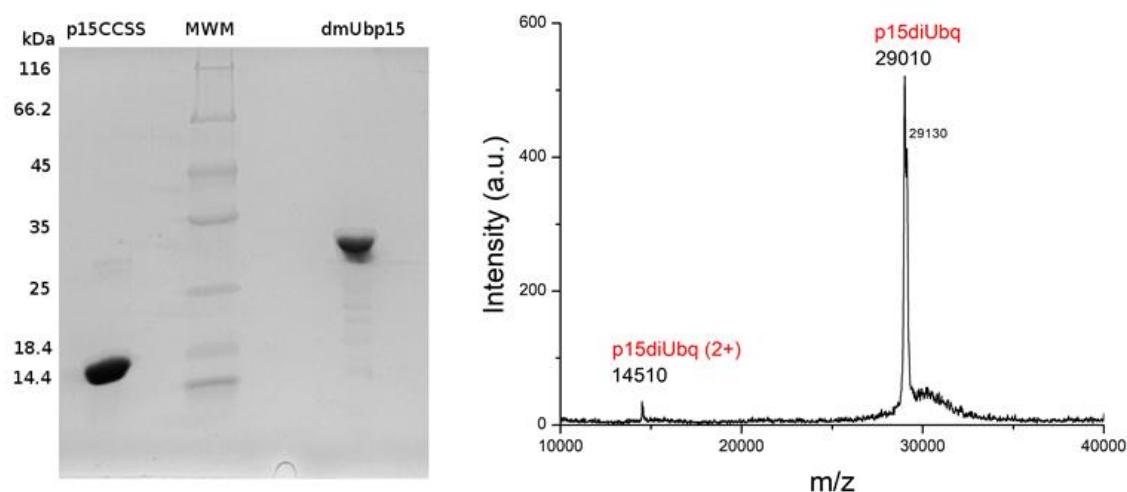### 4.6.1. *Supplementary figures*

Figure S4.1. (**Left**) SDS-PAGE analysis of purified p15CCSS mutant and it doubly monoubiquitinated form dmUbp15. MWM are the molecular weight markers with their weights indicated at the left. A 12 % gel was loaded with approximately 5 μg of p15CCSS or dmUbp15 in non-reducing loading buffer and was run at room temperature at 220 V for 45 min. The gel was stained with Coomassie brilliant blue and de-stained with 30% acetic acid in ethanol. (**Right**) MALDI-TOF spectrum of purified dmUbp15 showing the single and double charge peaks. The major peak corresponds to the protein without the initial methionine of the p15 moiety (the calculated mass for this molecule is 29,019 Da), and the minor one to the chain with the initial methionine included (calculated molar mass is 29,150 Da). A 2 μl protein sample was desalted using ZipTip® C4 micro-columns (Merck Millipore, Billerica, MA, USA) and eluted with 0.5 μl SA (sinapinic acid, 10 mg/ml in 7:3 Acetonitrile: Trifluoroacetic acid 0.1%) matrix onto a GroundSteel massive 384 target (Bruker Daltonics, Billerica, MA, USA). An Autoflex III MALDI-TOF/TOF spectrometer (Bruker Daltonics) was used in linear mode with the following settings: 5000-40000 Th window, linear positive mode, ion source 1: 20 kV, ion source 2: 18.5 kV, lens: 9 kV, pulsed ion extraction of 120 ns, high gating ion suppression up to 1000 Mr. Mass calibration was performed externally with protein 1 standard calibration mixture (Bruker Daltonics). Data acquisition, peak piking and subsequent spectra analysis was performed using FlexControl 3.0 and FlexAnalysis 3.0 software (Bruker Daltonics).
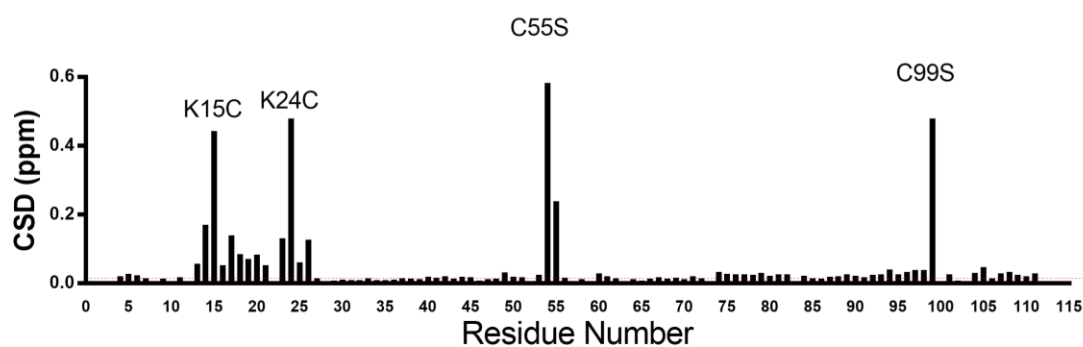
Figure S4.2. Chemical Shifts Deviations (CSD) between p15 and dmUbp15 measured in PBS pH 6.2 at 25°C (corresponding to the spectra shown in Figure 4.2B). The red dotted line indicates the experimental error estimated from the digital resolution of the spectra (0.006 ppm). The deviations were calculated as the combined difference in ^{1}H and ^{15}N chemical shifts. The location of the four mutations is indicated above the graph. Note that C15 and C24 and ligated to ubiquitin in dmUbp15.
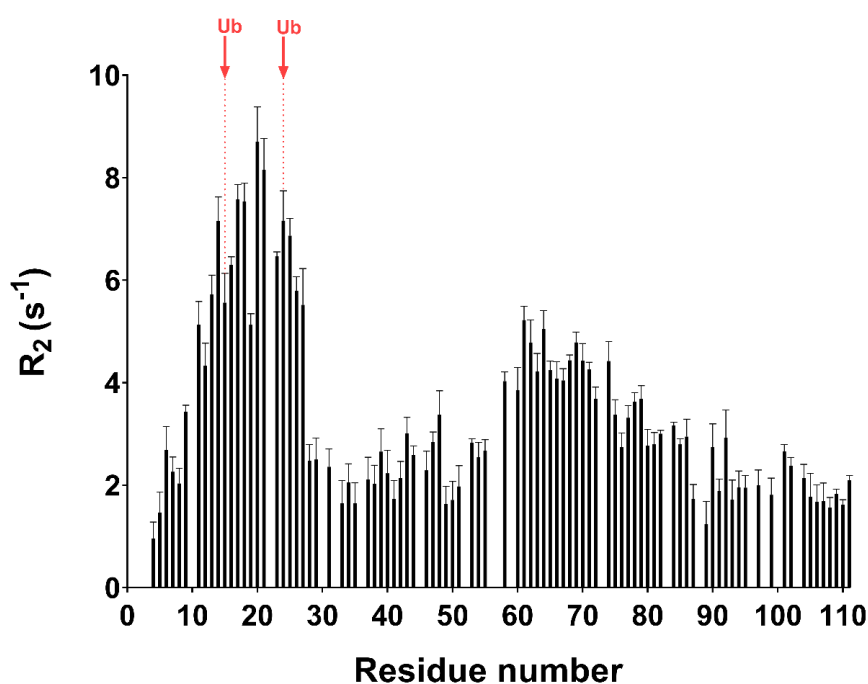

Figure S4.3. Transverse ^{15}N relaxation R_2 values (81.1 MHz, 25 ºC), shown against the dmUbp15 residue number. Sample: [U-^{15}N] 65 µM dmUbp15 in PBS pH 6.2 (137 mM NaCl, 2.7 mM KCl, 10 mM sodium phosphate, 2 mM potassium phosphate). The two ubiquitination sites are indicated in red.
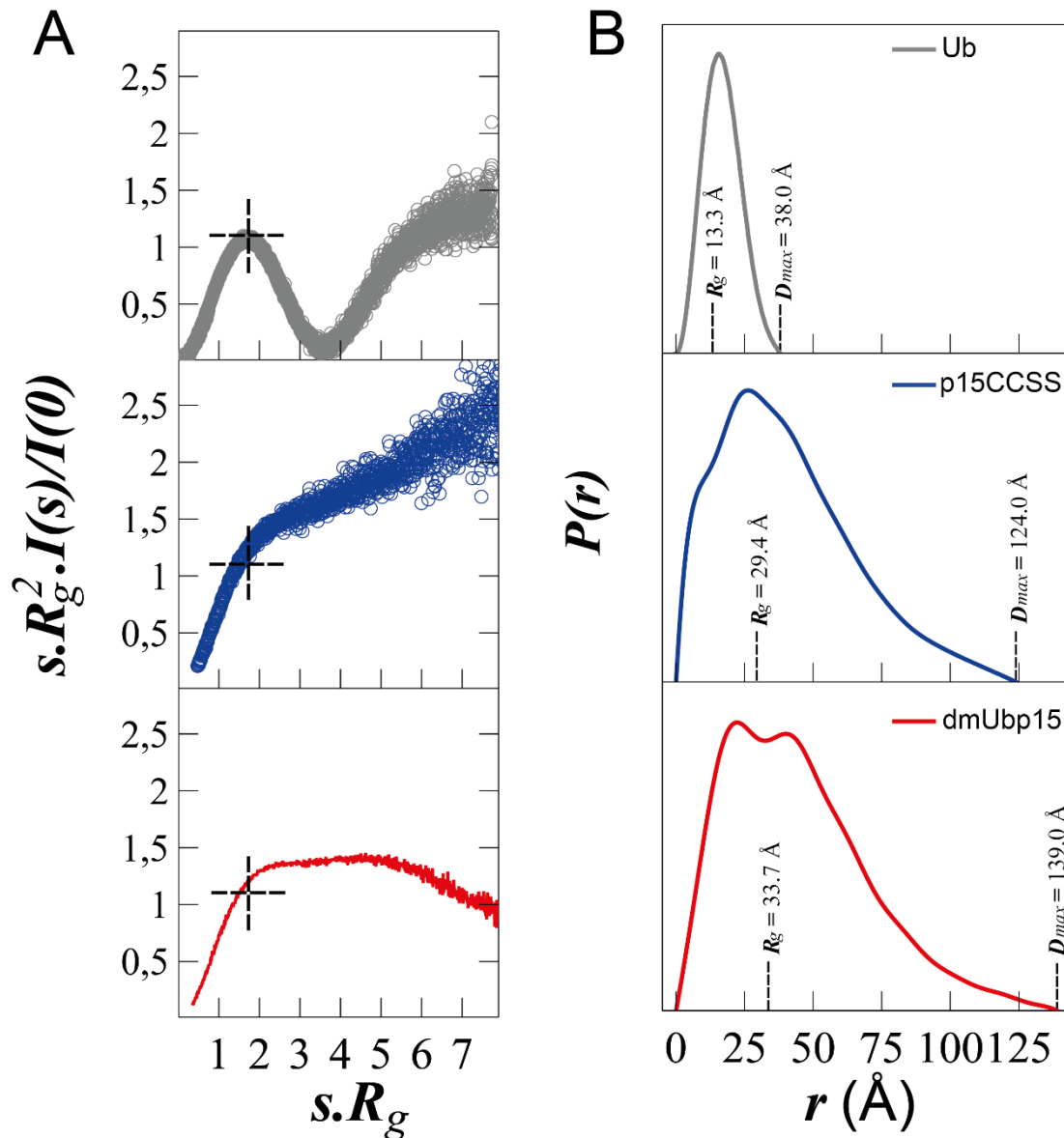
Figure S4.4. Comparative SAXS of p15CCSS and dmUbp15 to the monomeric globular ubiquitin protein. (A) Kratky representations of the SAXS patterns of ubiquitin (gray circles), p15CCSS (blue circles), and dmUbp15 (red) next to their respective *P(r)* versus *r* profiles (B) plotted using the same color code. The derived R_g and D_{max} values are indicated with dashed lines. The SAXS experimental set for free human Ub was obtained from the curated repository for scattering data SASDB[58] (www.sasbdb.org), with the entry code SASDAQ2[59].
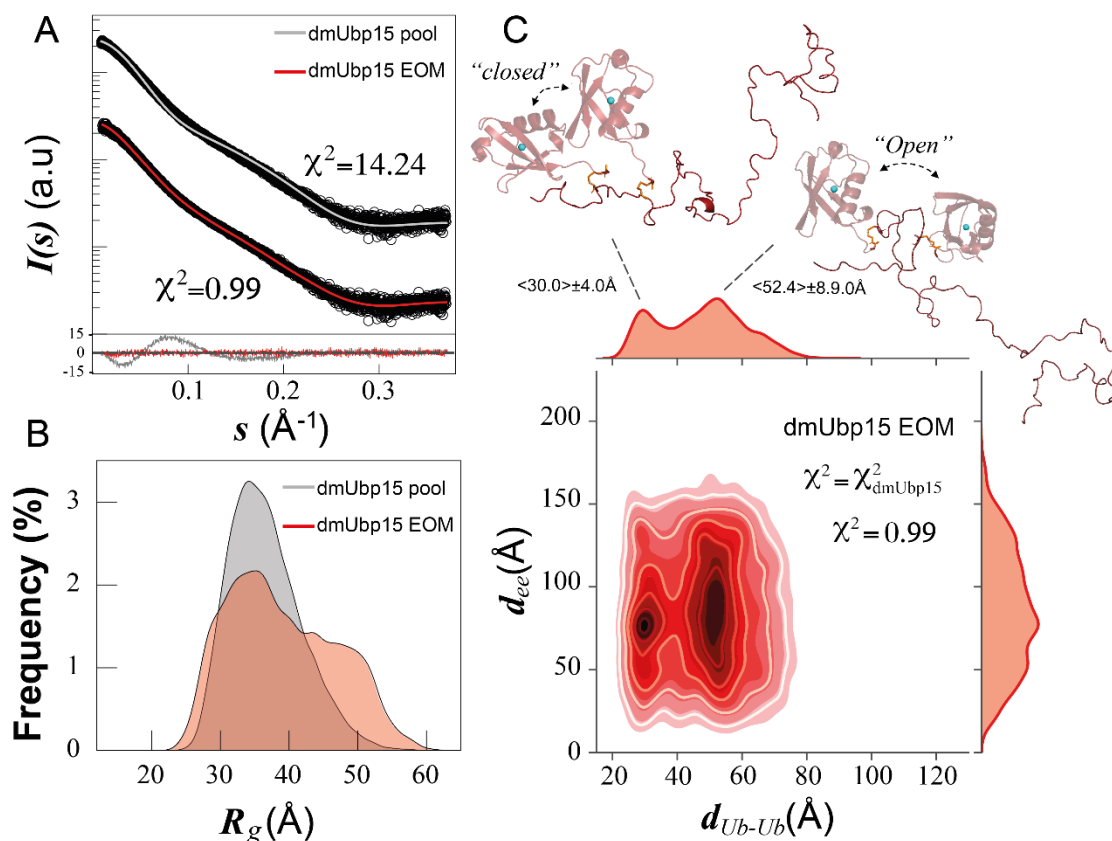
Figure S4.5. (A) Logarithmic-scale representation of scattering intensity, *I(s)*, as a function of the momentum transfer, *s*, measured for dmUbp15 (empty black circles). Solid lines are averaged back-calculated curves derived from the initial pool of dmUbp15 structures (grey) and EOM-select conformations (red). Residuals using absolute values are displayed at the bottom with the same color code, with a scale ranging from −15.0 to 15.0. (B) Comparison of the R_g distributions of the EOM-selected structures (red) and the initial pool of 12,000 random-coil structures of dmUbp15 (grey). As for p15CCSS, dmUbp15 refined ensemble is more extended than the statistically random pool (~18% of enrichment). (C) EOM-selected ensemble represented in the d_{ee} and d_{Ub-Ub} conformational space. Representative structural cartoons of closed and open states are displayed at the top. The center of mass of each ubiquitin moiety are indicated with cyan spheres.
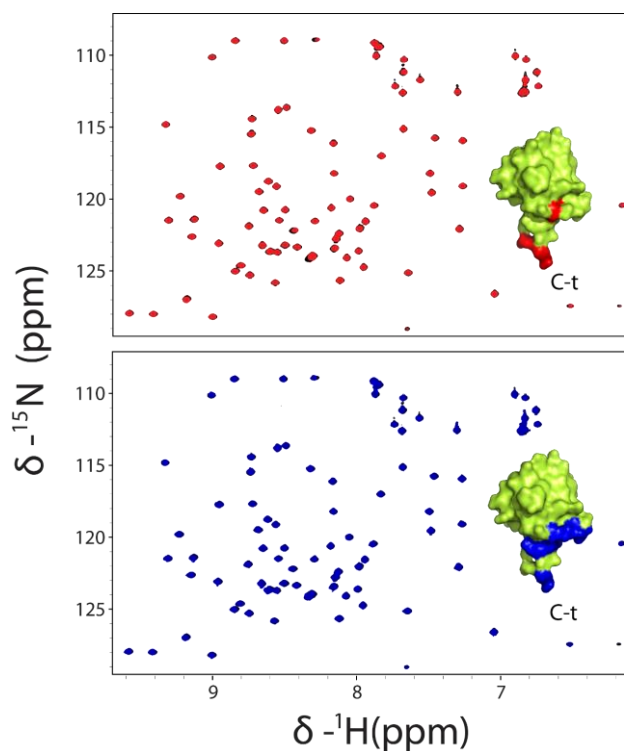
Figure S4.6. Overlay of ^{1}H-^{15}N HSQC spectra of a 17 µM sample of dmUbp15(black) against 17µM of mUbp15_15 in red (upper panel) and mUbp15_24 in blue (lower panel). The spectra were recorded in PBS at 25 ºC and pH 7.0. The p15 moiety is unlabeled while ubiquitin moieties are enriched in ^{15}N. The surface representations of ubiquitin inserted in the two panels show in red or in blue those residues experiencing a combined ^{1}H and ^{15}N chemical shift perturbation larger than the average plus one standard deviation.
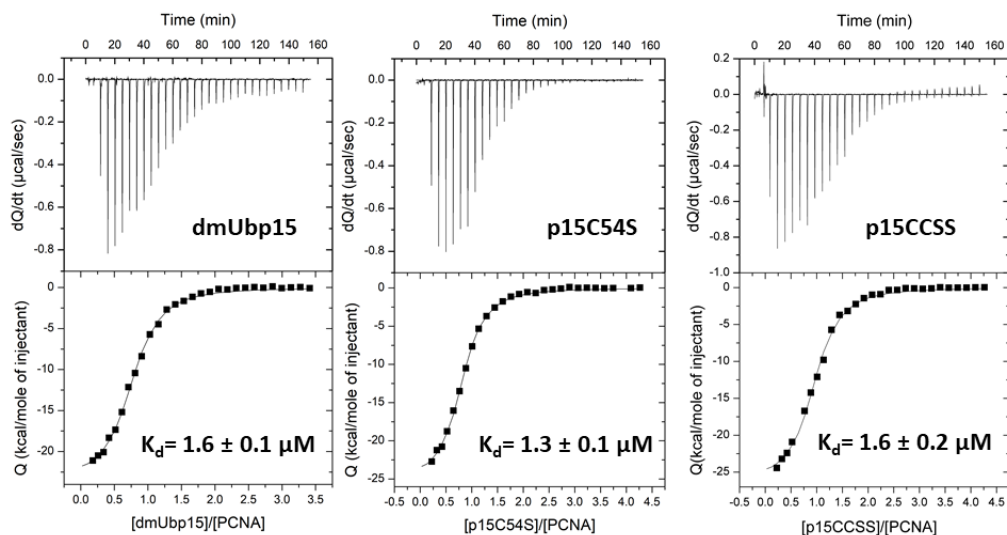
Figure S4.7. Calorimetric titrations of PCNA with dmUbp15 (left) or p15 mutant C54S (right). Upper panels represent the heat effect associated with the peptide injections and lower panels represent the ligand concentration dependence of the heat released upon binding. The symbols correspond to the experimental data and the continuous line the best fit to a model of one set of identical binding sites.
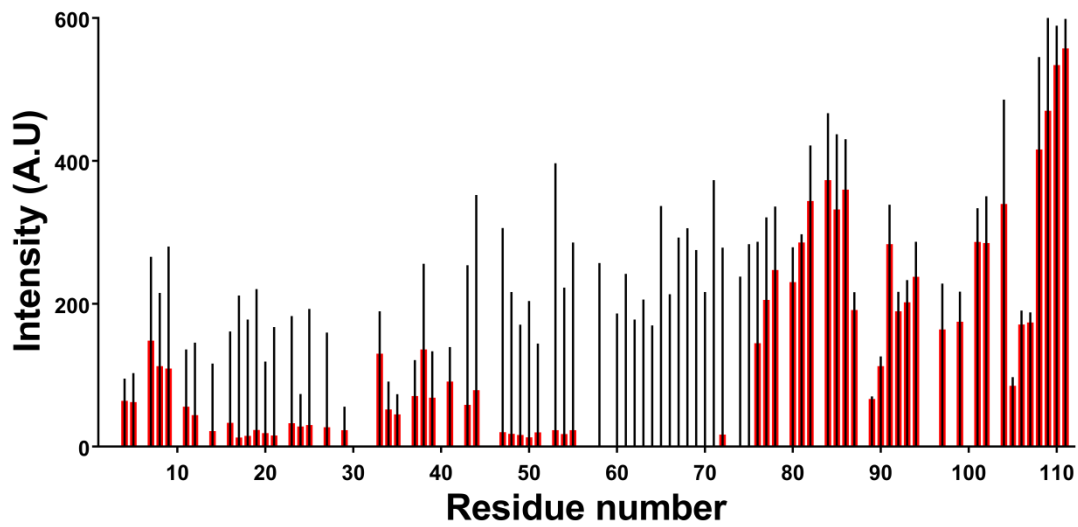
Figure S4.8. NMR signal intensities measured in the 2D ^{1}H,^{15}N HSQC spectra of dmUbp15 (^{15}N enriched only in the p15 moiety) in the absence (black) or presence (red) of PCNA. The intensity errors are below 0.2 % of the largest signal intensity observed in each spectrum.
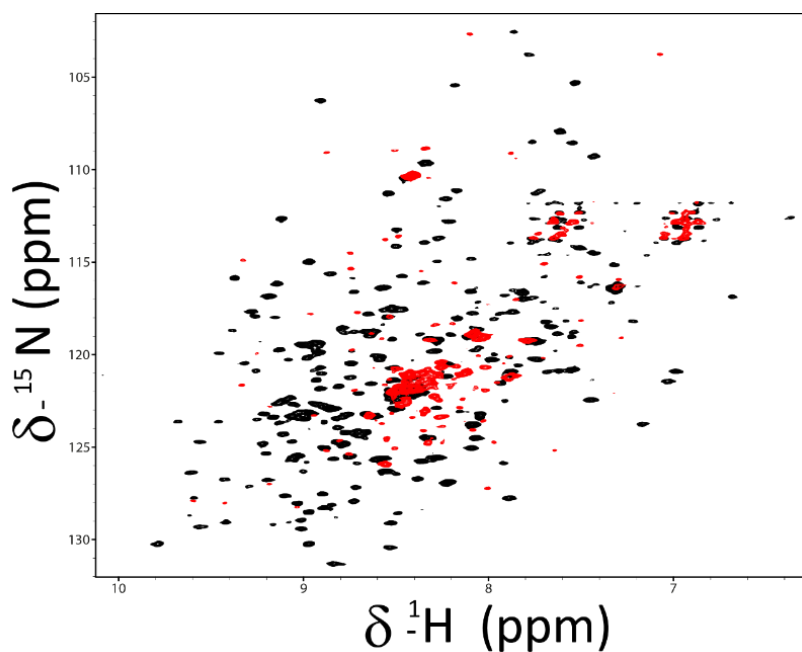

Figure S4.9. Overlay of ^{1}H,^{15}N TROSY spectra of perdeuterated PCNA [U-^{2}H,^{15}N] in the absence (black) or presence (red) of unlabeled dmUbp15 at 1:5 molar ratio. The experiments were recorded at 800 MHz and 25 ºC in PBS, pH 7.0, 1 mM DTT. The dispersed red resonances (which are just above the noise level) correspond to the natural abundance signals of the ubiquitin moieties.
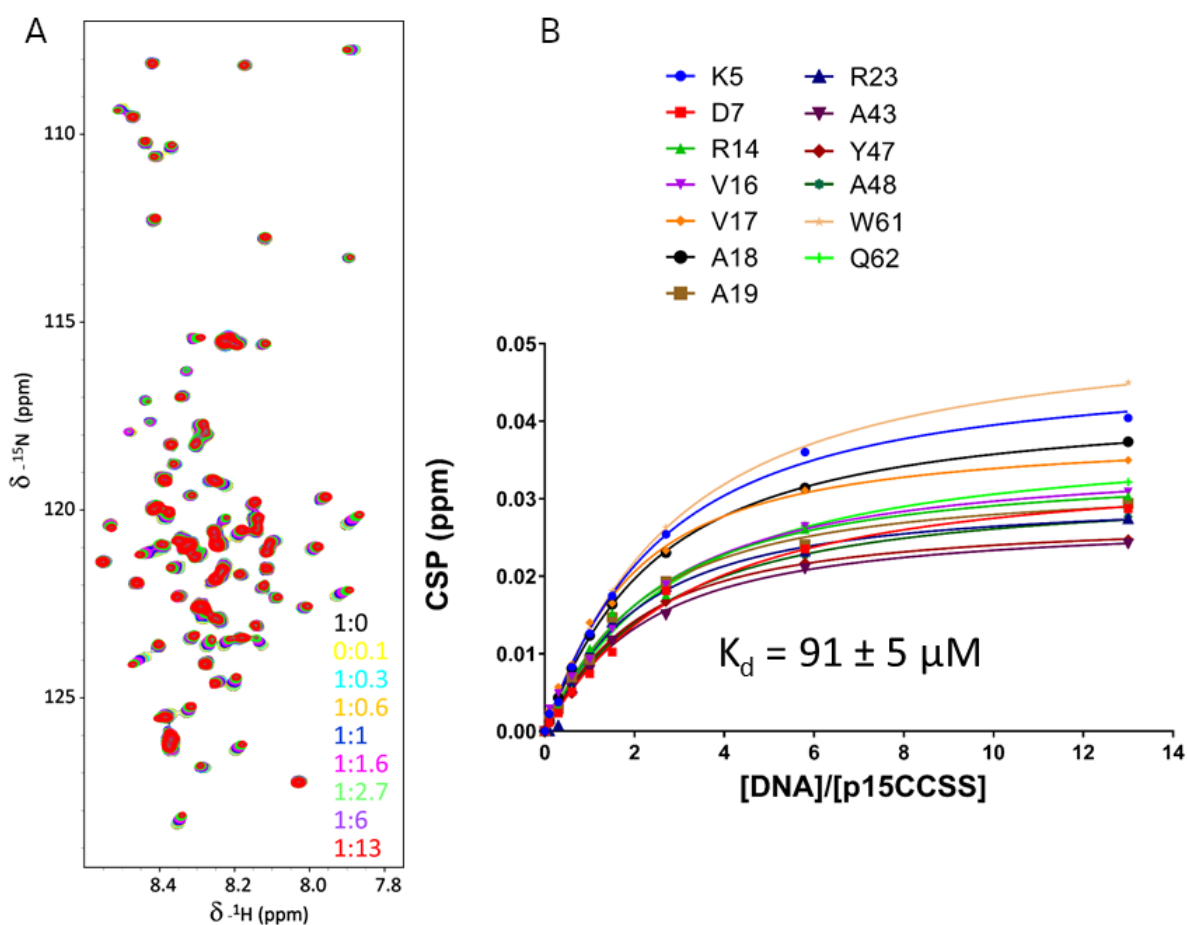
Figure S4.10. NMR analysis of p15CCSS (at 65 µM) binding to a 24 bp DNA duplex in PBS, 10 mM DTT, pH 6.3, 25 ºC. The (**A**) Overlay of ^{1}H,^{15}N HSQC spectra of U-^{15}N p15CCSS at increasing DNA:p15CCSS molar ratios with colors as indicated. (**B**) Chemical shift perturbation, CSP, versus DNA:dmUbp15 molar ratio for residues with CSP larger than the average plus one standard deviation. The indicated curves show fitting to a single-site binding model and yield *per residue* K_d values. The mean value of the dissociation constant and the standard error are indicated.
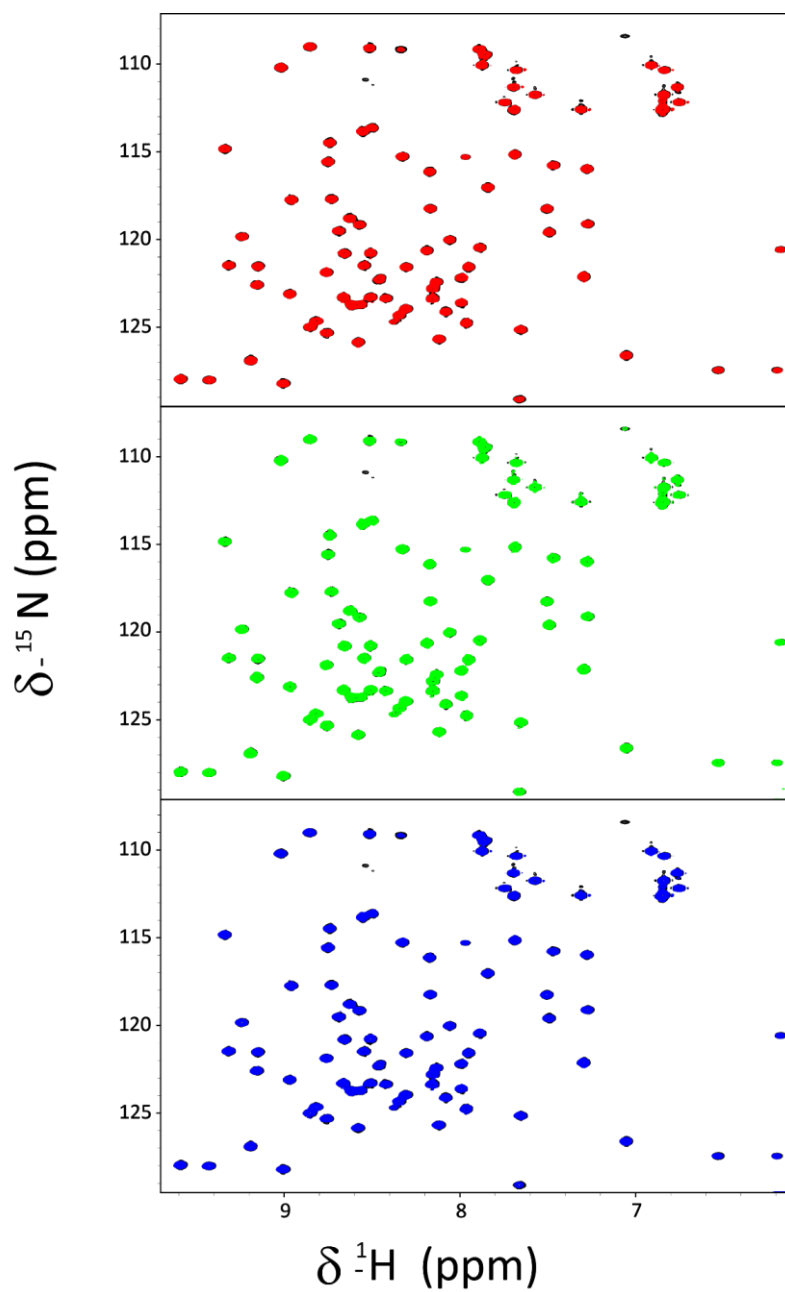
Figure S4.11. Overlay of ^{1}H-^{15}N HSQC spectra of a sample containing 50 µM of U-^{15}N isolated ubiquitin in the absence (black) or in the presence of dsDNA (red), p15 (green) or PCNA (blue) in a four-fold excess. The spectra were recorded in PBS pH 7.0 at 25 °C.
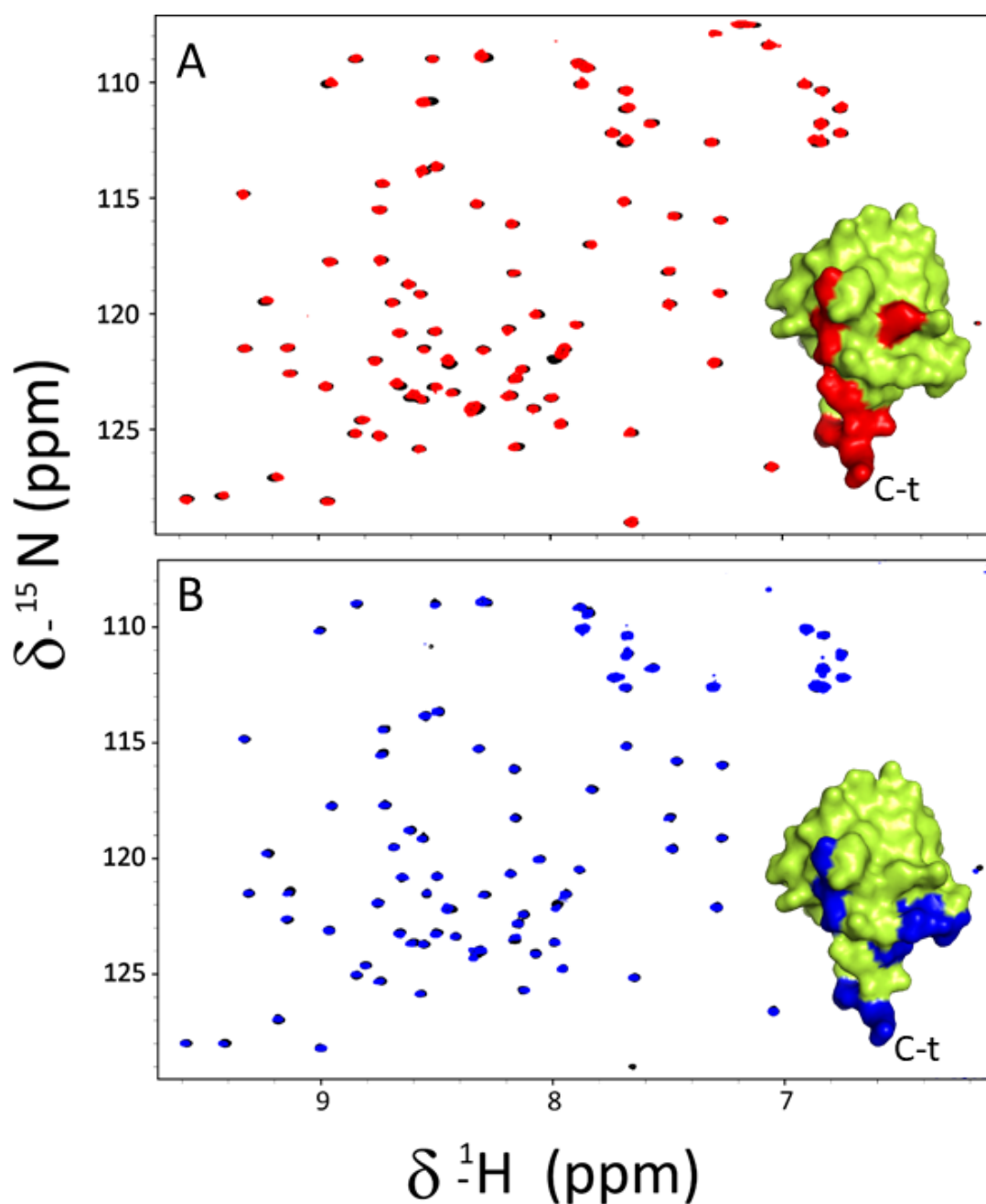
Figure S4.12. Overlay of ^{1}H-^{15}N HSQC spectra of a 21 μM sample of dmUbp15 with ^{15}N enrichment in the ubiquitin moieties only, in the absence (black) or in the presence of dsDNA (red), or PCNA (blue) in a four-fold excess. The spectra were recorded in PBS at 25 °C and pH 6.2 (A) or 7.0 (B). The surface representations of ubiquitin inserted in the two panels show in red or in blue those residues experiencing a combined ^{1}H and ^{15}N chemical shift perturbation larger than the average plus one standard deviation.
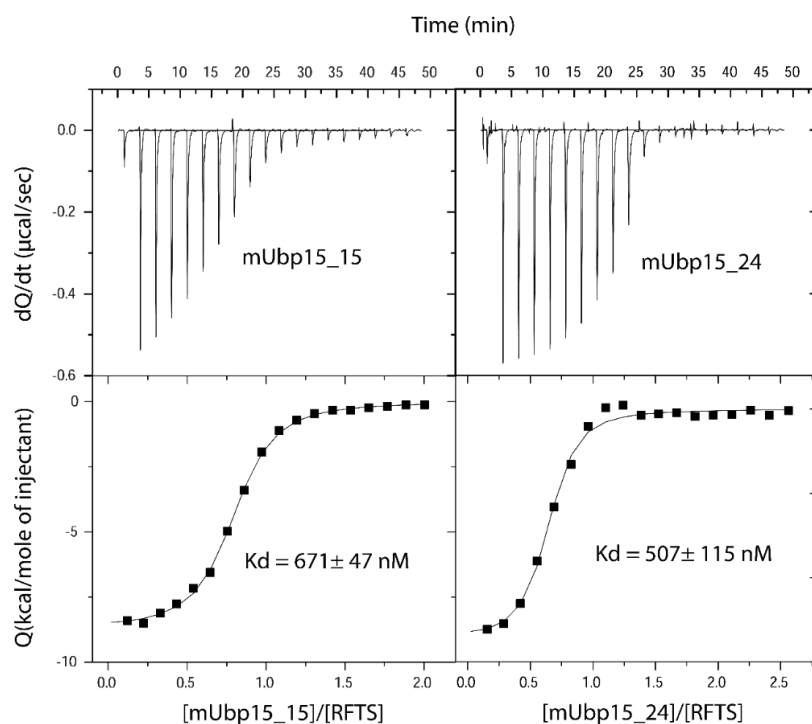
Figure S4.13. Calorimetric titrations of mUbp15C24K (left) and mUbp15C15K (right) with RFTS. Upper panels represent the heat effect associated with the peptide injections and lower panels represent the ligand concentration dependence of the heat released upon binding. The symbols correspond to the experimental data and the continuous line the best fit to a model of one set of identical binding sites. The error in the K_d values is the fitting error.

4.6.2. *Supplementary Tables*

Table S4.1. Thermodynamic parameters of the binding of p15 molecules to PCNA at 25 ºC as measured by ITC in PBS pH 7.0, 2 mM TCEP (except in the case of dmUbp15 where no TCEP was present).

p15 molecule	K_d [a] (μM)	K_a [b] ($10^5 \cdot M^{-1}$)	ΔG [c] (kcal/mol)	ΔH [d] (kcal/mol)	$-T\Delta S$ [e] (kcal/mol)	N [f]
dmUbp15	1.6 ± 0.1	6.3 ± 0.5	-7.90 ± 0.04	-23.9 ± 0.4	16.0 ± 0.4	0.78 ± 0.01
p15CCSS	1.6 ± 0.2	6.2 ± 0.9	-7.90 ± 0.08	-26.5 ± 0.7	18.4 ± 0.8	0.96 ± 0.01
p15C54S	1.3 ± 0.1	6.6 ± 0.3	-8.00 ± 0.03	-26.3 ± 0.3	17.3 ± 0.3	0.81 ± 0.01

[a]Dissociation constant.

[b]Association constant.

[c]Standard free energy change of the association reaction.

[d]Standard enthalpy change of the association reaction.

[e]Standard entropy change of the association reaction multiplied by the absolute temperature.

[f]Stoichiometry (the number of p15 binding sites on a single PCNA protomer), which was an adjustable parameter in the fitting of the experimental data.

The errors in the table are the fitting errors.

Table S4.2. Thermodynamic parameters of the binding of RFTS to different p15 molecules at 25 ºC as measured by ITC in PBS pH 7.0

p15 molecule	Kd [a] (µM)	Ka [b] ($10^7 \cdot$M-1)	ΔG[c] (kcal/mol)	ΔH[d] (kcal/mol)	$-T\Delta S$[e] (kcal/mol)	N[f]
dmUbqp15	0.029±0.008	3.6 ±1.1	-10.26± 0.17	-15.83 ± 1.14	5.5 ± 1.22	0.96 ±0.08
mUbqp15_15	0.67±0.05	0.15 ± 0.01	-8.39 ± 0.01	-8.72 ± 0.09	0.3 ± 0.1	0.767± 0.004
mUbqp15_24	0.50±0.1	0.20 ±0.01	-8.56 ± 0.01	-13.70 ± 0.46	5.1 ± 0.5	0.60 ±0.01

[a]Dissociation constant.

[b]Association constant.

[c]Standard free energy change of the association reaction.

[d]Standard enthalpy change of the association reaction.

[e]Standard entropy change of the association reaction multiplied by the absolute temperature.

[f]Stoichiometry which was an adjustable parameter in the fitting of the experimental data.

The errors in the table are the fitting errors except for dmUbp15 which are reported as the standard deviations from three replicates (and the corresponding magnitude value is the average of the three measurements).

4.7. References

1. Bah, A.; Forman-Kay, J. D., Modulation of Intrinsically Disordered Protein Function by Post-translational Modifications. J Biol Chem 2016, 291, 6696-6705.

2. Yu, P.; Huang, B.; Shen, M.; Lau, C.; Chan, E.; Michel, J.; Xiong, Y.; Payan, D. G.; Luo, Y., p15(PAF), a novel PCNA associated factor with increased expression in tumor tissues. Oncogene 2001, 20, 484-489.

3. Moldovan, G. L., Pfander, B. & Jentsch, S. PCNA, the maestro of the replication fork. *Cell* **129**, 665-679, doi:S0092-8674(07)00594-6 [pii], 10.1016/j.cell.2007.05.003 (2007).

4. De Biasio, A. *et al.* Structure of p15(PAF)-PCNA complex and implications for clamp sliding during DNA replication and repair. *Nat Commun* **6**, 6439, doi:10.1038/ncomms7439 (2015).

5. Povlsen, L. K. *et al.* Systems-wide analysis of ubiquitylation dynamics reveals a key role for PAF15 ubiquitylation in DNA-damage bypass. *Nat Cell Biol* **14**, 1089-1098, doi:ncb2579 [pii],10.1038/ncb2579 (2012).

6. Karg, E. *et al.* Ubiquitome Analysis Reveals PCNA-Associated Factor 15 (PAF15) as a Specific

Ubiquitination Target of UHRF1 in Embryonic Stem Cells. *J Mol Biol* **429**, 3814-3824, doi:10.1016/j.jmb.2017.10.014 (2017).

7   Nishiyama, A. *et al.* Uhrf1-dependent H3K23 ubiquitylation couples maintenance DNA methylation and replication. *Nature* **502**, 249-253, doi:10.1038/nature12488 (2013).

8   Qin, W. *et al.* DNA methylation requires a DNMT1 ubiquitin interacting motif (UIM) and histone ubiquitination. *Cell Res* **25**, 911-929, doi:10.1038/cr.2015.72 (2015).

9   Hosokawa, M. *et al.* Oncogenic role of KIAA0101 interacting with proliferating cell nuclear antigen in pancreatic cancer. *Cancer Res* **67**, 2568-2576, doi:67/6/2568 [pii], 10.1158/0008-5472.CAN-06-4356 (2007).

10   Kato, T. *et al.* Overexpression of KIAA0101 predicts poor prognosis in primary lung cancer patients. *Lung Cancer* **75**, 110-118, doi:S0169-5002(11)00324-2 [pii], 10.1016/j.lungcan.2011.05.024 (2012).

11   Liu, L. *et al.* Variant 1 of KIAA0101, overexpressed in hepatocellular carcinoma, prevents doxorubicin-induced apoptosis by inhibiting p53 activation. *Hepatology* **56**, 1760-1769, doi:10.1002/hep.25834 (2012).

12   Wang, X. *et al.* PAF-Wnt signaling-induced cell plasticity is required for maintenance of breast cancer cell stemness. *Nat Commun* **7**, 10633, doi:10.1038/ncomms10633 (2016).

13   Ong, D. S. T. *et al.* PAF promotes stemness and radioresistance of glioma stem cells. *Proc Natl Acad Sci U S A* **114**, E9086-E9095, doi:10.1073/pnas.1708122114 (2017).

14   De Biasio, A. *et al.* p15(PAF) Is an Intrinsically Disordered Protein with Nonrandom Structural Preferences at Sites of Interaction with Other Proteins. *Biophys J* **106**, 865-874, doi:S0006-3495(14)00072-1 [pii], 10.1016/j.bpj.2013.12.046 (2014).

15   Ishiyama, S. *et al.* Structure of the Dnmt1 Reader Module Complexed with a Unique Two-Mono-Ubiquitin Mark on Histone H3 Reveals the Basis for DNA Methylation Maintenance. *Mol Cell* **68**, 350-360 e357, doi:10.1016/j.molcel.2017.09.037 (2017).

16   Shah, N. H., Dann, G. P., Vila-Perello, M., Liu, Z. & Muir, T. W. Ultrafast protein splicing is common among cyanobacterial split inteins: implications for protein engineering. *J Am Chem Soc* **134**, 11338-11341, doi:10.1021/ja303226x (2012).

17   Abeywardana, T., Lin, Y. H., Rott, R., Engelender, S. & Pratt, M. R. Site-specific differences in proteasome-dependent degradation of monoubiquitinated alpha-synuclein. *Chem Biol* **20**, 1207-1213, doi:10.1016/j.chembiol.2013.09.009 (2013).

18   Faggiano, S. & Pastore, A. The challenge of producing ubiquitinated proteins for structural studies. *Cells* **3**, 639-656, doi:10.3390/cells3020639 (2014).

19   Morimoto, D. *et al.* The unexpected role of polyubiquitin chains in the formation of fibrillar

aggregates. *Nat Commun* **6**, 6116, doi:10.1038/ncomms7116 (2015).

20   De March, M. *et al.* p15PAF binding to PCNA modulates the DNA sliding surface. *Nucleic Acids Res* **46**, 9816-9828, doi:10.1093/nar/gky723 (2018).

21   Schubert, M., Labudde, D., Oschkinat, H. & Schmieder, P. A software tool for the prediction of Xaa-Pro peptide bond conformations in proteins based on 13C chemical shift statistics. *J Biomol NMR* **24**, 149-154 (2002).

22   Cho, M. K. *et al.* Amino acid bulkiness defines the local conformations and dynamics of natively unfolded alpha-synuclein and tau. *J Am Chem Soc* **129**, 3032-3033, doi:10.1021/ja067482k (2007).

23   Tamiola, K., Acar, B. & Mulder, F. A. Sequence-specific random coil chemical shifts of intrinsically disordered proteins. *J Am Chem Soc* **132**, 18000-18003, doi:10.1021/ja105656t (2010).

24   Cordeiro, T. N. *et al.* Small-angle scattering studies of intrinsically disordered proteins and their complexes. *Curr Opin Struct Biol* **42**, 15-23, doi:10.1016/j.sbi.2016.10.011 (2017).

25   Bernado, P. Effect of interdomain dynamics on the structure determination of modular proteins by small-angle scattering. *Eur Biophys J* **39**, 769-780, doi:10.1007/s00249-009-0549-3 (2010).

26   Bernado, P., Mylonas, E., Petoukhov, M. V., Blackledge, M. & Svergun, D. I. Structural characterization of flexible proteins using small-angle X-ray scattering. *J Am Chem Soc* **129**, 5656-5664, doi:10.1021/ja069124n (2007).

27   Liu, Z. *et al.* Noncovalent dimerization of ubiquitin. *Angew Chem Int Ed Engl* **51**, 469-472, doi:10.1002/anie.201106190 (2012).

28   Zhou, H. X. Quantitative account of the enhanced affinity of two linked scFvs specific for different epitopes on the same antigen. *J Mol Biol* **329**, 1-8 (2003).

29   Smith, Z. D. & Meissner, A. DNA methylation: roles in mammalian development. *Nat Rev Genet* **14**, 204-220, doi:10.1038/nrg3354 (2013).

30   Chuang, L. S. *et al.* Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1. *Science* **277**, 1996-2000 (1997).

31   Schermelleh, L. *et al.* Dynamics of Dnmt1 interaction with the replication machinery and its role in postreplicative maintenance of DNA methylation. *Nucleic Acids Res* **35**, 4301-4312, doi:10.1093/nar/gkm432 (2007).

32   Dyson, H. J. & Wright, P. E. Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol* **6**, 197-208, doi:nrm1589 [pii], 10.1038/nrm1589 (2005).

33   Tompa, P. & Fuxreiter, M. Fuzzy complexes: polymorphism and structural disorder in protein-

protein interactions. *Trends Biochem Sci* **33**, 2-8, doi:S0968-0004(07)00285-X [pii], 10.1016/j.tibs.2007.10.003 (2008).

34   Sharma, R., Raduly, Z., Miskei, M. & Fuxreiter, M. Fuzzy complexes: Specific binding without complete folding. *FEBS Lett* **589**, 2533-2542, doi:10.1016/j.febslet.2015.07.022 (2015).

35   Wang, Y., Bugge, K., Kragelund, B. B. & Lindorff-Larsen, K. Role of protein dynamics in transmembrane receptor signalling. *Curr Opin Struct Biol* **48**, 74-82, doi:10.1016/j.sbi.2017.10.017 (2018).

36   Ibanez de Opakua, A. *et al.* The metastasis suppressor KISS1 is an intrinsically disordered protein slightly more extended than a random coil. *PLoS One* **12**, e0172507, doi:10.1371/journal.pone.0172507 (2017).

37   Boeynaems, S. *et al.* Protein Phase Separation: A New Phase in Cell Biology. *Trends Cell Biol* **28**, 420-435, doi:10.1016/j.tcb.2018.02.004 (2018).

38   Cordeiro, T. N. *et al.* Disentangling polydispersity in the PCNA-p15PAF complex, a disordered, transient and multivalent macromolecular assembly. *Nucleic Acids Res* **45**, 1501-1515, doi:10.1093/nar/gkw1183 (2017).

39   Jimenji, T., Matsumura, R., Kori, S. & Arita, K. Structure of PCNA in complex with DNMT1 PIP box reveals the basis for the molecular mechanism of the interaction. *Biochem Biophys Res Commun* **516**, 578-583, doi:10.1016/j.bbrc.2019.06.060 (2019).

40   Syeda, F. *et al.* The replication focus targeting sequence (RFTS) domain is a DNA-competitive inhibitor of Dnmt1. *J Biol Chem* **286**, 15344-15351, doi:10.1074/jbc.M110.209882 (2011).

41   Takeshita, K. *et al.* Structural insight into maintenance methylation by mouse DNA methyltransferase 1 (Dnmt1). *Proc Natl Acad Sci U S A* **108**, 9055-9059, doi:10.1073/pnas.1019629108 (2011).

42   Sparky - NMR Assignment and Integration Software v. 3.0 (San Francisco, 2008).

43   Wishart, D. S. *et al.* [1]H, ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR. *J. Biomol. NMR* **6**, 135-140 (1995).

44   Solyom, Z. *et al.* BEST-TROSY experiments for time-efficient sequential resonance assignment of large disordered proteins. *J Biomol NMR* **55**, 311-321, doi:10.1007/s10858-013-9715-0 (2013).

45   Jung, Y. S., Sharma, M. & Zweckstetter, M. Simultaneous assignment and structure determination of protein backbones by using NMR dipolar couplings. *Angew Chem Int Ed Engl* **43**, 3479-3481, doi:10.1002/anie.200353588 (2004).

46   Farrow, N. A. *et al.* Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation. *Biochemistry* **33**, 5984-6003 (1994).

47   Palacios, A. *et al.* Solution structure and NMR characterization of the binding to methylated histone tails of the plant homeodomain finger of the tumour suppressor ING4. *FEBS Lett* **580**, 6903-6908 (2006).

48   Bernadó, P. *et al.* A structural model for unfolded proteins from residual dipolar couplings and small-angle x-ray scattering. *ProcNatlAcadSciUSA***102**, 17002-17007, doi:0506202102 [pii], 10.1073/pnas.0506202102 (2005).

49   Ozenne, V. *et al.* Flexible-meccano: a tool for the generation of explicit ensemble descriptions of intrinsically disordered proteins and their associated experimental observables. *Bioinformatics* **28**, 1463-1470, doi:bts172 [pii], 10.1093/bioinformatics/bts172 (2012).

50   Eyal, E., Najmanovich, R., McConkey, B. J., Edelman, M. & Sobolev, V. Importance of solvent accessibility and contact surfaces in modeling side-chain conformations in proteins. *J Comput Chem* **25**, 712-724, doi:10.1002/jcc.10420 (2004).

51   Hess, B., Kutzner, C., van der Spoel, D. & Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J Chem Theory Comput* **4**, 435-447, doi:10.1021/ct700301q (2008).

52   Lindorff-Larsen, K. *et al.* Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **78**, 1950-1958, doi:10.1002/prot.22711 (2010).

53   Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J Chem Phys* **79**, 926 (1983).

54   Rambo, R. P. ScÅtter, a JAVA-based application for basic analysis of SAXS datasets. *Diamond Light Source* **Version 3.1R**, UK (2017).

55   Svergun, D. I. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J Appl Cryst* **25**, 495-503 (1992).

56   Svergun, D. I., Barberato, C. & Koch, M. H. J. CRYSOL - a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *J Appl Cryst* **28**, 768-773 (1995).

57   Palacios, A. *et al.* Molecular basis of histone H3K4me3 recognition by ING4. *J Biol Chem* **283**, 15956-15964, doi:M710020200 [pii], 10.1074/jbc.M710020200 (2008).

58   Valentini, E., Kikhney, A. G., Previtali, G., Jeffries, C. M. & Svergun, D. I. SASBDB, a repository for biological small-angle scattering data. *Nucleic Acids Res* **43**, D357-363, doi:10.1093/nar/gku1047 (2015).

59   Mylonas, E. & Svergun, D. I. Accuracy of molecular mass determination of proteins in solution by small-angle X-ray scattering. *J Appl Cryst* **40**, s245-s249 (2007).

5. The p12 subunit of human polymerase δ uses an atypical PIP-box for molecular recognition of proliferating cell nuclear antigen (PCNA)

Amaia Gonzalez-Magaña[1], Alain Ibáñez de Opakua[1], Miguel Romano-Moreno[1], Javier Murciano-Calles[2], Nekane Merino[1], Irene Luque[2], Adriana L. Rojas[1], Silvia Onesti[3], Francisco J. Blanco[1,4]* and Alfredo De Biasio[3,5]*

5.1. Abstract

Human DNA polymerase δ is essential for DNA replication and acts in conjunction with the processivity factor proliferating cell nuclear antigen (PCNA). In addition to its catalyticc subunit (p125), pol δ comprises three regulatory subunits (p50, p68 and p12). PCNA interacts with all of these subunits, but only the interaction with p68 has been structurally characterized. Here, we report solution NMR-, isothermal calorimetry-, and X-ray crystallography- based analyses of the p12–PCNA interaction, which takes part in the modulation of the rate and fidelity of DNA synthesis by pol δ.

We show that p12 binds with micromolar affinity to the classical PIP-binding pocket of PCNA via a highly atypical PIP box located at the p12Nterminus. Unlike the canonical PIP box of p68, the PIP box of p12 lacks the conserved glutamine; binds through a 2-fork plug made of an isoleucine

and a tyrosine residue at +3 and +8 positions, respectively; and is stabilized by an aspartate at +6 position, which creates a network of intramolecular hydrogen bonds. These findings add to growing evidence that PCNA can bind a diverse range of protein sequences that may be broadly grouped as PIP-like motifs as has been previously suggested.

5.2. Introduction

Three eukaryotic DNA polymerases (pols), α, δ and ε participate in chromosomal DNA replication [1], with the latter two possessing the proofreading exonuclease activity required to replicate DNA with high fidelity. Human pol δ consists of a catalytic subunit (p125, harboring the polymerase and exonuclease activities), associated with three regulatory subunits (p50, p68, also known as p66, and p12) needed for optimal holoenzyme function [2,3], and there is evidence of different context-specific subassemblies of pol δ *in vivo* [3-5]. In particular, DNA damage or replication stress triggers the degradation of p12, a 12 kDa polypeptide of unknown structure, resulting in the formation of a three-subunit enzyme with an increased capacity for proofreading [3,4,6]. The processive activity of pol δ in DNA replication (*i.e.*, the ability of the polymerase to synthesize hundreds of base pairs without detaching from the

template) is conferred by its association with the sliding clamp PCNA, a ring-shaped homotrimer that encircles and slides on DNA [6-8].

Structurally, little is known on how the four subunits of mammalian pol δ interact with each other and PCNA. The p125 catalytic subunit and the p50 subunit form a tight heterodimer, which constitutes the core enzyme [9]. Biochemical analysis showed that the p68 subunit is attached to the core enzyme via an interaction between its N-terminal domain and p50, while p12 bridges the p125 and p50 subunits [10,11]. However, structural information is limited to the p50–p68 interaction only [12]. The p125 subunit, which is composed of a catalytic domain and a C-terminal domain containing an iron-sulfur cluster [13], has been reported to directly interact with PCNA [14,15], but the p125–PCNA complex showed negligible processivity *in vitro* [16], suggesting that the interaction is weak. Similarly, the p50–PCNA interaction, if any, is very weak [10,17]. On the other hand, the interactions between the p68 and p12 subunits and PCNA seem tighter [10,18], and especially p68 is critical for pol δ–PCNA processivity [16]. For both p68 and p12, examination of reconstituted holoenzymes in which the PCNA binding motifs (the PIP-box) have been mutated or inactivated has been performed [3,10,19], and it was found that both PIP-boxes were necessary for optimal pol δ activity. The PIP-box strict consensus sequence is *Qxxhxxaa*, where *h* is a hydrophobic, *a* is an aromatic, and *x* is any residue. The crystal structure of a peptide spanning the canonical PIP-box of p68 ([456]QVSITGFF[462]) bound to PCNA has been determined, showing the PIP-box interacting through the prototypical molecular surface observed in other PCNA-interacting partners [20]. The p68 PIP-box is located at the C-terminal region, which is predicted to be disordered (Supplementary Fig. 1).

Upon binding, the PIP-box forms a 3_{10} helix and the conserved hydrophobic trident inserts into a hydrophobic patch located between the N- and C-terminal domains of the PCNA protomer, whereas the glutamine binds in the so called "Q-pocket" [20]. By contrast, the PCNA-binding site of p12 is located in an N-terminal region [10], also predicted to be intrinsically disordered, and the proposed PIP-box, while highly conserved in mammals, is non-canonical ([4]KRLITDSY[11]; Fig. 1). Mutation of p12 residues I7, S10 and Y11 in the putative PIP-box results in defective binding of recombinant pol δ to PCNA, as well as in reduced pol δ processivity [10].
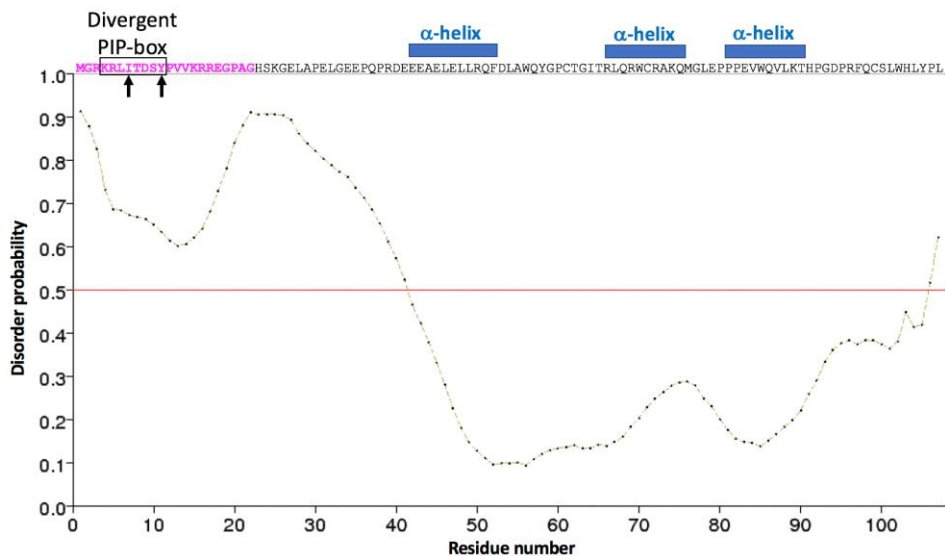
Figure 5.1. *Amino acid sequence of p12 together with disorder and secondary structure predictions*. Residues encompassing the peptide used for crystallization and biophysical characterization (p12[1-19]) with PCNA are indicated in bold pink characters, those belonging to the divergent PCNA interacting motif are boxed. The consensus residues in the PIP-like motif are indicated by arrows. The solid line shows the disorder prediction by the PrDOS server [21], and the red line the threshold of 0.5. Secondary structure elements (helices) predicted by JPred4 [22] are indicated above the sequence.

Because the p12 PIP-box significantly diverges from the strict PIP-box consensus sequence, we wondered whether it may encode for a novel PCNA binding motif with exclusive structural specificities. We therefore analyzed the p12–PCNA interaction by crystallography, NMR and isothermal calorimetry. Our data shows that a 19-residue p12 peptide containing the PCNA binding site recognizes PCNA via its divergent PIP-box, which adopts a characteristic 3_{10} helical fold. In the p12 PIP-box, the absence of the glutamine and the aromatic residue at +7 position and their associated stabilizing inter-molecular interactions is counterbalanced by an intra-molecular hydrogen-bonding network centered on the aspartate at +6 position, which stabilizes the 3_{10} helical conformation. Based on our data, p12 and p68 subunits contribute to the molecular recognition of PCNA by pol δ with different structural specificities and affinities. Surprisingly, we have found that the affinity of the p12–PCNA interaction is higher than that measured for the canonical PIP-box of the human DNA helicase RecQ5. These results reinforce the emerging idea of the existence of a broader class of PCNA interacting sequences, which may be called "PIP-like" motifs [23].

5.3. Results

5.3.3. *NMR and ITC analysis of the p12–PCNA interaction*

We first observed and characterized the interaction of PCNA with p12[1-19] by solution NMR. ^{2}H-^{15}N-labeled PCNA was titrated with unlabeled p12 peptide and chemical shift perturbations of PCNA backbone amide signals analyzed (Figure 5.2A). We identified two groups of perturbed residues: those whose signals gradually shift along the titration, implying a fast exchange regime on the NMR time scale (Figure 5.2B), and those residues whose signals broaden and disappear (due to signal attenuation below the noise level or untraceable shifting), indicating an intermediate exchange regime. When plotting the chemical shift perturbations (CSP) along the PCNA sequence (Figure 5.2C) a similar pattern as for p21 binding is observed [24], suggesting a similar mode of binding. The titration of the signals from those residues with CSP values larger than the average plus one standard deviation yielded an average dissociation constant of 130 ± 30 µM at 35 °C (Figure 5.2D).

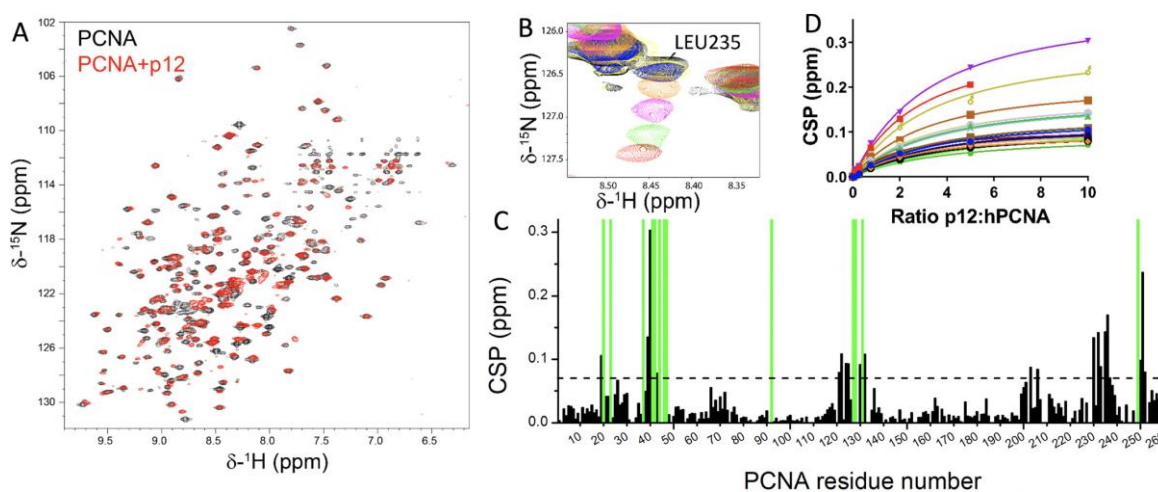**Figure 5.2**. *NMR analysis of the p12 peptide interaction with PCNA*. (A) Superposition of ^{1}H-^{15}N TROSY spectra of 51 µM PCNA in the absence (black) and presence (red) of a 10-fold molar excess of p12[1-19] peptide. Spectra were acquired at 35˚C in PBS pH 7.0, 1mM DTT. (B) Region of the NMR spectra of PCNA in the presence of increasing amounts of p12 peptide (from black to red) showing the titration of L235 signal. (C) Chemical shift perturbations (CSP) of PCNA backbone amide ^{1}H and ^{15}N NMR resonances induced by p12[1-19]. The dashed line indicates the average plus one standard deviation. The green bars indicate the position of residues that disappear upon peptide addition. (D) Chemical shift perturbation of the amide signal of residues with CSP larger than the average plus one standard deviation at different p12:PCNA ratios. The symbols correspond to the experimental data and the continuous line to the best fit to a model of one set of identical binding sites.

Isothermal titration calorimetry (ITC) measurements could be fitted well to a model of independent p12 peptide binding to equivalent sites in the PCNA trimer (N=0.87 ±0.02), with a dissociation constant of 38 ± 4 µM at 25 °C (Figure 5.3). Thus, three p12 peptides bind to the three equivalent PCNA protomers, in agreement with the NMR spectra, which show a single set of signals along the titration. The enthalpic term is negative (ΔH = -8.1 ± 0.3 kcal/mol) and is the driving force for the favorable Gibbs free energy, while the entropic term is unfavorable ($-T\Delta S$ = 2.1 ± 0.4 kcal/mol) and is in line with a loss of conformational freedom of the p12 peptide upon binding to PCNA.
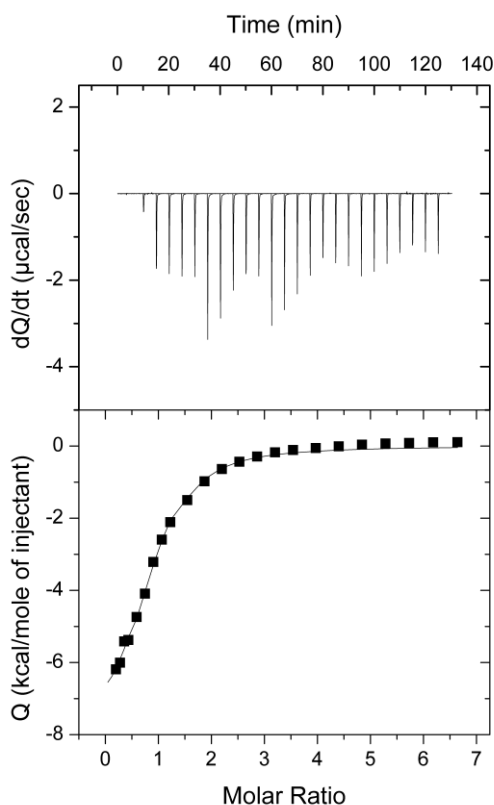

Figure 5.3. *Isothermal calorimetric titration of PCNA with p12 peptide.* The upper panel represents the heat effect associated with the variable volume peptide injection and lower panel represents the ligand concentration dependence of the heat released upon binding, after normalization and correction for the heat of dilution. The molar ratio is that of p12:PCNA protomer. The symbols correspond to the experimental data and the continuous line to the best fit to a model of one set of identical binding sites.

5.3.4. *Crystal structure of the p12–PCNA complex*

Crystals of PCNA bound to the p12 peptide diffracted to 2.1 Å resolution (Table 1), and the Fourier difference map calculated after placing and refining the PCNA ring alone in the asymmetric unit showed peaks of positive electron density across all three canonical PIP-box sites on PCNA, arising from three bound p12 peptides (Supplementary Figure 5.2). The structure of the PCNA ring is virtually identical to the previously determined structure of native human PCNA (PDB: 1VYM [25]; $RMSD_{C\alpha}$ = 0.67 Å). The three p12 peptides in the three PIP-box sites have nearly identical conformations and occupancies. Small differences were observed in the length of the modeled peptides according to their visibility in the electron density map, and vary from 12 (K4–K15) to 13 residues (R3–K15). p12 residues 7-10 adopt a 3_{10} helical conformation (Figure 5.4A), stabilized by an intramolecular hydrogen-bonding network in which the side chain carboxyl group and main chain nitrogen atom of D9 form hydrogen bonds with the main chain nitrogen and oxygen atoms of L6, respectively, while the main chain nitrogen atoms of both S10 and Y11 interact with the carbonyl oxygen of I7 (Figure 5.4B). The 3_{10} helix inserts into the hydrophobic cavity under the interdomain connector loop (IDCL) of PCNA (Figures 5.4B and C). The side chains of I7 and Y11 insert into the hydrophobic cavity, where the former is fully buried in the pocket lined with hydrophobic side chains of M40, L47 and Y250, and the latter is caged by the side chains of I128, P234 and Y250.

Differently from the canonical PIP-box, the p12 PIP-box lacks the terminal glutamine and an aromatic residue at +7 position, and thus lacks the stabilizing hydrophobic interactions mediated by those residues [20]. The p12–PCNA interaction is further stabilized by four inter-molecular hydrogen bonds involving the backbone groups of p12 residues I7, P12, V14 and K15 and those of PCNA residues H44, Q125 and G127, and one salt bridge between p12 residue R5 and residue D232 on one exposed loop of PCNA (Figure 5.4C). Notably, p12 lacks the stabilizing interactions established with the C-terminus of PCNA observed in the p68–PCNA structure [20](Figure 5.4D).

Table 1.1 – Crystallographic Statistics

|	PDB: 6HVO
Data collection	
Space group	P212121
Cell dimensions	
a, b, c (Å)	71.97 83.89 154.98
α, β, γ (°)	90.0 90.0 90.0
Resolution (Å)	77.5-2.1 (2.16-2.10)
R_{meas}	0.17(2.84)
R_{merge}	0.17(2.9)
$CC_{1/2}$ (%)	99.6 (44.7)
$I/\sigma I$	10.7 (1.3)
Completeness (%)	100.0 (100.0)
Redundancy	12.9 (13.2)
Refinement	
Resolution (Å)	73.9-2.1
No. reflections	55,825
R_{work}/ R_{free}	19.5/24.4
No. atoms	
Protein	5,928
Ligand/ion	320/35
Water	243
B-factors	
Protein	50.22
Ligand/ion	61.40/65.64
Water	51.61
R.mean square deviations	
Bond lengths (Å)	0.009
Bond angles (º)	1.49
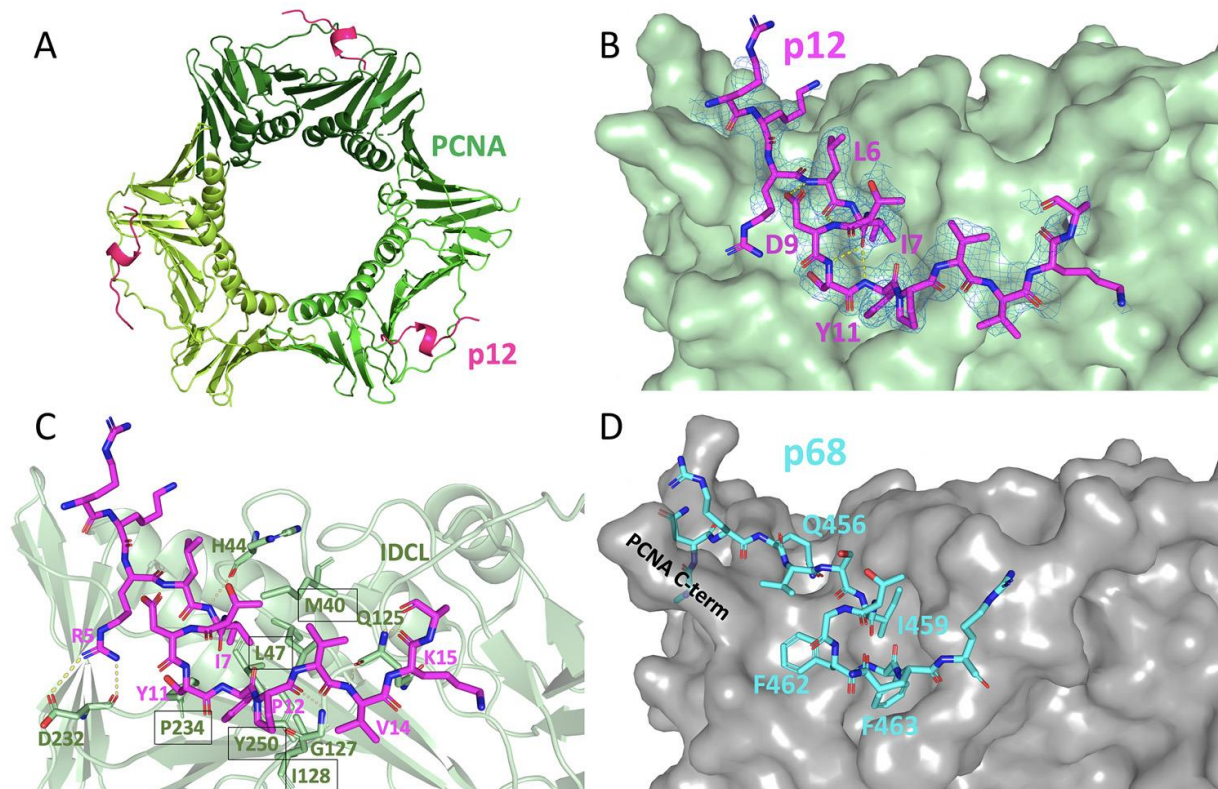
Figure 5.4. *Crystal structure of the p12–PCNA complex, and comparison with p68–PCNA structure*. (A) Overall structure of trimeric human PCNA (green) bound to the peptide derived from p12 (p12[1-19], magenta). Only p12 residues 3-15 are seen in the electron density map. (B) p12–PCNA binding site highlighting p12 intra-molecular interactions. PCNA surface is shown in pale green, p12 peptide in magenta with stick representation and interactions as yellow dotted lines. Peptide residues involved in the interactions are labelled. *2Fo-Fc* map around the p12 peptide is shown in blue contoured at 1σ(C) p12–PCNA binding site highlighting p12 inter-molecular interactions. PCNA is shown in green with ribbon representation, p12 in magenta with stick representation, and interactions as yellow dotted lines. PCNA and p12 interacting residues are labelled. Residues of PCNA involved in hydrophobic interactions are boxed. (D) p68–PCNA binding site (PDB:1U76 [20]). PCNA surface is shown in dark grey, p68 (residues 453-465) in turquoise stick representation. The p68 PIP-box consensus residues are labelled. *C-term,* C-terminus.

5.3.5. *Comparison of p12 and RecQ5 interactions with PCNA*

We have also characterized by NMR the interaction of PCNA with a peptide derived from the RecQ5 helicase (Figure 5A). RecQ5 plays an important role in the resolution between replication and transcription machineries [26], partly by a direct interaction with PCNA mediated by a C-terminal canonical PIP-box ([964]QNLIRHFF[971]) [27]. RecQ5 PIP-box drew our attention because the residues preceding the aromatic residues are basic, instead of the more

common neutral or acidic ones. The NMR titration of ^{2}H-^{15}N-labeled PCNA with unlabeled RecQ5 peptide shows a pattern of CSP similar to that of p12 (Figure 5B), but a higher peptide:PCNA ratio is necessary to approach a saturation level similar to p12 (Figure 5C). Those residues with CSP values larger than the average plus one standard deviation yielded an average dissociation constant of 210 ± 50 µM at 35 ºC. Surprisingly, the affinity of binding of this canonical PIP-box is lower than that measured for the divergent PIP-box of p12 under identical conditions (Figure 4C).
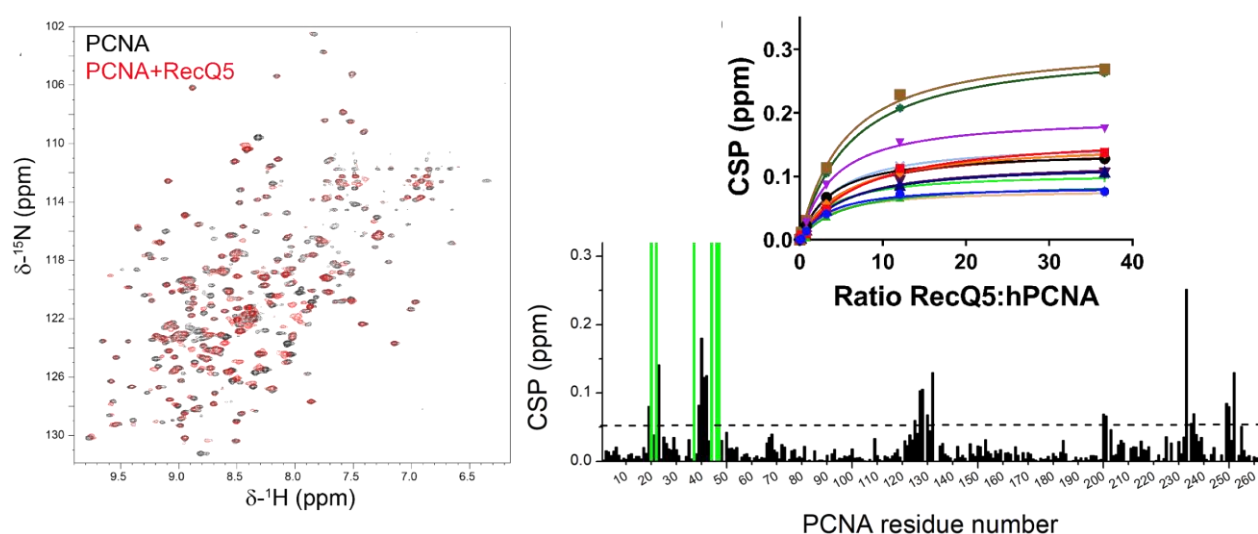
Figure 5.5. *NMR analysis of the RecQ5 peptide interaction with PCNA*. (A) Superposition of ^{1}H-^{15}N TROSY spectra of 50 µM PCNA in the absence (black) and presence (red) of a 37-fold molar excess of RecQ5 peptide. Spectra were acquired at 35˚C in PBS pH 7.0, 1mM DTT. (B) Chemical shift perturbations (CSP) of PCNA backbone amide ^{1}H and ^{15}N NMR resonances induced by RecQ5$^{952-979}$. The dotted line indicates the average plus one standard deviation. The green bars indicate the position of residues that disappear upon peptide addition. (C) Chemical shift perturbation of the amide signal of PCNA residues with CSP larger than the average plus one standard deviation at different RecQ5:PCNA ratios. The symbols correspond to the experimental data and the continuous line to the best fit to a model of one set of identical binding sites.

Our attempts to crystallize the complex formed by the RecQ5 peptide and PCNA were unsuccessful, but mapping the CSPs on the surface of PCNA shows that the RecQ5 binding site is located on the canonical PIP-box pocket of PCNA, and involves only a limited portion of the IDCL (Figure 6A). By contrast, the same mapping for p12 form a continuous surface centered on the PIP-box binding site and IDCL loop (Figure 6B), consistent with the p12–PCNA crystallographic interface and the higher affinity than RecQ5.
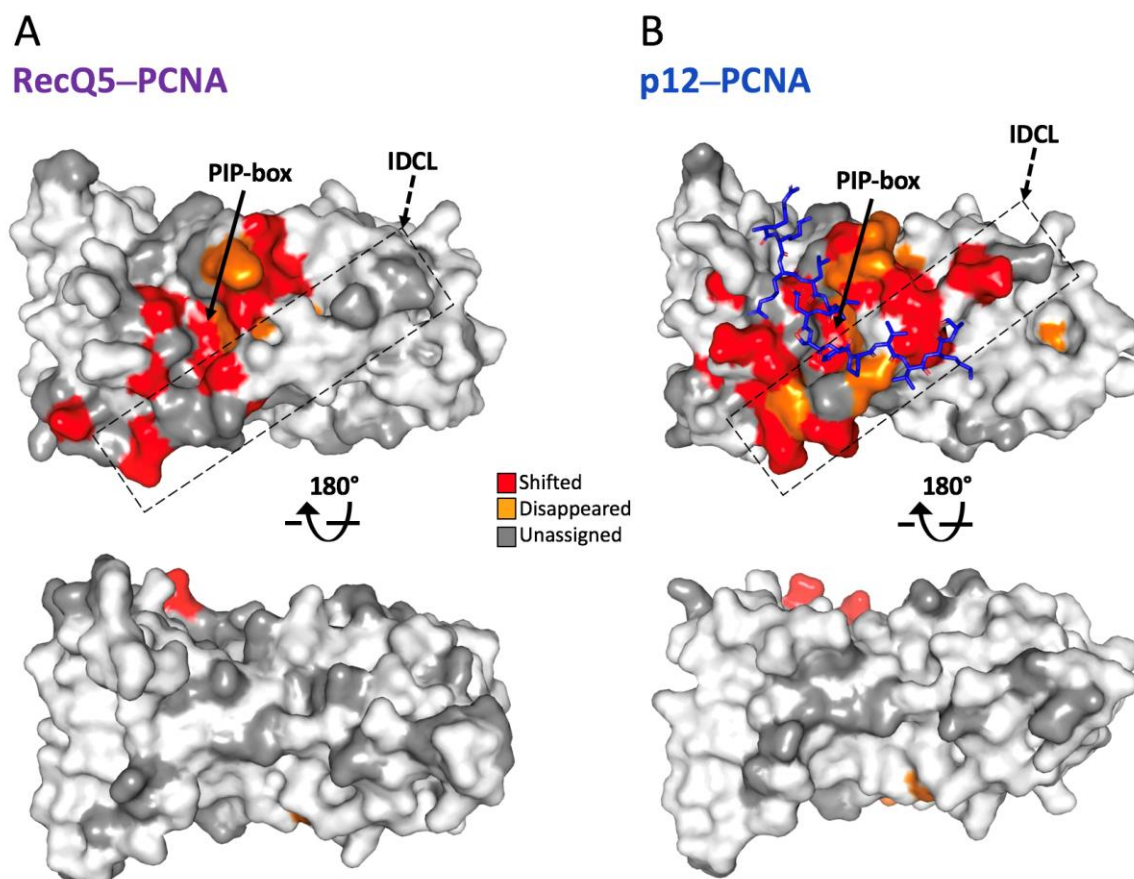
Figure 5.6. *NMR chemical shift mapping of the RecQ5* (A) *or p12* (B) *interaction site on PCNA*. Front- and back-face views of one of the PCNA protomers are shown in white surface representation, while p12 peptide in the crystallographic position is shown in blue sticks. PCNA residues whose amide signals significantly shift, disappear or remain unassigned in the titration with the RecQ5 or p12 peptides are painted in red, orange or grey, respectively.

5.4. Discussion

5.4.6. *Role of p12 subunit in the molecular recognition of PCNA by pol δ*

Our work reveals that the p12 and p68 subunits of pol δ interact differently with PCNA. Our crystal structure shows that, unlike p68, p12 recognizes PCNA through a highly divergent PIP-box that lacks the terminal glutamine as well as the aromatic residue at +7 position in the consensus sequence. The PIP-box of p12 binds the classical hydrophobic pocket on PCNA, through a 2-fork plug made of an isoleucine and a tyrosine residue at +3 and +8 positions, respectively. In the absence of the mentioned key consensus residues, the presence of an acidic residue at +6 position seems important to establish a network of intra-molecular

108

hydrogen bonds that stabilize the p12 peptide in the 3_{10} helical conformation. Such intra-molecular stabilizing effect, mediated by an aspartate at +6 position, was also observed in the interaction of p21 PIP-box with PCNA, which is the strongest interaction of a ligand with PCNA reported to date [24]. Based on our ITC analysis, we estimated a stoichiometry of binding of 1:1 (peptide-PCNA protomer) and a 38 μM affinity for the $p12^{1-19}$–PCNA interaction at 25˚C, which is 2.4-fold lower than the 16 μM affinity measured for the p68 PIP-box at the same temperature [20]. The affinity of the full pol δ complex for PCNA encircling DNA is much higher (dissociation constant <10 nM [28]), implying that the p125 and p50 subunits must also contribute to the formation of a tight pol δ–PCNA holoenzyme. In particular, p125 was shown to interact with PCNA via an N-terminal segment [14,15,29] which, based on our modeling, is buried within a structurally conserved region of the catalytic domain, and therefore rather inaccessible (Figure S5.3).

An alternative PCNA interaction site in p125 may be located at the flexible C-terminus of the catalytic domain, as observed in the crystal structure of *Pyrococcus furiosus* Polymerase B bound to PCNA [30] (Figure S5.3). The PCNA binding motifs of p68 and p12 are both located in regions predicted to be disordered (Figure 1 and S5.1), suggesting that flexibility is key to accommodate the two subunits on one PCNA ring, along with the bulkier p125 and p50 subunits (Figure 5.7).

Upon DNA damage or replication stress, the p12 subunit of pol δ is degraded, resulting in the formation of a three-subunit (p125–p50–p68) enzyme with enhanced proofreading capacity [3,5]. Thus, removal of the p12 subunit from PCNA is expected to lead to a rearrangement of the other subunits relative to each other, PCNA, and DNA, which may decrease the processivity of the enzyme in the context of higher probability of replication mistakes. Further analyses on the architecture, dynamics and activity of pol δ subassemblies are needed to shed light on this possibility. The degradation of p12 requires an intact ubiquitination system [31], and CRL4^{Cdt2}, a member of the Cullin family of E3 ligases, was identified as the ligase responsible for the UV- and DNA damage-induced degradation of p12 [32]. CRL4^{Cdt2} specifically recognizes substrates bound to PCNA and DNA through a specialized PIP-box (or PIP-degron, characterized by the Thr-Asp motif within the PIP-box, which confers high affinity to PCNA, and a basic amino acid 4 residues downstream of the PIP-box, which is required for recognition by the ligase) and triggers the degradation of several substrates, including p21 [33-

[35]. The structure of p21 PIP-box bound to PCNA suggests that an acidic patch on PCNA, centred on residues D122 and E124 on the IDCL, provides interaction for the basic residue in the PIP-degron (Figure 5.8A), and both D122 and E124 were shown to be required for CRL4^{Cdt2} recruitment [36]. Based on our X-ray structure, the basic residue of p12 degron (K14) points away from the acidic patch on the IDCL (Figure 5.4 and 5.8A). However, K14 is the last p12 C-terminal residue visible in the electron density map, and its side chain is poorly defined (Figure 5.4), suggesting that it is rather flexible and may possibly re-orient upon binding to CRL4^{Cdt2}, when p12 is targeted for degradation.
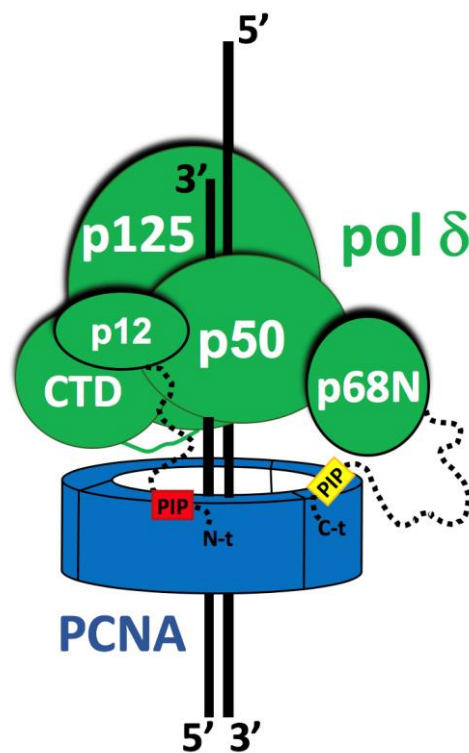

Figure 5.7. *Possible organization of the human pol δ–PCNA complex on primer/template DNA*. The N- and C-terminal PIP motifs of p12 and p68 subunits, connected to the folded domains by disordered regions (shown as dashed lines), are indicated with a red or yellow box, respectively, and binding two distinct PCNA subunits.

5.4.7. *The PCNA binding site of p12 belongs to the broad class of "PIP-like" motifs*

PCNA is a hub protein which physically interacts with dozens of protein partners, mostly involved in DNA transactions [7,37]. Most of the partners structurally characterized to date bind PCNA through canonical PIP-boxes [38] (Figure 5.8B, left). Notably, three X-ray crystal structures of PCNA bound to divergent PIP motifs have also been determined [39-41] (Figure 5.8B, right). This implies that it is highly problematic, or impossible at all, to identify PCNA-interacting motifs based on sequence analysis only [23]. Non-consensus PIP motifs lacking the canonical glutamine have been described for the TLS polymerases η and ι, in which the glutamine residue is replaced by methionine and arginine, respectively (28) (Table 5.2 and Figure 5.8). In both cases, an acidic residue is observed at position +5. The PIP motif of polι adopts a β-bend-like structure which poses the side chains of the consensus isoleucine and tyrosine and the non-consensus leucine residues at positions +4, +7 and +8 to insert into the canonical PIP-box pocket. The PIP-like motif of PARG (KDSK<u>I</u>TDH<u>F</u> [40]) shows striking similarities with the p12 motif described in this work, particularly for the lack of an aromatic residue at +7 position and the presence of an aspartate at +6 position, which creates a network of stabilizing intra-molecular interactions.

A second major class of PCNA interacting motifs named APIM (AlkB homologue 2 PCNA-interacting motif [42]) has been proposed, with consensus sequence (K/R)-(F/Y/W)-(L/I/V/A)-(L/I/V/A)-(K/R). However, the crystal structure [43,44] of an APIM motif (from the SWI/SNF helicase ZRANB3), bound to PCNA reveals a strong similarity between APIM and other atypical PIP-box motives in both their structures and their interaction with PCNA (Figure 5.8 and Table 5.2). In fact, it has been proposed that the PIP motif is not a distinct entity but rather part of a broad, loosely defined class of PIP-like motives together with the RIR (Rev1-interacting region) and the MIP (Mlh1-interacting protein) motives [23]. Therefore, it may be that the APIM motif is another variant of a PIP-like motif. Further crystallographic structures of APIM peptides bound to PCNA might support this hypothesis.

Surprisingly, our NMR data shows that the canonical PIP-box sequence of RecQ5 helicase (<u>Q</u>NL<u>I</u>RH<u>FF</u>) binds PCNA with lower affinity than the p12 divergent PIP-motif, and through a less extended surface of interaction. We propose that the low affinity of RecQ5 PIP-motif is due to the lack of the stabilizing acidic residue at +6 position. We conclude that an acidic residue at +6 position, besides the hydrophobic trident, is important to generate a high-

affinity PIP-box. This acidic residue is, however, not indispensable for binding (p68 PIP-box sequence does not have one: QVSITGFF), but at least the residue at this position should not be positively charged for a high affinity interaction.
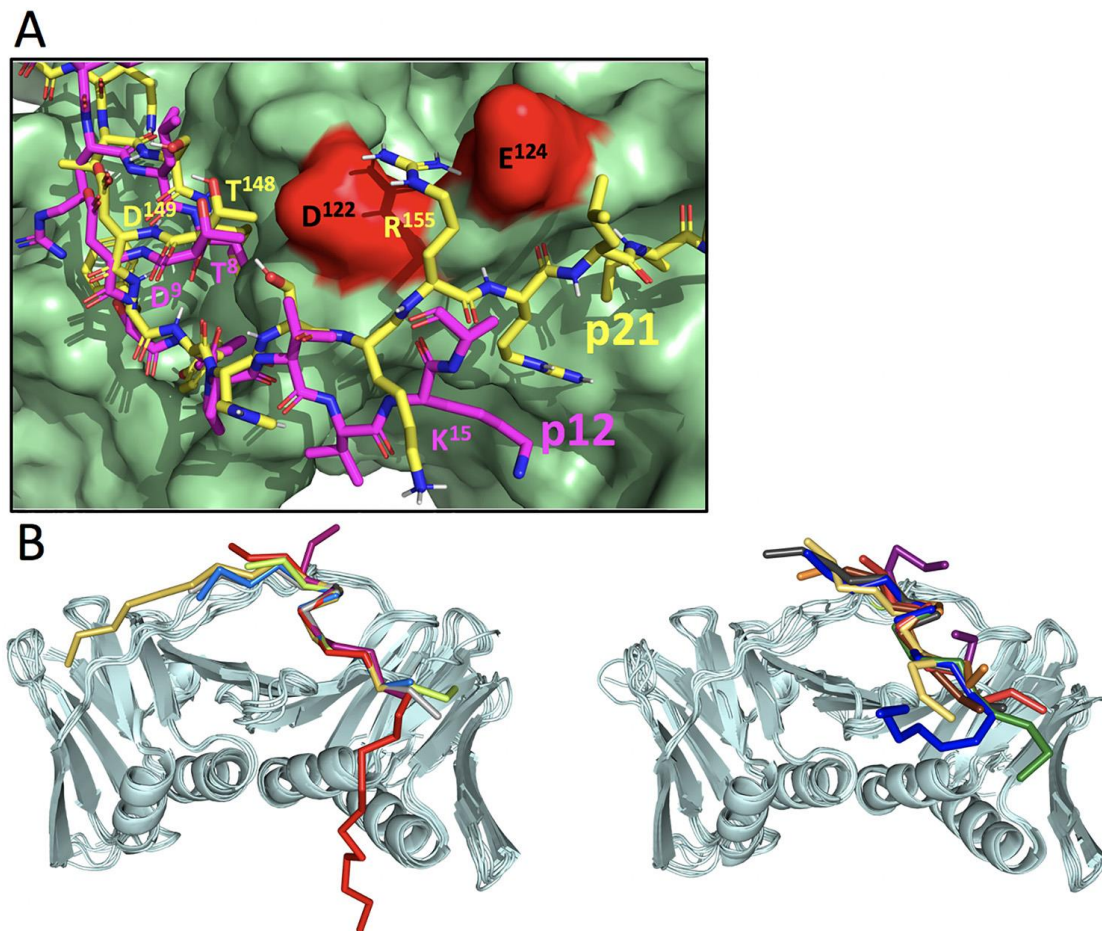

Figure 5.8. *(A) Comparison of p21 and p12 PIP-degrons interacting with PCNA. (B) Superposition of structures of canonical (left) and non-canonical (right) PCNA-interacting motifs bound to PCNA.* (A) p21– PCNA (PDB: 1AXC) [24] and p12–PCNA (PDB: 6HVO, current study) structures are aligned. p21 and p12 peptides are shown as yellow and magenta sticks, respectively. PCNA is shown as a green surface. The residues making up the acidic patch of the IDCL are coloured red. (B, left panel) The PCNA protomers are represented by ribbons and the peptides by their Cα-traces. Colour code is: p21 (yellow; PDB: 1AXC)[24], p15^PAF (red; PDB: 4D2G)[45], FEN-1 (blue; PDB: 1U7B)[20], p68 (green; PDB: 1U76)[20], ZRANB3-PIP (purple, PDB: 5MLO)[44], DVC1 (grey; PDB: 5IY4)[46]. (B, right panel) Colour code is: PARG (grey; PDB: 5MAV)[40], ZRANB3-APIM (brown; PDB: 5MLW)[43], pol η (blue; PDB: 2ZVK)[39], pol ι (purple; PDB: 2ZVM)[39], pol κ (yellow; PDB: 2ZVL)[39], RNH2B (green; PDB: 3P87)[47], TRAIP (red; PDB: 4ZTD) [48], p12 (orange; PDB: 6HVO).

Table 5.2. Sequence alignment of PCNA interacting protein fragments in crystal structures bound to PCNA. Consensus residues are highlighted[a].

_____

```
p21                          RQTSMTDFYHSKRRLIFS

p68                       KANRQVSITGFFQRK

p15             GNPVCVRPTPKWQKGIGEFFRL

FEN1                      TQGRLDDFFKVTG

ZRANB3(PIP)              KQHDIRSFFVP


p12                       RKRLITDSYPVVK

Pol ι                      KKGLIDYYLMPSL

Pol η                    GMQTLESFFKPLT

PARG                      KDSKITDHFMRL

Pol κ                    FKHTLDIFFKFLT

ZRANB3(APIM)          ASKHGSDITRFLVKK

RNH2B                 DKSGMKSIDTFF

TRAIP                    AFQAKLDTFLW
```

_____

[a]The residues shown in the alignment are those observed in the crystal structure, and do not include terminal disordered residues present in the peptides.

5.5. Experimental Procedures

5.5.1. *Protein expression and purification*

Human PCNA (UniProt: P12004) was produced in E. coli BL21(DE3) cells grown in LB medium to obtain protein with natural isotopic abundance, or in isotope enriched medium for uniform enrichment. A PCNA clone with N-terminal His$_6$-tag and HRV 3C protease cleavage site in a pET-derived plasmid was used. For NMR samples the protein was purified from the soluble fraction by Co^{2+}-affinity chromatography, cleaved by HRV 3C protease protease and polished by gel filtration chromatography [49]. All columns and chromatography systems used where from GE Healthcare. Protein elution was monitored by absorbance at 280 nm and confirmed by SDS-PAGE. The purified protein contained the extra sequence GPH- at the N-terminus. The PCNA sample for crystallization was obtained by introducing two additional purification steps [45]. The sample cleaved with HRV 3C protease was dialyzed against 50 mM sodium acetate pH

113

5.5, 100 mM NaCl. After separation of some precipitated material, the solution was loaded on a HiTrap Heparin HP column equilibrated with the same buffer. After column washing, the protein was eluted with a 0-100% gradient of 50 mM sodium acetate pH 5.5, 2 M NaCl in 20 column volumes (CV). The protein containing fractions of the major peak were dialyzed against 20 mM Tris-HCl buffer pH 7.6, 150 mM NaCl and injected into a HiTrap Chelating HP column loaded with Co^{2+} cations to remove uncleaved PCNA. The flowthrough was loaded on a HiTrap Q Sepharose column and eluted with a 0-60% gradient of 20 mM Tris-HCl pH 7.6, 1 M NaCl in 5 CV. The protein containing fractions were concentrated and polished using a Superdex 200 26/60 column equilibrated with PBS, pH 7.0, and then exchanged into the crystallization buffer (20 mM Tris-HCl, pH 7.5, 10% glycerol, 2 mM DTT) using a PD10 column. Stock solutions in PBS or crystallization buffer were flash-frozen in liquid nitrogen and stored at −80°C. The protein concentrations were measured by absorbance at 280 nm using the extinction coefficient calculated from the amino acid composition (15,930 M^{-1} cm^{-1}). All indicated concentrations of PCNA samples refer to protomer concentrations. The peptides were purchased as lyophilized powders from Apeptide company. The 19 residue long N-terminal fragment of p12 ([1]MGRKRLITDSYPVVKRREG[19]) was chosen to contain the PIP-box plus residues that could potentially interact with the IDCL of PCNA (by comparison with the p21-PCNA structure) and that would favor solubility at pH 7.0. The peptide concentration was measured by UV absorbance using the extinction coefficient of its tyrosine residue. The 28 residue long fragment of RecQ5 ([952]KTSPGRSVKEEAQNLIRHFFHGRARCES[979]) was chosen with similar criteria but elongated at the N-terminus to match the length of the p15[PAF] peptide, which also binds the inner channel of PCNA [45]. The peptide does not contain any tyrosine of tryptophan and the concentration was thus measured by UV absorbance at 205 nm [50].

5.5.2. *NMR spectroscopy*

^{1}H-^{15}N TROSY spectra were recorded at 35 °C on a Bruker Avance III 800 MHz (18.8 T) spectrometer equipped with a cryogenically cooled triple resonance z-gradient probe. A 400 μL sample of 51 μM U-[^{2}H,^{13}C, ^{15}N,] PCNA in PBS (10 mM phosphate, 140 mM chloride, 153 mM sodium y 4.5 mM potassium) pH 7.0, 20 μM DSS (4,4-dimethyl-4-silapentane-1-sulfonic acid), 1mM DTT, and 5% ^{2}H$_2$O was placed in a 5 mm Shigemi NMR tube (without plunger) and increasing volumes of the p12 peptide stock solution at 4.9 mM were added and mixed (by capping and inverting the NMR tube), causing a 7% PCNA dilution at the last point of the titration. The peptide solution was prepared in the same buffer as the PCNA samples (except that no NaN$_3$, DSS or ^{2}H$_2$O was added). For that purpose, and to remove unwanted salts from the synthetic peptide, the lyophilized powder was dissolved in PBS pH 7.0 and passed through a PD-10 Minitrap G25 column. BEST-^{1}H-^{15}N-TROSY spectra were measured with 256 indirect points for a total duration of 21.3 h. The p12-PCNA sample remained clear during the 6-day long titration. The titration with the peptide allowed for an extensive transfer of NMR signal assignments from the free PCNA to the p12-bound PCNA spectra (with a coverage of 73% of non-proline residues). The CSP caused by the peptide were computed as the weighted average distance between the backbone amide ^{1}H and ^{15}N chemical shifts in the free and bound states [49], and the estimated error in the calculated CSP is ±0.005 ppm. The fitting of the CSP changes (for those residues with CSP larger than the average plus one standard deviation) was performed using a single-site binding model, and the reported *Kd* is the average over all selected residues with the standard deviation as an estimate of its uncertainty of the error. The NMR titration of PCNA with the peptide from RecQ5 was done in the same conditions as p12 except that the peptide stock solution was 3.5 mM and that the intermediate titration points were monitored with ^{1}H-^{15}N HMQC spectra using 124 indirect points and a total duration of 10.6 h. The assignment coverage was 83% of non-proline residues).

Assignment of backbone amide NMR resonances of human PCNA bound to p12 and RecQ peptides are deposited in the Biological Magnetic Resonance database BMRB under accession code 27661 and 27662, respectively.

### 5.5.3. *Isothermal Calorimetry*

For ITC measurements, we employed an ITC200 calorimeter with 190 µM PCNA in the cell. The PCNA protein solution (dialyzed against PBS pH 7.0, 2 mM TCEP) was titrated with a 6 mM stock solution of p12 prepared by dissolving the lyophilized material in the dialysis buffer and adjusting the pH to 7.0 with NaOH. A sequence of variable injection volumes was designed based on simulations (one injection of 0.3 µL, five injections of 0.5 µL, five injections of 1.0 µL, seven injections of 2.0 µL and 7 injections of 2.5 µL). The heat produced by the binding reaction was obtained as the difference between the heat of reaction and the corresponding heat of dilution, as obtained from an independent titration of the peptides into the buffer. The binding isotherm was analyzed by non-linear least-squares fitting of the experimental data to a model assuming a single set of equivalent sites [51], using Microcal Origin (OriginLab) and in-house developed software.

5.5.4. *p12–PCNA complex crystallization and structure determination*

Stocks of PCNA and p12 peptide solutions were mixed to final concentrations of 0.4 mM and 0.5 mM, respectively (1:1.2 protein monomer:peptide), and incubated at room temperature for 30 min before screening crystallization conditions using the hanging drop vapour diffusion method. Best diffracting co-crystals grew within 2 days at 18 °C in droplets obtained by mixing 1 µl of the complex solution and 1 µl of a solution containing 24% polyethylene glycol 3350 in 0.2 M lithium sulfate buffer, pH 6.5. Crystals were transferred to precipitant solution supplemented with 20% PEG 400 and flash frozen in liquid nitrogen. The best crystals from the p12–PCNA complex diffracted at 2.1 Å resolution on the ALBA beamline XALOC (Barcelona, Spain), and belonged to $P2_12_12_1$ space group. XDS [52] and the CCP4i suite [53] were used for data processing. Molecular replacement was used to place one hPCNA trimer (PDB ID: VYM) [25] in the asymmetric unit. Several cycles of refinement using REFMAC5 [54] and model building using COOT [55] were carried out before placing the three p12 chains into the F_o-F_c electron density map. Data collection and refinement statistics are listed in Table 1. All figures with molecular models were prepared using PyMOL (www.pymol.org). Atomic coordinates of p12–PCNA complex have been deposited in the Protein Databank under the accession code 6HVO.

5.6. Supplementary material

5.6.1. *Supplementary figures*
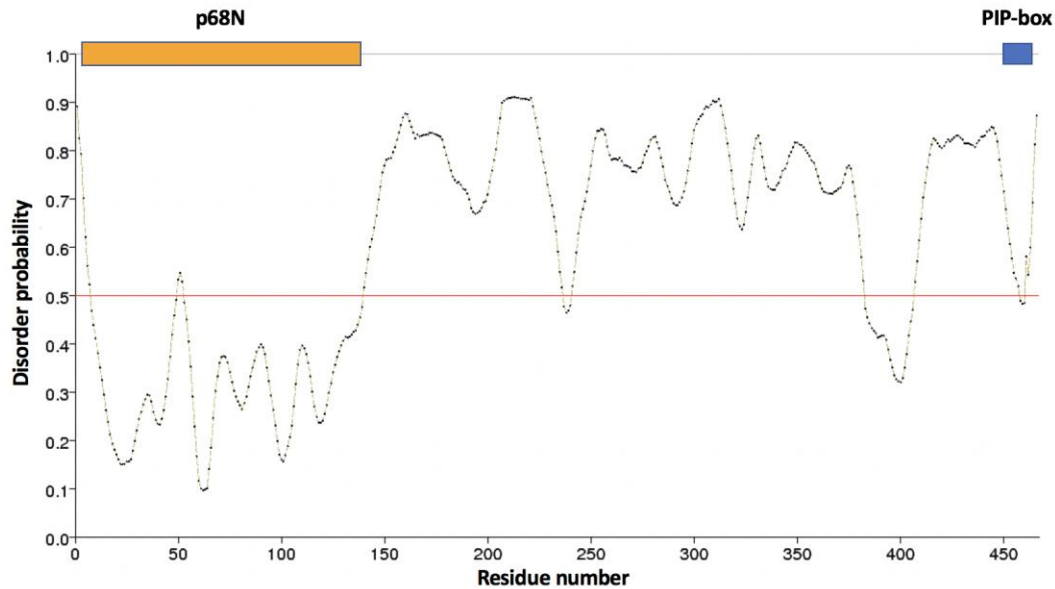
Supplementary Figure S 5.1. p68 subunit disorder prediction and motifs. The solid line shows the disorder prediction by the PrDOS server, using a threshold (red line) of 0.5. The N-terminal region (p68N) that was co-crystallized with p50 subunit of pol ∂∂PDB: 3E0J) [12] and the PIP-box motif that was crystallized with PCNA (PDB: 1U76) [20] are indicated.
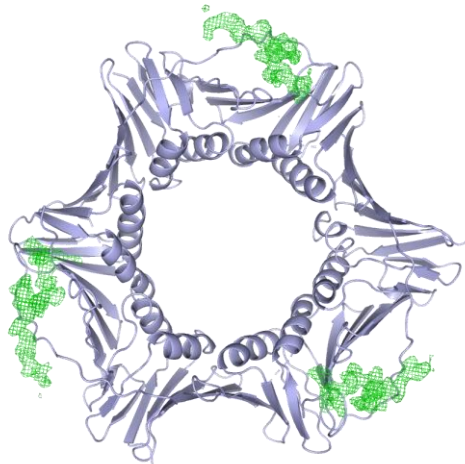
Supplementary Figure 5.2. View of the F_o-F_c omit map of the p12[1-19]–PCNA complex at the three PIP-box binding sites, refined without p12 peptide in the model and contoured at 2 σ.
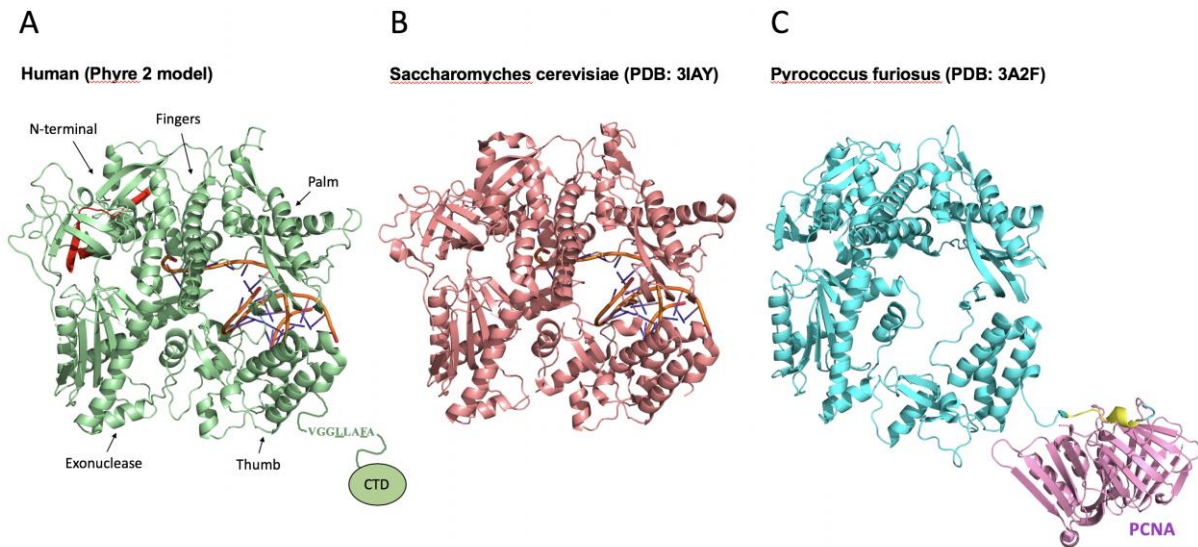
Supplementary Figure 5. 3. *Structure of the core catalytic domain of replicative polymerases and interaction with PCNA*. (A) The structure of the catalytic domain of the human p125 subunit was modelled with Phyre 2 [56] using the structure of the *Saccharomyches cerevisiae* homologue (3IAY) [57] as a template. The primer-template DNA is modelled in the position as in 3IAY. The segment coloured in red (residues 129-149) corresponds to the N-terminal region found to interact with PCNA based on biochemical studies [14,15]. The linker connecting the catalytic domain to the C-terminal domain (CTD) contains a putative divergent PIP motif (VGGLLAFA), in which only the hydrophobic and aromatic residues at +3 and +7 positions are conserved. Whether this motif mediates a direct interaction with PCNA remains to be determined. (B) Crystal structure of the catalytic domain of *Saccharomyches cerevisiae* pol δ bound to DNA [57]. (C) Crystal structure of *Pyrococcus furiosus* Polymerase B (turquoise) in complex with PCNA monomer (pink) [30]. The Pol B–PCNA interaction is mediated by a conserved PIP-box located at the Pol B C-terminus (colored in yellow).

5.7. References

1   Hubscher, U., Maga, G. & Spadari, S. Eukaryotic DNA polymerases. *Annu Rev Biochem* **71**, 133-163, doi:10.1146/annurev.biochem.71.090501.150041 (2002).

2   Lee, M., Wang, X., Zhang, S., Zhang, Z. & Lee, E. Y. C. Regulation and Modulation of Human DNA Polymerase delta Activity and Function. *Genes (Basel)* **8**, doi:10.3390/genes8070190 (2017).

3   Zhang, S. F. *et al.* A novel DNA damage response - Rapid degradation of the p12 subunit of DNA polymerase delta. *Journal of Biological Chemistry* **282**, 15330-15340, doi:10.1074/jbc.M610356200 (2007).

4   Meng, X. *et al.* DNA damage alters DNA polymerase delta to a form that exhibits increased discrimination against modified template bases and mismatched primers. *Nucleic Acids Res* **37**, 647-657, doi:10.1093/nar/gkn1000 (2009).

5 Meng, X., Zhou, Y., Lee, E. Y., Lee, M. Y. & Frick, D. N. The p12 subunit of human polymerase delta modulates the rate and fidelity of DNA synthesis. *Biochemistry-Us* **49**, 3545-3554, doi:10.1021/bi100042b (2010).

6 Tan, C. K., Castillo, C., So, A. G. & Downey, K. M. An auxiliary protein for DNA polymerase-delta from fetal calf thymus. *J Biol Chem* **261**, 12310-12316 (1986).

7 Choe, K. N. & Moldovan, G. L. Forging Ahead through Darkness: PCNA, Still the Principal Conductor at the Replication Fork. *Mol Cell* **65**, 380-392, doi:10.1016/j.molcel.2016.12.020 (2017).

8 Prelich, G., Kostura, M., Marshak, D. R., Mathews, M. B. & Stillman, B. The cell-cycle regulated proliferating cell nuclear antigen is required for SV40 DNA replication in vitro. *Nature* **326**, 471-475, doi:10.1038/326471a0 (1987).

9 Lee, M. Y., Tan, C. K., Downey, K. M. & So, A. G. Further studies on calf thymus DNA polymerase delta purified to homogeneity by a new procedure. *Biochemistry-Us* **23**, 1906-1913 (1984).

10 Li, H. *et al.* Functional roles of p12, the fourth subunit of human DNA polymerase delta. *J Biol Chem* **281**, 14748-14755, doi:10.1074/jbc.M600322200 (2006).

11 Xie, B. *et al.* Reconstitution and characterization of the human DNA polymerase delta four-subunit holoenzyme. *Biochemistry-Us* **41**, 13133-13142 (2002).

12 Baranovskiy, A. G. *et al.* X-ray structure of the complex of regulatory subunits of human DNA polymerase delta. *Cell Cycle* **7**, 3026-3036, doi:10.4161/cc.7.19.6720 (2008).

13 Tahirov, T. H. Structure and function of eukaryotic DNA polymerase delta. *Subcell Biochem* **62**, 217-236, doi:10.1007/978-94-007-4572-8_12 (2012).

14 Zhang, S. J. *et al.* A conserved region in the amino terminus of DNA polymerase delta is involved in proliferating cell nuclear antigen binding. *J Biol Chem* **270**, 7988-7992 (1995).

15 Zhang, P. *et al.* Direct interaction of proliferating cell nuclear antigen with the p125 catalytic subunit of mammalian DNA polymerase delta. *J Biol Chem* **274**, 26647-26653 (1999).

16 Zhou, Y., Meng, X., Zhang, S., Lee, E. Y. & Lee, M. Y. Characterization of human DNA polymerase delta and its subassemblies reconstituted by expression in the MultiBac system. *PLoS One* **7**, e39156, doi:10.1371/journal.pone.0039156 (2012).

17 Wang, Y. *et al.* P50, the small subunit of DNA polymerase delta, is required for mediation of the interaction of polymerase delta subassemblies with PCNA. *PLoS One* **6**, e27092, doi:10.1371/journal.pone.0027092 (2011).

18 Mo, J., Liu, L., Leon, A., Mazloum, N. & Lee, M. Y. Evidence that DNA polymerase delta isolated by immunoaffinity chromatography exhibits high-molecular weight characteristics and is associated with the KIAA0039 protein and RPA. *Biochemistry-Us* **39**, 7245-7254 (2000).

19   Rahmeh, A. A. *et al.* Phosphorylation of the p68 subunit of Pol delta acts as a molecular switch to regulate its interaction with PCNA. *Biochemistry-Us* **51**, 416-424, doi:10.1021/bi201638e (2012).

20   Bruning, J. B. & Shamoo, Y. Structural and thermodynamic analysis of human PCNA with peptides derived from DNA polymerase-delta p66 subunit and flap endonuclease-1. *Structure* **12**, 2209-2219, doi:S0969212604003533 [pii] 10.1016/j.str.2004.09.018 (2004).

21   Ishida, T. & Kinoshita, K. PrDOS: prediction of disordered protein regions from amino acid sequence. *Nucleic Acids Res* **35**, W460-464, doi:10.1093/nar/gkm363 (2007).

22   Drozdetskiy, A., Cole, C., Procter, J. & Barton, G. J. JPred4: a protein secondary structure prediction server. *Nucleic Acids Res* **43**, W389-394, doi:10.1093/nar/gkv332 (2015).

23   Boehm, E. M. & Washington, M. T. R.I.P. to the PIP: PCNA-binding motif no longer considered specific: PIP motifs and other related sequences are not distinct entities and can bind multiple proteins involved in genome maintenance. *Bioessays* **38**, 1117-1122, doi:10.1002/bies.201600116 (2016).

24   Gulbis, J. M., Kelman, Z., Hurwitz, J., O'Donnell, M. & Kuriyan, J. Structure of the C-terminal region of p21(WAF1/CIP1) complexed with human PCNA. *Cell* **87**, 297-306, doi:S0092-8674(00)81347-1 [pii] (1996).

25   Kontopidis, G. *et al.* Structural and biochemical studies of human proliferating cell nuclear antigen complexes provide a rationale for cyclin association and inhibitor design. *Proc Natl Acad Sci U S A* **102**, 1871-1876, doi:0406540102 [pii] 10.1073/pnas.0406540102 (2005).

26   Urban, V. *et al.* RECQ5 helicase promotes resolution of conflicts between replication and transcription in human cells. *J Cell Biol* **214**, 401-415, doi:10.1083/jcb.201507099 (2016).

27   Kanagaraj, R., Saydam, N., Garcia, P. L., Zheng, L. & Janscak, P. Human RECQ5beta helicase promotes strand exchange on synthetic DNA structures resembling a stalled replication fork. *Nucleic Acids Res* **34**, 5217-5231, doi:10.1093/nar/gkl677 (2006).

28   Hedglin, M., Pandey, B. & Benkovic, S. J. Characterization of human translesion DNA synthesis across a UV-induced DNA lesion. *Elife* **5**, doi:ARTN e19788 10.7554/eLife.19788 (2016).

29   Xu, H., Zhang, P., Liu, L. & Lee, M. Y. A novel PCNA-binding motif identified by the panning of a random peptide display library. *Biochemistry-Us* **40**, 4512-4520 (2001).

30   Nishida, H. *et al.* Structural determinant for switching between the polymerase and exonuclease modes in the PCNA-replicative DNA polymerase complex. *Proc Natl Acad Sci U S A* **106**, 20693-20698, doi:10.1073/pnas.0907780106 (2009).

31   Zhang, S. *et al.* A novel DNA damage response: rapid degradation of the p12 subunit of dna polymerase delta. *J Biol Chem* **282**, 15330-15340, doi:10.1074/jbc.M610356200 (2007).

32   Zhang, S. *et al.* A novel function of CRL4(Cdt2): regulation of the subunit structure of DNA polymerase delta in response to DNA damage and during the S phase. *J Biol Chem* **288**, 29550-29561, doi:10.1074/jbc.M113.490466 (2013).

33   Havens, C. G. & Walter, J. C. Docking of a specialized PIP Box onto chromatin-bound PCNA creates a degron for the ubiquitin ligase CRL4Cdt2. *Mol Cell* **35**, 93-104, doi:10.1016/j.molcel.2009.05.012 (2009).

34   Abbas, T. *et al.* PCNA-dependent regulation of p21 ubiquitylation and degradation via the CRL4Cdt2 ubiquitin ligase complex. *Genes Dev* **22**, 2496-2506, doi:10.1101/gad.1676108 (2008).

35   Michishita, M. *et al.* Positively charged residues located downstream of PIP box, together with TD amino acids within PIP box, are important for CRL4(Cdt2) -mediated proteolysis. *Genes Cells* **16**, 12-22, doi:10.1111/j.1365-2443.2010.01464.x (2011).

36   Havens, C. G. *et al.* Direct role for proliferating cell nuclear antigen in substrate recognition by the E3 ubiquitin ligase CRL4Cdt2. *J Biol Chem* **287**, 11410-11421, doi:10.1074/jbc.M111.337683 (2012).

37   Moldovan, G. L., Pfander, B. & Jentsch, S. PCNA, the maestro of the replication fork. *Cell* **129**, 665-679, doi:S0092-8674(07)00594-6 [pii], 10.1016/j.cell.2007.05.003 (2007).

38   De Biasio, A. & Blanco, F. J. Proliferating cell nuclear antigen structure and interactions: too many partners for one dancer? *Adv Protein Chem Struct Biol* **91**, 1-36, doi:B978-0-12-411637-5.00001-9 [pii], 10.1016/B978-0-12-411637-5.00001-9 (2013).

39   Hishiki, A. *et al.* Structural Basis for Novel Interactions between Human Translesion Synthesis Polymerases and Proliferating Cell Nuclear Antigen. *Journal of Biological Chemistry* **284**, 10552-10560, doi:10.1074/jbc.M809745200 (2009).

40   Kaufmann, T. *et al.* A novel non-canonical PIP-box mediates PARG interaction with PCNA. *Nucleic Acids Res* **45**, 9741-9759, doi:10.1093/nar/gkx604 (2017).

41   Armstrong, A. A., Mohideen, F. & Lima, C. D. Recognition of SUMO-modified PCNA requires tandem receptor motifs in Srs2. *Nature* **483**, 59-63, doi:10.1038/nature10883 (2012).

42   Olaisen, C. *et al.* The role of PCNA as a scaffold protein in cellular signaling is functionally conserved between yeast and humans. *FEBS Open Bio* **8**, 1135-1145, doi:10.1002/2211-5463.12442 (2018).

43   Hara, K. *et al.* Structure of proliferating cell nuclear antigen (PCNA) bound to an APIM peptide reveals the universality of PCNA interaction. *Acta Crystallogr F Struct Biol Commun* **74**, 214-221, doi:10.1107/S2053230X18003242 (2018).

44   Sebesta, M., Cooper, C. D. O., Ariza, A., Carnie, C. J. & Ahel, D. Structural insights into the

121

function of ZRANB3 in replication stress response. *Nat Commun* **8**, 15847, doi:10.1038/ncomms15847 (2017).

45   De Biasio, A. *et al.* Structure of p15(PAF)-PCNA complex and implications for clamp sliding during DNA replication and repair. *Nat Commun* **6**, 6439, doi:10.1038/ncomms7439 (2015).

46   Wang, Y., Xu, M. & Jiang, T. Crystal structure of human PCNA in complex with the PIP box of DVC1. *Biochem Biophys Res Commun* **474**, 264-270, doi:10.1016/j.bbrc.2016.04.053 (2016).

47   Bubeck, D. *et al.* PCNA directs type 2 RNase H activity on DNA replication and repair substrates. *Nucleic Acids Res* **39**, 3652-3666, doi:gkq980 [pii], 10.1093/nar/gkq980 (2011).

48   Hoffmann, S. *et al.* TRAIP is a PCNA-binding ubiquitin ligase that protects genome stability after replication stress. *J Cell Biol* **212**, 63-75, doi:10.1083/jcb.201506071 (2016).

49   De Biasio, A. *et al.* Proliferating Cell Nuclear Antigen (PCNA) Interactions in Solution Studied by NMR. *Plos One* **7**, doi:ARTN e48390, 10.1371/journal.pone.0048390 (2012).

50   Anthis, N. J. & Clore, G. M. Sequence-specific determination of protein and peptide concentrations by absorbance at 205 nm. *Protein Sci* **22**, 851-858, doi:10.1002/pro.2253 (2013).

51   Palacios, A. *et al.* Molecular basis of histone H3K4me3 recognition by ING4. *J Biol Chem* **283**, 15956-15964, doi:M710020200 [pii], 10.1074/jbc.M710020200 (2008).

52   Kabsch, W. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr D Biol Crystallogr* **66**, 133-144, doi:S0907444909047374 [pii], 10.1107/S0907444909047374 (2010).

53   Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**, 235-242, doi:10.1107/S0907444910045749 (2011).

54   Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* **67**, 355-367, doi:10.1107/S0907444911001314 (2011).

55   Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501, doi:S0907444910007493 [pii], 10.1107/S0907444910007493 (2010).

56   Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* **10**, 845-858, doi:10.1038/nprot.2015.053 (2015).

57   Swan, M. K., Johnson, R. E., Prakash, L., Prakash, S. & Aggarwal, A. K. Structural basis of high-fidelity DNA synthesis by yeast DNA polymerase delta. *Nat Struct Mol Biol* **16**, 979-986, doi:10.1038/nsmb.1663 (2009).

6.General discussion

The structural characterization of doubly monoubiquitinated p15 shows that double monoubiquitination does not alter the conformational properties of p15, which remains an intrinsically disordered protein. It´s mode of binding PCNA remains also unaltered, with the central region tightly bound to PCNA and the N- and C-terminal tails disordered, likely emerging at opposites side of the ring. However, dmUbp15 binds DNA with 5-fold reduced affinity relative to the non-ubiquitinated p15, which supports the hypothesis of p15 modulating the sliding velocity of PCNA on DNA: when replicative polymerase encounters a lesion in the DNA, the replication stalls and dmUbp15 is degraded and replaced by its non-ubiquitinated form and, since p15 binds DNA more tightly than dmUbp15 by its basic disordered N-terminus, this binding should slow down the sliding velocity of PCNA on DNA facilitating the exchange between replicative and TLS polymerase η (Figure 1.9). Whether p15 binding slows down PCNA sliding could be examined by single molecule experiments similar to those used to measure PCNA diffusion on DNA [1]. A recent work suggests that p15 binding to PCNA reduces the available surface for the DNA and may act as a belt tethering the DNA to the clamp. This study also shows that the binding of p15 inhibits η which becomes idle and dissociates from the replication fork to restore normal replication. Further work needs to be performed to clarify the role of p15 in DNA repair [2].

Calorimetry experiments revealed that dmUbp15 is recognized by the RFTS domain of DNMT1 with high affinity whereas p15 or ubiquitin are not. Based on these results and on the recently published structure of histone H3 doubly monoubiquitinated at K18 and K23 in complex with the RFTS domain of DNMT1 [3], a likely structural model of the RFTS-dmUbp15 consist of the two ubiquitins bound to the RFTS domain whereas the C-lobe of the RFTS interact with the N-terminus tail of p15 (Figure 1.9). It can be proposed that during normal replication UHRF1 recognizes the hemimethylated DNA and ubiquitinates PCNA-bound p15 (Figure 6.1A), which is recognized by the RFTS domain. This binding could lead to a spatial rearrangement of the two lobes in the RFTS domain resulting in the activation of the enzyme and the methylation of the newly synthetized strand, (Figure 6.1B), as occurs with the ubiquitinated Histone H3. Further investigations regarding the conformational change that might undergo the DNMT1 upon dmUbp15 binding are required. At the functional level, the exact effect of p15 binding to RFTS in DNA methylation should be studied. Recently, the structure of PCNA bound to a peptide of the DNMT1 PIP box has also also been recently reported, suggesting that multiple

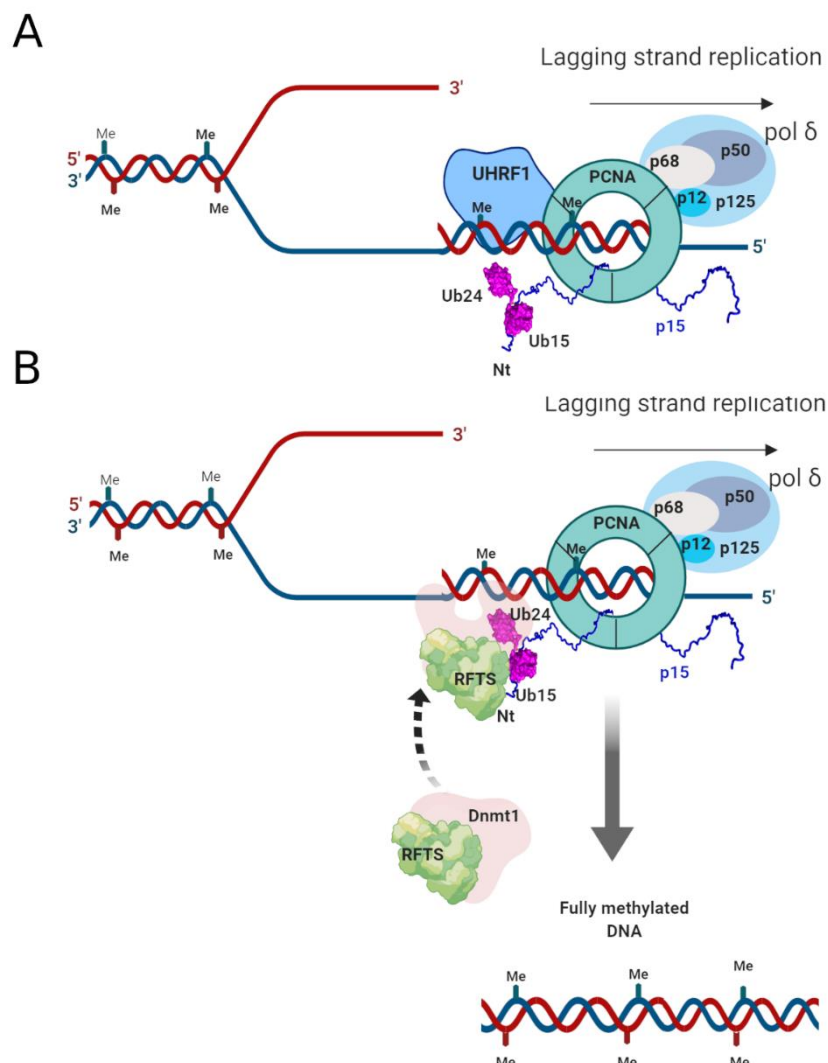players may act concomitantly to recruit DNMT1 to DNA replication sites during the S phase of the cell cycle [4].

Figure 6.1. Proposed model of dmUbp15 binding to RFTS domain of Dnmt1 methyl transferase. (A) UHRF1 recognizes the hemimethylated DNA and ubiquitinates PCNA-bound p15. (B) dmUbp15 is recognized by the RFTS domain of Dnmt1 methyl transferase and is recruited to the replication fork. The binding might induce the rearrangement of Dnmt1 that leads to the opening of the active site and methylation of the newly synthesized DNA strand.

A key component of the eukaryotic replisome is DNA polymerase δ, which also binds PCNA. There is no high resolution information on the spatial organization of the four subunits of human pol δ and how the whole holoenzyme interacts with PCNA. Recently, the structure of the *S. cerevisae* polymerase δ bound to DNA has been solved by cryo-electron microscopy at 3.2 Å resolution [5]. Since eukaryotic polymerases present a high degree of sequence homology (specially at the catalytic subunit), they likely exhibit features similar to human polymerase δ [6]. The yeast cryoEM structure reveals a unique arrangement not previously observed in other

polymerases, with the regulatory subunits (Pol31 and Pol32) positioned next to the exonuclease domain of Pol3 but not involving the DNA (Table 1.2) (Figure 6.2). The Pol31-Pol32N part resembles the human homolog p50-p66N structure (PDB code: 3E0J) although with differences in the relative orientation of the subunits as well as in certain loops [7]. However, the yeast structure lacks the fourth regulatory subunit that is present in human pol δ (p12).
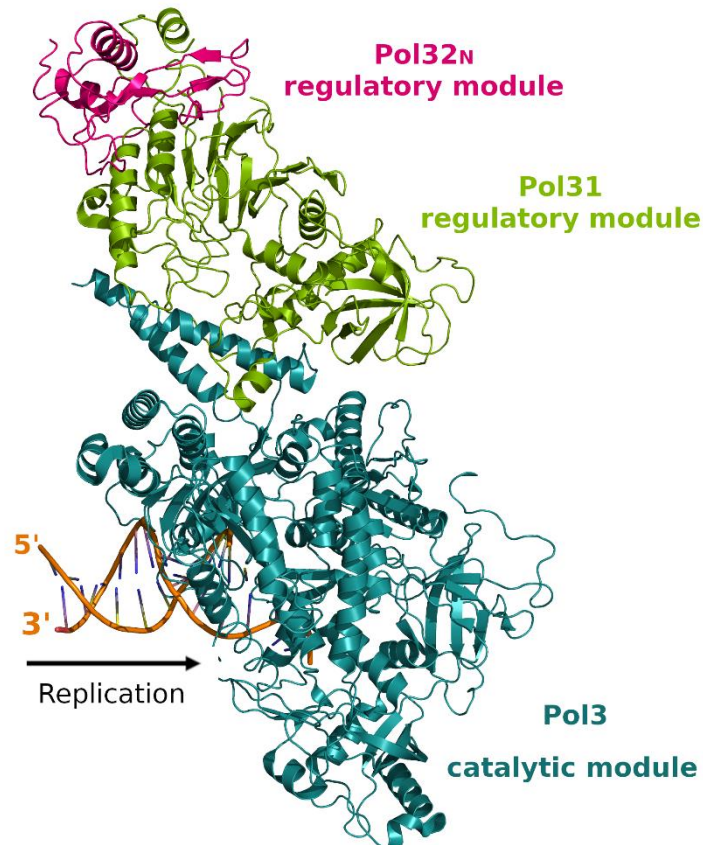

Figure 6.2. Cryo-EM structure of eukaryotic DNA polymerase δ with the DNA of *S.cerevisae* (PDB code: 6P1H). The structure represents two discrete modules: a globular 'catalytic module' consisting of Pol3 amino acids 1–985 and a flat 'regulatory module' consisting of Pol31 and Pol32N-terminus.

The structure of the p12 N-terminal peptide bound to PCNA shows that the interaction occurs via a very divergent PIP-box of p12 that lacks the first glutamine and one of the aromatic residues. The crystallographic structure reveals that the presence of an acidic residue at position +6 appear to stabilize the 3_{10} helix through a network of intramolecular hydrogen bonds and might be required for a high affinity interaction (Figure 5.4). In particular, p12 harbors a PIP degron, which is characterized by a TD motif in the PIP box and a basic residue (K in this case) at +4 position from the last aromatic residue, which is also present in the high

affinity p21. Interestingly, it binds with higher affinity than the PIP fragment of RecQ5 helicase, which has a canonical PIP sequence but not the TD degron motif. The comparative study of canonical and non-canonical PIP boxes shows that a number of highly divergent PIP boxes have already been reported to bind PCNA, suggesting that there are aspects that contribute to generate a high affinity PIP box other than the canonical sequence. These data supports the previously proposed idea of the need of a broader classification for PCNA binding proteins that extend the PIP box [8] [9] (Figure 5.8).

The global affinity of pol δ holoenzyme for PCNA is reported to be in the nM range, much higher than the affinity measured for p12 and p68 alone (37 and 1.5 μM respectively; Table 1.1)[10], however, p125 has also been reported to interact with PCNA (Figure S.5.1). In vivo, multiple context-specific subassemblies of polymerase δ have been described and it is suggested that the degradation of p12 leads to a spatial rearrangement of the remaining three subunits with PCNA [11]. The pol δ3 form of the enzyme exhibits a high proofreading capacity with the corresponding decrease in processivity and is present not only under DNA damage conditions but also appears to be the predominant form during normal replication. The pol δ4 form it has been described to be essential for HR [12]. Future work would help to unravel the spatial arrangement of the human polymerase δ and its subunits in complex with PCNA and DNA as well as the functional roles of its different forms.

6.1. References

1. Kochaniak, A. B. *et al.* Proliferating cell nuclear antigen uses two distinct modes to move along DNA. *J. Biol. Chem.* **284**, 17700–17710 (2009).

2. De March, M. *et al.* p15 PAF binding to PCNA modulates the DNA sliding surface. 1–13 (2018). doi:10.1093/nar/gky723

3. Ishiyama, S. *et al.* Structure of the Dnmt1 Reader Module Complexed with a Unique Two-Mono-Ubiquitin Mark on Histone H3 Reveals the Basis for DNA Methylation Maintenance. *Mol. Cell* **68**, 350-360.e7 (2017).

4. Jimenji, T., Matsumura, R., Kori, S. & Arita, K. Structure of PCNA in complex with DNMT1 PIP box reveals the basis for the molecular mechanism of the interaction. *Biochem. Biophys. Res. Commun.* **516**, 578–583 (2019).

5.   Jain, R. *et al.* Cryo-EM structure and dynamics of eukaryotic DNA polymerase δ holoenzyme. *Nat. Struct. Mol. Biol.* **26**, (2019).

6.   Lin, B. *et al.* Architecture of the Saccharomyces cerevisiae Replisome. *Adv Exp Med Biol.* **1042**, 207–228 (2017).

7.   Baranovskiy, A. G. *et al.* X-ray structure of the complex of regulatory subunits of human DNA polymerase delta. *Cell cycle* **7**, 3026–3036 (2008).

8.   Boehm, E. M. & Washington, M. T. R.I.P. to the PIP: PCNA-binding motif no longer considered specific: PIP motifs and other related sequences are not distinct entities and can bind multiple proteins involved in genome maintenance. *BioEssays* **38**, 1117–1122 (2016).

9.   Prestel, A. *et al.* The PCNA interaction motifs revisited : thinking outside the PIP - box. *Cell. Mol. Life Sci.* (2019). doi:10.1007/s00018-019-03150-0

10.   Hedglin, M., Pandey, B. & Benkovic, S. J. Characterization of human translesion DNA synthesis across a UV-induced DNA lesion. 1–18 (2016). doi:10.7554/eLife.19788

11.   Meng, X. *et al.* DNA damage alters DNA polymerase δ to a form that exhibits increased discrimination against modified template bases and mismatched primers. *Nucleic Acids Res.* **37**, 647–657 (2009).

12.   Zhang, S. *et al.* Loss of the p12 subunit of DNA polymerase delta leads to a defect in HR and sensitization to PARP inhibitors. *DNA Repair (Amst).* **73**, 64–70 (2019).

7.Conclusions

- The p15 chain remains disordered and monomeric in its doubly monoubiquitinated form at lysine residues K15 and K24.

- dmUbp15 binds PCNA through the central part with the same stoichiometry, affinity and mode of binding as non-ubiquitinated p15.

- dmUbqp15 binds DNA with a dissociation constant of 24 ± 3 µM at 25 °C. This is a 5-fold lower affinity than non-ubiquitinated p15, supporting the hypothesis that p15 ubiquitination may modulate the sliding velocity of p15-bound PCNA on DNA.

- The ubiquitin moieties in dmUbp15 form transient dimers which might facilitate its recognition by the RFTS domain of DNMT1.

- dmUbqp15 binds the RFTS domain of DNMT1 with a dissociation constant of 29 ± 8 nM at 25 °C, while p15 or ubiquitin do not. Singly monoubiquitinated p15 molecules bind the RFTS domain with approximately 20-fold reduced affinity and impaired stoichiometry. These results strongly suggest that dmUbp15 is involved in the maintenance of DNA methylation during chromosomal replication by recruitment of DNMT1.

- The N-terminal fragment p12$^{1-19}$ subunit of human polymerase δ binds PCNA through an atypical PIP box, located at the N-terminus, with a dissociation constant of 38 ± 4 µM at 25°C. This affinity is approximately 5 times lower than the affinity of the canonical RecQ5 PIP peptide.

- The p12 fragment folds into a 3_{10} helix that inserts into the hydrophobic PIP-box cavity of PCNA under the IDCL. Unlike canonical PIP boxes, the PIP box of p12 lacks the conserved glutamine; binds through a 2-fork plug made of an isoleucine and a tyrosine residue at +3 and +8 positions, respectively; and is stabilized by an aspartate at +6 position, which creates a network of intramolecular hydrogen bonds.

- The PIP box of p12 belongs to the growing class of PIP-like motifs which bind similarly to canonical PIP-boxes but with a more divergent three–dimensional structure outside the 3_{10} helix.

8.Appendix I: Other publications

p15PAF binding to PCNA modulates the DNA sliding surface

Matteo De March[1],[†], Susana Barrera-Vilarmau[2],[†], Emmanuele Crespan[3],[†], Elisa Mentegari[3], Nekane Merino[4], Amaia Gonzalez-Magaña[4], Miguel Romano-Moreno[4], Giovanni Maga[3], Ramon Crehuet[2], Silvia Onesti[1], Francisco J. Blanco[4],[5] and Alfredo De Biasio [1],[6],[*]

[1]Structural Biology Laboratory, Elettra-Sincrotrone Trieste S.C.p.A., Trieste 34149, Italy, [2]Institute of Advanced Chemistry of Catalonia (IQAC), CSIC, Jordi Girona 18–26, 08034, Barcelona, Spain, [3]Institute of Molecular Genetics, IGM-CNR, via Abbiategrasso 207, 27100 Pavia, Italy, [4]CIC bioGUNE, Parque Tecnológico de Bizkaia Edificio 800, 48160 Derio, Spain, [5]IKERBASQUE, Basque Foundation for Science, Bilbao, Spain and [6]Leicester Institute of Structural & Chemical Biology and Department of Molecular & Cell Biology, University of Leicester, Lancaster Rd, Leicester LE1 7HB, UK

ABSTRACT

p15PAF is an oncogenic intrinsically disordered protein that regulates DNA replication and lesion bypass by interacting with the human sliding clamp PCNA. In the absence of DNA, p15PAF traverses the PCNA ring via an extended PIP-box that contacts the sliding surface. Here, we probed the atomic-scale structure of p15PAF–PCNA–DNA ternary complexes. Crystallography and MD simulations show that, when p15PAF occupies two subunits of the PCNA homotrimer, DNA within the ring channel binds the unoccupied subunit. The structure of PCNA-bound p15PAF in the absence and presence of DNA is invariant, and solution NMR confirms that DNA does not displace p15PAF from the ring wall. Thus, p15PAF reduces the available sliding surfaces of PCNA, and may function as a belt that fastens the DNA to the clamp during synthesis by the replicative polymerase (pol δ). This constraint, however, may need to be released for efficient DNA lesion bypass by the translesion synthesis polymerase (pol η). Accordingly, our biochemical data show that p15PAF impairs primer synthesis by pol η–PCNA holoenzyme against both damaged and normal DNA templates. In light of our findings, we discuss the possible mechanistic roles of p15PAF in DNA replication and suppression of DNA lesion bypass.

INTRODUCTION

Sliding clamps are ring-shaped proteins that tether polymerases and other factors of the replisome to the genomic template, enabling DNA replication and repair. The molecular architecture of sliding clamps is conserved in all domains of life. Proliferating Cell Nuclear Antigen (PCNA)—the eukaryotic sliding clamp—is an 86-kDa homotrimeric ring with a six-fold pseudosymmetric rotation axis running through the centre of the clamp and a central channel lined with lysine and arginine-rich α-helices through which DNA is threaded (1–6). Experimental and computational analyses of the human PCNA–DNA complex showed that the DNA in the channel is tilted and its phosphates transiently interact with a set of basic residues forming a right-hand spiral that matches the DNA pitch (5,7). This dynamic interaction may allow the clamp to slide by rotationally tracking the DNA helix, or by a linear motion uncoupled from the helical pitch, or by a combination of the two modes (8). A recent report based on coarse-grained MD simulations (9) supports that the coupling between rotation and translation in PCNA sliding is weak, and that the translational motion is much faster than the rotational one, suggesting that PCNA slides on the double helix like a washer on a screw, rather than a nut on a screw. Importantly, growing evidence confirms the earlier observation that the PCNA–DNA interaction is critical for the function of the polymerase bound to the clamp (10). A recent report showed that acetylation of conserved lysine residues at the sliding surface of yeast PCNA is induced by DNA lesions and stimulates repair by homologous recombination (11). Acetylation of K20 negatively affects the processivity of the replicative polymerase δ (pol δ), but not that of the translesion synthesis (TLS) polymerase η (pol η), spe-

cialized in traversing DNA lesions such as thymine dimers or cisplatin adducts, suggesting that the modulation of the PCNA–DNA interaction can regulate the function of polymerases.

The front face of the PCNA ring contains the site of interaction of polymerases and other proteins, named the PIP-box binding site (3,12). The back face contains the sites of ubiquitylation and sumoylation (13,14). PIP-box interacting partners bind PCNA through a short consensus sequence with the pattern *QXXhXXaa*, where *h* is an aliphatic-hydrophobic residue, *a* aromatic-hydrophobic and *X* any residue (3,12).

The PCNA–associated factor p15^PAF (hereafter named p15) is an oncogenic, 11 kDa intrinsically disordered protein that regulates DNA replication and lesion bypass via a PIP-box interaction with PCNA (4,15–18). p15 co-localizes with PCNA in the nucleus of proliferating cells mainly in the S phase of the cell cycle (18–20). Depletion of p15 significantly decreases DNA synthesis (17–20), suggesting that p15 modulates the function of PCNA as a processivity factor. Co-immunoprecipitation from pancreatic cancer cell lines suggests that p15 is part of a DNA-replication complex with PCNA, pol δ and the endonuclease Fen-1 in replication foci (18). During unperturbed DNA replication, PCNA-bound p15 is mono-ubiquitylated at two N-terminal lysines (18) (K14 and K25). UV-induced replication stalling triggers the recruitment of pol η to the damaged site and the degradation of ubiquitylated p15 (18), and Povlsen and co-workers proposed that p15 may compete with pol η for binding to PCNA. However, the molecular mechanisms underlying the function of p15 in DNA replication and DNA lesion bypass remain unclear.

We have previously shown that, in absence of DNA, up to three p15 molecules bind the human trimeric PCNA ring at a site that extends from the PIP-box binding pocket to the clamp inner channel (4), and that the disordered p15 N-termini exit the clamp from the back face and directly interact with DNA (4,21). Negative stain electron microscopy of a p15–PCNA–DNA ternary complex showed particles with DNA in the clamp channel (4), but the molecular details of this assembly remained undefined.

In this work, we characterized the structure of ternary complexes composed of PCNA, DNA and the PCNA-interacting region of p15, by combining experimental and computational approaches. The co-crystal structure of PCNA in complex with two p15 molecules and a 10 bp primed DNA substrate, solved at 3.2 Å resolution, shows the duplex portion of DNA in the PCNA channel leaning towards the subunit not occupied by p15. The p15 residues N-terminal to the PIP-box contact four helices on the ring inner wall, partly shielding the DNA binding site of two subunits. This molecular arrangement is recapitulated by MD simulations of PCNA in complex with two p15 fragments with longer N-termini and a 40 bp DNA duplex. Solution NMR experiments show that DNA does not displace p15 from the inner rim of a PCNA ring in which the three subunits are occupied by p15. Accordingly, when PCNA is co-crystallized with three p15 peptides and DNA, the electron density map does not show features of DNA in the clamp channel. Thus, p15 outcompetes DNA for a common binding site in the clamp channel, and the stoichiom-

etry of binding dictates the available sliding surfaces. Given the inhibitory activity of p15 in TLS (18), we hypothesized that the constraint imposed by p15 to DNA may need to be released for efficient DNA lesion bypass by the TLS polymerase pol η. This agrees with our data showing that p15 downregulates the activity of pol η–PCNA holoenzyme in bypassing a cisplatin-induced DNA lesion and in extending the corresponding undamaged template. Based on our findings, we discuss the possible mechanistic roles of p15 in DNA replication and lesion bypass.

MATERIALS AND METHODS

Protein expression and DNA duplexes

Human PCNA (UniProt: P12004) was produced in *Escherichia coli* BL21(DE3) cells grown in appropriate culture media to obtain protein with natural isotopic abundance or uniform enrichment using a clone with N-terminal His6-tag and PreScission protease cleavage site in a pET-derived plasmid. For NMR samples the protein was purified from the soluble fraction by Co^{2+}-affinity chromatography, cleaved by PreScission protease and polished by gel filtration chromatography (22). All columns and chromatography systems used where from GE Healthcare. Protein elution was monitored by absorbance at 280 nm and confirmed by SDS-PAGE. The purified protein contained the extra sequence GPH- at the N-terminus. The PCNA sample for crystallization was obtained by introducing two additional purification steps (4). The sample cleaved with PreScission protease was dialyzed against 50 mM sodium acetate pH 5.5, 100 mM NaCl. After separation of some precipitated material, the solution was loaded on a HiTrap Heparin HP column equilibrated with the same buffer. After column washing, the protein was eluted with a 0–100% gradient of 50 mM sodium acetate pH 5.5, 2 M NaCl in 20 column volumes (CV). The protein containing fractions of the major peak were dialyzed against 20 mM Tris–HCl buffer pH 7.6, 150 mM NaCl and injected into a HiTrap Chelating HP column loaded with Co^{2+} cations to remove uncleaved PCNA. The flowthrough was loaded on a HiTrap Q Sepharose column and eluted with a 0–60% gradient of 20 mM Tris–HCl pH 7.6, 1 M NaCl in 5 CV. The protein containing fractions were concentrated and polished using a Superdex 200 26/60 column equilibrated with PBS, pH 7.0, and then exchanged into the crystallization buffer (20 mM Tris–HCl, pH 7.5, 10% glycerol, 2 mM DTT) using a PD10 column. Stock solutions in PBS or crystallization buffer were flash-frozen in liquid nitrogen and stored at −80°C. The protein concentrations were measured by absorbance at 280 nm using the extinction coefficient calculated from the amino acid composition (15 930 M^{−1} cm^{−1}). All indicated concentrations of PCNA samples refer to protomer concentrations. dsDNA and pDNA duplexes were obtained by mixing equimolar amounts of the appropriate oligonucleotides, at 93°C for 2 min with subsequent annealing by slow cooling at room temperature.

PCNA complexes crystallization and structure determination

p15^{50–77}–PCNA–pDNA complex. Stocks of PCNA, p15^{50–77} and pDNA solutions were mixed to final con-

centrations of 0.4, 0.5 and 1.1 mM, respectively (1:1.2:2.7 protein monomer:peptide:pDNA molar ratio), and incubated at room temperature for 30 min before screening crystallization conditions using the hanging drop vapour diffusion method. Best diffracting co-crystals grew within 2 days at 4°C in droplets obtained by mixing 1 μl of the complex solution and 1 μl of a solution containing 10% polyethylene glycol 3350 in 0.1 M sodium acetate buffer, pH 4.5. The best crystals from the p15$^{50-77}$–PCNA–DNA complex diffracted at 3.2 Å resolution on the ESRF-ID29 beamline, and belonged to $P2_1$ space group. XDS (23) and the CCP4i suite (24) were used for data processing. Molecular replacement was used to place one hPCNA trimer (PDB ID: 4D2G) in the asymmetric unit after removing p15$^{50-77}$ molecule and solvent. Several cycles of refinement using REFMAC5 (25) and model building using COOT (26) were carried out before placing the two p15$^{50-77}$ chains into the $F_o – F_c$ electron density map. NCS and TLS restraints were used. Inspection of the resulting unbiased difference Fourier's map inside the PCNA ring showed some electron density for one DNA strand of the double helix. Due to the disorder and/or partial occupancy the second DNA strand was only partially visible. However, modeling and refining the DNA in the same position as in the PCNA–DNA binary complex (5) gave rise to reasonable crystallographic parameters (i.e. model statistics, *B*-factor values and quality of the $2F_o – F_c$ map) and was consistent with the result from the MD simulations (see text for further details). Data collection and refinement statistics are listed in Supplementary Table ST2. Stereo view of $2F_o – F_c$ electron density map around the p15$^{50-77}$ peptide with higher occupancy is displayed in Supplementary Figure S1. All figures with molecular models were prepared using PyMOL (www.pymol.org). Atomic coordinates and structure factors of p15$^{50-77}$–PCNA–DNA complex have been deposited with PDB ID: 6EHT.

p15$^{41-72}$–PCNA complex. Crystals of p15$^{41-72}$–PCNA complex were obtained by hanging-drop vapor diffusion method at 4°C. Cubic crystals were grown on precipitant solution containing 28% polyethylene glycol 400, 0.2 M CaCl$_2$ in 0.1 M Hepes pH 7.0 buffer. Stocks of PCNA, p15$^{41-72}$ and DNA solutions were mixed to final concentrations of 0.5, 0.5 and 0.5 mM, respectively (1:1:1 protein monomer: peptide: DNA duplex molar ratio), and incubated at room temperature for 30 min before crystallization. Best diffracting co-crystals grew within 1 day and were flash-frozen directly, and diffracted to 2.9 Å resolution on the Elettra XRD1 beamline, and belonged to P2$_1$ space group (Supplementary Table ST2). XDS (23) and the CCP4i suite (24) were used for data processing. Molecular replacement was used to place one hPCNA trimer (PDB ID: 1VYM) in the asymmetric unit. Several cycles of refinement using REFMAC5 (25) and model building using COOT (26) were carried out before placing the three p15$^{41-72}$ chains into the $F_o – F_c$ electron density map. Jelly-Body restraints were used. Inspection of the resulting unbiased difference Fourier's showed no electron density attributable to DNA inside the PCNA ring.

MD simulations

Two 300 ns MD simulation replicas were performed for the same system. The system is a ternary complex composed of PCNA, two p15 peptides spanning residues 47–70 and a 40 bp DNA. The initial MD model was built by combining two different crystal structures: the p15$^{50-77}$–PCNA complex (4) and the PCNA–dsDNA complex (5). The p15 peptide was designed based on the p15$^{50-77}$–PCNA complex structure (4). The seven C-terminal residues are flexible and do not interact with PCNA and were thus excluded, while five extra residues at the flexible N-terminus were added as they are located in the clamp channel, and may transiently interact with DNA, resulting in a final segment p15$^{47-70}$. The DNA segment, with 10 bp in the crystallographic structure, was also extended by 15 bp in each direction of the helical axis. Extension of all fragments was performed using COOT (26). The system was completed by adding TIP3P water molecules in a truncated dodecahedron box at least 1.5 nm away from the DNA or protein atoms. Cl- and Na+ ions were added for charge neutralization and to mimic experimental conditions of 100 mM salt concentration. The system was minimized and equilibrated for 0.1 ns in the NVT ensemble and then for 0.1 ns in the NPT ensemble. All calculations were performed using Gromacs 5.1 (27) and parmBSC1 force field (28), the trajectories were analyzed using MDTraj package (29) and plots were generated using Matplotlib Python Library (30). The stability of the simulations was checked by visual inspection of the trajectories and the RMSD with respect to the starting structure as plotted in Supplementary Figure S2. To track the evolution of the DNA position inside the PCNA ring (Figure 2B), we have followed this procedure: First, the PCNA chains are superimposed for all frames, to eliminate global rotation and translation. Second the centre of each PCNA chain is calculated for all trajectory frames. Third, DNA base-pairs 17–21 for chain F and 20–24 for chain G are selected as those being inside the PCNA ring. Its centre is also calculated for each frame. Fourth, to project into a 2D space, a Principal Component Analysis is performed for the PCNA chain centres, and DNA centres are projected into the first two components of this subspace.

NMR spectroscopy

^{1}H–^{15}N TROSY spectra were recorded at 35°C on a Bruker Avance III 800 MHz (18.8 T) spectrometer equipped with a cryogenically cooled triple resonance z-gradient probe. A 400 ul sample of 100 μM U-[^{2}H,^{13}C, ^{15}N,] PCNA in 20 mM sodium phosphate buffer, 50 mM NaCl, pH 7.0, 20 μM 2,2-dimethyl-2-silapentane-5-sulfonate, 0.01% NaN3, 1 mM DTT and 5% ^{2}H$_2$O was placed in a 5 mm Shigemi NMR tube (without plunger) and increasing volumes of p15$^{50-77}$ or dsDNA stock solutions were added and mixed (by capping and inverting the tube). The peptide and the DNA stocks solutions were prepared in the same buffer as the PCNA samples (except that no NaN$_3$, DSS or ^{2}H$_2$O was added). For that purpose, and to remove unwanted salts from the synthetic peptide and oligonucleotides, they were dissolved in 20 mM sodium phosphate buffer, 50 mM NaCl, pH 7.0 and desalted on a PD-10 Minitrap G25 column. For

duplex formation, equimolar amounts were mixed and annealed (2 min at 95°C in a thermoblock followed by slow cooling down to room temperature). The duplex and the peptide were concentrated by ultrafiltration up to 20.84 mM (dsDNA) or 9.52 mM (p15$^{50–77}$) and concentrated DTT was added up to 1 mM. Small volumes of the stock peptide solution were added stepwise to the PCNA samples, causing a 7% PCNA dilution. TROSY spectra were measured with 144 or 256 indirect points (alternating between 8 and 14 h total duration). The PCNA–p15$^{50–77}$ sample remained clear during the 6-day long titration. When the peptide was present at an excess molar ratio of 6.4, the observed changes in the PCNA spectrum were judged to be within the experimental error with respect to the previous addition, and PCNA was considered to be saturated with the peptide. Then a volume of dsDNA stock was added to a 1:3 molar ratio (PCNA trimer:DNA duplex). Further additions of DNA did not cause further changes in the PCNA signals. The structural integrity of the DNA duplex was assessed from the imino signals observed in one-dimensional proton spectra, whose intensities increased upon duplex addition. The pH of the PCNA samples was measured at the beginning and at the end of the titrations inside the NMR tubes and found to deviate by less than 0.1 units. Therefore, the small measured shifts are not caused by differences in pH. The titration with the peptide allowed for an extensive transfer of NMR signal assignments from the free PCNA to the p15$^{50–77}$-bound PCNA spectra (with a coverage of 72% of non proline residues). For the p15$^{50–77}$-DNA-bound PCNA the assignment transfer covered 69% of the PCNA signals. The CSP caused by the peptide or the dsDNA were computed as the weighted average distance between the backbone amide ^{1}H and ^{15}N chemical shifts in the free and bound states (31,32).

DNA synthesis assays

Chemicals. Deoxynucleotides were purchased from Gene-Spin (Milan, Italy).

Oligonucleotides. The 24-mer template oligonucleotide containing the cis-PtGG adduct was a kind gift from S.J. Sturla (ETH, Zürich) and was prepared and purified as described previously (Nilforoushan, 2015). All other DNA oligonucleotides, all HPLC purified, were synthesized by Biomers.net (Germany). The 18mer primer oligonucleotide was 5′-labeled with carboxyfluorescin (FAM) group. The labeled primer was mixed to the complementary template oligonucleotide at 1:1 (M/M) ratio in the presence of 150 mM Hepes–KOH pH 7.4, 500 mM KCl, 10 mM MgCl$_2$, 250 mM NH$_4$Ac, heated at 95°C for 5 min and then slowly cooled at room temperature.

Human recombinant pol η. pJM879 (33), expressing N-terminal His-tagged human pol η, was a kind gift from R. Woodgate (NIH, USA). Human recombinant pol η was expressed and purified with a modified protocol: BL21 DE3 competent *E. coli* cells were transformed with pJM879. Plates containing 30 mg/ml kanamycin were used to identify kanamycin-resistant colonies that were picked and grown overnight at 37°C. DNA purified from bacterial cultures using NucleoSpin® Plasmid (NoLid) kit

(MACHEREY-NAGEL, Düren, Germany) was digested to confirm the presence of the gene expressing pol η. Transformed *E. coli* cells were used to inoculate a 50 ml starter culture. After overnight growth at 37°C, 12 ml of the starter culture were put into 1 l of LB medium (containing 30 mg/ml kanamycin). After 6 h growth at 37°C, cells were harvested by centrifugation and pellet frozen at –80°C. The pellet was resuspended in 40ml of lysis buffer (50mM Tris–HCl pH 7.5, 0.3 M NaCl, 20 mM imidazole, 10% glycerol, 10 mM b-mercaptoethanol [BME], 1× lysozime (Eurobio, Courtaboeuf, France), 1× ethylenediaminetetraacetic acid [EDTA]-free protease inhibitor (SIGMAFAST™ Protease Inhibitor Cocktail Tablets, Sigma-Aldrich), 1 mM phenyl-methane sulfonyl fluoride [PMSF] and lysed through sonication. After ultracentrifugation at 98 000g at 4°C for 1.5 h, the supernatant was loaded onto a 1-ml Ni-NTA column (HisTrap™ HP, GE Healthcare). Column washing was performed with 3 ml of W1 buffer (20 ml Tris–HCl pH 7.5, 1 M NaCl, 20 mM imidazole, 10% glycerol, 10 mM BME) followed by 3 ml of W2 buffer (10 mM Na-phosphate pH 7.7, 0.3 M NaCl, 20 mM imidazole, 10% glycerol, 10 mM BME). Elution of target protein was obtained using buffer H (10 mM Na-phosphate pH 7.7, 0.3 M NaCl, 200 mM imidazole, 10% glycerol, 10 mM BME). The pol η positive fractions (0.5 ml each) were pooled and overnight dialyzed with buffer M (20 mM Na-phosphate pH 7.3, 0.1 M NaCl, 10% glycerol, 10 mM BME), in dialysis cassette (Slide-A-Lyzer™ Dialysis Cassettes, 3.5K MWCO, Thermo Scientific). Pool was then loaded onto a 1-ml cation exchanger column (HiTrap SP, Pharmacia Biotech). Buffer N (20 mM Na-phosphate pH 7.3, 0.5 M NaCl, 10% glycerol, 10 mM BME) was used to elute pol η with a linear gradient. The pol η positive fractions (0.5 ml each) were aliquoted and stored at –80°C in 20% final glycerol.

Recombinant ScRF-C was obtained as a kind gift from the laboratory of Alessandro Costa (Francis Crick Institute), and was prepared following the procedure described in (34).

Enzymatic assays. All reactions were performed in a 10 μl final volume using the following conditions: 40 mM Tris pH 8, 1 mM dithiotreitol (DTT), 0.25 mg/ml bovine serum albumin (BSA), 10 mM Mg^{2+} (unless otherwise stated in the figures or figure legends). Enzymes and DNA substrates concentrations are indicated in figure legends. Reactions were incubated at 37°C for 15 min, unless otherwise stated. Reaction mixtures were stopped by addition of standard denaturing gel loading buffer (95% formamide, 10 m methylenediaminetetraacetic acid, xylene cyanol and bromophenol blue), heated at 95°C for 5 min and loaded on a 7-M urea 12% polyacrylamide (PA) gel.

Electronic image manipulation. Linear transformations have been applied in some instance to the whole images using the exposure/brightness filters of Adobe Photoshop CS6 with the sole purpose of reducing excessive background. No masking/enhancement was applied to any specific feature of the images.

RESULTS

Crystallographic evidence for p15–PCNA–DNA interactions

We co-crystallized human PCNA with a 10-bp primed DNA duplex (pDNA, as seen in published crystallographic analysis with both bacterial clamp (35) and *Saccharomyces cerevisiae* PCNA (2) and either the p15 fragment that was previously co-crystallized with PCNA alone (p15$^{50-77}$, comprising the extended PIP-box) (4), or a longer fragment that includes nine additional N-terminal residues (p15$^{41-72}$). In both crystals, incorporation of DNA is confirmed by their blue color, due to the presence of a Cy5 probe attached to the DNA (Figure 1A).

Crystals including the p15$^{50-77}$ peptide diffracted to 3.2 Å resolution, and the Fourier difference map calculated after placing and refining the PCNA ring alone in the asymmetric unit showed two PIP-box sites occupied by the p15$^{50-77}$ peptide and electron density features in the channel that may be attributed to one strand of the DNA duplex. However, electron density from the second strand is weak and the 4-base single stranded overhang invisible (Figure 1B). This suggests partial disorder and/or low occupancy of DNA. We then aligned the previously determined structure of PCNA bound to a 10 bp dsDNA (5) onto the current model and found that the DNA duplex fits reasonably well the residual electron density in the ring channel (Figure 1B). This DNA position, which leans towards the PCNA subunit not occupied by p15$^{50-77}$, also results in the best model statistics, lowest *B*-factors of DNA and best quality of the $2F_o - F_c$ electron density map (Figure 1C, and Supplementary Table ST2) and, importantly, agrees with the results from MD simulations (see below). The electron density for the DNA is still weak, indicating low occupancy and/or the presence of multiple conformations, but allows to propose a model, which shows the duplex portion of the DNA substrate in a similar position to that seen in the PCNA–dsDNA binary model, in the presence of two p15 peptides (Figure 1C). As in the p15$^{50-77}$–PCNA binary structure (4), the stoichiometry of p15$^{50-77}$ binding to PCNA is defined by the crystal packing, where a symmetry-related PCNA molecule occludes the peptide binding site on one subunit. The two p15$^{50-77}$ peptides show different occupancies. The peptide with higher occupancy has its PIP-box (Q62-F69) sitting on at the PCNA front face forming a 3_{10} helical turn, with a type-I β-turn at its N-terminus (P59-Q62) that positions residues P52-T58 to contact PCNA helices αA2 and αB2 on the clamp inner wall (Figure 1C and Supplementary Figure S1). Features of the peptide with lower occupancy are comparable yet weaker, and the peptide N-terminus could be modelled up to V53. Because of the partial disorder of DNA and the presence of a symmetry related loop that plugs the top of the PCNA channel (and may potentially affect both DNA positioning and occupancy, Figure 1B), we resorted to MD simulations to corroborate the crystallographic results and gain further structural insights on the ternary assembly.

Co-crystals of PCNA with p15$^{41-72}$ and pDNA were also obtained and diffracted to 2.9 Å resolution (Supplementary Table ST2). In the electron density map, however, no significant density was observed that may arise from DNA (Fig-

ure 1D), suggesting that, although incorporated in the crystal, the DNA is not sufficiently ordered to generate a structured signal. The electron density map showed three p15 fragments spanning residues 50–72 in the corresponding PIP-box sites, with a conformation analogous to that observed in the co-structure with the p15$^{50-77}$ peptide (Figure 1E and Supplementary Figure S1). Notably, due to steric hindrance, the location of the p15$^{41-72}$ peptides would interfere with DNA binding to PCNA in the orientation observed in the PCNA–dsDNA binary structure (Figure 1E), suggesting that p15 may outcompete DNA for binding to the PCNA inner rim.

MD simulation of PCNA in complex with two p15 PIP-boxes and a 40 bp DNA

Two replicas of a 300 ns MD simulation of a ternary complex composed of PCNA, two p15 peptides spanning residues 47–70 and a 40 bp DNA (Supplementary Table ST1) were performed. The initial MD model was built by combining features of two different crystal structures: the p15$^{50-77}$–PCNA complex (4) and the PCNA–dsDNA complex (5) by extending the DNA segment by 15 bp in each direction of the helical axis. The p15 peptide for this model was designed based on the p15$^{50-77}$–PCNA complex structure, by deleting the seven p15 C-terminal residues that are flexible and do not interact with PCNA and adding three extra residues at the flexible N-terminus, as they are located inside the clamp channel, and may transiently interact with DNA. Before starting the simulation, the DNA was moved away from its binding site on the inner wall of the clamp channel to a central position with minimal contacts with PCNA. Along the trajectory, the p15 peptides stay anchored to their binding sites on two PCNA subunits, while DNA rotates and tilts towards the wall of the subunit that is not occupied by p15 (Figure 2A and B, Supplementary Movie S1 and Supplementary Figure S2). At the end of the simulation, the DNA segment within the channel has a position similar to that observed in a previous 250 ns simulation of PCNA bound to a 30 bp DNA in absence of p15, where DNA simultaneously interacts with two adjacent sets of DNA-helix matching residues located on two PCNA subunits, as well as with residues on the clamp back face (5) (Figure 2A). Importantly, the topology of polar interactions between DNA and the PCNA subunit not occupied by p15 coincides with that observed in the crystal structure of the PCNA–dsDNA complex (Figure 2C). These results are consistent with the crystal structure of the p15$^{50-77}$–PCNA–DNA complex presented in this work and the proposed position for the partially disordered and/or incompletely occupied DNA. Altogether, these data suggest that, in the presence of two p15 molecules, a DNA duplex longer than 10 bp may still bind one of the three PCNA sliding surfaces.

Distance analysis of the intermolecular contacts along the MD trajectory shows that p15 residues N-terminal to the PIP-box (residues 52–61) are stably anchored to the inner wall of the PCNA ring, while the extreme N-termini (residues 47–51) remain flexible and thread the channel (Supplementary Figure S3 and Supplementary Movie S1). Overall, the p15 peptides establish limited contacts with
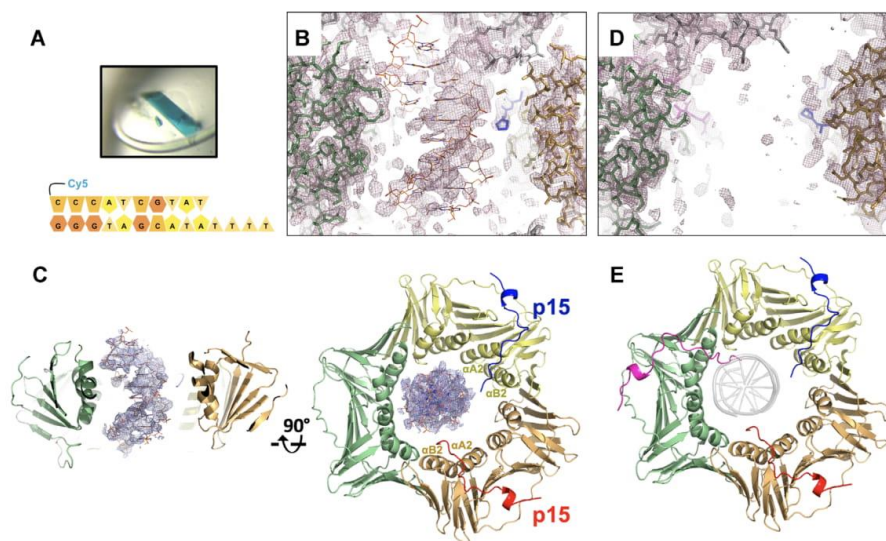
Figure 1. Crystal structures of human PCNA bound to p15 fragments and DNA. (**A**) Blue crystals of p15$^{50–77}$–PCNA–DNA complex. Co-crystals of PCNA mixed with p15$^{41–72}$ and DNA were also blue, confirming incorporation of DNA in the crystal lattice. The cartoon below shows the sequence of the DNA substrate. (**B**) Side view of the $2F_o – F_c$ omit map of the p15$^{50–77}$–PCNA–DNA complex refined without DNA in the model, contoured at 0.7 σ, showing the PCNA central channel. PCNA subunits (green and wheat) and p15$^{50–77}$ peptide in the background (blue) are in stick representation. The loop of a symmetry related PCNA molecule is shown in grey. The DNA, modeled as in the PCNA–dsDNA binary structure (5), is shown in orange. (**C**) Side- and top views of the refined p15$^{50–77}$–PCNA–DNA complex structure. PCNA and p15$^{50–77}$ are shown in ribbon representation, and the protein and peptide chains colored differently. The DNA, shown in orange, is modeled as in the PCNA–dsDNA binary structure. The $2F_o – F_c$ map around DNA is shown contoured at 0.7σ. (**D**) Side view of the $2F_o – F_c$ map of the p15$^{41–72}$–PCNA complex contoured at 0.7σ, showing the PCNA central channel as in (B). (**E**) Top view of p15$^{41–72}$–PCNA complex structure, color-coded as in (C). The DNA shown as a grey transparent ribbon in the same position as in (C) would cause a steric clash with the N-terminus of the p15 peptide on the third PCNA subunit.

DNA. In particular, polar contacts between peptide Y47 and G49 and DNA phosphates are detected. Conversely, DNA shows extensive interactions with the clamp (Figure 2A), and the side chains of many basic residues at the interface can randomly switch between adjacent DNA phosphates on a sub-nanosecond time scale (Supplementary Figure S4), as was observed in the MD simulation of the binary complex (5).

NMR analysis of PCNA binding to the p15 PIP-box and a 10 bp DNA

We characterized the interaction of PCNA with p15$^{50–77}$ and a 10 bp dsDNA (Supplementary Table ST1) by solution NMR. ^{2}H–^{15}N-labeled PCNA was firstly titrated with unlabeled p15$^{50–77}$ and chemical shift perturbations of PCNA backbone amide signals analyzed. Two groups of interacting residues were identified: residues whose signals gradually shift along the titration (Figure 3A and Supplementary Figure S5), implying a fast exchange regime on the NMR time scale, and residues whose signals broaden and disappear (due to signal attenuation below the noise level or untraceable shifting) at substoichiometric concentrations of peptide (Figure 3A and B), indicating an intermediate exchange regime. For the residues of the first group, a dissociation constant of ~35 μM at 35°C was derived (Supplementary Figure S6), at the same order of magnitude as the 12.5 μM constant previously measured at the same tempera-

ture by isothermal calorimetry (4). For some residues of the latter group, new signals appearing at saturation could be tentatively assigned to the bound form (Figure 3A). When projected onto the PCNA surface (Figure 3C), the residues whose signals disappear at substoichiometric peptide concentration strongly overlap with those at the interface in the p15$^{50–77}$–PCNA–DNA crystal structure (Figure 1C), indicating that they interact tightly with PCNA, while the residues whose persisting signals significantly shift are located next to the main binding site (Supplementary Figure S5). Signals of K77, K80 and H153 disappear, while signal of K217 persists but is significantly perturbed (Figure 3B, C and Supplementary Figure S5). These are four of the five residues at the PCNA–DNA interface in the p15$^{50–77}$–PCNA–DNA crystal structure (Figure 1B). This is consistent with the partial overlap between the p15 and DNA binding sites seen in the crystal structure, and suggests that, in solution, p15 may compete with DNA binding.

In order to map the interaction site of DNA onto p15-bound PCNA, a labeled PCNA sample saturated with unlabeled p15$^{50–77}$ was used for a second titration with dsDNA. Signal shift saturation with dsDNA was achieved at 1:3 molar ratio (PCNA trimer:DNA duplex) (Figure 3A). The fact that, in an analogous titration of PCNA with dsDNA in the absence of p15$^{50–77}$, only ~10% of complex was formed at this molar ratio (5) indicates that the presence of p15$^{50–77}$ increases the apparent DNA affinity for PCNA.
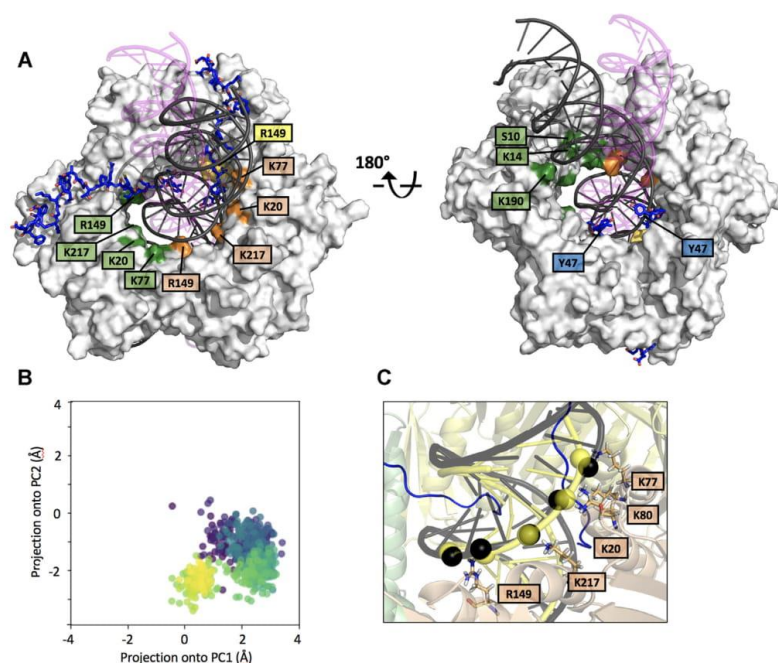
Figure 2. MD simulation of PCNA bound to two p15[47–70] peptides and a 40 bp DNA (**A**) Superposition of the initial and equilibrium states of the MD trajectory. PCNA is shown as a gray surface and DNA as a ribbon. The DNA in magenta (with transparency), and black correspond to the initial and equilibrium states of the simulation, respectively. PCNA residues whose side chains are engaged in polar contacts with DNA phosphates are labeled. Residues of different PCNA subunits are colored in green, yellow and wheat. (**B**) Principal Component Analysis of the evolution of the DNA position inside the PCNA ring (see Methods section for details). The centre of DNA in each trajectory frame was projected onto the first 2 components of the subspace composed of the centres of the 3 PCNA subunits. Each frame is coloured using the viridis colormap, which goes from dark purple for the first frames to yellow for the last ones. In the initial frame, DNA is close to (0,0), the centre of the three PCNA chains, and it quickly translates to a non-centered position. The final position is retained due to the stabilizing interactions reported in Supplementary Figure S4. (**C**) Close-up of the equilibrium state of the MD trajectory showing the PCNA–DNA interface. Interacting PCNA side chains and DNA phosphates (interatomic side chain nitrogen – DNA phosphorus distance < 4 Å) are shown as sticks and black spheres, respectively. DNA in yellow corresponds to the position in the crystallographic PCNA–dsDNA binary structure (5), with interfacial phosphates shown as spheres.

This result is consistent with the fact that no DNA binding affinity can be biochemically measured for PCNA alone, whereas a weak but detectable affinity for DNA has been previously measured (4) by fluorescence polarization in the presence of p15[50–77]. Like in PCNA alone titrated with ds-DNA or pDNA, backbone amide signal shifts are small (CSP < 0.06 ppm), suggesting that the interaction involves amino acids with long side chains (Figure 3B). The DNA-induced perturbations map to residues within the channel as well as residues that line the p15[50–77] binding site on the front face of the PCNA ring (Figure 3C). The peptide, however, remains anchored to its binding site since the signals of PCNA residues that disappear in the presence of substoichiometric p15[50–77] are not recovered by DNA addition (Figure 3A and B). Overall, these data suggest that DNA may thread through the PCNA channel when p15[50–77] saturates the three PIP-box sites, but that p15[50–77] remains anchored to the inner wall of the ring. The perturbations near the front face of the ring may arise from transient contacts that the threaded DNA makes with p15[50–77], slightly altering the position of the side chains. Based on this data, how-

ever, we cannot discard the possibility that DNA only partially penetrates the PCNA channel saturated with p15[50–77], either from the front or back face.

p15-induced inhibition of pol η–PCNA holoenzyme activity

Given the importance of p15 in regulating the activity of pol η during TLS shown in cell-based experiments (18), we performed biochemical studies using purified proteins to gain further mechanistic insight in light of our new structural data. Thus, we first probed the effects of p15 on the activity of pol η–PCNA holoenzyme in extending a DNA primer across a site-specific cisplatin lesion. A time course of pol η bypass in the presence or absence of PCNA and p15 was performed with a DNA template bearing a cisPt(GG) adduct at positions +1 and +2 (Figure 4A). Data shows that pol η alone (40 nM) was able to complete the bypass, resulting in the incorporation of two dCMPs opposite both Gs in the adduct (lanes 6–7, Figure 4A). At increasing times, the +1 product was reduced, being converted into +2 product. Further elongation past the lesion was minimal, as expected from the highly distributive nature of pol η, especially in
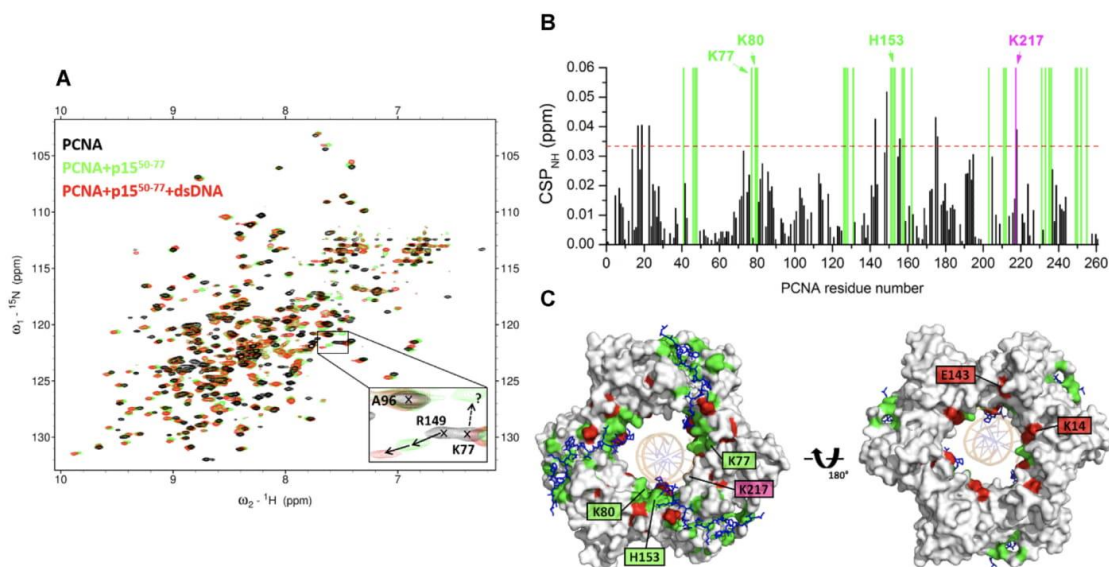
Figure 3. NMR analysis of PCNA binding to p15$^{50-77}$ and a 10 bp dsDNA. (**A**) Superposition of ^{1}H-^{15}N TROSY spectra of 95 μM PCNA in the absence (black) and presence (green) of 606 μM of p15$^{50-77}$ and (red) of 92 μM dsDNA (left) generated with oligonucleotides 3–4 in Supplementary Table S1. Spectra were acquired at 35°C on samples in 20 mM sodium phosphate, 50 mM NaCl, pH 7.0. The expansion shows signals of three representative residues. A96 signal is not perturbed by the addition of either p15$^{50-77}$ or DNA. R149 signal persists upon p15$^{50-77}$ addition, and shifts significantly by the sequential addition of DNA. Conversely, K77 signal disappears at substoichiometric concentrations of p15$^{50-77}$, and is not recovered by DNA addition. The dotted arrow points to a signal that is tentatively assigned to K77 in the p15$^{50-77}$-bound form. (**B**) Chemical shift perturbations (CSP) of backbone amide ^{1}H and ^{15}N NMR resonances induced by DNA. The dotted line indicates the average plus two standard deviations. The green bars indicate the position of residues that disappear upon addition of substoichiometric p15$^{50-77}$, and are not drawn to scale. The residues perturbed by p15$^{50-77}$ and that also appear at the interface of the p15$^{50-77}$–PCNA–DNA crystal structure are labeled. (**C**) Front- and back-face views of PCNA surface. PCNA residues whose amide signals disappear in the presence of substoichiometric p15$^{50-77}$, or are significantly perturbed by DNA are colored green or red, respectively. p15$^{50-77}$ at the three PCNA PIP-box sites is shown in sticks, and DNA in the crystallographic position is shown as and orange ribbon.

replicating damaged templates (36). Addition of a 10-fold excess of PCNA (lanes 4–5) did not significantly affect pol η activity. Such effect is not surprising: while PCNA has been shown to stimulate pol η activity on DNA substrates with blocked ends (36,37) or circular templates (38), this stimulation may not be captured on a DNA substrate with free ends as the one in our assay, because of the rapid turnover of PCNA across the substrate. In addition, if PCNA stimulation of pol η results from an increased affinity for the incorporated nucleotide (38), the saturating nucleotide concentrations used in our experiment would mask the stimulation. On the other hand, addition of equimolar amounts of PCNA and p15 (lanes 2–3) reduced nucleotide incorporation at +2 and +3 positions reproducibly also in independent experiments (Supplementary Figure S7). To rule out the possibility that the inhibitory effect of p15 on pol η activity is due to a defective loading of PCNA onto the primer-template (P/T) junction of the DNA substrate, the TLS experiment was repeated in the presence of *Saccharomyces Cerevisiae* Replication Factor C (ScRF-C), which is able to load human PCNA on DNA (39). Our results (Supplementary Figure S8) show that p15 delays TLS by pol η–PCNA across a cis-Pt lesion even in the presence of RF-C.

The replication experiment was repeated with the same DNA substrate without the lesion (Figure 4B). Again, while pol η alone (lanes 10–13) or in the presence of PCNA (lanes 6–9) showed equal processivity, the addition of PCNA and p15 reduced nucleotide incorporation at position +2 at the initial time point (lane 2), and slowed down the synthesis of the full-length product (lanes 3–5). To confirm that the inhibitory effect of p15 is mediated by the interaction with PCNA and not by a direct interaction with pol η, the activity of pol η in replicating the substrate was tested in the absence or presence of p15 alone, showing that p15 alone does not affect the activity of the polymerase (Supplementary Figure S9). To assess whether a p15 fragment spanning the region interacting with PCNA is sufficient to induce the inhibitory effect on pol η observed with full length p15, pol η activity was tested on the undamaged DNA substrate (Figure 4C) in the presence of PCNA and in the absence (lanes 2 and 3) or in the presence of p15$^{41-72}$ peptide (lanes 4,5) or full length p15 (lanes 6,7). Addition of p15$^{41-72}$ or p15 in combination with PCNA caused a reduction of DNA synthesis with respect to pol η and PCNA alone, which stopped at positions +1 and +2, corresponding to incorporation opposite the first two Gs of the template. These results indicate that p15 or p15$^{41-72}$ in conjunction with PCNA are able to reduce pol η synthesis at comparable levels.

Considering the concentrations of pol η, PCNA and p15 in the assays in Figure 4A and B, and the dissociation con-
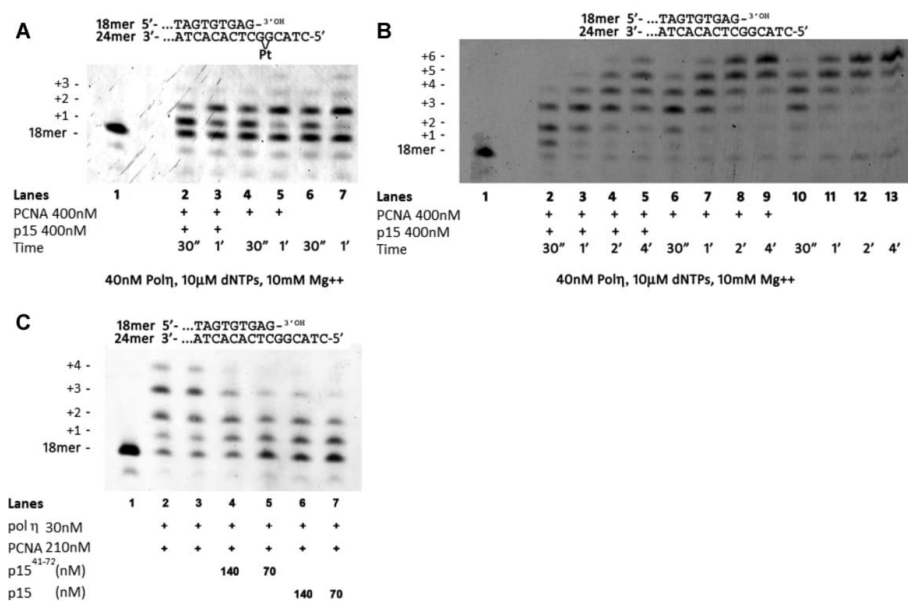
145

Figure 4. Inhibition of pol η holoenzyme by p15 (**A**) Time course of the reaction of pol η in the presence of PCNA/p15 at equimolar concentrations (Lanes 2 and 3), in the presence of PCNA (lanes 4 and 5), or with pol η alone (lanes 6 and 7) on a cisPt(GG) template (10 nM), with all four dNTPs at the indicated concentration. (**B**) Time course of the reaction of pol η on the template without the lesion (10 nM), in the presence of PCNA/p15 at equimolar concentrations (lanes 2–5), in the presence of PCNA (lanes 6–9), or pol η alone (lanes 10–13), with all four dNTPs at the indicated concentration. (**C**) Reaction of pol η replicating the undamaged template in the presence of PCNA and in the absence or presence of p15[41–72] peptide or full length p15. Reactants at the indicated concentrations were incubated at 37°C for 30 s and the reaction was stopped by addition of standard denaturing gel loading buffer. In all these experiments, PCNA was not ubiquitylated. These experiments show that p15 downregulates the activity of pol η–PCNA holoenzyme in bypassing a cisplatin lesion as well as in replicating a normal DNA substrate.

stants for the pol η–PCNA ($K_d = 0.4$ μM, measured by surface plasmon resonance at 25°C) (40) and p15–PCNA ($K_d = 1.1$ μM, measured by ITC at 25°C) (4) binding equilibria, we estimated the relative populations of binary complexes assuming that binding of pol η and p15 to PCNA is mutually exclusive. Under this assumption, 42% of pol η and 23% of p15 are saturated with PCNA. Thus, each dNTP insertion step is carried out by a combination of pol η alone and pol η holoenzyme. However, if p15 did inhibit binding of pol η to PCNA, a drop of pol η activity in the presence of p15 should not be observed, because the latter would favor the formation of free pol η, which shows full activity. Furthermore, the relatively high (30 nM) (36) and low (5 μM) (4) affinity of pol η and p15 for DNA, respectively, rules out that p15 may prevent the access of pol η to the DNA P/T junction. This suggests that inhibition of pol η processivity in the presence of PCNA and p15 is due to the formation of an impaired ternary p15–PCNA–pol η holoenzyme, deficient in primer synthesis against both damaged and normal templates.

DISCUSSION

Topology and stoichiometry of p15 binding to PCNA loaded on DNA

The data reported in this study reveals that DNA can thread through the PCNA ring along with two p15 chains, and that the structure of the p15 segment interacting with PCNA is invariant in the absence or presence of DNA. In our MD simulations, the disordered N-termini of the p15 fragments exit the PCNA back face, a topology analogous to that of full length p15 bound to PCNA in the absence of DNA (4,21). The limited number of contacts between p15[50–77] and DNA observed in our computational analysis suggests that p15 mostly operates as a passive steric obstacle constraining the DNA in the clamp channel. Critically, p15 shields key residues at the PCNA sliding surfaces, confining DNA in discrete positions, which depend on the p15 stoichiometry of binding (Figure 5A–D). This mode of binding, with DNA partially competing with p15 for a single binding site on the clamp inner ring, is also supported by (i) our NMR study showing that DNA does not disengage p15 from the inner wall of PCNA saturated with p15 and (ii) our structure of PCNA co-crystallized with three p15 peptides and DNA, showing that DNA does not occupy the central channel.

Because p15 is stably associated with PCNA on chromatin during the S phase of the cell cycle (18), it is likely that during replication polymerases and other DNA-editing enzymes bind one or two PCNA sites only, so that the free sites(s) would be available for p15 binding. Indeed, the catalytic subunit of pol δ (p125) and Fen-1 were co-precipitated with p15 and PCNA from pancreatic cell lysates (19). While Fen-1 is a monomeric enzyme that binds PCNA through a
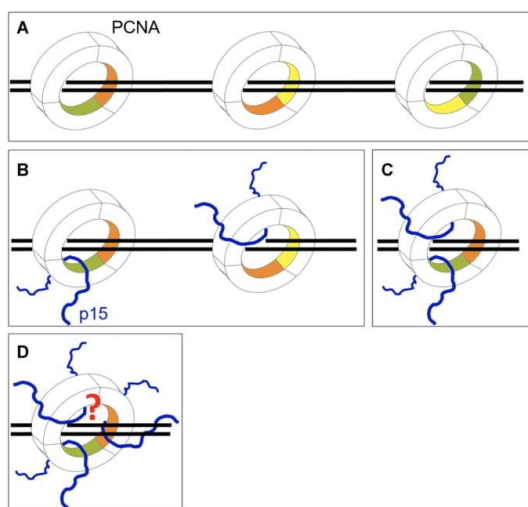
Figure 5. Possible effects of p15 on PCNA sliding (**A**) PCNA can diffuse on DNA contacting three equivalent sliding surfaces, each composed of two homologous sets of basic residues spanning across the interface of two subunits (the 3 PCNA subunits are colored green, yellow and wheat). (**B–D**) The stoichiometry of p15 binding to the PCNA homotrimer defines the available surfaces for clamp sliding. Whether a configuration where PCNA simultaneously binds three p15 chains and DNA, can be achieved, and whether it completely or partially hinders sliding, remains to be determined.

single PIP-box (41), human pol δ consists of four subunits (p125, p66, p50 and p12) and all of them are required for optimal holoenzyme activity (42). Although all four subunits contain potential PIP-box sites, examination of reconstituted holoenzymes in which the PCNA binding motifs have been mutated or inactivated have only been tested for p12 and p66 (43–45), and there may exist multiple subassemblies of pol δ *in vivo* (46,47). Thus, considering its small size and high flexibility, p15 may coexist with pol δ on the same PCNA homotrimer in a replicating cell. Likewise, a comparison of the Fen-1–PCNA (41) (PDB ID: 1UL1) and p15$^{50–77}$–PCNA (4) (PDB ID: 4D20) crystal structures suggests that Fen-1 and p15 may both be accommodated on a single PCNA ring.

The 78-kDa pol η is less bulky than pol δ and binds PCNA mainly through a single PIP-box located at its flexible C-terminus (40). Considering the comparable PIP-box affinities for PCNA (4,40), pol η and p15 may co-exist on one PCNA homotrimer. The recently determined negative stain EM structure of ubiquitylated PCNA bound to pol η and DNA (48), where two PIP-box sites would be free for p15 binding, supports this hypothesis (Figure 6A).

Possible roles of p15 in the holoenzymes with the replicative and TLS polymerases

We propose that p15 is part of the human holoenzyme that replicates the DNA lagging strand, and may function to fasten the DNA within the clamp channel by reducing the accessible sliding surfaces. We and others have shown that the

PCNA–DNA interaction is weak and transient (2,4,5,7,9), and recent evidence showed that human pol δ maintains a loose association with PCNA while replicating (49). Because the polymerase needs to maintain a fixed position relative to the phosphodiester backbone at the P/T junction, a constrained orientation of PCNA with respect to the helical pitch may improve the overall stability of the holoenzyme. Such stabilizing role of p15 would explain the negative effects of p15 knockdown on DNA synthesis (19,20).

While a high-resolution structure of a PCNA–polymerase–DNA complex is still awaiting, both the medium resolution EM structure (50) and MD simulations of *Pyrococcus furiosus* (Pfu) PCNA bound to PolB and DNA (51) show that DNA within the clamp is tilted. Particularly, the MD model of the complex in polymerizing mode shows features at the clamp–DNA interface analogous to those observed in the p15$^{50–77}$–PCNA–DNA model presented here, where five conserved positively charged residues, matching the dsDNA B-helix architecture, interact with five consecutive phosphates of one DNA strand, suggesting that this key determinant of the interaction is conserved and may be present in the pol δ holoenzyme. Indeed, mutation of residues at the PCNA–DNA interface impairs both initiation of DNA synthesis (10) and processivity of pol δ (11), suggesting that the PCNA–DNA interactions control both clamping and sliding activities of PCNA in processive DNA replication. In their computational work, Ivanov and co-workers (51) also showed that the repositioning of the PolB core during the conformational switch from polymerizing to editing modes forces the DNA to tilt from one side of the PfuPCNA channel to the other. A 30° change in DNA tilt within the clamp in the catalytic core of the bacterial replisome from polymerizing to editing modes was also observed in the recent cryo-EM work by Lamers and colleagues (52,53). Perhaps, in the human pol δ holoenzyme, p15 plays a role in guiding DNA through the PCNA inner rim in between DNA synthesis and editing steps of the polymerase.

Single molecule experiments suggested that PCNA may slide by rotationally tracking the DNA helix or by a less frequent translational mode uncoupled from the helical pitch (8), while a recent computational work predicts that the coupling between rotation and translation is weak (9), suggesting that the translational mode is prevalent. Thus, p15 binding to PCNA may increase the rotation-translation coupling by reducing the available sliding surfaces (Figure 5). This, together with the DNA binding activity of the p15 disordered N-terminus, may result in a slower diffusion of PCNA on DNA. Therefore, p15 might regulate the sliding velocity of PCNA, and this function may be required for the DNA damage response to prevent a rapid drift of PCNA from stalled forks in between polymerase swapping events.

Upon encounter of a DNA lesion, pol δ dissociates from PCNA, which becomes ubiquitylated, and is replaced by pol η that replicates DNA past the lesion (54,55). While the ubiquitin moieties of ubiquitylated-PCNA may interact with the C-terminal ubiquitin binding motif (UBZ) of pol η, a large body of data argues that ubiquitylation of PCNA is not strictly necessary for pol η recruitment and activity in TLS (38,56,57). Notably, a recent report unambiguously demonstrates that the binding of pol η to PCNA, and DNA
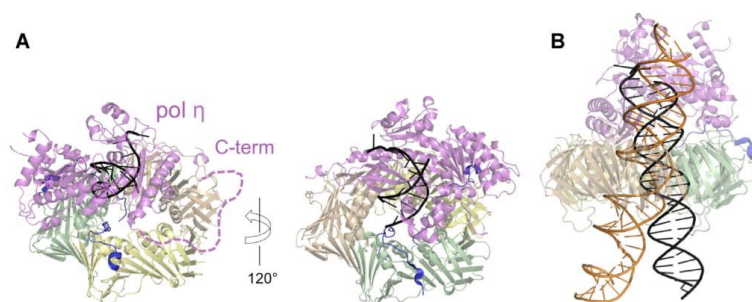
Figure 6. Structural models of pol η-PCNA holoenzymes with p15 and DNA. (**A**) The PCNA trimer of the p15$^{50-77}$–PCNA crystal structure (PDB ID: 4D2G) was superposed to PCNA of the low-resolution structure of human pol η–PCNA–DNA generated from EM data (PDB ID: 3JA9 and 3JAA) (33). DNA is shown in black. The vacant PIP-box site on PCNA (subunit wheat) was occupied by the C-terminal PIP-box of pol η using the crystal structure of human PCNA bound to pol η residues 700–710 (chain W of PDB ID 2ZVK) (40). The dashed line indicates the flexible pol η C-terminus (residues 433–699). (**B**) The PCNA trimer of the structure of p15$^{50-77}$–PCNA–DNA complex (PDB: 6EHT) was superposed to PCNA of the pol η–PCNA–DNA (48) complex (PDB: 3JA9 and and 3JAA). The DNA of the first complex (elongated to 40 bp) is shown as an orange ribbon, that of the latter (elongated to 25 bp) as a black ribbon. According to these models, it is possible that p15 may co-exist with pol η on the same PCNA ring. However, the constraint on DNA within the clamp channel imposed by p15 may hinder the translocation of pol η holoenzyme on DNA.

synthesis by a pol η holoenzyme are both independent of PCNA monoubiquitylation (36).

During unperturbed replication, p15 is monoubiquitylated at K15 and K24 and is degraded by the proteasome after UV irradiation or cisplatin treatment (18). Degradation of ubiquitylated p15 upon DNA damage is required for the recruitment of pol η to the replication foci and efficient lesion bypass, and the authors suggested that p15 may prevent the binding of pol η to PCNA. In this work, we showed that p15 has an inhibitory activity on the pol η–PCNA holoenzyme in synthesizing past the 5′dG of a 1,2-d(GpG) cisplatin DNA adduct and in extending the undamaged template. Importantly, a p15 fragment spanning only the region of interaction with PCNA is sufficient to inhibit pol η activity. Rather than preventing pol η from binding to PCNA, our data suggests that p15 inhibits pol η activity by associating to the holoenzyme, a possibility supported by structural considerations (Figure 6A). Although the DNA in the EM map of pol η bound to ubiquitylated PCNA and DNA is not fully defined (48), the DNA duplex within PCNA lies close to one of the clamp subunits, in a position different from that observed in our p15–PCNA–DNA complex structure (Figure 6B). We propose that the constraint imposed to DNA by p15 in the central channel of PCNA may hinder the advancement of DNA in the pol η active site required for the incorporation of the nucleotide opposite to the 5′dG of the DNA template. Thus, after the insertion of the first dCMP, the polymerase may become 'idle' and dissociate from PCNA. Further high-resolution structural studies on pol η holoenzyme will shed light on this possibility.

DATA AVAILABILITY

Atomic coordinates of p15$^{50-77}$–PCNA–DNA and p15$^{41-72}$–PCNA complexes have been deposited in the Protein Databank under the accession codes 6EHT and 6GWS, respectively. Assignments of backbone amide NMR resonances of human PCNA bound to p15$^{50-77}$ and

DNA are deposited in the Biological Magnetic Resonance database BMRB under accession code 27558.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

REFERENCES

1. Krishna,T.S., 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.
2. McNally,R., Bowman,G.D., Goedken,E.R., O'Donnell,M. and Kuriyan,J. (2010) Analysis of the role of PCNA-DNA contacts during clamp loading. *BMC Struct. Biol.*, **10**, 3.
3. Mailand,N., Gibbs-Seymour,I. and Bekker-Jensen,S. (2013) Regulation of PCNA-protein interactions for genome stability. *Nat. Rev. Mol. Cell Biol.*, **14**, 269–282.
4. De Biasio,A., de Opakua,A.I., Mortuza,G.B., Molina,R., Cordeiro,T.N., Castillo,F., Villate,M., Merino,N., Delgado,S., Gil-Carton,D. *et al.* (2015) Structure of p15(PAF)-PCNA complex and implications for clamp sliding during DNA replication and repair. *Nat. Commun.*, **6**, 6439.
5. De March,M., Merino,N., Barrera-Vilarmau,S., Crehuet,R., Onesti,S., Blanco,F.J. and De Biasio,A. (2017) Structural basis of human PCNA sliding on DNA. *Nat. Commun.*, **8**, 13935.
6. De March,M. and De Biasio,A. (2017) The dark side of the ring: role of the DNA sliding surface of PCNA. *Crit. Rev. Biochem. Mol. Biol.*, **52**, 663–673.
7. Ivanov,I., Chapados,B.R., McCammon,J.A. and Tainer,J.A. (2006) Proliferating cell nuclear antigen loaded onto double-stranded DNA: dynamics, minor groove interactions and functional implications. *Nucleic Acids Res.*, **34**, 6023–6033.
8. Kochaniak,A.B., Habuchi,S., Loparo,J.J., Chang,D.J., Cimprich,K.A., Walter,J.C. and van Oijen,A.M. (2009) Proliferating cell nuclear antigen uses two distinct modes to move along DNA. *J. Biol. Chem.*, **284**, 17700–17710.
9. Daitchman,D., Greenblatt,H.M. and Levy,Y. (2018) Diffusion of ring-shaped proteins along DNA: case study of sliding clamps. *Nucleic Acids Res.*, **46**, 5935–5949.
10. 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. *J. Biol. Chem.*, **270**, 22527–22534.
11. Billon,P., Li,J., Lambert,J.P., Chen,Y., Tremblay,V., Brunzelle,J.S., Gingras,A.C., Verreault,A., Sugiyama,T., Couture,J.F. *et al.* (2017) Acetylation of PCNA sliding surface by Eco1 promotes genome stability through homologous recombination. *Mol. Cell*, **65**, 78–90.
12. De Biasio,A. and Blanco,F.J. (2013) Proliferating cell nuclear antigen structure and interactions: too many partners for one dancer? *Adv. Protein Chem. Struct. Biol.*, **91**, 1–36.
13. Freudenthal,B.D., Gakhar,L., Ramaswamy,S. and Washington,M.T. (2010) Structure of monoubiquitinated PCNA and implications for translesion synthesis and DNA polymerase exchange. *Nat. Struct. Mol. Biol.*, **17**, 479–484.
14. Freudenthal,B.D., Brogie,J.E., Gakhar,L., Kondratick,C.M. and Washington,M.T. (2011) Crystal structure of SUMO-modified proliferating cell nuclear antigen. *J. Mol. Biol.*, **406**, 9–17.
15. De Biasio,A., Ibanez de Opakua,A., Cordeiro,T.N., Villate,M., Merino,N., Sibille,N., Lelli,M., Diercks,T., Bernado,P. and Blanco,F.J. (2014) p15PAF is an intrinsically disordered protein with nonrandom structural preferences at sites of interaction with other proteins. *Biophys. J.*, **106**, 865–874.
16. Yu,P., Huang,B., Shen,M., Lau,C., Chan,E., Michel,J., Xiong,Y., Payan,D.G. and Luo,Y. (2001) p15(PAF), a novel PCNA associated factor with increased expression in tumor tissues. *Oncogene*, **20**, 484–489.
17. Emanuele,M.J., Ciccia,A., Elia,A.E. and Elledge,S.J. (2011) Proliferating cell nuclear antigen (PCNA)-associated KIAA0101/PAF15 protein is a cell cycle-regulated anaphase-promoting complex/cyclosome substrate. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 9845–9850.
18. Povlsen,L.K., Beli,P., Wagner,S.A., Poulsen,S.L., Sylvestersen,K.B., Poulsen,J.W., Nielsen,M.L., Bekker-Jensen,S., Mailand,N. and Choudhary,C. (2012) Systems-wide analysis of ubiquitylation
19. dynamics reveals a key role for PAF15 ubiquitylation in DNA-damage bypass. *Nat. Cell Biol.*, **14**, 1089–1098.
19. Hosokawa,M., Takehara,A., Matsuda,K., Eguchi,H., Ohigashi,H., Ishikawa,O., Shinomura,Y., Imai,K., Nakamura,Y. and Nakagawa,H. (2007) Oncogenic role of KIAA0101 interacting with proliferating cell nuclear antigen in pancreatic cancer. *Cancer Res.*, **67**, 2568–2576.
20. Chang,C.N., Feng,M.J., Chen,Y.L., Yuan,R.H. and Jeng,Y.M. (2013) p15(PAF) is an Rb/E2F-regulated S-phase protein essential for DNA synthesis and cell cycle progression. *PLoS One*, **8**, e61196.
21. Cordeiro,T.N., Chen,P.C., De Biasio,A., Sibille,N., Blanco,F.J., Hub,J.S., Crehuet,R. and Bernado,P. (2017) Disentangling polydispersity in the PCNA-p15PAF complex, a disordered, transient and multivalent macromolecular assembly. *Nucleic Acids Res.*, **45**, 1501–1515.
22. De Biasio,A., Sanchez,R., Prieto,J., Villate,M., Campos-Olivas,R. and Blanco,F.J. (2011) Reduced stability and increased dynamics in the human proliferating cell nuclear antigen (PCNA) relative to the yeast homolog. *PLoS One*, **6**, e16600.
23. Kabsch,W. (2010) Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. D. Biol. Crystallogr.*, **66**, 133–144.
24. Winn,M.D., Ballard,C.C., Cowtan,K.D., Dodson,E.J., Emsley,P., Evans,P.R., Keegan,R.M., Krissinel,E.B., Leslie,A.G., McCoy,A. *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. D. Biol. Crystallogr.*, **67**, 235–242.
25. Murshudov,G.N., Skubak,P., Lebedev,A.A., Pannu,N.S., Steiner,R.A., Nicholls,R.A., Winn,M.D., Long,F. and Vagin,A.A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D. Biol. Crystallogr.*, **67**, 355–367.
26. Emsley,P., Lohkamp,B., Scott,W.G. and Cowtan,K. (2010) Features and development of Coot. *Acta Crystallogr. D. Biol. Crystallogr.*, **66**, 486–501.
27. Abraham,M.J., Murtola,T., Schulz,R., Pail,S., Smith,J.C., Hess,B. and Lindhal,E. (2015) GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, **1–2**, 19–25
28. Ivani,I., Dans,P.D., Noy,A., Perez,A., Faustino,I., Hospital,A., Walther,J., Andrio,P., Goni,R., Balaceanu,A. *et al.* (2016) Parmbsc1: a refined force field for DNA simulations. *Nat. Methods*, **13**, 55–58.
29. McGibbon,R.T., Beauchamp,K.A., Harrigan,M.P., Klein,C., Swails,J.M., Hernandez,C.X., Schwantes,C.R., Wang,L.P., Lane,T.J. and Pande,V.S. (2015) MDTraj: A modern open library for the analysis of molecular dynamics trajectories. *Biophys. J.*, **109**, 1528–1532.
30. Hunter,J.D. (2007) Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.*, **9**, 90–95.
31. Sanchez,R., Torres,D., Prieto,J., Blanco,F.J. and Campos-Olivas,R. (2007) Backbone assignment of human proliferating cell nuclear antigen. *Biomol. NMR Assign.*, **1**, 245–247.
32. De Biasio,A., Campos-Olivas,R., Sanchez,R., Lopez-Alonso,J.P., Pantoja-Uceda,D., Merino,N., Villate,M., Martin-Garcia,J.M., Castillo,F., Luque,I. *et al.* (2012) Proliferating Cell Nuclear Antigen (PCNA) interactions in solution studied by NMR. *PLoS One*, **7**, e48390.
33. Frank,E.G., McDonald,J.P., Karata,K., Huston,D. and Woodgate,R. (2012) A strategy for the expression of recombinant proteins traditionally hard to purify. *Anal. Biochem.*, **429**, 132–139.
34. Yeeles,J.T.P., Janska,A., Early,A. and Diffley,J.F.X. (2017) How the eukaryotic replisome achieves rapid and efficient DNA replication. *Mol. Cell*, **65**, 105–116.
35. Georgescu,R.E., Kim,S.S., Yurieva,O., Kuriyan,J., Kong,X.P. and O'Donnell,M. (2008) Structure of a sliding clamp on DNA. *Cell*, **132**, 43–54.
36. Hedglin,M., Pandey,B. and Benkovic,S.J. (2016) Characterization of human translesion DNA synthesis across a UV-induced DNA lesion. *Elife*, **5**, e19788.
37. Haracska,L., Unk,I., Prakash,L. and Prakash,S. (2006) Ubiquitylation of yeast proliferating cell nuclear antigen and its implications for translesion DNA synthesis. *Proc. Natl. Acad. Sci.*, **103**, 6477–6482.
38. Haracska,L., Johnson,R.E., Unk,I., Phillips,B., Hurwitz,J., Prakash,L. and Prakash,S. (2001) Physical and functional interactions of human DNA polymerase eta with PCNA. *Mol. Cell. Biol.*, **21**, 7199–7206.

REFERENCES

1. Krishna,T.S., 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.
2. McNally,R., Bowman,G.D., Goedken,E.R., O'Donnell,M. and Kuriyan,J. (2010) Analysis of the role of PCNA-DNA contacts during clamp loading. *BMC Struct. Biol.*, **10**, 3.
3. Mailand,N., Gibbs-Seymour,I. and Bekker-Jensen,S. (2013) Regulation of PCNA-protein interactions for genome stability. *Nat. Rev. Mol. Cell Biol.*, **14**, 269–282.
4. De Biasio,A., de Opakua,A.I., Mortuza,G.B., Molina,R., Cordeiro,T.N., Castillo,F., Villate,M., Merino,N., Delgado,S., Gil-Carton,D. *et al.* (2015) Structure of p15(PAF)-PCNA complex and implications for clamp sliding during DNA replication and repair. *Nat. Commun.*, **6**, 6439.
5. De March,M., Merino,N., Barrera-Vilarmau,S., Crehuet,R., Onesti,S., Blanco,F.J. and De Biasio,A. (2017) Structural basis of human PCNA sliding on DNA. *Nat. Commun.*, **8**, 13935.
6. De March,M. and De Biasio,A. (2017) The dark side of the ring: role of the DNA sliding surface of PCNA. *Crit. Rev. Biochem. Mol. Biol.*, **52**, 663–673.
7. Ivanov,I., Chapados,B.R., McCammon,J.A. and Tainer,J.A. (2006) Proliferating cell nuclear antigen loaded onto double-stranded DNA: dynamics, minor groove interactions and functional implications. *Nucleic Acids Res.*, **34**, 6023–6033.
8. Kochaniak,A.B., Habuchi,S., Loparo,J.J., Chang,D.J., Cimprich,K.A., Walter,J.C. and van Oijen,A.M. (2009) Proliferating cell nuclear antigen uses two distinct modes to move along DNA. *J. Biol. Chem.*, **284**, 17700–17710.
9. Daitchman,D., Greenblatt,H.M. and Levy,Y. (2018) Diffusion of ring-shaped proteins along DNA: case study of sliding clamps. *Nucleic Acids Res.*, **46**, 5935–5949.
10. 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. *J. Biol. Chem.*, **270**, 22527–22534.
11. Billon,P., Li,J., Lambert,J.P., Chen,Y., Tremblay,V., Brunzelle,J.S., Gingras,A.C., Verreault,A., Sugiyama,T., Couture,J.F. *et al.* (2017) Acetylation of PCNA sliding surface by Eco1 promotes genome stability through homologous recombination. *Mol. Cell*, **65**, 78–90.
12. De Biasio,A. and Blanco,F.J. (2013) Proliferating cell nuclear antigen structure and interactions: too many partners for one dancer? *Adv. Protein Chem. Struct. Biol.*, **91**, 1–36.
13. Freudenthal,B.D., Gakhar,L., Ramaswamy,S. and Washington,M.T. (2010) Structure of monoubiquitinated PCNA and implications for translesion synthesis and DNA polymerase exchange. *Nat. Struct. Mol. Biol.*, **17**, 479–484.
14. Freudenthal,B.D., Brogie,J.E., Gakhar,L., Kondratick,C.M. and Washington,M.T. (2011) Crystal structure of SUMO-modified proliferating cell nuclear antigen. *J. Mol. Biol.*, **406**, 9–17.
15. De Biasio,A., Ibanez de Opakua,A., Cordeiro,T.N., Villate,M., Merino,N., Sibille,N., Lelli,M., Diercks,T., Bernado,P. and Blanco,F.J. (2014) p15PAF is an intrinsically disordered protein with nonrandom structural preferences at sites of interaction with other proteins. *Biophys. J.*, **106**, 865–874.
16. Yu,P., Huang,B., Shen,M., Lau,C., Chan,E., Michel,J., Xiong,Y., Payan,D.G. and Luo,Y. (2001) p15(PAF), a novel PCNA associated factor with increased expression in tumor tissues. *Oncogene*, **20**, 484–489.
17. Emanuele,M.J., Ciccia,A., Elia,A.E. and Elledge,S.J. (2011) Proliferating cell nuclear antigen (PCNA)-associated KIAA0101/PAF15 protein is a cell cycle-regulated anaphase-promoting complex/cyclosome substrate. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 9845–9850.
18. Povlsen,L.K., Beli,P., Wagner,S.A., Poulsen,S.L., Sylvestersen,K.B., Poulsen,J.W., Nielsen,M.L., Bekker-Jensen,S., Mailand,N. and Choudhary,C. (2012) Systems-wide analysis of ubiquitylation dynamics reveals a key role for PAF15 ubiquitylation in DNA-damage bypass. *Nat. Cell Biol.*, **14**, 1089–1098.
19. Hosokawa,M., Takehara,A., Matsuda,K., Eguchi,H., Ohigashi,H., Ishikawa,O., Shinomura,Y., Imai,K., Nakamura,Y. and Nakagawa,H. (2007) Oncogenic role of KIAA0101 interacting with proliferating cell nuclear antigen in pancreatic cancer. *Cancer Res.*, **67**, 2568–2576.
20. Chang,C.N., Feng,M.J., Chen,Y.L., Yuan,R.H. and Jeng,Y.M. (2013) p15(PAF) is an Rb/E2F-regulated S-phase protein essential for DNA synthesis and cell cycle progression. *PLoS One*, **8**, e61196.
21. Cordeiro,T.N., Chen,P.C., De Biasio,A., Sibille,N., Blanco,F.J., Hub,J.S., Crehuet,R. and Bernado,P. (2017) Disentangling polydispersity in the PCNA-p15PAF complex, a disordered, transient and multivalent macromolecular assembly. *Nucleic Acids Res.*, **45**, 1501–1515.
22. De Biasio,A., Sanchez,R., Prieto,J., Villate,M., Campos-Olivas,R. and Blanco,F.J. (2011) Reduced stability and increased dynamics in the human proliferating cell nuclear antigen (PCNA) relative to the yeast homolog. *PLoS One*, **6**, e16600.
23. Kabsch,W. (2010) Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. D. Biol. Crystallogr.*, **66**, 133–144.
24. Winn,M.D., Ballard,C.C., Cowtan,K.D., Dodson,E.J., Emsley,P., Evans,P.R., Keegan,R.M., Krissinel,E.B., Leslie,A.G., McCoy,A. *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. D. Biol. Crystallogr.*, **67**, 235–242.
25. Murshudov,G.N., Skubak,P., Lebedev,A.A., Pannu,N.S., Steiner,R.A., Nicholls,R.A., Winn,M.D., Long,F. and Vagin,A.A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D. Biol. Crystallogr.*, **67**, 355–367.
26. Emsley,P., Lohkamp,B., Scott,W.G. and Cowtan,K. (2010) Features and development of Coot. *Acta Crystallogr. D. Biol. Crystallogr.*, **66**, 486–501.
27. Abraham,M.J., Murtola,T., Schulz,R., Pail,S., Smith,J.C., Hess,B. and Lindhal,E. (2015) GROMACS: High performance molecular simulations through multi- level parallelism from laptops to supercomputers. *SoftwareX*, **1–2**, 19–25
28. Ivani,I., Dans,P.D., Noy,A., Perez,A., Faustino,I., Hospital,A., Walther,J., Andrio,P., Goni,R., Balaceanu,A. *et al.* (2016) Parmbsc1: a refined force field for DNA simulations. *Nat. Methods*, **13**, 55–58.
29. McGibbon,R.T., Beauchamp,K.A., Harrigan,M.P., Klein,C., Swails,J.M., Hernandez,C.X., Schwantes,C.R., Wang,L.P., Lane,T.J. and Pande,V.S. (2015) MDTraj: A modern open library for the analysis of molecular dynamics trajectories. *Biophys. J.*, **109**, 1528–1532.
30. Hunter,J.D. (2007) Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.*, **9**, 90–95.
31. Sanchez,R., Torres,D., Prieto,J., Blanco,F.J. and Campos-Olivas,R. (2007) Backbone assignment of human proliferating cell nuclear antigen. *Biomol. NMR Assign.*, **1**, 245–247.
32. De Biasio,A., Campos-Olivas,R., Sanchez,R., Lopez-Alonso,J.P., Pantoja-Uceda,D., Merino,N., Villate,M., Martin-Garcia,J.M., Castillo,F., Luque,I. *et al.* (2012) Proliferating Cell Nuclear Antigen (PCNA) interactions in solution studied by NMR. *PLoS One*, **7**, e48390.
33. Frank,E.G., McDonald,J.P., Karata,K., Huston,D. and Woodgate,R. (2012) A strategy for the expression of recombinant proteins traditionally hard to purify. *Anal. Biochem.*, **429**, 132–139.
34. Yeeles,J.T.P., Janska,A., Early,A. and Diffley,J.F.X. (2017) How the eukaryotic replisome achieves rapid and efficient DNA replication. *Mol. Cell*, **65**, 105–116.
35. Georgescu,R.E., Kim,S.S., Yurieva,O., Kuriyan,J., Kong,X.P. and O'Donnell,M. (2008) Structure of a sliding clamp on DNA. *Cell*, **132**, 43–54.
36. Hedglin,M., Pandey,B. and Benkovic,S.J. (2016) Characterization of human translesion DNA synthesis across a UV-induced DNA lesion. *Elife*, **5**, e19788.
37. Haracska,L., Unk,I., Prakash,L. and Prakash,S. (2006) Ubiquitylation of yeast proliferating cell nuclear antigen and its implications for translesion DNA synthesis. *Proc. Natl. Acad. Sci.*, **103**, 6477–6482.
38. Haracska,L., Johnson,R.E., Unk,I., Phillips,B., Hurwitz,J., Prakash,L. and Prakash,S. (2001) Physical and functional interactions of human DNA polymerase eta with PCNA. *Mol. Cell. Biol.*, **21**, 7199–7206.

39. Dzantiev,L., Constantin,N., Genschel,J., Iyer,R.R., Burgers,P.M. and Modrich,P. (2004) Defined human system that supports bidirectional mismatch-provoked excision. *Mol. Cell*, **15**, 31–41.

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

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

42. Xie,B., Mazloum,N., Liu,L., Rahmeh,A., Li,H. and Lee,M.Y.W.T. (2002) Reconstitution and characterization of the human DNA polymerase delta four-subunit holoenzyme. *Biochemistry*, **41**, 13133–13142.

43. Li,H., Xie,B., Zhou,Y.J., Rahmeh,A., Trusa,S., Zhang,S.F., Gao,Y., Lee,E.Y.C. and Lee,M.Y.W.T. (2006) Functional roles of p12, the fourth subunit of human DNA polymerase delta. *J. Biol. Chem.*, **281**, 14748–14755.

44. Rahmeh,A.A., Zhou,Y.J., Xie,B., Li,H., Lee,E.Y.C. and Lee,M.Y.W.T. (2012) Phosphorylation of the p68 subunit of Pol delta acts as a molecular switch to regulate its interaction with PCNA. *Biochemistry*, **51**, 416–424.

45. Zhang,S.F., Zhou,Y.J., Trusa,S., Meng,X., Lee,E.Y.C. and Lee,M.Y.W.T. (2007) A novel DNA damage response - rapid degradation of the p12 subunit of DNA polymerase delta. *J. Biol. Chem.*, **282**, 15330–15340.

46. Meng,X., Zhou,Y., Zhang,S., Lee,E.Y., Frick,D.N. and Lee,M.Y. (2009) DNA damage alters DNA polymerase delta to a form that exhibits increased discrimination against modified template bases and mismatched primers. *Nucleic Acids Res.*, **37**, 647–657.

47. Meng,X., Zhou,Y.J., Lee,E.Y.C., Lee,M.Y.W.T. and Frick,D.N. (2010) The p12 subunit of human polymerase delta modulates the rate and fidelity of DNA synthesis. *Biochemistry*, **49**, 3545–3554.

48. Lau,W.C., Li,Y., Zhang,Q. and Huen,M.S. (2015) Molecular architecture of the Ub-PCNA/Pol eta complex bound to DNA. *Sci. Rep.*, **5**, 15759.

49. Hedglin,M., Pandey,B. and Benkovic,S.J. (2016) Stability of the human polymerase delta holoenzyme and its implications in lagging strand DNA synthesis. *Proc. Natl. Acad. Sci.*, **113**, E1777–E1786.

50. Mayanagi,K., Kiyonari,S., Saito,M., Shirai,T., Ishino,Y. and Morikawa,K. (2009) Mechanism of replication machinery assembly as revealed by the DNA ligase-PCNA-DNA complex architecture. *Proc. Natl. Acad. Sci.*, **106**, 4647–4652.

51. Xu,X., Yan,C., Kossmann,B.R. and Ivanov,I. (2016) Secondary interaction interfaces with PCNA control conformational switching of DNA polymerase PolB from polymerization to editing. *J. Phys. Chem. B*, **120**, 8379–8388.

52. Fernandez-Leiro,R., Conrad,J., Yang,J.C., Freund,S.M., Scheres,S.H. and Lamers,M.H. (2017) Self-correcting mismatches during high-fidelity DNA replication. *Nat. Struct. Mol. Biol.*, **24**, 140–143.

53. Fernandez-Leiro,R., Conrad,J., Scheres,S.H. and Lamers,M.H. (2015) cryo-EM structures of the E. coli replicative DNA polymerase reveal its dynamic interactions with the DNA sliding clamp, exonuclease and tau. *Elife*, **4**, e11134.

54. Biertumpfel,C., Zhao,Y., Kondo,Y., Ramon-Maiques,S., Gregory,M., Lee,J.Y., Masutani,C., Lehmann,A.R., Hanaoka,F. and Yang,W. (2010) Structure and mechanism of human DNA polymerase eta. *Nature*, **465**, 1044–1048.

55. Waters,L.S., Minesinger,B.K., Wiltrout,M.E., D'Souza,S., Woodruff,R.V. and Walker,G.C. (2009) Eukaryotic translesion polymerases and their roles and regulation in DNA damage tolerance. *Microbiol. Mol. Biol. Rev.*, **73**, 134–154.

56. 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.

57. O'Donnell,M., Langston,L. and Stillman,B. (2013) Principles and concepts of DNA replication in bacteria, archaea, and eukarya. *Cold Spring Harb. Perspect. Biol.*, **5**, a010108.

Supplementary Information for

p15^{PAF} binding to PCNA modulates the DNA sliding surface

Matteo De March[1], Susana Barrera-Vilarmau[2], Emmanuele Crespan[3], Elisa Mentegari[3], Nekane Merino[4], Amaia Gonzalez-Magaña[4], Miguel Romano-Moreno[4], Giovanni Maga[3], Ramon Crehuet[2], Silvia Onesti[1], Francisco J. Blanco[4,5] and Alfredo De Biasio[1,6]

[1] Structural Biology Laboratory, Elettra-Sincrotrone Trieste S.C.p.A., Trieste 34149, Italy;

[2] Institute of Advanced Chemistry of Catalonia (IQAC), CSIC, Jordi Girona 18-26, 08034, Barcelona, Spain;

[3] Institute of Molecular Genetics, IGM-CNR, via Abbiategrasso 207, 27100 Pavia, Italy

[4] CIC bioGUNE, Parque Tecnològico de Bizkaia Edificio 800, 48160 Derio, Spain;
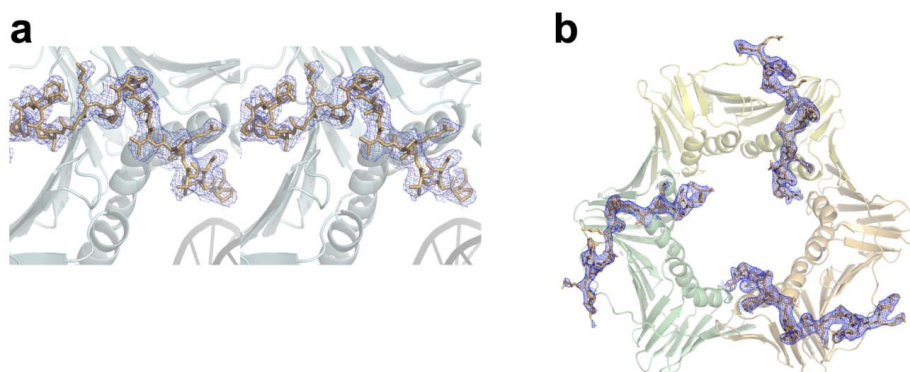
[5] IKERBASQUE, Basque Foundation for Science, Bilbao, Spain

[6] Leicester Institute of Structural & Chemical Biology and Department of Molecular & Cell Biology, University of Leicester, Lancaster Rd, Leicester LE1 7HB, UK

This file includes:
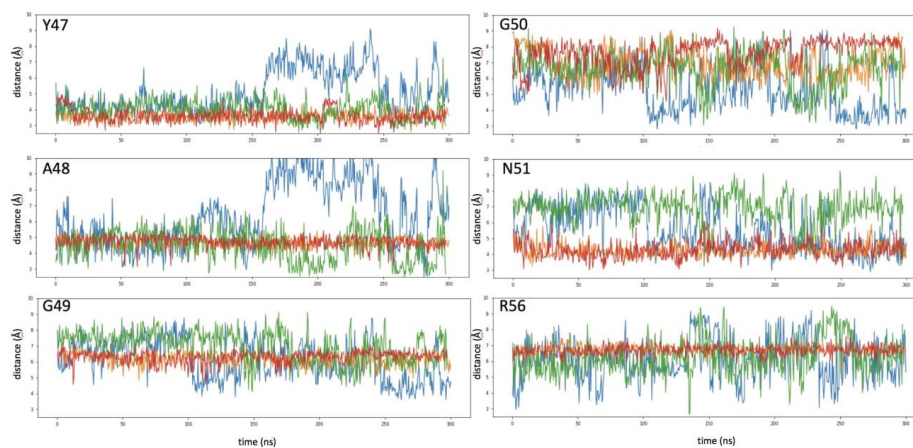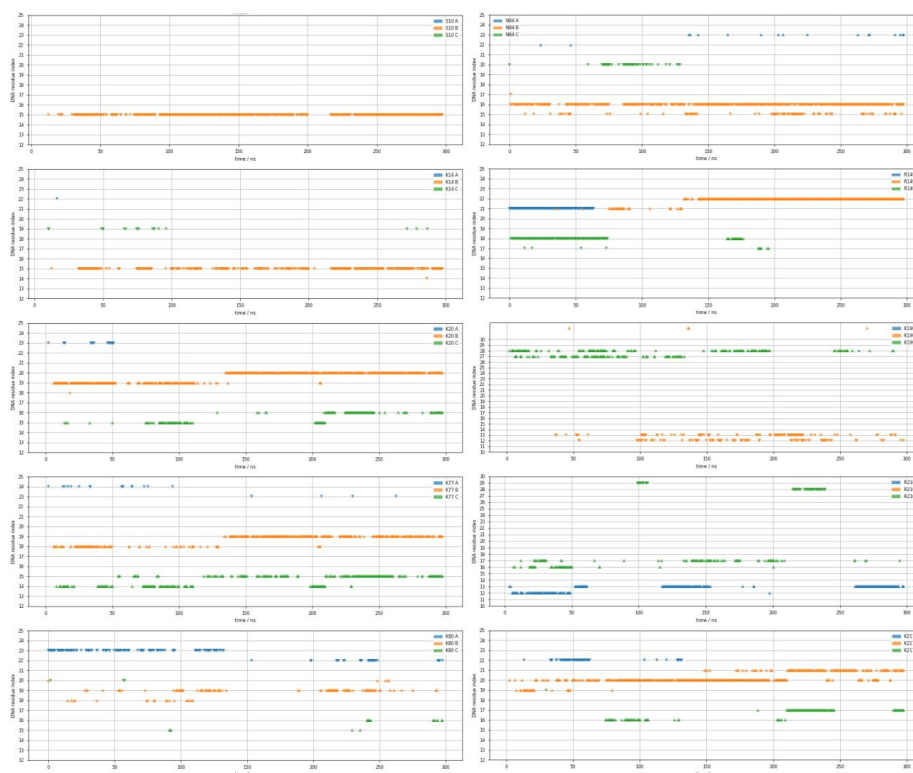
Supplementary Figures S1-9

Supplementary Tables T1-2

Supplementary Figure S1. (**a**) Stereo view of the *2F$_o$-F$_c$* electron density map around the p15$^{50-77}$ peptide with higher occupancy in the p15$^{50-77}$–PCNA–DNA complex, contoured at 1 σ (**b**) View of the *2F$_o$-F$_c$* map of the p15$^{41-72}$–PCNA complex, contoured at 1 σ around the three p15$^{41-72}$ peptides. In both panels, p15 peptides are in stick representation, PCNA in ribbon representation.
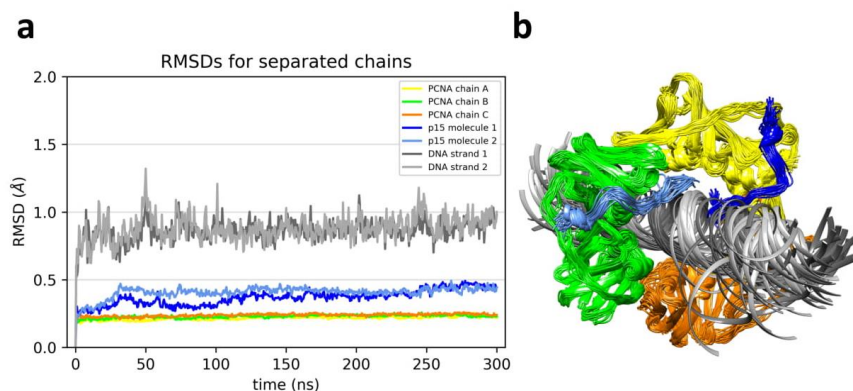

Supplementary Figure S2. (**a**) Root Mean Square Deviation (RMSD) calculated for the backbone heavy atoms of the PCNA protein, the DNA duplexes, and p15 peptides along the MD trajectories. In each case, the system has been superimposed onto the initial minimized structure that was built from the crystal structure, as explained in the methods section. (**b**) Superposition of the ternary complex structures from MD replica 1, with colour code as in (**a**). One can see that most of DNA flexibility arises from its terminal regions. The results for replica 2 are visually similar.
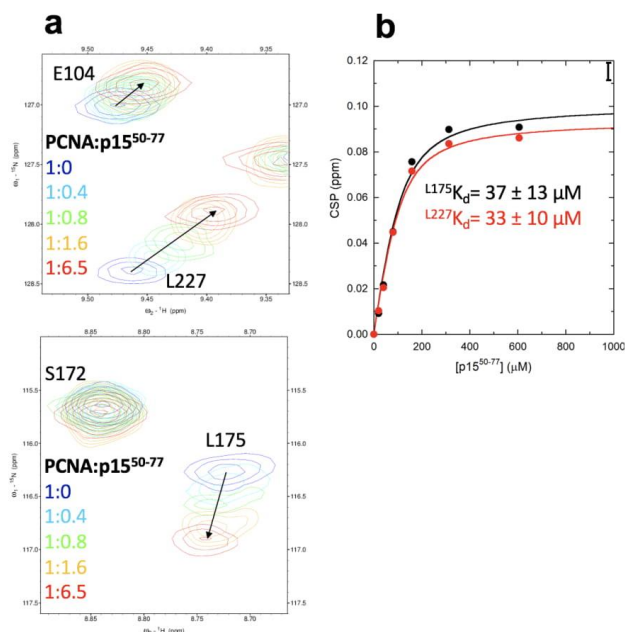
Supplementary Figure S3. *Analysis of the interaction of the p15 with DNA in the MD simulation of PCNA bound to two p15$^{47-70}$ peptides and a 40 bp DNA duplex*. Evolution of distances between p15$^{47-70}$ peptide residues (backbone or side-chain atoms) to DNA phosphorus (P) in the MD trajectory for the indicated residues (red and orange traces correspond to peptide 1 in MD replica 1 and 2, respectively; blue and green traces correspond to peptide 2 in replica 1 and 2, respectively).

Supplementary Figure S4. *Analysis of the MD simulation of PCNA bound to two p15$^{47-70}$ peptides and a 40 bp DNA duplex*. Time evolution of contacts (interatomic side chain nitrogen or oxygen – DNA phosphorus distance < 4.7 Å) between PCNA interfacial residues and DNA phosphorus atoms. Contacts with residues belonging to PCNA subunit A, B and C are shown as triangles colored in blue, orange and green, respectively. The interacting nucleotides of the two strands of the 40 bp dsDNA are consecutively numbered, and triangles with vertexes pointing up and down correspond to contacts with DNA strand 1 and 2, respectively.
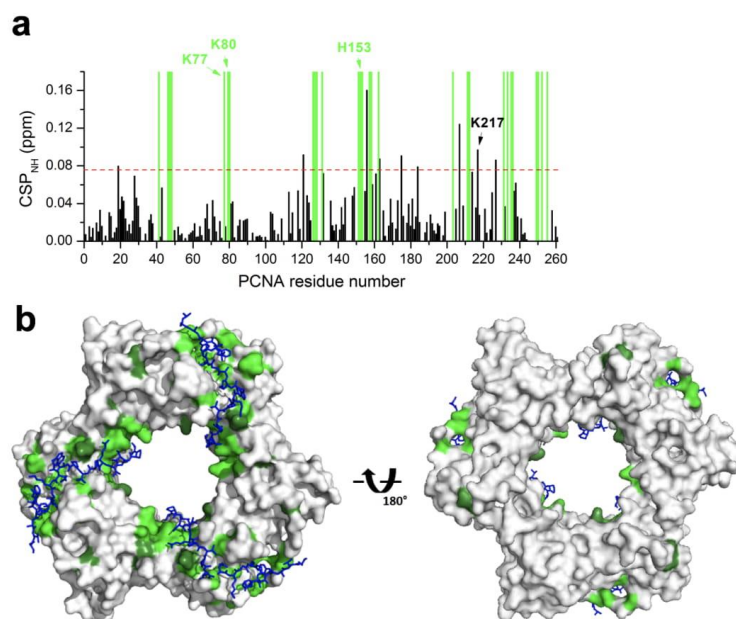
Supplementary Figure S5 (a) Chemical shift perturbations (CSP) of PCNA backbone amide ^{1}H and ^{15}N NMR resonances induced by p15$^{50-77}$. The dotted line indicates the average plus two standard deviations. The green bars indicate the position of residues that disappear upon addition of substoichiometric p15$^{50-77}$, and are not drawn to scale. The residues perturbed by p15$^{50-77}$ and that also appear at the PCNA–DNA interface in the p15$^{50-77}$–PCNA–DNA crystal structure are labeled **(b)** Front- and back-face views of PCNA surface. PCNA residues whose amide signals disappear in the presence of substoichiometric p15$^{50-77}$, or whose signals persist but shift significantly along the titration are colored light or dark green, respectively. p15$^{50-77}$ bound to the three PCNA PIP-box sites is shown in sticks.

Supplementary Figure S6. *NMR titration analysis of PCNA binding to p15$^{50-77}$.* (**a**) Overlay of ^{1}H-^{15}N TROSY spectra of 100 μM PCNA in the presence of increasing concentrations of unlabeled p15$^{50-77}$, as indicated by the PCNA protomer : p15 peptide molar ratios. For the sake of clarity only five points along the titration are plotted. Spectra were acquired at 35 °C on samples in 20 mM sodium phosphate, 50 mM NaCl, pH 7.0. The selected region shows significantly perturbed residues (L175, L227) and two unperturbed residues (S172, E104). Arrows indicate the peak center of the residue at the last titration point. The steady shifts of the signals imply a fast exchange between free and bound PCNA with respect to the NMR chemical shift time scale (**b**) Analysis of NMR CSP of two PCNA residues at increasing concentrations of p15$^{50-77}$, using a single-site binding model. The errors in the K_D values are fitting errors. The CSP experimental error estimated from the digital resolution of the spectra is 0.009 ppm, which is the size of the bar at the right upper corner of the plot.
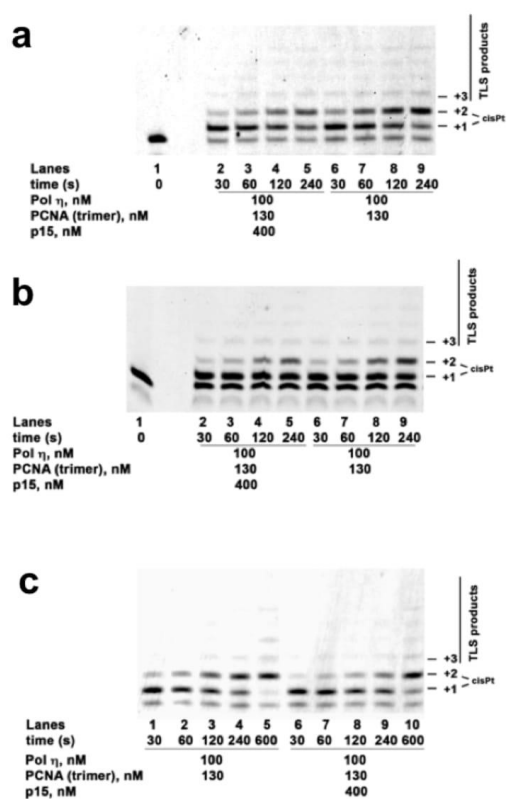
Supplementary Figure S7. (a-c) Three independent time course repeats of the reaction of pol η

in the presence of PCNA/p15 at the indicated concentrations on the cisPt(GG) template (10nM).
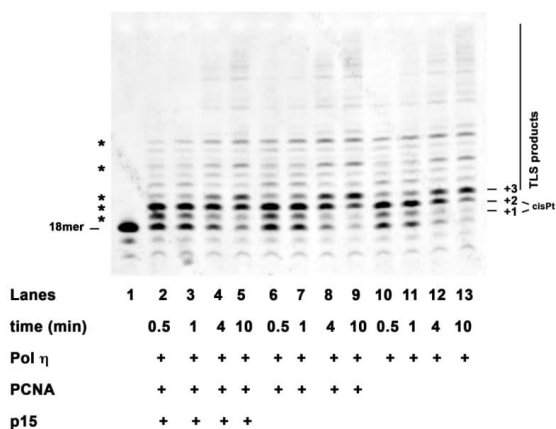
Supplementary Figure S8.

Inhibition of pol η activity by PCNA-p15 in the presence of RF-C. Pol η was tested on the 18/24mer primer-template substrate bearing a cis-Pt adduct at the two guanines at positions +1 and +2 of the template strand, in the presence of RF-C and in the absence (lanes 10-13) or in the presence of PCNA (lanes 6-9) or PCNA and p15 (lanes 2-5). Reactant concentrations were: 50 μM dNTPs, 0.5 mM ATP, 150 nM ScRF-C, 400 nM Pol η, 400 nM p15, 50 nM PCNA, 40 nM Cis-Pt 18/24mer. Time course of nucleotide incorporation showed that PCNA did not significantly change the translesion synthesis products with respect to pol η alone. However, addition of p15 in combination with PCNA, caused an accumulation of +1 and +2 products, corresponding to incorporation opposite the two G's of the cis-Pt adduct, and a delay in the appearance of longer products, particularly at short incubation times (compare lanes 2,3 with lanes 6,7 and 10, 11), as can be seen comparing the products marked with asterisks across the different lanes. Thus, these results indicated that the PCNA/p15 complex was able to delay TLS across a cis-Pt lesion even in the presence of RF-C.

24/18mer control 10 nM

Lanes	1	2	3	4	5	6	7	8	9
time (min)	0	0.5	1	2	4	0.5	1	2	4
p15 (nM)		400				-			
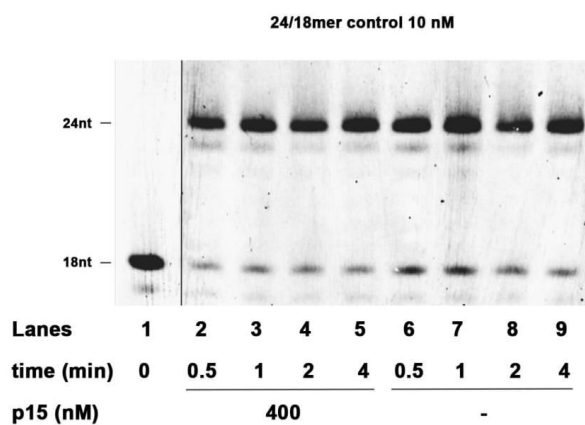
Supplementary Figure S9.

Time course of the reaction of pol η (70 nM) alone (Lanes 6-9) or in the presence of p15 (400 nM) (Lanes 2-5) on the 24/18mer control template (10 nM). Bar indicates were two portions of the same gel have been moved close for clarity.

Supplementary Table T1. Sequences of the oligonucleotides used in this study.

DNA oligo	Sequence (5′-3′)
1	Cy5-CCCATCGTAT
2	TTTTATACGATGGG
3	CCCATCGTAT
4	ATACGATGGG
5	ATACGATGGGATACGATGGGATACGATGGGATACGATGGG
6	CCCATCGTATCCCATCGTATCCCATCGTATCCCATCGTAT
7	CTAC**GG**CTCACACTATCTCACACT
8	CTACGGCTCACACTATCTCACACT
9	AGTGTGAGATAGTGTGAG

The sequence of oligonucleotides 1 and 2 were used to form the primed DNA and correspond to the substrate used in the crystal structure of β-clamp bound to pDNA, reported by Georgescu *et al.*[1]. Oligonucleotides 3 and 4 were used to form the dsDNA duplex for co-crystallization with PCNA and p15$^{50-77}$ and for NMR titration. Oligonucleotides 5 and 6 were used to form the 40 bp DNA for the MD simulation. Oligonucleotides 7-9 were used for the DNA synthesis studies with pol η. The bold letters in oligo 7 indicate the position of the cisPt(GG) crosslink. Oligo 8 represents the 24-mer undamaged template, and the sequence underlined is the one complementary to oligo 9 (18-mer primer).

[1] Georgescu, R.E., Kim, S.S., Yurieva, O., Kuriyan, J., Kong, X.P. and O'Donnell, M. (2008) Structure of a sliding clamp on DNA. *Cell*, **132**, 43-54.

Supplementary Table T2: Data collection and refinement statistics.

	PDB: 6EHT p15$^{50-77}$–PCNA–pDNA	PDB: 6GWS p15$^{41-72}$– PCNA
Data collection		
Space group	P2$_1$	P2$_1$
Cell dimensions		
a, b, c (Å)	75.99, 42.30, 141.83	79.24, 89.75, 85.13
α, β, γ (°)	90, 102.70, 90	90, 117.25, 90
Resolution (Å)	46.12 - 3.20 (3.37 - 3.20)	75.69 - 2.90 (3.08 - 2.90)
R_{merge}	0.105 (0.320)	0.11 (1.113)
R_{PIM}	0.087 (0.257)	0.087 (0.730)
$I / \sigma I$	4.5 (2.0)	4.77 (1.6)
CC(1/2)	0.987 (0.916)	0.986 (0.763)
Completeness (%)	95.3 (98.8)	92.2 (91.3)
Redundancy	2.2 (2.3)	1.8 (1.8)
Refinement		
Resolution (Å)	40.484 - 3.2	44.9 - 2.9
No. reflections	13440	23688
R_{work} / R_{free}	0.26 / 0.33	0.20 / 0.26
No. atoms		
Protein	5229	5882
Peptide	277	538
DNA	410	-
Water	9	10
B-factors (Å^2)		
Protein	60.6	99.04
Peptide	70.6	98.99
DNA	230.1	-
Water	60.5	70.65
R.m.s. deviations		
Bond lengths (Å)	0.0102	0.0030
Bond angles (°)	1.3961	1.2740