Dissertation zur Erlangung des Doktorgrades

der Fakultät für Chemie und Pharmazie

der Ludwigs-Maximilians-Universität München

Structural and Biochemical Characterisation of Minimal Translesion Synthesis Complexes

Giulia Chiapparini

aus

Chiavenna, Italien

2020

Erklärung

Diese Dissertation wurde im Sinne von §7 der Promotionsordnung vom 28. November 2011 von Frau Prof. Elena Conti, PhD betreut.

Eidesstattliche Versicherung

Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfe erarbeitet.

München, den 07.12.2020

Giulia Chiapparini

Dissertation eingereicht am 07.12.2020

1. Gutachterin : Prof. Dr. Elena Conti

2. Gutachter : Prof. Dr. Julian Stingele

Mündliche Prüfung am 10.02.2021

A mio fratello Diego.

Contents

Summary

Translesion DNA synthesis (TLS) is a fundamental DNA damage tolerance pathway. Although many sophisticated DNA repair pathways have evolved, DNA lesions can be still present during DNA replication. When a replication fork encounters a bulky lesion (e.g. UV light-induced pyrimidine dimers) it is in danger to stall, as replicative high-fidelity DNA polymerases are not able to process such obstacles. However, replicating across the lesion and avoiding replication fork stalling is of utmost importance to prevent the onset of major problems like fork instability, formation of double strand breaks or missing the correct timing of replication fork convergence. Here, specialized TLS polymerases (in eukaryotes Polη, Polι, Polκ, REV1 and Pol ζ) step in and bypass the lesion, often at the expense of accuracy. Polymerases are anchored to the replisome by the DNA sliding clamp PCNA. Mono-ubiquitination of PCNA on K164 residue is the trigger which leads to the assembly of TLS complexes and specialised polymerases recruitment on damaged DNA templates. Interactions between TLS polymerases and recruitment of several proteins on PCNA trimer (PCNA tool belt model) have been investigated for the last decades, but many fundamental questions are still open. Additionally, the role of ubiquitination is still not clear as in the literature there is diverging evidence between *in vitro* and *in vivo* studies. In the presented work, recombinant human TLS complexes were extensively characterized from a biochemical point of view, with a special focus on Polη, Polι and their interaction with PCNA or mono-UbPCNA. Complexes could be efficiently reconstituted, and Polη and Polι showed different occupancies on PCNA trimers. Moreover, I characterized the regulatory disordered C-terminus of Polη and its role in the interaction with PCNA and UbPCNA, with a focus on newly identified PCNA interacting peptides (PIP boxes) and ubiquitin binding zinc finger (UBZ). Firstly, analytical pulldown assays showed that PCNA ubiquitination strongly favours the recruitment of Polη. Secondly, I observed that PIP boxes have a key role in the interaction of Polη with either unmodified or ubiquitinated PCNA, and that the sole UBZ is necessary but not sufficient to recruit UbPCNA. Moreover, results indicated a possible recruitment of both Polη and Polι to UbPCNA, but not to unmodified PCNA, suggesting a critical role of PCNA ubiquitination in the simultaneous binding of different types of TLS polymerases. Taking advantage of cryo electron microscopy (cryo-EM) to carry out structural analyses of complexes and single proteins, provided a molecular view of PCNA and UbPCNA and the flexibility of ubiquitin on the sliding clamp. Furthermore, this thesis presents a first report of a cryo-EM reconstruction of PCNA-Polι-DNA complex as well

as full-length Polι in complex with DNA. The cryo-EM reconstruction of a UbPCNA-Polη complex at 9 Å resolution showed that ubiquitin relocated to the front side of PCNA, bridging the polymerase to the DNA junction and making additional contacts with the flexible regulatory C-terminus of Polη. Altogether, the structural and biochemical results presented here give new insights into the role of PCNA ubiquitination for the recruitment of the flexible C-terminal regulatory regions of TLS polymerases. The presented work provides the basis for further structural and biochemical studies to understand the regulation and mechanisms of translesion synthesis.

1. Introduction

1.1 DNA damage and DNA damage tolerance

DNA, the molecule that enshrines genetic information, has to be accurately duplicated and its integrity has to be maintained during the cell cycle. Ironically, the process of living itself generates reactive metabolites that can cause DNA damage. Moreover, cells are exposed to a vast array of exogenous stresses that can directly or indirectly lead to DNA damage (Waters et al., 2009). Primarily, spontaneous reactions that naturally occur in the cells, such as hydrolysis, derive by the intrinsic chemical nature of DNA in an aqueous solution and create abasic sites (Lindahl, 1993). Furthermore, cellular metabolism generates molecules which damage DNA, such as reactive oxygen and nitrogen species, endogenous alkylating agents, estrogen and cholesterol metabolites, lipid peroxidation products and reactive carbonyl species (De Bont, 2004). DNA is constantly damaged by exogenous physical and chemical agents like ultraviolet and ionizing radiations and chemical carcinogens. Endogenous DNA damages occur more frequently compared to exogenous damages, but the type of damage produced by cellular processes or exogenous stimulation are very similar (Jackson and Loeb, 2001). In both normal and cancer cells, tens of thousands of DNA lesions per day are formed (Jackson and Bartek, 2009). Some types of damage, such as oxidative damage to DNA bases, arise, and are repaired, as often as 10^5 lesions per cell each day (Mehta and Haber, 2014). Single-strand breaks and spontaneous base losses in nuclear DNA occur at an estimated daily rate of 10^4 lesions per cell. This type of lesions, added to other types of spontaneous damages, reaches the total rate of 10^5 lesions per cell per day.

This huge plethora of DNA damages has compelled the evolution of a multitude of cellular responses, generally referred as DNA damage response (Figure 1.1) DNA damage response (DDR) is a complicated and entangled network of mechanisms, including multiple DNA repair pathways, damage tolerance processes, and cell-cycle checkpoints which safeguard genomic integrity (Giglia-Mari et al., 2011). Specific DNA repair pathways deal with multitudes of DNA damages in several phases of the cell cycle. Repair mechanisms bound to solve incorrect bases or damages at single nucleotide level are excision repair (BER), nucleotide excision repair (NER) and mismatch repair (MMR). In cases of most complex situations, like single or double strand breaks (SSB and DSB) the activation of homologous recombination (HR) and non-

homologous end joining (NHEJ) pathways is required. Specific lesions or crosslinks can also be removed by direct chemical reversal and interstrand crosslink (ICL) repair. Overall, all these repair processes are key to maintaining genetic stability in cells (Chatterjee and Walker, 2017).

Although numerous repair mechanisms efficiently repair different types of DNA lesions, it is likely that the replication machinery would still encounter lesions that are mis-repaired, or newly occurring lesions not repaired at all. Replication through persisting lesions would interfere with normal DNA replication and result in high frequency of fork collapse and genome instability, eventually leading to cell death (Ghosal and Chen, 2013). Fortunately, cells possess mechanisms to temporarily tolerate DNA damage until DNA repair processes can successfully remove it. Such mechanisms go under the name of DNA damage tolerance (DDT). Conceptually, DNA damage tolerance is quite different from DNA repair in that, rather than restoring DNA to its proper sequence and structure, the lesion is still present in the DNA at the conclusion of DNA damage tolerance process. The function of DNA damage tolerance is to temporarily bypass a DNA lesion rather than to regenerate the original sequence, therefore DDT mechanisms are optimized to allow survival by promoting the completion of DNA replication (Waters et al., 2009). DNA damage tolerance pathways promote the bypass of single-stranded DNA lesions encountered by DNA polymerases during DNA replication, which prevents the stalling of DNA replication and completion of DNA synthesis in a timely and coordinated manner.

In eukaryotes, two mechanistically distinct DNA damage tolerance branches have been characterized: template switching (TS) and translesion synthesis (TLS) (Bi, 2015). In template switching, the stalled replicative polymerase finds an alternative template, most commonly the newly synthesised strand on the sister chromatid, that can use as template to get around the lesion and restart replication (Andersen et al., 2008). Translesion synthesis is a complex mechanism where the lesion bypass is achieved by a process called “polymerase switch” where replicative polymerase is temporarily substituted by a specialised TLS polymerase, which can incorporate nucleotides against the DNA lesion. DNA polymerases besides having a pivotal role in DNA replication, also play an crucial role in protecting the cell against the effects of DNA damage, being involved in most of the DNA repair processes and DNA damage tolerance (Lange et al, 2011). Translesion synthesis and polymerase switch are the main topic in my PhD thesis, and will be deeply described in the next paragraphs, with a major focus on TLS polymerase (section 1.3) interactions with DNA sliding clamp (section 1.4).

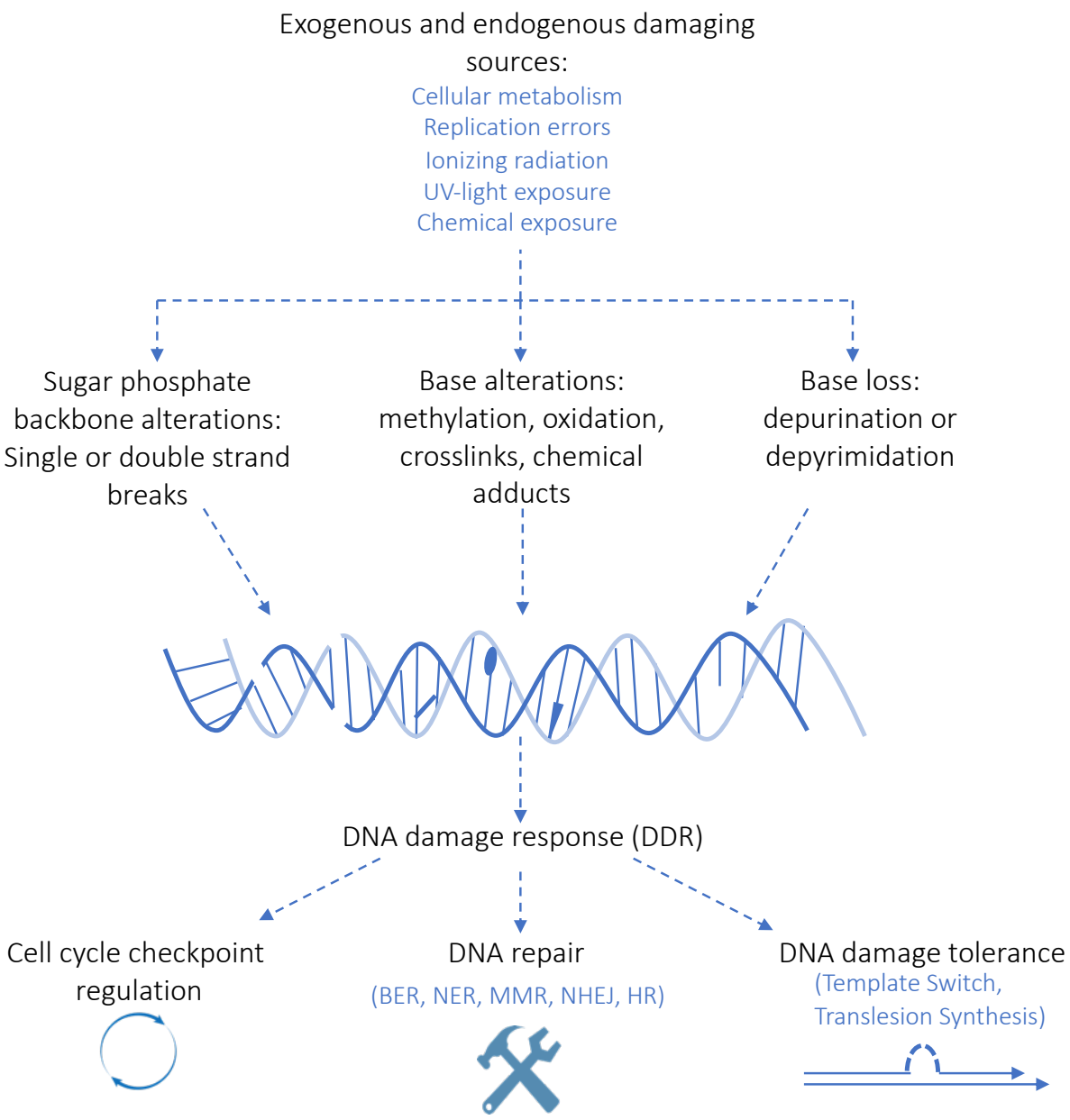


Figure 1.1 DNA damage and DNA damage response (DDR). Schematic representation of causes and effects of DNA damage in human cells. DNA damage is induced by various exogenous and endogenous sources. Common exogenous DNA damaging agents are UV-radiation, chemicals and chemotherapeutic drugs. Endogenous damaging agents originate from the chemical instability of the nucleotides, cellular metabolites, reactive oxygen species and stalled replication forks. For the different lesions types (SSB, DSB, base methylation, base loss, oxidation, intrastrand crosslink, DNA-protein crosslink etc). Several pathways have evolved to counteract DNA damage and undergo the large umbrella of DNA damage response (DDR). DDR is an intricated network that leads to repair of different lesions by coordinated cell cycle control mechanisms, DNA repair, chromatin modifications and DNA damage tolerance pathways.

1.2 Translesion DNA synthesis: an error prone DDT pathway

In eukaryotic cells, nuclear DNA is predominantly replicated by polymerases Polα, Polδ and Polε, all belonging to the B-family (table 1.2) (Haracska et al., 2005). Each of these replicative polymerases feature highly selective active sites that enable them to synthesize new DNA with high fidelity (McCulloch and Kunkel, 2008). However, their restrictive active sites cannot accommodate distortions in the DNA structure or bulky base adducts derived from UV-light irradiation. Therefore, replicative polymerases are incapable of duplicating damaged templates and are prone to stall upstream of DNA lesions (Beard and Wilson, 2003). These barriers must be overcame to ensure the complete duplication of DNA prior to cell division, as well as to prevent formation of lethal double-strand breaks (DSBs) due to replication fork collapse following their prolonged stalling (Saintigny, 2001). Translesion DNA synthesis is a DNA damage tolerance (DDT) mechanism which leads to bypass of a DNA lesion, by exploiting specialised DNA polymerases. These specialised TLS polymerases, thanks to their broader active sites and low nucleotide selectivity, are capable of incorporating nucleotides across the lesion, at the expenses of accuracy.

Replicative DNA polymerases are coordinated to the replisome by Proliferating Cell Nuclear Antigen (PCNA), a heterotrimeric ring-shaped DNA sliding clamp. Polymerase interaction with PCNA prevents dissociation from DNA, and its ability to slide along the double helix ensures high processivity. Notably, PCNA is a fundamental hub, which interacts with many proteins involved in DNA metabolism and the interactions are regulated via post-translational modifications. The mono-ubiquitination of PCNA is an important step that triggers the exchange between replicative polymerases and specialised TLS pols (Sale, 2013). Following replication fork stalling, the accumulation of replication protein A (RPA) triggers the recruitment of the Rad6 (E2)/Rad18 (E3) complex to sites of damage, which sequentially transfers a single ubiquitin moiety to the lysine residue 164 of PCNA (K164) (Davies et al., 2008) (Hedglin et al., 2019) (Figure 1.2 b.). This modification, resulting in the formation of mono-ubiquitinated PCNA (UbPCNA), has a key role in TLS, providing a binding platform for specialised TLS polymerases, which are recruited to the lesion site after the temporary displacement of the replicative polymerase (Figure 1.2). In spite of some debate as to whether PCNA mono-ubiquitination is strictly essential for initiating translesion synthesis (Hendel et al., 2011), numerous groups have demonstrated that PCNA mono-ubiquitination strongly promotes TLS polymerase recruitment to the replication fork by providing an additional means

through which these polymerases can compete for PCNA binding (Kannouche et al.) (Guo et al., 2006).

The principal DNA polymerases responsible for TLS during replication belong to the Y-family (Sale et al., 2012), respectively Polη, Polι, Polκ and REV1, together with the B-family enzyme Polζ (Gan et al., 2008). While providing a benefit through their translesion synthesis ability Y-family polymerases have a 10- to 1000-fold higher mutational rate than replicative polymerases (Kunkel, 2009). In some specific instances, like in the case of Polι, mis-pair formation is favoured over correct nucleotide insertion (Tissier et al., 2010). Thus, access of TLS DNA polymerases to primer termini during replication of genomic DNA in human cells must be tightly regulated, to prevent accumulation of mutations in the genome. The selection of the most appropriate specialised TLS polymerase for recruitment to the lesion site is still a poorly understood mechanism and is dictated probably by a number of factors. One obvious factor is the identity of the DNA lesion, as specialised polymerases recognise a set of cognate lesions (table 1.1), which are preferentially accommodated in their active site. For instance, Polη can efficiently bypass *cis–syn* cyclobutane dimer (CPDs), correctly incorporating two As across the lesion. REV1 is known to be a template-dependent dCMP transferase that can insert a C residue opposite to an abasic site (Lin et al., 1999) (Gibbs et al., 2005). Polκ is most efficient in translesion synthesis across oxidation derivatives of DNA bases, such as thymine glycol and 8-oxoguanine, as well as bulky DNA adducts (Zhang et al.) (Avkin et al., 2002) (Lior-Hoffmann et al., 2012). Polι is another member of Y-family with a peculiar active site that allows Hoogsteen base pairing instead of Watson-Crick base pairing, and can incorporate nucleotides across bulky adducts with very low fidelity (Nair et al., 2004). A second potential factor is the interactions of specialized polymerases with hub proteins such as PCNA and REV1. Other potential factors may include the availability of TLS polymerases in the vicinity of stalled replication forks within replicative factory depending on the cell cycle state and transcription regulation or protein degradation.

Y-family polymerase	Cognate lesion(s)
Polη	Thymine-thymine dimers and 8-oxoguanine
Polκ	8-oxoguanine and thymine glycols
Polι	Minor groove-adducted purines and bulky major groove-adducted purines
REV1	Abasic sites and damaged G templates

Table 1.1 Y-family polymerase cognate lesions. Specialised polymerases have a set of preferred lesions (cognate lesion) that can by efficiently bypassed. Across these lesions specific polymerases can accurately incorporate nucleotides, depending on conformations of their active site and chemical selectivity of DNA adducts. Table adapted from (Boehm, 2016)

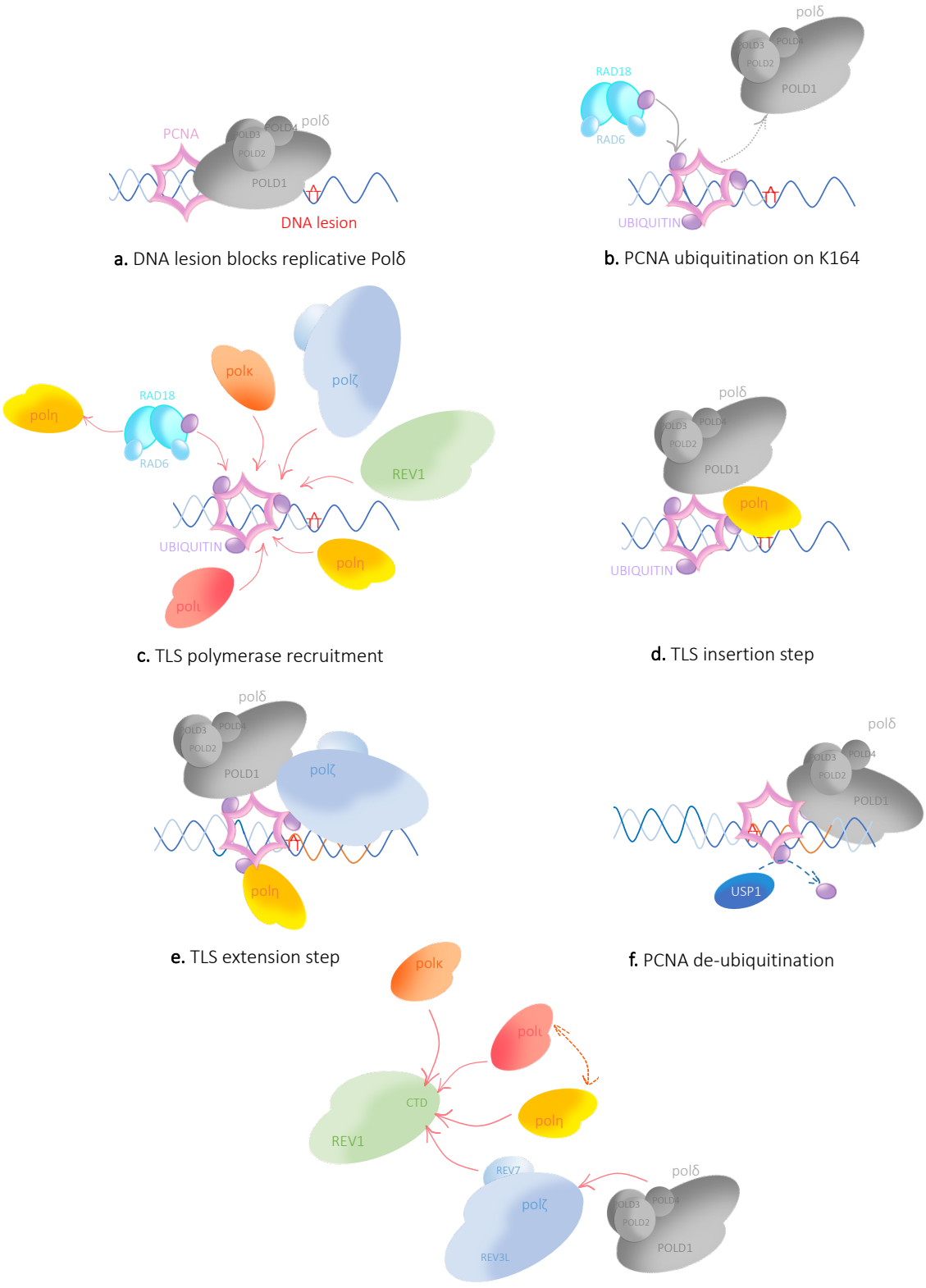
TLS polymerases play an important role in creating mutations, which can be desirable in specific cases. Somatic hypermutation for variability of antigen receptors is an essential mechanism by which the immune system adapts to foreign antigens. In order to accomplish this task, targeted mutations in the variable regions of immunoglobulin genes must be generated. One of the pathways by which mutations are created involves DNA synthesis by error-prone DNA polymerases. All TLS polymerases are considered as potential suspects responsible for the mutagenesis of antibody genes but only REV1 and Polη have been unequivocally proven to play a central role in somatic hypermutation (Zeng et al., 2001) (Jansen et al., 2006) (Krijger et al., 2013).

The mutagenic nature of specialised DNA polymerases is also playing a critical role in maintaining balance between completion of DNA replication without the risk of fork collapse and introduction of mutations in the DNA sequence itself. Therefore such potentially mutagenic process has to be strictly regulated. To this end, TLS polymerases possess structural features that often help them extract the correct code from diversely damaged templates. Coupled with control mechanisms ensuring that their access to DNA templates is strictly regulated, helps limit the opportunities for these polymerases to cause undesirable mutagenesis (Sale, 2013).

1.2.1 Translesion synthesis and polymerase switch mechanism

Translesion synthesis can require more than one polymerase exchange, or polymerase switch on the sliding clamp. The replicative bypass of the vast majority of DNA lesions requires a coordinated action of several TLS DNA polymerases in a two-step process. In the first step, an "inserter" TLS polymerase incorporates one or two nucleotides opposite the specific DNA lesion. The insertion is typically performed by Polη, Polι, or Polκ, (Livneh et al., 2010) (Shachar et al., 2009) although for some specific lesions, the nucleotide insertion is carried out by REV1 (Nelson et al., 1996a) or Polζ (Nelson et al., 1996b) (Lin et al., 2014). In the second step, the "inserter" polymerase is replaced by an "extender", which extends the aberrant primer-template distorted by a DNA lesion. Some cell-based data revealed that this step is exclusively carried out by a master "extender" B-family TLS polymerase Polζ, (Livneh et al., 2010) (Shachar et al., 2009) although Polκ has also demonstrated the ability to extend primer-templates *in vitro* (Washington et al., 2002) (Washington et al., 2004). DNA polymerases involved in TLS are organized as a multiprotein TLS complex, as they all form an intricate interaction network stabilized by the two central scaffold proteins: PCNA and REV1. The C-terminal domain (CTD) of REV1, in spite of its relatively small size (11 kDa), uses two independent interaction

interfaces to simultaneously bind the accessory REV7 subunit of the "extender" polymerase Polζ, and a REV1-interacting region (RIR) (Ohashi et al., 2009) from one of the "inserter" Y-family polymerases (Figures 1.2 and 1.3).



a. DNA lesion blocks replicative Polδ

b. PCNA ubiquitination on K164

c. TLS polymerase recruitment

d. TLS insertion step

e. TLS extension step

f. PCNA de-ubiquitination

g. TLS polymerases interactions

Figure 1.2 Mechanism of translesion synthesis and polymerase switch. Cartoon representation of each translesion DNA synthesis step. (a) DNA lesion (T^T) is encountered and the replicative polymerase (Polδ) is blocked. (b) Signalling cascade activates mono-ubiquitination of PCNA on K164, performed by Rad6-Rad18 E2-E3 ubiquitin ligase. (c) PCNA ubiquitination triggers the first 'polymerase switch' where replicative Polδ is displaced from DNA and one (or more) of the TLS polymerases is recruited. (d) First step of TLS insertion, in this case Polη is incorporating directly across the lesion. (e) A second polymerase switch might occur, and an extender polymerase (here Polζ) is recruited to extend the primer a few more nucleotides. (f) Once the lesion is left behind PCNA is de-ubiquitinated by USP1. (g) Simplified interaction network of all likely polymerases involved in TLS: all Y-family members (Pol, Polη, Polι, Polκ) as well as REV7 subunit of B-family Polζ interact with CTD of REV1. Polζ shares POLD2 and POLD3 regulatory subunits with replicative Polδ. Additionally, upon ubiquitination Polη and Polι can interact. Molecular details of TLS and polymerases interactions are still under investigation.

1.3 DNA polymerases, an overview

DNA polymerases act as key players in DNA metabolism. These enzymes are the only biological macromolecules able to duplicate the genetic information stored in the DNA and are absolutely required every time this information has to be copied. Since the discovery of the first DNA replicating enzyme, *E.coli* DNA polymerase I, by the Kornberg group in 1956, the number of known DNA polymerases has mushroomed. DNA polymerases are present in all domains of life, as well as viruses, and they have been divided into A-, B-, C-, D-, X-, Y-, RT-, and PrimPol- families (Ito and Braithwaite, 1991) (Rudd et al., 2014) based on sequence similarity (table 1.2). Except for C- and D-families, which exist only in bacteria and archaea (Timinskas et al., 2014)(Matsui et al., 2013) respectively, human DNA polymerases are present in all six other families. Most members of families A- and B- are replicative polymerases and are associated with proofreading 3′–5′ exonuclease activities (Bruck et al., 2003). X-family polymerases mainly participate in short gap repair synthesis (sGRS) associated with BER and NHEJ (Moon et al., 2007) (Nick McElhinny et al., 2005). Telomerase uses its own RNA subunit to template DNA synthesis of many telomere repeats (Blackburn et al., 2006). PrimPol is a very unique polymerase since it can synthesize DNA primers *de novo* and continue DNA synthesis without an RNA primer (Rudd et al., 2014). Y-family members specialize in TLS by incorporating nucleotides directly opposite to the damaged bases (Yang, 2014) (Vaisman and Woodgate, 2017). In addition, DNA polymerases of the X- and A- family are capable of TLS in particular cases. A common feature among X-, Y-, RT-, and PrimPol-families is the absence of a proofreading 3′-5′ exonuclease activity (table 1.2) which is instead present in most of replicative polymerases, conferring them high synthesis accuracy and low mutagenesis.

1. Introduction

By latest account there are 17 DNA polymerases in human cells. Of those, only a handful (Pol α, γ, δ, ε, and telomerase) are in charge of replicating the bulk of nuclear and mitochondrial DNA. The rest and majority of DNA polymerases exhibit low fidelity and processivity and are specialised in translesion and repair synthesis. The presence of low fidelity polymerases in all domains of life, and indeed their expansion in higher organisms, suggests an essential evolutionarily conserved role (Moldovan et al., 2007).

DNA polymerases from different families greatly differ in terms of sequence homology but they all share a common structural topology of the catalytic core. Due to the vivid imagination of crystallographers who first solved the polymerase structure (Ollis et al., 1985) and thanks its peculiar shape, the catalytic domain of a typical polymerase is described as a right hand with fingers, palm, and thumb subdomain. The active site of a polymerase resides in the palm domain, where three conserved acidic residues, essential for the catalysis, coordinate the essential metal ions required for the nucleotidyl transferase reaction. The thumb subdomain plays a central role in DNA binding and determining the enzyme's processivity and translocation. The fingers domain is mainly responsible for the positioning of the DNA template required for optimal nucleotide pairing (Vaisman and Woodgate, 2017). Notwithstanding the lack of apparent sequence homology, the palm domain of the A-, B-, and Y-family polymerases, as well as RT-, is highly conserved, and three carboxylates essential for the catalysis are located on identical structural elements (Yang, 2005). Though the secondary structures vary broadly in the thumb and fingers domains across the different polymerase families, their location in the tertiary structure and roles in interaction with DNA and nucleotide substrate are conserved among all polymerases. Noticeably, the thumb and fingers are distinctively smaller in Y-family polymerase, resulting in a broader active site (Yang and Woodgate, 2007).

Family	Function	Exonuclease 3'-5'	Error rate	Human pols	PDB identification
A	Replication	Yes	10^{-5}–10^{-7}	Pol γ (mitochondrial)	5C53
	Repair	No	10^{-3}–10^{-4}	Pol θ (TMEJ) Pol ν (end processing)	4X0Q 4XVK, 4XVM
B	Replication	Yes	10^{-4}–10^{-7}	Pol α (primer extension) Pol δ (lagging strand, large gap filling, lGRS) Pol ε (leading strand)	4QCL, 5EXR 3IAY 4M8O, 4PTF
	TLS extension	Yes/no	10^{-3}–10^{-6}	Pol ζ (TLS extension)	3K5O, 3K5M
C	Replication	Yes	10^{-5}–10^{-7}	NA	
	Repair	ND	10^{-2}–10^{-3}	NA	
D	Replication	Yes	10^{-4}–10^{-5}	NA	
X	BER BER/sGRS sGRS sGRS	No	10^{-2}–10^{-4}	Pol β (BER, sGRS) Pol λ (BER, NHEJ) Pol μ (NHEJ) TdT (NHEJ)	2FMS, 4KLH 2PFO 4M04 4QZ9
Y	TLS insertion	No	10^{-2}–10^{-4} 3–10^{-4} 10^{-3}–10^{-4} 10^{-3}–10^{-4}	Pol η (CPD, SHM) Pol ι Pol κ (steroid, BPDE) Rev1 (AP, SHM)	3MR3, 4ECV 5ULW 1T94, 2OH2 3GQC
RT	Replication	No	10^{-3}–10^{-4}	HIV-1 RT TERT (telomerase)	3KK2 3KYL
PrimPol	Repair	No	10^{-2}–10^{-4}	PrimPol	5L2X

Table 1.2 DNA polymerases families. DNA Polymerases can be grouped in different functional families, based on sequence homology. In this table only human members of polymerases families are given as examples. A-family includes polymerases involved in mitochondrial DNA replication (Polγ) and non-homologous end joining (NHEJ) repair pathways (as Polθ TMEJ: theta mediated end joining). B-family includes the eukaryotic polymerases involved in replication of genomic DNA, the primase Polα, and leading and lagging strand Pol ε and Polδ, as well as the TLS extender polymerase Polζ. X-family members, Polμ, Polβ, Polλ and TdT are involved in base excision repair (BER), NHEJ and short gap repair synthesis (sGRS).RT is reverse transcriptase, TERT is the catalytic subunit of Telomerase. PrimPol is a novel identified repair polymerase. Y-family members (Polη, Polι, Polκ and REV1) are mainly responsible for translesion synthesis and have lesion preferences. Polη and REV1 and also involved in somatic hypermutation (SHM). Error rates are indicated as frequency of wrongly incorporated nucleotides. Some examples of PDB structures are reported.

In high fidelity polymerases, during the polymerization reaction, the 3′-OH terminus of a primer DNA strand and the α-phosphate of a dNTP have to be placed adjacent to each other and correctly oriented for the phosphoryl transfer reaction. Upon the presence of a correct dNTP for the template–primer duplex within the catalytic core of high-fidelity enzymes, the finger domain undergoes a large conformational change and the active site becomes "closed."(Yang and Woodgate, 2007) In this closed structure, the 3′-OH and α-phosphate of dNTP are then aligned with the catalytically essential metal ions and carboxylates residues for the nucleophilic attack and phosphodiester bond formation. The fidelity of incorporation is thus accomplished mostly in two steps: a large conformational change of the finger domain and alignment of

nucleotides and catalysts in the active site. A wrong incoming dNTP and therefore a mismatched replicating base, inhibits both these highly selective steps (Rothwell et al., 2005).

Besides the conserved polymerase domains, DNA polymerases in general often have additional domains and regulatory subunits that have evolved to fulfil specific functions. Thus, most replicative polymerases contain a 3′-5′ exonuclease domain that proofreads newly synthesized DNA and corrects mismatched base pairs. In contrast, the majority of specialized polymerases lack the 3′-5′ exonuclease domain exhibiting less fidelity in coping templates DNA (Vaisman and Woodgate, 2017). The next paragraph will focus on specialized polymerases involved in TLS, especially Y-family members.

1.3.1 Y-Family DNA polymerases

DNA Polymerases which take part in translesion synthesis are members of the Y-family, respectively Pol eta (Polη) , Pol iota (Polι), Pol kappa (Polκ) and REV1, along with the B-family enzyme Pol zeta (Polζ). Y-family polymerases are comprised of a structured N-terminal catalytic domain and an unstructured C-terminal regulatory domain. The catalytic domain is similar to the classical polymerase structure, with the right-hand organisation of a palm, finger, and thumb domain. Unlike replicative DNA polymerases, Y- family enzymes have smaller thumb and finger domains and form less contacts with dNTP and DNA, resulting in more spacious active sites capable of accommodating distorted lesion-containing DNA substrates. A characteristic domain present only in Y-family polymerases is the little finger (LF) (Ling et al., 2001), or polymerase-associated domain (PAD) (Trincao et al., 2001).The LF domain is connected to the catalytic core by a flexible linker and mediates the contact of the polymerase with DNA. While the palm, thumb and finger domains of the Y-family polymerases are highly conserved, the LF domain is unique for each family member and plays an significant role in determining the biochemical properties and biological function of each individual polymerase (Boudsocq et al., 2004) (Wilson et al., 2013).

The active site of Y-family polymerases is pre-formed before substrate binding, in contrast to that of the A-, B-, and RT-family polymerases (Zhou et al., 2001) (Yang, 2003). Because of the small finger and thumb, the active site is also remarkably solvent-exposed and not as geometrically constrained to reject non-Watson–Crick base pairs, leading to a decreased selectivity for correct nucleotides. While the spaciousness of the active site makes erroneous base pairing possible, it allows TLS polymerases to accommodate bulky DNA lesions, which

would not fit in the narrow active site of replicative polymerases. The finger domain, which interacts with a nascent base pair, contributes to different lesion preferences and biased dNTP selection. Despite the lack of finger domain movement, the LF may be flexible and creates a structural gap capable of accommodating bulky aromatic adducts. The flexible LF domain and the spacious active site, which readily accepts damaged or mis-paired DNA, make the alignment of a dNTP, DNA and two metal ions for catalysis challenging and even more difficult in the presence of mismatched and damaged DNA. Consequently, the Y-family polymerases depend more on hydrogen bonds between a replicating base pair than its size or shape (Mizukami et al., 2006) (Johnson, 2010) and are catalytically less efficient than the A- and B-family enzymes (Beard et al., 2002). With no conformational selection or proofreading activity, the accuracy of Y-family polymerases is achieved by chemical selection alone and can worsen significantly (Yang and Gao, 2018). There is a considerable redundancy in TLS DNA polymerase functions and a given DNA lesion can be copied over by multiple TLS enzymes, albeit with different efficiencies and accuracies. Nevertheless, there is also a significant diversity in active site structures and detailed mechanisms of TLS polymerases, with different TLS enzymes specialized to bypass distinct types of DNA lesions, so called cognate lesions (table 1.1).

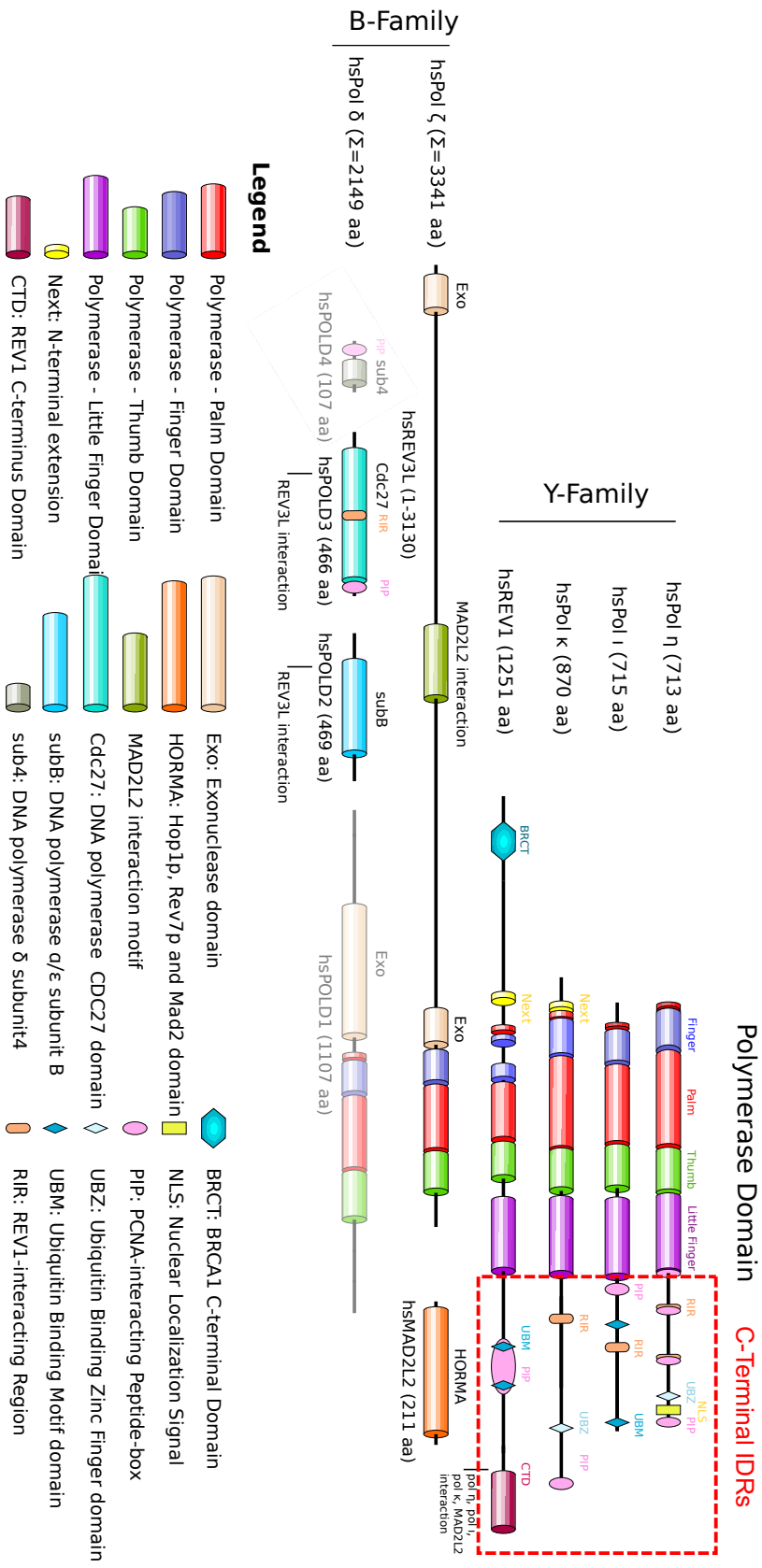


Figure 1.3 Domains organisation of translesion polymerases. DNA polymerases involved in TLS are mainly members of the Y-family (Polη, Polι, Polκ and REV1). Additionally, the B-family Polζ (REV3L and REV7) is depicted as it participates in TLS. For clarity also B-family replicative Polδ is depicted, and greyed out, as it can share POLD3 and POLD4 subunits with Polζ. In the scheme is depicted the domain organisation of polymerases.

Y-family members are characterized by finger, palm, thumb and little finger subdomains within the polymerase domain. All Y-family pols possess a intrinsically disordered regulatory C-terminus (IDR), which contains several motifs for interactions with other proteins (RIR: REV1 interacting region, PIP: PCNA interacting peptide, UBZ: ubiquitin binding zinc finger, UMB: ubiquitin binding motif, NLS: nuclear localisation signal). REV1 has a CTD which binds all other Y-family pols through their RIR and also the REV7 (MAD2L2) subunit of Polζ. Structurally ordered domains, depicted as coloured domains, accurate description can be found in the Legend. Unstructured parts are indicated by the black line.

The C-termini of Y-family polymerases are regulatory, flexible and largely unstructured region, spanning roughly half of the protein in length. This regulatory C-terminal region is an intrinsically disordered region (IDR), which contains motifs for protein-protein interaction. As depicted in Figure 1.3, several domains responsible for PCNA binding (PCNA interacting peptides boxes- PIP boxes), ubiquitin interactions (Ubiquitin binding zinc fingers -UBZ-, or motifs, UBM), as well as binding interfaces of other polymerases (F1,RIR, CDT), are located at the C-terminus of Y-family members. These C-terminal protein binding domains are involved in protein-protein interactions with important regulatory roles. The next paragraphs will focus on the two polymerases that were the focus of this study Polη and Polι in more detail, highlighting their specialised catalytic activities in lesion bypass.

1.3.2 DNA polymerase eta (Polη)

Polη was initially discovered in yeast as product of the RAD30 gene and identified as a bona fide DNA polymerase able to bypass thymidine-thymidine dimer efficiently (Johnson et al., 1999). One of the major lesions generated by exposure to sunlight is *cis–syn* cyclobutane dimer (CPDs). The loss of Polη results in a reduced efficiency to copy DNA containing these lesions (Johnson, 1999) (Masutani et al., 1999). In humans, Polη is encoded by the *POLH* gene, and mutations in the human *POLH* gene are associated to variant of Xeroderma Pigmentosum syndrome (XPV). A fraction of patients suffering of XP disease did not carry any mutation in NER XP genes, which are associated with the XP syndrome, but were shown to possess mutations in the *POLH* gene (Masutani et al., 1999; Johnson et al., 1999a). XPV patients account for almost 20% of XP cases and show UV-light sensitivity, and strong predisposition to skin cancer (Cleaver, 1972). XPV cells have normal rates of nucleotide excision repair, but have a defect in DNA synthesis, and are hypermutable, after UV-irradiation (Lehman et al., 1975) In the absence of Polη, mutations in XPV cells accumulate at CPD sites, as a result of error-prone bypass by another TLS polymerase, possibly Polι (Gueranger et al., 2008) (Stary et al., 2003) (Wang et al., 2007). Thus, error-free TLS by Polη plays a key role in preventing

mutation fixation at CPD sites, the most common UV-induced lesion, and reduces the incidence of UV-induced skin cancer (McCulloch et al., 2004) (Johnson et al., 1999). In addition to its main biological role in bypass of UV-induced CPDs, recent interest has focused on the role of Polη in bypass of lesions induced by cisplatin-based chemotherapeutic drugs, since damage tolerance by TLS may contribute to the resistance of tumours cells to chemotherapy. Two intrastrand purines crosslinked by cisplatin resemble a thymine dimer, and human Polη has been proven able to incorporate dCTPs opposite crosslinked dG bases quite efficiently (Masutani et al.) (Zhao et al., 2012). Full bypass of a cisplatin crosslinked lesion, however, requires Polζ to extend the primer beyond the damaged site (Shachar et al., 2009). In addition in being proficient in bypassing both UV and cisplatin-induced lesions, Polη also can carry out efficient and accurate replication past 8-oxoguanine (8-oxoG) lesion *in vitro* (Haracska et al., 2000). 8-oxoG results from exposure of cells to oxidative stress; by promoting error-free replication through the 8-oxoG lesion, Polη may contribute to reducing mutagenesis and carcinogenesis that could result from mutagenic bypass of this lesion by replicative DNA polymerases (Haracska et al., 2000).

The ability of Polη to replicate across bulky lesions arises from a relatively spacious active site, and structural information provided by X-ray crystallography on human Polη (Biertümpfel et al., 2010) (PDB: 3MR2) contributed to shed light on the catalytic mechanism. The structures solved by Biertümpfel and co-workers showed that human Polη contains four domains -palm, finger, thumb and little finger (LF)- with the active site in the palm domain and DNA bound between thumb and LF. In contrast to the structures of yeast Polη–cisplatin complexes (Alt et al., 2007), the finger domain of human Polη is closed and contacts the replicating base pair, and the LF domain interacts with both template and primer in the major groove extensively. The finger, palm and incoming dNTP in human Polη superimpose well with those in other TLS polymerases (the archaeal Dpo4, and human Polι and Polκ) Distinctively, in human Polη however, two bases of the template strand, instead of one, are in the active site and the 3′-T of the CPD is correctly base-paired with the incoming non-hydrolysable nucleotide (Biertümpfel et al., 2010). Translocation of two template bases into the active site was previously observed with Dpo4, however in this case, the 5′-T of the CPD paired with the dNTP, resulting in a misalignment (Ling et al.). In human Polη, a slight shift of LF relative to the finger domain enlarges the active site, and allows it to accommodate two template bases without misalignment. The thumb domain instead only makes contacts to the DNA primer (Biertümpfel et al., 2010).

The accurate bypass of *cis-syn* pyrimidine dimers Polη depends on two uniquely conserved residues (R61 and Q38 in humans) (Alt et al., 2007) (Masutani et al.). R61 adopts different rotamer conformations and is hydrogen bonded with the N7 of purines (A or G) or the phosphates, favouring the *anti*-conformation of dNTP and preventing Hoogsteen base pairing (Ling and Yang, 2003). Q38 is important because it is the sole contact of the CPD template base, which are not base stuck with the rest of the template to be kept in place and aligned with the incoming dNTP (Biertümpfel et al., 2010). In addition, these two residues also play a key role in the efficient mis-incorporation of dGTP opposite a dT in the context of undamaged DNA, particularly when the mismatched dT:dGTP replicating base pair is immediately preceded by an AT base pair, (Zhao et al., 2013) in the DNA sequence called the WA motif. The mis-incorporation of dGTP by human pol η at the WA motif in undamaged DNA contributes to the creation of variable sequences in the process of somatic hypermutation (Zeng et al., 2001) (Zeng et al., 2004). Importantly, human Polη was shown to keep the distorted CPD-containing DNA in a straight and rigid B-form conformation. This characteristic earned Polη the nickname of "molecular splint" (Biertümpfel et al., 2010). The molecular splint function of Polη is achieved by a continuous and highly positively charged DNA-binding surface that interacts extensively with the four template nucleotides immediately upstream of the active site. Specifically a β-strand in LF (amino acids from 316 to 324) is nearly parallel to the template strand, and each phosphate forms additional hydrogen bonds with side chains of several residues, mainly arginine, lysine, tyrosine and threonine . Additionally, the side chains of R93 and R111 extend from the palm domain to the template strand and further strengthen this molecular splint (Biertümpfel et al., 2010).

Apart from its intensively characterised N-terminally catalytic domain (Biertümpfel et al., 2010), human Polη contains a C-terminal flexible region, which spans from amino acids 440 to 713 and is intrinsically disordered, making structural studies very challenging. This IDR is the most interesting part of the protein, fundamental for Polη protein-protein interaction, regulation, post-translational modifications and cellular localisation (Figure 1.4). The C-terminus of Polη contains the nuclear localization signal (NLS, 682-694) important for compartmentalisation to the nucleus, as well as a number of structural features that play key roles in the recruitment of the DNA polymerase to sites of DNA damage. These include the non-canonical C-terminal PCNA-interacting peptide (PIP) domain (702–708) (Haracska et al., 2001), as well as another putative PIP box (432-444) identified right at the C-terminus of the catalytic domain (Freudenthal et al., 2008) (Despras et al., 2012). Another important region is the C_2H_2 ubiquitin-binding zinc finger domain (UBZ, 633-655) (Bienko, 2005) which mediates ubiquitin binding

and interaction with ubiquitinated PCNA. Additionally, other motifs have been shown to play a role in Polη interaction with other TLS. Notably, Polη interacts with the C-terminal domain (CTD) of REV1 via its REV1 interacting region (RIR, 476-491) (Yang and Woodgate, 2007) which partially overlaps with the F1 motifs for interaction with POLD2 subunit of Polδ (Baldeck et al., 2015).

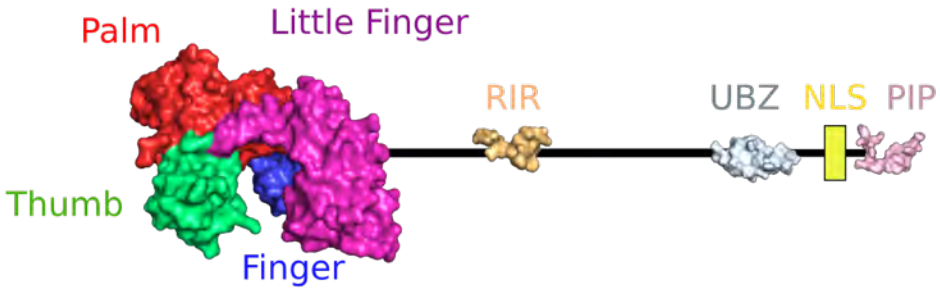


Figure1.4 Polη. Schematic structural representation of human Polη available structures, determined by either X-ray crystallography or NMR. Polymerase domain from aa 1-432 (PDB: 3MR2) is depicted as finger (blue) palm (red) thumb (green) and little finger (magenta) subdomains. REV1 interacting region is depicted in gold (RIR aa 524-539 PDB: 2LSK). Ubiquitin binding zinc finger (aa 628–662, UBZ PDB: 2I5O) is coloured in grey, and in pink is depicted the very C-terminus (aa 694–713) containing the PIP box (PDB: 2ZVK). NLS: nuclear localization signal. The figure was created in PyMOL and Inkscape.

1.3.3 DNA polymerase iota (Polι)

DNA polymerase iota (Polι), originally named as Rad30B, was discovered as the second human homolog of Rad30 protein (Polη) (McDonald et al.). The *POLI* gene is found in only higher eukaryotes, and is supposed to have arisen from a gene duplication from Polη, therefore Polι is considered a paralog of Polη. Interestingly, unlike other Y-family polymerases, homologs of Polι have not been identified in prokaryotes, yeast, nematodes and most plants (Vaisman et al., 2002) (McIntyre, 2020).

The role of Polι in TLS is poorly understood. Its activity largely overlaps with that of Polη, and likely functions as a backup or complementary polymerase. However, despite high sequence similarity, overexpressed Polι is not able to complement defects in human XPV cells caused by Polη loss (Kannouche et al., 2002). The architecture of the catalytic domain of Polι follows the Y-family polymerase classical scheme (Figure 1.3 and 1.5). The N-terminal catalytic core of

Polι contains two partly overlapping catalytic domains: one responsible for DNA polymerase activity and one that harbours 5′-deoxyribose phosphate-lyase (dRP-lyase) activity. The DNA polymerase activity domain, with the fingers, palm, thumb and little finger sub-domains, is located between amino acids 24-414 while the domain with dRP-lyase activity is located between amino acids 79-445. What differentiates Polι form other Y-family enzymes is its rather narrow active site, which stimulates formation of non-canonical interactions between nucleotides. The palm domain, containing catalytic residues D34, D126, and E127 located at the bottom of the DNA binding groove, as well as the fingers domain that positions over the template base and incoming nucleotide are critical for Polι activity (Nair et al., 2006) (Makarova et al., 2011). The thumb domain makes a few contacts with the minor groove of DNA, while the LF domain makes multiple contacts with the major groove of DNA (Makarova et al., 2011) (Makarova and Kulbachinskiy, 2012). In comparison to other Y-family polymerases, the little finger domain of Polι has greater mobility, which may be a reason for the low processivity of DNA synthesis by this polymerase (Nair et al., 2009). Increased mobility is caused by a smaller number of intramolecular contacts of the little finger with other domains, in particular with the finger domain.

Polι, like other Y-family polymerases, lacks a 3′-5′ exonuclease activity, and exhibits high rate of mis-incorporation, especially in presence of undamaged templates. Opposite template A, Polι is the most accurate and mis-incorporates A, G, or C with a frequency of 10^{-4} nucleotides, which is moderate for Y-family polymerases. Opposite template G and C, Polι is less accurate and mis-incorporates roughly one in 100-1000 nucleotides. Interestingly, the fidelity of Polι is lost opposite template T, where the polymerase exhibits 3-10-fold higher incorporation of incorrect G compared to correct A (Johnson et al., 2005).

The crystal structure, presented by the Prakash group, (PDB: 1T3N) of human Polι catalytic domain elucidated the mechanism behind this high template dependent infidelity (Nair et al., 2004). The structure showed that, unlike other Y-family members, Polι binds template bases in a *syn* conformation. This favours the formation of Hoogsteen base-pairing, rather than Watson-Crick, with the incoming nucleotide and explains the high fidelity and efficiency of Polι opposite template A and low fidelity and efficiency opposite template T (Nair et al., 2004). Remarkably, based on crystallographic analysis, it appears that upon dNTP binding, Polι does not undergo a conformational change, but instead, its active site and the incoming nucleotide, cause template A and G bases to convert from the *anti* to *syn* conformation. Remarkably,

Hoogsteen base-pairing provides Polι with the ability to incorporate nucleotides opposite certain DNA lesions (Johnson et al., 2005)

Against the most common UV lesion, *cis–syn* cyclobutane dimer (CPD), efficiently and accurately bypassed by Polη, Polι generally shows difficulties (Johnson et al., 2000). *In vitro* data (Tissier et al., 2000) showed that extended incubation times and excess amounts of enzyme allow for insertion of 1 or 2 nucleotides opposite the lesion by Polι, with mis-incorporation rates higher than that for Polη. The second most frequent UV lesion, 6-4 pyrimidine photoproducts (6-4PPs), although presenting a more pronounced structural distortion, is less problematic for Polι. In fact, Polι efficiently inserts all four types of nucleotides opposite the 3′-T of this lesion, with a preference for correct A residues (Johnson et al., 2005) (Tissier et al., 2000) Polι is also capable to bypass one of the most abundant oxidative lesions in DNA, 8-oxoguanine (8-oxoG). When bypassing 8-oxoG, Polι predominantly incorporates correct C or G, exhibiting the highest fidelity among all Y-family polymerases (Zhang et al.). While Polι can incorporate nucleotides opposite many lesions, it is generally incapable to extend resulting products, thus the bypass of these lesions often requires the combined action of a second extender polymerase, most likely Polζ or Polκ (Vaisman et al., 2002) (Washington et al., 2004).

Similarly to other TLS polymerases, the C-terminal half of Polι is predicted to be an intrinsically disordered region (IDR) playing a regulatory role and being involved in protein-protein interactions. The C-terminus is mostly disordered except for short motifs corresponding to protein binding domains (Ohmori et al., 2009). Polι interacts with PCNA through the single PIP box centrally located immediately downstream of the LF domain, spanning from amino acid 420 to 427 (Vidal et al., 2004) (Haracska et al., 2005). The interaction through the PIP box of Polι, is rather weak, as it is for other Y-family polymerases, and PCNA-Polι interaction can be significantly strengthened through PCNA mono-ubiquitination, which forms an additional surface for interaction. Polι has two functional ubiquitin binding motifs, UBM1 (496-524) and UBM2 (681-709) (Bienko, 2005) (Plosky et al., 2006). Interestingly, while most known UBDs require the presence of conserved I44 in the hydrophobic patch of ubiquitin for binding, Polι instead requires the unorthodox interaction with L8 (Bomar et al., 2010). Additionally, Polι has a REV1-interacting region (RIR, 539-554 aa) required for the interaction with the scaffold TLS polymerase REV1, through its CTD (Ohashi et al., 2009).

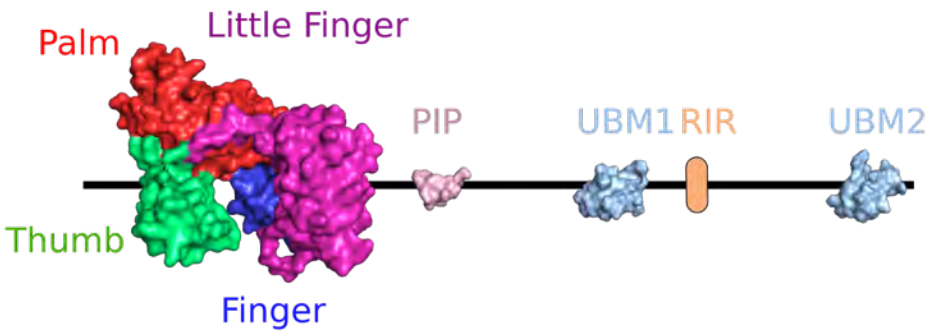


Figure1.5 Polι. Schematic structural representation of human Polι available structures, determined by either X-ray crystallography or NMR. Polymerase domain aa 27–414 (PDB: 1T3N) is depicted as finger (blue) palm (red) thumb (green) and little finger (magenta) subdomains. The PIP box containing region (419–434 PDB: 2ZVM) is depicted in pink. Ubiquitin binding motifs are depicted in light blue (aa 491–530 for UBM1 PDB: 2MBB) (aa 674–715 for UBM2 PDB: 2KHU). The figure was assembled in PyMOL and Inkscape.

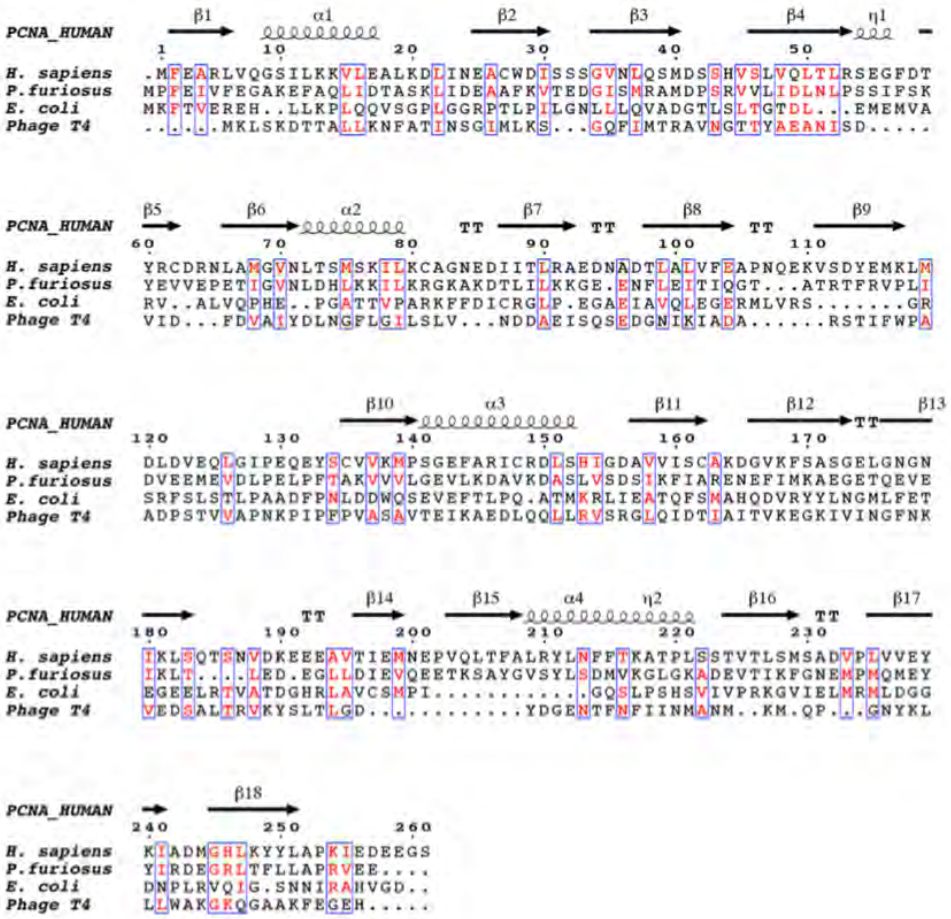
1.4 PCNA structure and function: sliding clamp roles in DNA metabolism

1.4.1 PCNA structure

DNA sliding clamps are essential proteins found in all three domains of life, from archaea to eukaryotes, functionally and structurally similar in all living organisms, including some viruses. They can assemble in homodimeric, homotrimeric or heterotrimeric rings. Despite little sequence homology, these proteins show similar overall structures and architectures (as depicted in Figure 1.6 a.) they adopt a pseudo-six fold symmetry structure resulting in 'donut'-shaped proteins that encircle double- stranded DNA and are capable of diffusing along it. DNA sliding clamps are important mobile platforms on the DNA and recruit many of the enzymes involved in DNA replication, repair, and recombination such as polymerases, ligases, exonucleases etc, conferring to each of these factors higher DNA affinity and processivity. In eukaryotes, the DNA sliding clamp is a homotrimer formed by three identical monomers, each of them composed of a C-terminal and N-terminal domain, connected by the intradomain connecting loop (IDCL). The outer layer is composed of 6 β-sheets and the inner layer is composed of 12 α-helices facing the central channel. In this section, the human Proliferating Cell Nuclear Antigen (PCNA) and its role in DNA replication and DNA damage tolerance will

be discussed. PCNA has such a crucial role in DNA replication and repair that has been described as the "maestro" of the replication fork (Moldovan et al., 2007).

a.



b.

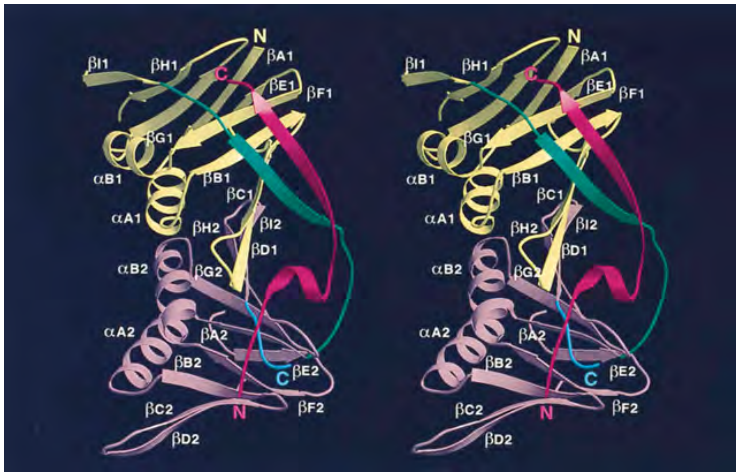


Figure 1.6 PCNA alignments and first crystal structure. (a) Structure-based sequence alignment of DNA sliding clamps from *H. sapiens, P. furiosus, E. coli*, and the gp45 gene of bacteriophage T4. Similar residues are coloured in red. Secondary structure elements corresponding to human proliferating cell nuclear antigen are shown above the alignment, β-strands are indicated as arrows, α-helices as spirals, and β-turns as (PCNA), β-strands are indicated as arrows, α-helices as spirals, and β-turns as TT. (González-Magaña and Blanco, 2020) (b) Structure of one monomer of the PCNA–p21 peptide complex, looking directly onto the face of the ring. The peptide (colored red) runs alongside the interdomain connecting loop (green). The N-terminal domain of PCNA is drawn in yellow and the C-terminal domain in plum, with five residues at the terminus shaded blue (PDB: 1AXC) (Gulbis et al.)

Human PCNA was discovered by two separate groups: in 1978, Miyachi and colleagues identified an auto-antigen in the sera of patients with systemic lupus erythematosus (Miyachi et al., 1978). Given that the protein was detected in the nuclei of dividing cells, it was later named Proliferating Cell Nuclear Antigen (PCNA) (Kelman, 1997). Few years later, another group (Bravo and Celis, 1980) identified a protein which was synthesized during the S-phase cell cycle and called it cyclin. Further experiments showed that both were the same protein. Human PCNA (NP_002583) is a 261 amino acids protein, with a monomeric molecular weight of ~29 kDa. In aqueous solutions it forms stable trimeric assembly where each monomer is connected head to tail to the next one, resulting in a ring-shaped protein with a pseudo hexameric fold symmetry and a molecular weight of *circa* 86 kDa.

The 86 kDa PCNA trimer is opened and loaded around DNA by the pentameric complex, RFC (Majka and Burgers, 2004). RFC is an arc-shaped complex of five similarly structured essential proteins (AAA+ type ATPases) and associates with PCNA like a screw cap with a bottle. RFC specifically recognizes template-primer 3′ ends and loads PCNA to these sites. ATP binding is required for the formation of a stable PCNA-RFC complex and for its loading to primed DNA (Moldovan et al., 2007). DNA binding in turn activates the ATP hydrolysis activity of RFC, apparently leading to its dissociation from the loaded clamp (Gomes and Burgers, 2001). RFC binds to the so-called C-side of PCNA (termed so because the C-termini of the PCNA monomers protrude from this face, also referred as front face - Figure 1.7 b.) and loads it with this side positioned toward the 3′ end of the elongating DNA. This ensures that polymerases, which also bind to the front of PCNA at the level of the interdomain connecting loop, are oriented toward the growing end.

PCNA is an ancillary factor of replicative Polδ and Polε, increasing their processivity by tethering them and sliding along the double helix. Particularly important is its role in the synthesis of the lagging strand, where PCNA acts as a platform where Pol δ, flap endonuclease 1 (FEN1), and DNA ligase I (LIG1) bind to synthesize, process, and join Okazaki fragments (Slade, 2018). 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 (Slade, 2018). Overall, in the last decade, more than 80 different binding partners of PCNA have been identified, all playing a role in cell cycle regulation, DNA metabolism and maintenance of genome integrity (McNally et al., 2010) (Boehm et al., 2016a). Not all PCNA partners can bind simultaneously and switching may be triggered by different mechanisms: affinity-competition, proteolysis, or post-translational

modifications. The underlying mechanisms of recruitment and coordination of different binding partners has been the focus of several investigations in the last decades but still remains obscure.

The first structural information about human PCNA were obtained by X-ray crystallography in 1996, when Gulbis and colleagues solved the structure of hsPCNA bound to the C-terminal region of the cell-cycle checkpoint protein $p21^{WAF1/CIP1}$ (PDB: 1AXC) (Gulbis et al.). Subsequent nuclear magnetic resonance (NMR) studies (Sánchez et al., 2007) confirmed that the trimeric structure exists in solution, and that it is not an artefact of crystallization. Moreover, size exclusion chromatography coupled to multiangle light scattering (SEC-MALS) measurements on human PCNA and the yeast homologue showed that the trimeric form is predominant even at high concentrations, confirming that PCNA behaves as one single trimer in solution (De Biasio et al., 2011).

Two mutations were identified in yeast PCNA that hamper translesion synthesis in cells (Zhang et al., 2006). These substitutions (G178S and E113G) are of residues in the β-strands that constitute the subunit interface, and the X-ray crystal structures of both mutant proteins reveal perturbations that reduce the number of hydrogen bonds between these strands (Freudenthal et al., 2008) (Freudenthal et al., 2009), leading to trimer instability. The mutant PCNA trimers are less stable than wt PCNA, and they are unable to stimulate the catalytic activity of non-classical polymerases (Dieckman and Washington, 2013). There are two potential explanations for the inability of these mutant proteins to stimulate these polymerases. Firstly, decreased PCNA trimer stability might lead to transient opening and closing of the ring that disrupts the activity of non-classical polymerases to a greater extent than it does other PCNA-interacting enzymes. Second, the PCNA subunit interface may serve as an additional point of contact for non-classical polymerases. Further investigation is necessary to distinguish between these possible scenarios.

Gulbis and co-workers (Gulbis et al.) describe the structure of human PCNA as following: a set of three molecules of PCNA, each consisting of two similar domains, is assembled into a circular ring with a central hole wide enough to encircle duplex DNA. The central pore has an internal diameter of approximately 35 Å, large enough to accommodate the 24 Å wide double helix of B-form DNA. The overall diameter of the trimer is approximately 80 Å. As can be noticed in Figure 1.6 b. antiparallel interactions between strands βD2 and βI1 in adjacent PCNA subunits are the key elements in the intermolecular interface and lead to the formation of a total of three nine-stranded β sheets that form a contiguous surface across each intermolecular

boundary (Gulbis et al.). Likewise, antiparallel interactions between strands βD1 and I2 within one molecule result in the formation of three similar β-sheets across the interdomain boundaries within one molecule. Together, these six curved β-sheets form an outer scaffold that supports an inner lining of 12 α-helices that are oriented approximately perpendicular to the ring face (Gulbis et al.). There are two distinct faces of the PCNA ring: the front face, also known as the C-face (since the carboxy-terminal ends of the protomers are located there), contains a hydrophobic pocket next to the IDCL (the interdomain connecting loop, βI1–βA2 which connects the N-terminal and the C-terminal domain). The front face is the side where most of PCNA partners bind.

The back face has number of prominent loops connecting adjacent antiparallel β-strands protrude into solvent (in particular βD2–βE2 and βH1–βI1), and is the target for post-translational modifications that alter the clamp's properties (Kelch, 2016) (Gulbis et al.). Calculation of the electrostatic potential of human PCNA shows that the surface of the trimer has a highly negative potential, except in the central pore where a region of positive potential is created by several lysine and arginine side chains (Gulbis et al.). The positive charge in the central region is thought to be important for stabilizing the PCNA trimer on the negatively charged phosphate backbone of DNA (Krishna and Gary).

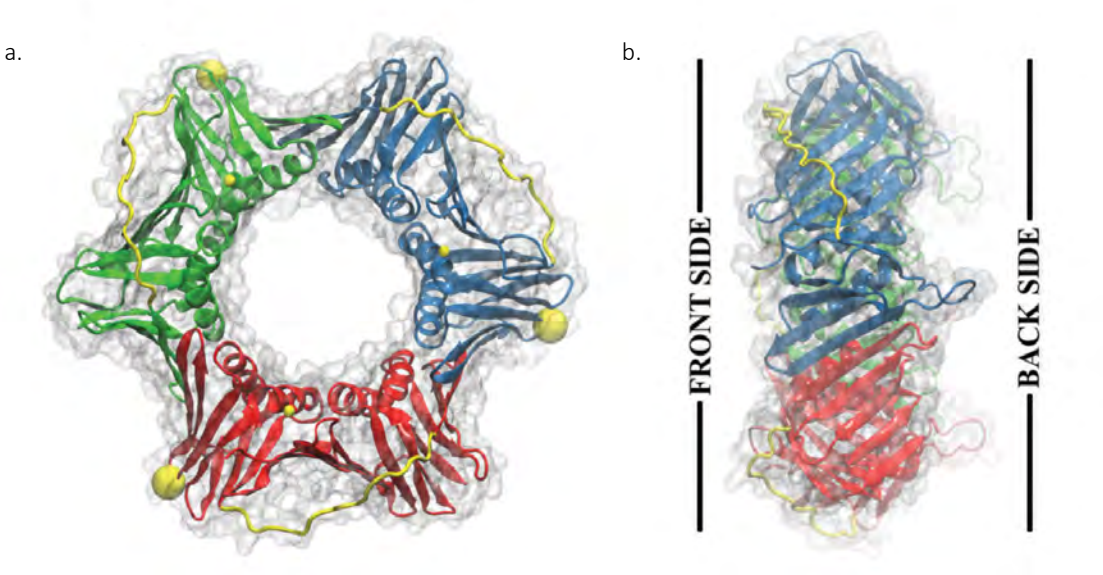


Figure 1.7 PCNA structure and features. (a) Front view of the three-dimensional structure of human PCNA, different subunits are coloured red, green and blue, interdomain connecting loop (IDCL) is coloured yellow, and important amino acid residues are designated with yellow balls (big yellow balls indicate position of K164). (b) Side view of human PCNA showing a smooth front side and visible protrusions in the back side. Adapted from (Stoimenov and Helleday, 2009).

1.4.2 PCNA sliding on DNA

Some decades have passed since a crystal structure of human PCNA bound to DNA was obtained. 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 along DNA (Armstrong et al., 2012). However, the lack of sequence specificity and the weakness of PCNA-DNA interactions made this task difficult. In 2017, the De Biasio group published a crystal structure of human PCNA bound to a 10 bp DNA duplex, showing DNA inside the channel, tilted by an angle of 15 degrees relative to the three-fold rotation axis of the ring (De March et al., 2017). The PCNA-DNA interface comprises six conserved basic residues (K20, K77, K80 R149, H153 and K217) spread along five α-helices of two protomers that establish polar contacts with five consecutive phosphates of one of the DNA strands. In order to understand the PCNA sliding mechanism on DNA, molecular dynamics simulations were combined with available structural data (De March and De Biasio, 2017). Molecular dynamics simulations presented a more evident tilting of the DNA, at approximately 30°, as compared to the crystal structure and a larger portion of DNA interacting simultaneously with two protomers of the trimer. Nevertheless, the basic residues within the PCNA channel made transient interactions with the phosphates of DNA, and switched between adjacent phosphates in a non-coordinated manner. The data suggest a "cogwheel diffusion" mechanism for PCNA sliding along the DNA: when a sufficient number of polar contacts are made with adjacent phosphate groups of DNA, the ring rotates on the DNA, resulting in advancement by one base pair. Therefore, the conserved patch of basic residues on the inner wall enables PCNA to rotationally track the helical pitch by spiral diffusion, which is named "cogwheel diffusion". This process maintains DNA-protein contacts and keeps the clamp orientation invariant relative to the DNA (Yao and O'Donnell, 2017). These results are consistent with previous single molecule studies, which showed two modes of PCNA sliding along DNA. In one mode, the clamp tracks the helical pitch of the DNA duplex, resulting in a rotation of the protein around the DNA (cogwheel diffusion). In the second mode, which is less common, the protein translationally diffuses at higher rates making no contacts with the DNA (Kochaniak et al., 2009). The major sliding mechanism, cogwheel sliding, confers the PCNA ring, a tilted orientation with respect to the DNA, which is supposed to be important for productive encounters with polymerases and enzymes involved in Okazaki fragments maturation.(González-Magaña and Blanco, 2020) The mechanisms of PCNA diffusion along

DNA during TLS still remain obscure and further investigation is needed to elucidate whether the primary cogwheel-sliding mode is maintained when PCNA is bound to different partners.

1.4.3 Post translational modifications of PCNA

Being such a central player in DNA replication, cell cycle regulation and DNA repair, coordinating a huge number of enzymes working on DNA, PCNA is subjected to strict regulation, mainly by post translational modifications (PTMs). Many post-translational modifications of human PCNA have been described, but the functional significance of many of them remain to be elucidated (Leung et al., 2018). One of the most studied post-translational modifications of PCNA is ubiquitination. First described in yeast and later found in mammalian cells, ubiquitination is targeted on highly conserved lysine residues. Of particular biological importance for translesion synthesis is ubiquitination, occurring on the conserved lysine residue in position 164 (Figure 1.7 a. big yellow ball). Two distinct types of ubiquitination could occur: attachment of a single ubiquitin moiety (mono-ubiquitination) and building of a polyubiquitin chain via an 'unusual' K63 inter-ubiquitin linkage (K63 poly-ubiquitination) (Hoege et al., 2002). The nature of the ubiquitin modification determines which damage tolerance pathway will be initiated, with mono-ubiquitination leading to translesion synthesis and poly-ubiquitination leading to template switching (Hoege et al., 2002) (Stelter and Ulrich, 2003) (Branzei et al., 2004).

PCNA mono-ubiquitination is achieved by classical E1-E2-E3 ubiquitination cascade as consequent action of an ubiquitin-activating enzyme E1 (UbaI in humans), specific ubiquitin-conjugation enzyme E2 (in humans either Rad6A or Rad6B) and E3 ubiquitin ligase. Rad18 is the principal E3 ubiquitin ligase that mono-ubiquitinates PCNA in human, but several laboratories demonstrated the presence of residual mono-ubiquitinated PCNA in Rad18-deficient cells (Arakawa et al., 2006) (Huang et al., 2006), indicating the existence of alternative E3 ligases. RNF8 (RING finger protein 8) and CRL4(Cdt2) (Cullin-4-RING ligase (CRL4)-Ddb1-Cdt2) were later identified as two E3 ligases that can catalyse the mono-ubiquitination of PCNA, although their contribution is considered rather minor, compared to Rad18. RNF8 in concert with the E2 conjugating enzyme, UbcH5c, was shown to mono-ubiquitinate PCNA in response to UV radiation (Zhang et al., 2008). Additionally UbcH5c was described to catalyse *in vitro* mono-ubiquitination on K164 by the Sixma lab (Hibbert and Sixma, 2012) and in this work. Mono-ubiquitination of PCNA by the Rad6/Rad1 E2-E3 system is a process that can be

reversed by the deubiquitinating enzyme ubiquitin-specific protease 1 (USP1), which catalyses the de-ubiquitination of K164 residue, once the DNA lesion is by-passed by specialized TLS polymerases. PCNA mono-ubiquitination on K164 is a crucial modification, which results in a new surface property of PCNA and induces a change in the interaction partners. This is a key concept in the proposed polymerase-switch model during translesion synthesis.

PCNA K164 polyubiquitination via K63 chains requires the heterodimeric ubiquitin-conjugation enzyme Ubc13–Mms2 and a specific RING-finger-containing E3 ubiquitin ligase. In humans, three different enzymes are shown to polyubiquitylate PCNA on position 164 by forming K63 linkages, SHPRH (SNF2 histone linker PHD RING helicase), HLTF (helicase-like transcription factor) and RNF8 (RING finger protein 8), at least *in vitro* (Motegi et al., 2006) (Zhang et al., 2008). Moreover, a quantitative proteomics study reported the mono-ubiquitination of K117 in human cells, although the exact function of this modification is still not fully understood (Elia et al., 2015)

Other post translational modifications of PCNA play a key role in its regulation. Another important modification occurring on lysine is SUMOylation. SUMO stands for Small Ubiquitin-like Modifier (or SUMO). These proteins are similar to ubiquitin and are members of the ubiquitin-like protein family. SUMOylation of PCNA can occur at several lysine sites, and SUMOylation (predominantly with SUMO1) at K164, has been shown to prevent replication fork collapse at double stranded breaks (Arakawa et al., 2006). Another SUMOylation site was described in human PCNA at the level of K254, although its role has not yet been addressed (Gali et al., 2012). ISGylation and NEDDYlation are other type of modifications which use ubiquitin-like proteins: respectively ISG15 and NEDD8 ubiquitin-like modifiers. The ISGylation of PCNA was shown to occur on both K164 and K168, playing a role in the termination of error-prone TLS, to prevent excessive mutagenesis (Park et al., 2014). Similarly, NEDDYlation at K164 was proposed to antagonize ubiquitination, therefore regulating the recruitment of Polη in response to oxidative DNA damage (Guan et al., 2017). As different modifications on the same residue are mutually exclusive, the dynamic switching between several post translational modified PCNA states is an efficient way of regulating PCNA interactions with other proteins. Additionally, other types of modifications like the attachment of methyl, acetyl or phosphoryl groups, can regulate PCNA function. The methylation on K248 by Histone Lysine Methyltransferase SETD8 stabilizes PCNA protein levels and enhances PCNA's interaction with FEN1 (Takawa et al., 2012). Studies in

mammalian cells show that PCNA is also di-methylated at K110 by the EH2Z enzyme, which appears to stabilize the trimeric form of PCNA (A et al., 2018). The acetylation of PCNA can occur at several lysine residues (K13, K14, K77, and K80) by CREB-binding protein (CBP), and less efficiently by p300, which promotes the removal of chromatin-bound PCNA and its proteasomal degradation during nucleotide excision repair synthesis (Cazzalini et al., 2014). Several phosphorylation sites have been identified on PCNA (Slade, 2018). Y211-phosphorylation increases the half-life of chromatin-bound PCNA and facilitates cell proliferation (Wang et al., 2006). Y211-phosphorylated PCNA facilitates error-prone DNA replication and the suppression of the mismatch repair mechanism, which enhances cancer development and progression (Peng et al., 2019) (Ortega et al., 2015).

1.4.4 Structural studies on mono-UbPCNA

An efficient method to produce sufficient quantities of mono-ubiquitinated PCNA has been for many years challenging and has hindered the possibility to perform structural and biochemical studies on mono-UbPCNA (UbPCNA for simplicity). In 2010, Freudenthal and colleagues (Freudenthal, 2010) were able to produce large quantities of the yeast UbPCNA by splitting the protein into two polypeptides. Specifically, yeast PCNA was split into two fragments between residues 163 and 164, and a ubiquitin molecule was fused via its C-terminus and a linker (two glycine, for total of four glycine residues) to the N-terminus of the second fragment of the split PCNA (amino acids 164-261). This Ub-PCNA fusion construct was co-expressed together with the N-terminal fragment of PCNA (1-163) allowing the two polypeptides to self-assemble *in vivo*. With this method they were able to obtain the first crystal structure of "split-UbPCNA". In their structure, Freudenthal and co-workers observe that the ubiquitin molecule is located on the back of PCNA, not covering the hydrophobic pocket on the front. Ubiquitin moiety uses its canonical hydrophobic surface to interact specifically but also interacts weakly via electrostatic and hydrogen-bond interactions. Furthermore, they proved that attachment of ubiquitin does not alter the conformation of PCNA, suggesting that there is no or minimal conformational change of PCNA upon ubiquitin conjugation (Freudenthal, 2010).

A year later, Tsutakawa and co-workers performed in-solution structural studies to compare the same yeast split-UbPCNA and a chemically crosslinked version of UbPCNA, obtained using cysteine crosslinking with a K164C mutant (Tsutakawa et al., 2011). They used a combination of small angle X-ray scattering (SAXS) and multiscale molecular modelling to describe the behaviour of yeast UbPCNA in solution (Tsutakawa et al., 2011). The small angle X-ray

scattering (SAXS) analysis provided a predicted shape, in which a single ubiquitin was shown projecting away from the PCNA surface, indicating that ubiquitin did not only occupy the position observed in the crystal structure described by Freudentahl and colleagues, but also can adopt other positions. In the same study, computational modelling of UbPCNA using tethered Brownian dynamics simulation was used to generate a large number of conformations of a ubiquitin molecule tethered to a PCNA trimer. This showed a range of favoured conformations that lie to the side of PCNA in the equatorial plane up to the front face of PCNA, indicating high flexibility of ubiquitin moieties with respect to PCNA. In both the split-UbPCNA and the cross-linked UbPCNA SAXS studies, the ubiquitin was about 25 to 30 % in the crystallographic position (back side), about 40 to 50% in the computationally determined positions, and about 25% to 30% flexible in solution. This SAXS analysis supported a model of segmental flexibility, where ubiquitin can occupy, relative to PCNA, both discrete and flexible positions, moving from one to the other in dynamic equilibrium (Tsutakawa et al., 2011). The same group (Tsutakawa et al., 2015) a few years later performed additional SAXS studies on chemically crosslinked UbPCNA which revealed several different conformations of ubiquitin on the sliding clamp. Among the many possible conformation there is one with ubiquitin extending away from PCNA, as well as a ‘‘docked’’ conformation, in which ubiquitin interacts extensively with the PCNA surface, forming a contiguous binding interface. It is possible that these different conformations allow ubiquitinated PCNA the necessary plasticity to interact with different binding partners.

Structural information on human mono-ubiquitinated PCNA was obtained for the first time in 2012 when, in parallel, two groups published crystal structure (Zhang et al., 2012) and in solution studies (Hibbert and Sixma, 2012) of native human mono-ubiquitinated PCNA, obtained with enzymatic method. Zhang and co-workers (Zhang et al., 2012) prepared UbPCNA by utilizing E3 ubiquitin ligase RNF8 and the E2 ubiquitin-conjugating enzyme UbcH5, which were previously shown to catalyse K164 mono-ubiquitination of PCNA (Zhang et al., 2008). They obtained a crystal structure of human native UbPCNA in high ionic salt, where only two ubiquitin molecules were visible of PCNA trimer, due to crystal packing. The two ubiquitin are projected radially away from the trimeric ring of PCNA at an angle of about 30° to the equatorial plane, toward the back face of PCNA. In this crystal structure, there are no other contacts between ubiquitin and PCNA aside from the isopeptide linkage. Additionally, Zhang and co-workers showed that ubiquitin has negatively and positively charged surfaces located on opposite sides of the molecule, oriented almost perpendicular to the plane of the PCNA trimer. The acidic face is adjacent to an acidic patch on PCNA, resulting in the formation

of a continuous acidic surface, and the basic face is similarly adjacent to a positively charged patch on PCNA (Zhang et al., 2012).

In-solution studies on native human UbPCNA, obtained in an analogous way by *in vitro* reaction, were carried out by the Sixma lab (Hibbert and Sixma, 2012). They described for the first time, high yield production of native UbPCNA using an enzymatic method where the sole E2 ubiquitin conjugating enzyme UbcH5c was surprisingly ubiquitinating PCNA, even in the absence of an E3 enzyme. They showed good mono-ubiquitination reaction yield in presence of wt UbcH5c and showed that a S22R point mutant, which does not form ubiquitin chains (Brzovic et al., 2006), was even more efficient in mono-ubiquitinating PCNA on K164. This native *in vitro* mono-ubiquitinated PCNA was subjected to small-angle X-ray scattering (SAXS), size exclusion chromatography coupled multiangle light scattering (SEC-MALS) and nuclear magnetic resonance (NMR) studies, which revealed intrinsic flexibility of the protein. Hibbert and Sixma, considering the radii of gyration and molecular weight obtained from SEC-MALS and SAXS, describe four possible molecular models of UbPCNA. Each model proposed showed the ubiquitin moieties in different orientations: one orientation fits the crystal structure of split PCNA (Freudenthal, 2010) with ubiquitin on the back side of PCNA, another model shows a closed conformation where all ubiquitins sit in the groove at the PCNA subunit interface, a third model has an extended structure with ubiquitin completely extended outside and the last model shows an intermediate structure where ubiquitin is not completely extended. The final two models represent flexible states because they show no significant contacts between ubiquitin and PCNA to stabilize the conformations. In this work, both the split and closed UbPCNA models were in poor agreement with the experimental data, whereas structures that represent flexibility (third and fourth model) fit the data considerably better. This was interpreted as a symptom of ubiquitin flexibility with respect to PCNA, which was furthermore confirmed by NMR studies. Traditional NMR is sensitive to the rate of motion and hence the rotational correlation time of residues in a protein: large proteins have broad NMR peaks. Surprisingly, and despite UbPCNA being a large protein for NMR application (112 kDa), the peaks in the spectra are sharp (Hibbert and Sixma, 2012). This is a direct evidence that ubiquitin is much more mobile than would be expected if there was a stable interaction between the PCNA and ubiquitin: the ubiquitin is able to move independently of PCNA and hence has sharp peaks in the NMR spectra described by (Hibbert and Sixma, 2012). In conclusion, all these structural studies on UbPCNA highlight the great flexibility of ubiquitin with respect to PCNA and the existence of several conformations of UbPCNA in solution, pointing to a physiologically relevant role of flexibility for recruitment of TLS polymerases.

1.4.5 PIP boxes: canonical and non-canonical binding to PCNA

The hydrophobic pocket near the intradomain connecting loop (IDCL), on the front face of PCNA, is the region where polymerases and other proteins bind via their PCNA interacting peptides (PIP box). Most interacting partners bind PCNA through a conserved sequence named the PIP (PCNA-interacting peptide) box. The consensus PIP box sequence is Q-x-x-Ψ-x-x-ϑ-ϑ, in which Ψ is a moderately hydrophobic amino acid (L, V, I, or M), ϑ is an aromatic residue (Y or F) and x can be any residue. The crystal structure of PCNA bound to a consensus PIP box from the cell-cycle checkpoint protein p21 described by Gulbis and colleagues (section 1.4.1) reveals how PCNA recognizes consensus PIP motifs (Gulbis et al.). p21 PIP box is a canonical PIP box, composed of ^{144}QTSMTDFY^{151}. The consensus glutamine in position 1 (Q144) binds in a small pocket on the front face of the PCNA ring called the "Q-pocket." Moreover, residues from position 4 to 8 (M147, T148, D149, F150, Y151) form a short 3_{10} helix in which the consensus methionine in position 4 (M147), the consensus phenylalanine residue in position 7 (F150), and the consensus tyrosine in position 8 (Y151) insert into a large hydrophobic pocket on the front of PCNA called the "three-forked plug" because of its resemblance to an electrical outlet. In addition, p21 PIP box shows residues flanking its C-terminal side (^{152}HSKRRLIFS^{160}), which form an extended anti-parallel β-sheet with residues of the extended interdomain connector loop of PCNA. Besides the presence of consensus PIP boxes among several PCNA interacting proteins, many proteins have degenerate sequences missing some of the core amino acids.

In the last decades, several crystal structures of PCNA bound to non-consensus PIP boxes have been determined (Hishiki et al., 2009a) (Armstrong et al., 2012). These structures show how PCNA accommodates PIP boxes that diverge from the consensus sequence. For example, the specialised Polη has a non-canonical C-terminal PIP box that presents the unusual sequence ^{701}MQTLESFF^{708} at the very C-terminus of the protein. An evident difference compared to canonical PIP-sequences is that Polη PIP box lacks a glutamine in position 1, substituted by a non- consensus methionine that can still be accommodated in the Q-pocket on the front of PCNA. Likewise, positions 4 through 8 still retain the typical 3_{10} helix and the consensus residues in positions 4, 7, and 8 still insert into the three-forked plug.

In this work, I identified by sequence comparison and similarity three additional putative Polη PIP boxes (see table 1.3) besides the very C-terminal one (701-708) here referred to as PIP box 3, all spanning at Polη C-terminus (see Figure 1.8). Their sequence is diverging form the non-

canonical PIP box at 701-708 and they partially overlap with REV1 interaction region (RIR motif) or POLD2 subunit binding motif (F1 motif) (see table 1.3) (Baldeck et al., 2015). These sequences can be identified as part of a large group of "PIP-like motifs" which are gaining more and more importance in characterising PCNA binding mechanisms (Boehm and Washington, 2016) as the definition of PIP box has been recently revisited. A more exhaustive inquiry about the broader role of PIP boxes in TLS polymerases will be reviewed in the following paragraphs.

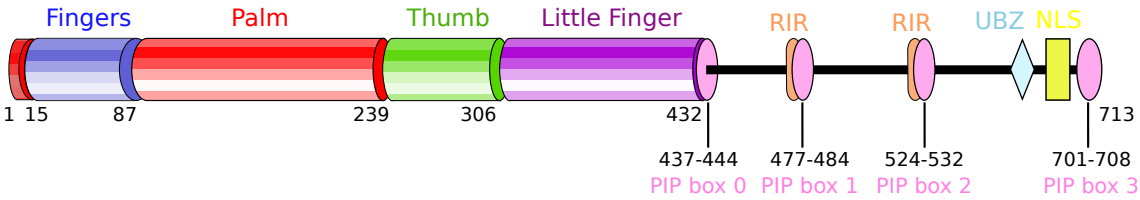


Figure 1.8 Polη domain organisation. Domain organisation of human Polη and putative PIP boxes described in this study. N-terminal polymerase domain composed of finger aa 15-87 (blue), palm aa 87-239 (red), thumb aa 239-306 (green) and little finger aa 306-432 (magenta) subdomains. Putative PIP boxes at the disordered C-terminus are depicted as pink ovals, respectively PIP box 0 (aa 437-444), PIP box 1(aa 477-484) partially overlapping with RIR motifs and F1 motif, PIP box 2 (aa 524-532) partially overlapping with RIR motifs, and PIP box 3 (aa 701-708). Additionally the UBZ domain is depicted in light blue and NLS in yellow.

Name in this work	Position	Sequence alignment	Motif references
PIP box 0	(437-444)	FLCATKFSASAPSSSTDITSFLSSDPSSLPK	This work
PIP box 1	(477-484)	GSGPAVTATKKATTSLESFFQKAAERQKVK	F1 motif (Baldeck et al., 2015) RIR motif
PIP box 2	(524-532)	NSPSKPSLPFQTSQSTGTEPFFKQKSLLLKQ	RIR motif (Boehm and Washington, 2016)
PIP box 3	(701-708)	LACTNKRPRPEGMQTLESFFKPLTH	(Hishiki et al., 2009a)

Table 1.3 PIP boxes identified in this study. Putative non-canonical PIP boxes identified at the disorder C-terminus of human Polη. Names refer to the one used in this work, sequence alignments shows the typical sequence of the non-canonical Polη PIP box (underlined sequence). Specifically, Q in position 1 is highlighted in blue, the moderately hydrophobic residues in the middle of the sequence (L/I/T) are in grey, the two aromatic residues (FF) or in case of PIP box 0 (FL) are in green. F1 and RIR motif additional residues are dotted-underlined.

1.4.6 Revisiting the concept of PIP box

Most of the proteins that interact with PCNA either are IDPs (intrinsically disordered proteins) or have IDRs (intrinsically disordered regions). IDPs lack defined secondary and tertiary structures under physiological conditions and their frequency of occurrence highly correlates with the complexity of organisms -being most abundant in eukaryotes- and regulation of several cellular pathways. Numerous evidences reveal a correlation between complex cell regulation

and the greater presence of IDPs/IDRs (Cordeiro et al., 2017). Moreover, most of the proteins associated with cancer have been identified as IDP or IDR-containing proteins (79%), which stresses the crucial roles they play in several cellular events that are altered in cancer, such as cell proliferation, DNA repair, and apoptosis (Iakoucheva et al., 2002). The evolutionary benefit of IDPs probably lies in their plasticity: flexibility allows IDPs or IDRs to interact with many ligands. Disordered proteins have large accessible surface areas, which increase their ability to interact with diverse binding partners through Short Linear Motifs (SLiMs) (Prestel et al., 2019). Many of the proteins that interact with PCNA display a characteristic SLiM known as PIP box (which was deeply described in section 1.4.5).

Besides the well described canonical PIP box sequences, present in many PCNA binding partners, several non-canonical PIP sequences have been reported in Y-family of DNA polymerases (Pol η, Pol ι, and Pol κ) and have been structurally characterized (Hishiki et al., 2009a). An interesting case is the non-canonical PIP box of Polι. The sequence ^{421}KGLIDYYL^{428} is located in the intrinsically disordered C-terminus and was structurally and thermodynamically described by Hishiki and colleagues. The consensus glutamine in position 1 is substituted by a lysine that, surprisingly, does not accommodate in the Q-pocket. Moreover, the PIP motif does not adopt the classical 3_{10} helix but instead forms a β-bend-like structure that allows the consensus isoleucine and tyrosine residues at positions 4 and 7, and the non-consensus leucine in position 8 to insert into the three-forked plug (Hishiki et al., 2009a).

Another non-canonical PIP box was describe in a recent structural study showing that the fourth regulatory subunit of the replicative Polδ p12 (POLD4) binds PCNA through a highly divergent PIP box lacking the glutamine and the first aromatic residue of the PIP consensus sequence (Gonzalez-Magaña et al., 2019). Unlike the canonical PIP box of p66 (POLD3), the PIP box of p12 lacks the conserved glutamine; binds through a two-fork plug made of an isoleucine and a tyrosine residue at position 3 and 8, respectively; and is stabilized by an aspartate at position 6, which creates a network of intramolecular hydrogen bonds (Gonzalez-Magaña et al., 2019). The most extreme example of a non-canonical PIP motif whose structure has been determined is that of the anti-recombinogenic helicase Srs2 (Armstrong et al., 2012). It lacks the two consensus aromatic residues in positions 7 and 8. The consensus glutamine in position 1 binds in the Q-pocket. However, instead of forming a 3_{10} helix, the residues spanning position 4 through 8 form an α-helix. This allows the consensus isoleucine at position 4 and the non-consensus glutamine and leucine residues at positions 7 and 8 to insert into the three-forked plug. A recent comparative analysis of the canonical and non-canonical PIP motifs revealed

that the presence of an acidic residue at position 6 seems to stabilize the 3_{10} helix through a network of intramolecular hydrogen bonds and might be required for a high affinity interaction. By contrast, the presence of positive charged residues inside the PIP sequence seems to hinder PCNA binding (Gonzalez-Magaña et al., 2019).

An extended version of PIP box was identified as PIP degron (Havens and Walter, 2009). The PIP degron motif targets the protein for degradation, upon binding to PCNA, and besides the classical PIP sequence, it also harbours a basic residue (K or R) four amino acids after the second aromatic residue, as well as a TD motif just before the aromatic residues (Havens and Walter, 2009). Additionally, a novel motif was identified and found to be present in several PCNA interacting proteins, named KA box, that consists of residues K-A-(A/L/I)-(A/L/Q)-x-x-(L/V) (Liang et al., 2011) (Xu et al., 2001). Moreover, an alternative PCNA binding motif was recently described: the AlkB homologue PCNA interacting motif (APIM), which comprises only five 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 (Lancey et al., 2020a). The crystal structure of PCNA in complex with APIM peptides showed a similar binding mode to the PIP motif, despite having a highly divergent sequence from the canonical PIP box (Hara et al., 2018). All these sequences, PIP degron, KA box and APIM motif go under the large umbrella term of PIP-like motif, which is gaining more biological relevance as more and more PCNA interactions mechanisms are investigated.

Few years ago, Baldeck and colleagues showed that Polη binds the POLD2 subunit of Polδ and that this interaction is mediated by two adjacent phenylalanine residues in a motif called F1(Baldeck et al., 2015). Interestingly this F1 motif is overlapping with one of the two previously identified RIR motifs of Polη (Ohmori et al., 2009). It has to be noticed that this is the first instance of any PIP-like motif, in this case a RIR motif, reported to mediate interactions with multiple target proteins, in this case REV1 and Polδ. This was the first hint that the different classes of PIP-like motifs may not be distinct entities. Furthermore, a recent piece of work points out that the PIP sequence has regions overlapping with RIR (REV1 interacting region) and MIP (Mlh1 interacting proteins) motifs (Boehm and Washington, 2016). These findings show that PIP motifs, RIR motifs, and F1 motifs have partially overlapping specificities, and the lines separating these motifs became blurred.

As evidence keeps emerging, a more general class of PIP-like motifs should be considered. Taken together, these studies show that PCNA can bind a wide range of PIP motifs that diverge

significantly from the consensus sequence. PCNA might be able to bind a wider 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. Emerging evidence shows that the canonical PIP box definition is becoming too narrow as other sequences diverging from the canonical PIP box binding PCNA are being described. This implies that identifying PIP motifs from sequence analysis alone is highly problematic and quite obsolete.

1.4.7 Simultaneous interaction with multiple partners: PCNA as a "tool belt"

The PCNA ring forms a stable trimer in solution and has three identical PIP-box binding sites, one on each protomer. Therefore, theoretically PCNA could bind three different ligands at the same time, coordinating a variety of functions in space and time. In the past decades, several crystal structures of PCNA bound to PIP box peptides were solved. Some relevant examples are the cyclin-dependent kinase inhibitor $p21^{WAF1/CIP1}$ (Gulbis et al.,1996), the PIP-peptides of human flap endonuclease 1 (FEN1) and of the third subunit (POLD3 or p66) of human replicative Polδ (Bruning and Shamoo, 2004), and the PIP box peptides of TLS polymerases Polη and Polι (Hishiki et al., 2009a). In the crystals structures of these complexes, the stoichiometry is three peptide molecules per PCNA trimer, with the peptides accommodated in the hydrophobic pocket of each protomer. In each complex, the three binding sites and the conformation of the three bound peptide molecules are structurally indistinguishable. For those peptides for which thermodynamics parameters have been measured by isothermal calorimetry (ITC), a model of a single set of sites fit the measurements, with no evidence of allosteric effects or cooperativity. These evidences indicate a general competitive binding for PCNA with a hierarchy that is probably dictated by the affinity of the ligand. Affinity could be modulated by different factors. Post translational modifications of PCNA or its binding partners, in different ways, can reorder the binding hierarchy. Direct interaction with DNA could also play a role in recruitment of DNA binding proteins to PCNA. The cell cycle regulator p21 is an intrinsically disordered protein and one of the PCNA binding partners showing the highest affinity, possibly because it needs to displace other proteins to block replication in response to DNA damage. This high affinity relies on efficient hydrophobic packing, as well as electrostatic interactions with the C-terminal region of PCNA where patches of basic residues at the N and C-termini surrounding the PIP box contribute to an increased binding (Xu et al.). Moreover, FEN1 endonuclease and p66 (POLD3 the third subunit of Polδ) exhibit fewer basic residues than p21, and accordingly exhibit lower affinity of 60 μM for FEN1 and 16 μM for p66, as measured by

ITC. (Bruning and Shamoo, 2004) Nevertheless, comparison of structure of PCNA with full-length FEN1 revealed additional contacts at the level the core domain and the C-terminal region of FEN1. Full-length FEN1 affinity of PCNA measured by surface plasmon resonance (SPR) showed a K_d of 60 nM, three orders of magnitude higher than the peptide affinity value measured by ITC (Chen et al., 1996) (Bruning and Shamoo, 2004). This indicates that binding affinities reported by different techniques (SPR, ITC) of single peptides might be very different form actual full-length protein affinities for PCNA.

For TLS polymerases object of this study, Polη and Polι, crystal structure of PIP peptides bound to PCNA was solved a decade ago (Hishiki et al., 2009a). In the study presented by Hishiki and colleagues each crystal stucture, Polη-PCNA and Polι-PCNA complexes, shows three PIP peptides bound to PCNA, one on each protomer. Additionally, in the study binding affinites of PIP peptides of Polη and Polι bound to PCNA were measured by SPR, and reported to be 0.4 and 0.39 μM respectively but, as for the FEN1 case, affinity of full-length polymerases might differ greatly from this values. Unfortunately, obtaining structures of full-length Polη and Polι in complex with PCNA still remains a big challenge, and was the main focus of this study.

Notwithstanding the great availability of crystal structures of PCNA bound to short peptides of interacting partners (Hishiki et al., 2009a) (Armstrong et al., 2012), the field still lacks structural information of full-length proteins bound to PCNA. The scarcity of full-length complexes makes difficult the validation of the proposed "PCNA tool belt" model, where PCNA coordinates different proteins on each protomers. For many years the only available high-resolution structure of a full-length protein bound to the sliding clamp was of PCNA-Flap Endonuclease 1 (FEN1) complex in the absence of DNA (PDB: 1UL1) (Sakurai et al., 2005). This crystal structure shows three endonuclease molecules attached to the ring, with different relative orientations (Sakurai et al., 2005). It was proposed that these conformations represent inactive complexes, with FEN1 required to swing through a flexible hinge to acquire a DNA-cleavage competent position. This was the only structure of a full-length protein bound to PCNA until the very recent cryo-EM reconstruction of PCNA-DNA-Pol δ and FEN1 complexes (Lancey et al., 2020a). The recent work of Lancey and colleagues is a good example of PCNA coordinating different partners, and to date is the only structure of PCNA coordinating two different full-length proteins. In the process of Okazaki fragment maturation, it is imaginable that different protomers of the PCNA trimer interact with Pol δ, FEN1, and DNA ligaseI; mediating DNA synthesis, flap cleavage, and ligation, respectively. This theoretical complex has been called the "PCNA tool belt" and has been experimentally demonstrated in the case of

yeast PCNA (Stodola and Burgers, 2016). However, mutant PCNA trimers with a single binding site are still capable of Okazaki fragment maturation (Dovrat et al., 2014), arguing that PCNA tool belts may not be absolutely required. The PCNA tool belt mechanism plays a role in Okazaki fragment maturation and was also proposed as a possible translesion synthesis mechanism, where PCNA coordinates more than one polymerase.

The cryo-EM reconstruction presented by Lancey and colleagues shows a defined FEN1 conformer comprising only ~20% of the particle dataset, while the map from the entire dataset displays no significant density at the FEN1 binding site. This indicates that FEN1 containing complexes may be a little minority. Additionally, the structure, with two of the PIP box sites occupied by the catalytic subunit of Polδ and by FEN1, indicates that DNA ligase I would be sterically excluded. This is in agreement with a sequential switching of binding partners proposed few years before for Okazaki fragment maturation (Dovrat et al., 2014). The mechanism of PCNA tool belts in translesion synthesis would allow the coordination of several polymerases on PCNA (or most likely UbPCNA). This setting would permit the competition with the replicative polymerase from the primer/template junction, ideally by keeping such polymerase still anchored to PCNA, and at the same time would permit the recruitment of one or more TLS polymerases. In fact, according to the two-polymerase TLS mechanism lesion bypass starts with the insertion of one or more nucleotides by an inserter polymerase (Polη, Polι or Polκ), followed by the extension of the primer by an extender polymerases (REV1 or Polζ).

1.5 Aim of the thesis

When the replication fork encounters a DNA lesion, replicative polymerases are unable to bypass it and the replication machinery stops, eventually leading to fork stalling. In these circumstances translesion synthesis (TLS), a DNA damage tolerance pathway, is activated. The DNA sliding clamp PCNA is a fundamental platform which coordinates polymerases during DNA replication. Mono-ubiquitination of PCNA on K164 is a central requirement for the mechanism of polymerase switch, where the replicative polymerase is replaced by specialised translesion polymerases. Specialised TLS polymerases are recruited to the lesion site where they can insert the right -or wrong- nucleotides opposite the lesion, bypassing the damaged DNA and restarting the replication fork. The rationale behind which specific polymerase is recruited and the process of polymerase switching are still not well understood. In fact, the intricated interaction network between sliding clamp and TLS polymerases remains quite elusive. The aim of my PhD project was to disentangle some of the molecular mechanisms of TLS, such as polymerase switch and their recruitment to sliding clamp. To this end I performed structural and biochemical analysis on minimal translesion synthesis complexes. I extensively characterized recombinant human complexes from a biochemical point of view, with a special focus on Polη and Polι and their interaction with PCNA or mono-UbPCNA. Furthermore, I characterized the regulatory disordered C-terminus of Polη and its roles in the interaction with PCNA and UbPCNA, with a focus on newly identified PCNA interacting peptides (PIP boxes) and ubiquitin binding zinc finger (UBZ). In parallel to biochemical studies, I exploited cryo electron microscopy (cryo-EM) to carry out extensive structural analysis of complexes and single proteins. Cryo-EM analysis provided a molecular view of PCNA and UbPCNA and the flexibility of ubiquitin on the sliding clamp. Furthermore, this thesis presented a first report of a cryo-EM reconstruction of PCNA-Polι-DNA complex as well as full-length Polι in complex with DNA. Finally, I obtained a cryo-EM reconstruction of a UbPCNA-Polη complex at 9 Å resolution which showed that ubiquitin relocates to the front side of PCNA, bridging the polymerase to the DNA junction and making additional contacts with the flexible regulatory C-terminus of Polη. Altogether, the structural and biochemical results presented here give new insights into the role of PCNA ubiquitination in the recruitment of TLS polymerases especially via their flexible C-terminal regulatory regions.

2. Results

2.1 Expression and purification of recombinant human TLS polymerases in bacterial expression system

In order to understand on a molecular level the interaction network of TLS polymerases and sliding clamp, it is needed to obtain full-length proteins and to reconstitute stable complexes. To this end, I devised large scale preparation of human TLS polymerases, specifically Polη and Polι as full-length enzymes, and successfully obtained preparative amounts of pure protein, amenable for biochemical characterisation. Polymerases were purified as recombinant tagged proteins form *E. coli*. Purification of human full-length polymerases has always been a challenging task, especially for structural studies. Polymerases tend to show low expression in bacterial systems and are often found in inclusion bodies. Moreover, they have a very flexible C-terminal domain, a highly disordered region, which makes them prone to degradation during purification and with high tendency to aggregate. Considering these challenges, I designed and optimised specific tags and purification procedures in order to obtain preparative milligram yields of stable, full-length and active proteins, at high degrees of purity. The polymerases activity was confirmed *in vitro* by activity assays, such as primer extension, electrophoretic mobility shift assay (EMSA) and analytical size-exclusion chromatography (SEC). Furthermore, purifications of several Polη C-terminal mutants designed to characterise Polη binding to PCNA and UbPCNA, were also performed. In order to obtain high-affinity Polη constructs I generated PIP boxes mutants of Polη: a total of 4 putative PIP boxes motifs were identified by sequence similarity at Polη C-terminus. For convenience the identified sequences were named PIP box 0 (437-444), PIP box 1 (477-484), PIP box 2 (524-532) and PIP box 3 (701-708) (see table1.3 for details). I created high-affinity Polη mutants by substituting some of the PIP box sequences with the PIP sequence of the cyclin-dependant kinase inhibitor $p21^{CIP1/WAF}$ (p21), which is described as the strongest interacting partner of PCNA (Bruning and Shamoo, 2004). These high-affinity PIP mutants were purified similarly to the Polη wild type and used for biochemical characterisation as well as sample optimisation in cryo-EM. Additionally, in order to compare binding affinities via pulldown assays another set of Polη constructs was purified. These constructs carried impairing mutations at the level of either PIP

boxes or UBZ, aimed to reduce binding to sliding clamps. All the above mentioned Polη mutants, either overactivated or impaired, were cloned with a N-terminal GFP tag, which was suitable for pulldown assays. A visual summary of all Polη constructs created and utilised in this study is shown in Figure 2.5. In the following paragraphs details of protein purifications will be discussed.

2.1.1 High-yields purification of full-length human Polι from bacterial expression system using His-SUMO tag

Human Polι is a 80 kDa protein consisting of 715 amino acids (GenBank Accession number AF140501.1, UniProt: Q9UNA4). Some reference sequences (e.g. NM_007195.2) describe Polι as a 740 amino acid protein with an additional 25 residues at the N-terminus, however, experimental evidence suggests that the 715 amino acid form of Polι is the predominant isoform (Frank, 2017). The sequence encoding for Polι predominant isoform (26-740) was cloned into our bax04 vector, carrying a N-terminal His-SUMO tag (see table 4.4). Small ubiquitin-related modifier (SUMO) tags are ~100 residues in length and 12kDa in mass. A SUMO fusion system can facilitate the efficient expression of target proteins in *E. coli* and enhance the solubility of proteins (Malakhov et al., 2004). The SUMO tag can be efficiently removed using a highly specific protease, SENP2, which recognizes the SUMO protein structure rather than a peptide specific sequence (Figure 2.3 a.). The SUMO system allowed purification of full-length, properly folded and active human Polι at high yields. Detailed procedures are discussed in section 4.2.3 of Material and Methods. Briefly, the lysate corresponding to 4 l *E.coli* cultures pellet equivalent, was applied onto a HisCOmplete column (Roche) and after extensive washing and an additional Chaperone wash step, the protein was eluted with imidazole. After overnight tag cleavage by SENP2 the protein was further purified via a Heparin column, which is negatively charged mimicking the phosphate of DNA, and works very well with nucleic acid-binding proteins. After a salt gradient elution from Heparin column fractions containing full-length Polι were pooled, concentrated and loaded onto a Superdex 200 16 600 size-exclusion chromatography column. The peak from the SEC elution profile was symmetric and contained high amounts of pure protein, as can be observed by the chromatogram and SDS-PAGE, although some minor degradation could be detected in the late eluting fractions (Figure 2.1). Fractions containing pure and full-length Polι were pooled and concentrated till a suitable working concentration and stored at -80˚C. A sample of the final protein was subjected to electrospray ionisation - time of flight (ESI-TOF) mass spectrometry, confirming the mass of

Polι (80560,5429 Da). Additionally, primer extension assay showed that Polι was active, and could incorporate nucleotides, though with low processivity, in presence of both $MgCl_2$ and $MnCl_2$, (Figure 2.2) validating previous observations (Frank, 2017) and confirming the good quality of the sample. From a standard purification I routinely obtained a final yield of ~3mg per litre culture of pure and active protein.

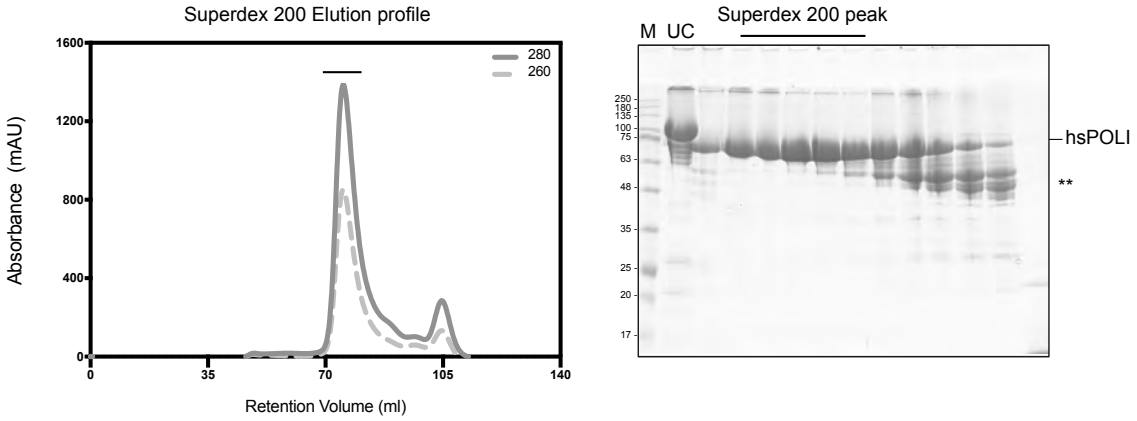


Figure 2.1 Purification of hsPolι. SEC chromatogram with absorbance at 260 nm (dotted line) and 280 nm (continuous line) shows elution of pure Polι around 75 ml retention volume, in a sharp symmetric peak (black line), with a slight tail on the right hand side. The corresponding SDS-PAGE shows pure full-length Polι (hsPOLI) in the peak fraction (line), around 80 kDa molecular weight and some degradation in the following fractions (**), UC: un-cleaved SUMO- Polι, M: molecular weight marker.

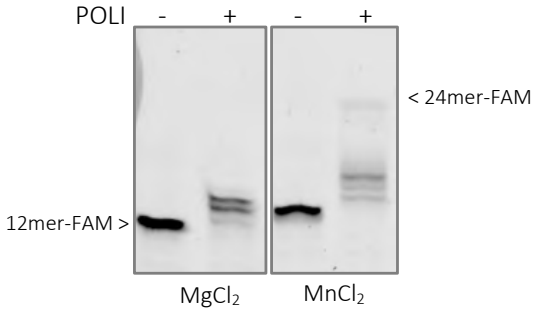


Figure 2.2 Primer extension activity of hsPolι. Polι (POLI) confirmed polymerase activity of the protein *in vitro* on a primer/template substrate (12/24 nt) at a concentration of 500 nM in presence of 50 μM of dNTPs. 500 nM Polι POLI (+) compared to reaction control (-) incorporates a few nucleotides (+1, +2) in presence of 5 mM $MgCl_2$ as divalent cation. In presence of 5 mM $MnCl_2$ Polι is more processive, incorporating also at +3 position and with additional faint bands visible up to full-length +12 product (24mer).

2.1.2 Purification of full-length human Polη from bacterial expression system using an optimised CPD-Twin-Strep-tag

The synthetic codon optimized gene encoding full-length hsPolη (1-713) was cloned into a bacterial expression vector (bax25) with an in-frame C-terminal tag composed of cysteine protease domain (CPD) and Twin-Strep-tag (see table 4.4). The C-terminal tag was needed to prevent degradation and to ensure that the protein survived the purification procedure intact. CPD Twin-Strep-tag is a great combination as it puts together the solubilisation and self-cleavage properties of the CPD tag and specificity of the Twin-Strep-tag. The CPD tag is an inducible auto-cleaving protease tag (Shen et al., 2009) adapted from a Vibrio cholerae MARTX toxin. The addition of inositol-6-phosphate (IP_6) leads to activation of the CPD protease that cleaves at the target protein-CPD junction (Figure 2.3 b.). Importantly, IP_6 is absent from bacterial cells, therefore the full-length CPD-Twin-Strep-tag fusion protein could be purified from lysates in a protease-inactive form using the Strep-Tactin affinity resin. Addition of IP_6, induces auto-processing: the protease domain recognises the amino acid sequence XXXL-AGGK and cleaves between the leucine and alanine residue leaving only a leucine residue at the C-terminus of the target protein. After binding the target protein to an affinity column, tag cleavage allows a very convenient separation between target protein and the remaining tag. Furthermore, it has been shown before, that the CPD tag increases expression, solubility and stability of target proteins (Shen et al., 2009). Moreover, CPD is very specific and has poor trans-cleavage efficiency, which limits fusion protein cleavage to the CPD-target protein junction. Using this strategy human Polη, with a residual Leu at the C-terminus, could be efficiently purified to homogeneity using a quick two step purification procedure, by affinity chromatography on Strep-Tactin Superflow column and gel filtration. Full-length Polη was purified starting from 6 l cultures, subjected to lysis and clarification as described in section 4.2.3 Material and Methods. Briefly, the cell lysate was applied to a Strep-Tactin Superflow column, washed extensively and eluted upon tag cleavage by IP_6. The fractions containing pure full-length protein were then concentrated and loaded onto a HighLoad Superdex 200 column. Protein eluted around 70 ml retention volume and showed an absorbance ratio (260/280 nm) of ~0.6 indicating pure protein, without additionally bound DNA (Figure 2.4). DNA contamination is generally a big issue in polymerases purification, however in this case as can be seen in the chromatogram (Figure 2.4) all the contaminant DNA eluted as a first void volume peak, and most of the protein appeared to be in a soluble and folded state, eluting in the main peak. Protein purity was assessed by total ESI-TOF spectrometry and the activity was

confirmed by primer extension assay (Appendix Figure A8). A major problem regarding Polη purification was its tendency to aggregate at higher concentrations. Therefore it was never possible to concentrate the protein to concentrations greater than 3mg/ml. Generally, from a standard purification I could obtain ~ 0.2 mg per each litre culture, which was a rather exiguous amount, but still enough to perform biochemistry and cryo-EM studies.

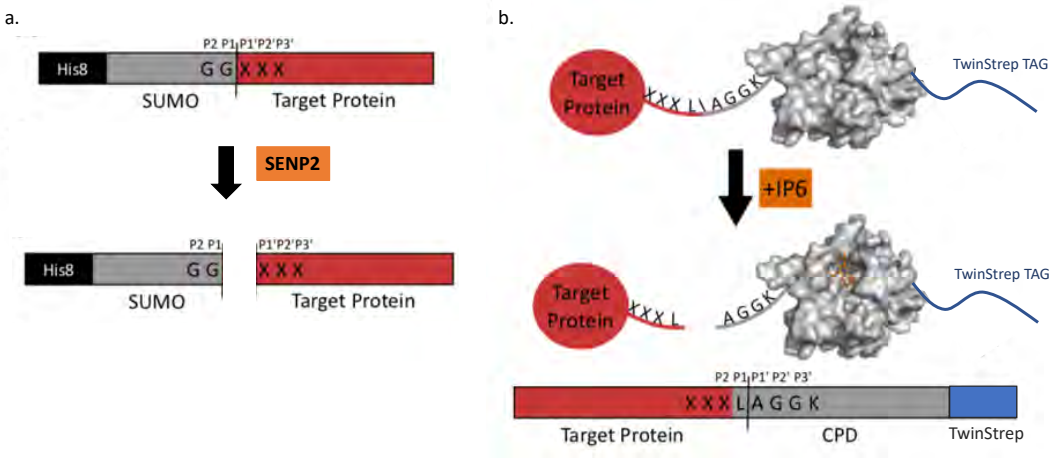


Figure 2.3 Tags used in purification systems. (a) The small-ubiquitin related modifier (SUMO)-tag serves as a solubility enhancer for recombinant protein and is a protease recognition site for the protease SENP2. (b) Cysteine protease domain (CPD) tag (PDB 3EEB) is covalently linked to the target protein. Upon addition of IP_6 the CPD protein undergoes cleavage at the P1 Leu site. The residues termed X are the C-terminal residues of the target protein. The auto cleavage site is marked by the black line between P1 and P1'.

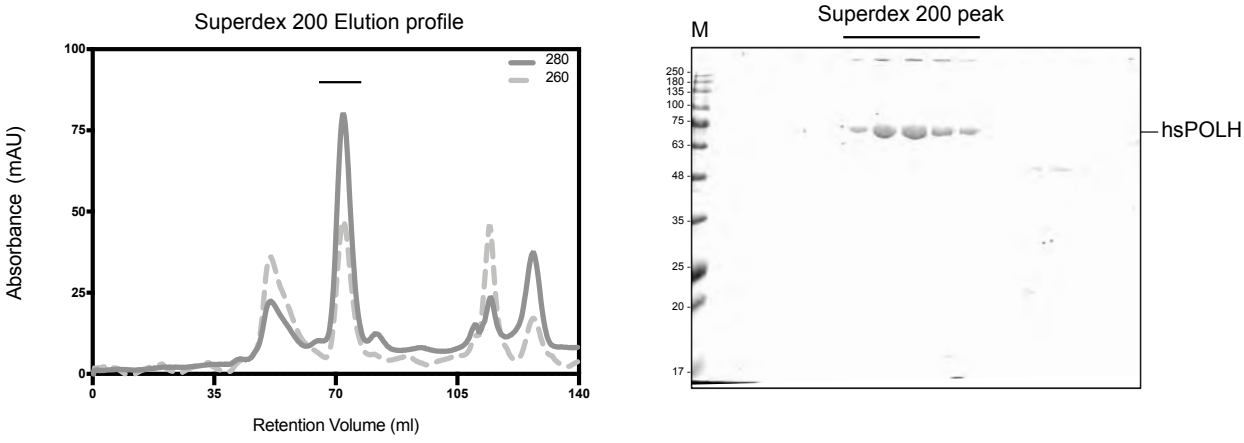


Figure 2.4 Purification of hsPolη wt. SEC chromatogram with absorbance at 260 nm (dotted line) and 280 nm (continuous line), shows elution of pure protein around 70 ml retention volume, in a sharp symmetric peak (indicated by black line). On the right side the corresponding SDS-PAGE: the line represents the peak fractions, loaded respectively on the gel. A band at ~79 kDa shows the presence of pure Polη (hsPOLH) in fractions eluting under the main peak. M: molecular weight marker.

2.1.3 Purification of Polη PIP mutants with higher PCNA binding affinity (pip3 and 3PIPS)

Given the supposed low affinity of Polη for PCNA or UbPCNA, I decided to increase the probability of stable complexes by mutating the non-canonical PIP boxes at Polη C-terminus. These putative PIP boxes, respectively at regions 477-484, 524-532 and 701-708 (see table 2.1) were identified by sequence analysis and comparison to the non-canonical PIP box sequence of Polη (M-x-x-Ψ-x-x-ϑ-ϑ). Polη PIP sequences differs from canonical PIP box consensus (Q-x-x-Ψ-x-x-ϑ-ϑ) where Ψ is hydrophobic and ϑ aromatic. I observed that the putative PIP boxes partially overlapped with other motifs, like RIR (REV1 interacting region) or F1 motif (Boehm and Washington, 2016), and were characterized by the presence of the two aromatic residues FF. I created constructs with mutations at each of these putative PIP boxes, or only the very C-terminal one, respectively. The mutations were substitutions of the natural Polη PIP sequences with the PIP sequence of p21, which is described as the strongest PCNA binding partner, with a K_d of 65 nM (Hishiki et al., 2009a). Such sequence is a canonical PIP box (Q-x-x-Ψ-x-x-ϑ-ϑ) and it is composed of: **Q**TS**M**TD**FY**HSKRR.

Name in this work	Position	Sequence alignment	Motif references
p21 PIP box	(144-160)	GPGDSQGRKRRQTSMTDFYHSKRRLIFS	(Gulbis et al.)
Polη PIP box 1	(477-484)	GSGPAVTATKKATTSLESFFQKAAERQKVK	F1 motif (Baldeck et al., 2015) RIR motif
Polη PIP box 2	(524-532)	NSPSKPSLPFQTSQSTGTEPFFKQKSLLLKQ	RIR motif (Boehm and Washington, 2016)
Polη PIP box 3	(701-708)	LACTNKRPRPEGMQTLESFFKPLTH	(Hishiki et al., 2009a)

Table 2.1 Polη PIP boxes mutated in pip3 and 3PIPS constructs. Putative non-canonical PIP boxes identified at the disorder C-terminus of human Polη. Sequence comparison shows the typical conserved residues. Specifically, Q in position 1 is highlighted in blue, the moderately hydrophobic residues in the middle of the sequence (L/M/T) are in grey, the two aromatic residues (FF/FY) are in green. F1 and RIR motif additional residues are dotted-underlined. PIP boxes names as refer to in this work (PIP box 1, 2 and 3) and p21 PIP box is also reported as reference (underlined continuous and dotted). p21 sequence (underlined continuous) was substituted in place of PIP box 3 (for the construct Polη pip3) or at level of PIP box 1, 2 and 3 (for the construct 3PIPS).

Following this strategy I created two high-affinity mutants. The first one, named 3PIPS for convenience, carried the p21 PIP substitution at 3 different PIP boxes: PIP box 1 (aa 477-484), PIP box 2 (aa 524-532) and PIP box 3 (aa 701-708). The second construct was named pip3 and carried the substitution only at the level of the very C-terminal PIP box PIP box 3 (aa 701-708). Both these Polη mutants were cloned into bax25, with a C-terminal CPD-Twin-Strep-tag and

purified following similar procedures as Polη wild type. Purification procedure is described in details in Material and Methods section 4.2.3. Both pip3 and 3PIPS constructs were purified successfully by the standard procedure of affinity chromatography using Strep-Tactin Superflow column followed by SEC. The yields were comparable, or a little higher than Polη wild type. Over time, especially in the case of the 3PIPS prep, I could notice some instability and degradation possibly due to the introduction of some degron sequences within the p21 PIP box (Havens and Walter, 2009). Nonetheless, the proteins were successfully subjected to analytical SEC for complex formation with PCNA and UbPCNA, and their activity was confirmed by primer extension assay (Appendix Figure A8) and EMSA (section 2.5). Furthermore, cryo-EM analysis was carried out and will be described later in section 2.4.

2.1.4 Purification of Polη constructs (pip3, 3PIPS, 8A, pip3 dead, UBZ dead and 2PIPS mutants) with N-terminal GFP tag used for pulldown assays

In parallel to structural studies (see section 2.4), I decided to biochemically characterise the nature of the interaction of Polη with PCNA and UbPCNA, with focused attention on the C-terminal disordered region and its PIP box and UBZ motifs. The unavailability of stable Polη at high concentrations, due to its aggregation behaviours, made nearly impossible to reproducibly measure binding affinities via isothermal calorimetry (ITC) analysis. Microscale thermophoresis (MST) was attempted but the proteins were sticking to the capillaries, even in presence of detergent, therefore the experimental approach was impossible. Additionally, the quantities of UbPCNA, obtained with an optimised *in vitro* enzymatic reaction (see section 2.2.2), were limiting to allow analysis via other biophysical methods as fluorescence anisotropy (FA) or isothermal calorimetry (ITC). Altogether, these reasons brought me to establish a very clear, reproducible and semi-quantitative pulldown assay, in order to tackle binding affinities of different Polη constructs to PCNA or UbPCNA. Different bait/preys combination, diverse kind of beads and several tags (CPD-Twin-Strep-tag, EPEA, His) were tested but none of these approaches could give reproducible results. The approach of choice was finally to use GFP as affinity tag for binding to beads. Specifically, the stationary phase used in the pulldown assays was magnetic agarose beads GFP-Trap®_MA (Chromotek). GFP-Trap beads worked very specifically in binding the bait proteins and did not show any unspecific binding of preys (see Appendix Figure A9). Upon testing different combinations of preys and baits, I observed that a setup in which Polη was the bait resulted in maximum reproducibility. The rationale was to

clone all Polη constructs with a N-terminal GFP tag, rather than the usual C-terminal CPD tag, in order to leave the C-terminus of Polη free to capture the sliding clamp out of the mobile phase. To this end, a set of new N-terminally GFP tagged polymerases had to be purified. All constructs were cloned into bax29 carrying a N-terminal GFP-Twin-Strep-tag. Specifically, the Polη mutants used in pulldowns (section 2.6) were Polη wt, the high-affinity mutants pip3 and 3PIPS, as well as the impaired mutants at the level of ubiquitin binding zinc finger (UBZ dead), at the very C-terminal PIP box 3 (pip3 dead) or all 4 putative PIP boxes (8A) (see Figure 2.5). All constructs were purified via the N-terminal Twin-Strep-tag, embedded in the GFP, via Strep-Tactin affinity and SEC. Unlike the Polη constructs expressed with the C-terminal CPD tag, these constructs showed a small level of degradation, due to the lack of a C-terminal protective tag. Nonetheless, almost all constructs were purified successfully to preparative amounts, analysed by fluorescent gels and confirmed by ESI-TOF mass spectrometry (see Appendix Figure A2).

However, purification of the high-affinity PIP mutants (pip3 and 3PIPS) lacking the protective C-terminal CPD tag proved challenging. I repetitively observed, via mass spectrometry analysis, that N-GFP-Polη-pip3 and N-GFP-Polη-3PIPS were losing the C-terminal p21 PIP box (see Appendix Figure A2). In these two constructs, carrying only N-GFP tag and lacking a protective C-terminal CPD tag, the already flexible C-termini, were possibly exposed to cellular proteases. A feasible reason for this C-terminal degradation might be the fact that the p21 PIP box could be making the protein more toxic to the host *E.coli*, therefore becoming a target for degradation. The specific degradation product, at the level of the last PIP box, could be possibly due to the introduction of some degron sequences within the p21 PIP box (Havens and Walter, 2009).

This problem of specific degradation at the level of the mutated PIP box3 was solved by re-introducing the C-terminal CPD tag. Therefore the pip3 and 3PIPS mutant were obtained from constructs carrying both N-GFP-His and C-CPD-Twin-Strep-tag (bax30-CPD vectors) and purified as the previously described bax25 constructs, leaving the N-terminal GFP tag un-cleaved (Appendix Figure A2 c. and e.). Surprisingly, I could take advantage of one of the degradation products, derived from the unsuccessful full-length purification of N-GFP- Polη-3PIPS. Upon mass spectrometry analysis (see Appendix Figure A2 d.) the degradation product was revealed to be Polη 3PIPS where the 25 residues were missing. The sequence in brackets (SPLAC^TNKRPRPEG**QTSMTDFYHSKRRLIFS)** depicts the cleavage site ^ and the bold sequence is the very C-terminal p21 PIP box sequence. This unintentional construct was an interesting one and was kept for biochemical analysis, and will be referred to as 2PIPS. Polη

2PIPS construct carried two ‘overactivated’ PIP boxes at the level of PIP box 1 and PIP box 2, but lacks completely the last PIP box3.

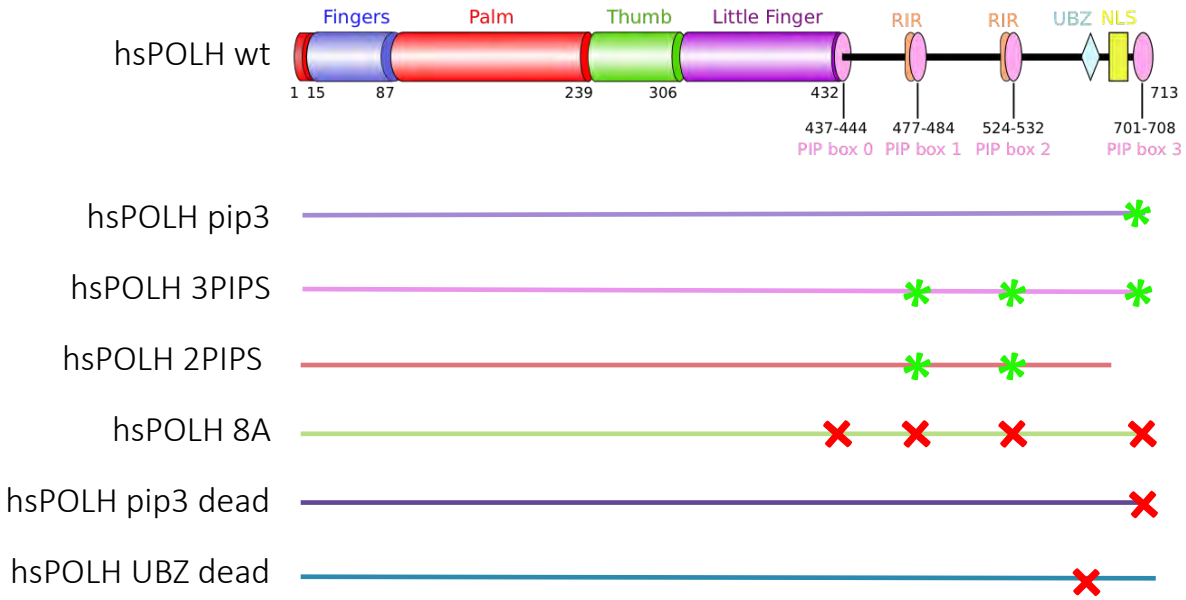


Figure 2.5 Polη constructs scheme. Depiction of Polη construct generated in this study, used to characterize the C-terminus region of Polη (POLH). Polη domain organisation, comprising polymerase domain (finger, palm, thumb and little finger domains), and C-terminal region with identified PIP boxes (box 0, box1, box2 and box3) partially overlapping with REV1 interacting region (RIR). NLS: nuclear localisation signal, UBZ: ubiquitin binding zinc finger. Substitution of natural Polη PIP box sequences at the level of putative PIP boxes by the p21 PIP box sequence (QTSMTDFYHSKRR) is indicated by a green star. The pip3 mutant (lilac) carries the substitution only at the level of PIP box 3 (701-708). 3PIPS mutant (pink) carries the substitution at the level of PIP box1 (477-484), PIP box 2(524-532) and PIP box 3 (701-708). The 2PIPS truncated mutant (salmon) has p21 PIP box sequence at the level of PIP box1 (477-484), PIP box 2(524-532) and lacks PIP box 3. POLH 8A mutant (green) with substitution of all 4 putative PIP boxes 0, 1, 2 and 3 (F443A, L444A, F483A, F484A, F531A, F532A, F707A, F708A respectively) which impair the PIP box binding capacity, is indicated by red crosses. POLH pip3 dead mutant (violet) has substitution at the level of F707A, F708A , which inactivates the PIP box3. POLH UBZ dead mutant (petroleum blue) with D652A mutation at the level of ubiquitin binding zinc finger.

2.2 Purification and enzymatic mono-ubiquitination of human PCNA

2.2.1 Recombinant hsPCNA can be efficiently purified at high-yields using a CPD-Twin-Strep-tag

Proliferating cell nuclear antigen (PCNA) is a ring-shaped DNA sliding clamp which is conserved in each domain of life. PCNA is known mainly for its roles in DNA replication and DNA repair: it is a crucial scaffold protein that coordinates polymerases and other proteins in various DNA metabolism processes. Purification of eukaryotic PCNA has been established few decades ago, as the protein was crystallised and characterised by different groups (Bauer and Burgers, 1988) (Krishna and Gary, 1994). Here, a fast and efficient method to purify on a preparative scale human PCNA from bacterial cells is described, exploiting the already described CPD-Twin-Strep-tag approach of bax25 (section 2.1.2). With this strategy hsPCNA was efficiently purified to homogeneity in a two-steps purification, via Strep affinity and size-exclusion chromatography. The stringent selectivity of the Strep-Tactin column allowed exclusive binding of PCNA to the resin, therefore permitting its quick isolation from lysate. The detailed purification is described in section 4.2.3 of Materials and Methods. A standard purification yielded very high amounts of extremely pure protein (Figure 2.6) approximately 8 mg per each litre of bacterial cells. The homogeneity and the molecular weight of the protein was confirmed by mass spectrometry. Moreover, PCNA was subjected to size-exclusion chromatography, native PAGE (Figure 2.7 c.) and native mass spectrometry which confirmed the actual trimeric quaternary structure of the protein, which in solution assembled in very stable homotrimers of ~86 kDa (Figure 2.8). PCNA was then utilised for further biochemical characterisation, subjected to complex formation with either Polι or Polη, analysed by cryo-EM and used in activity assays, like primer extension, EMSA or pulldowns. Additionally, PCNA was subjected to enzymatic mono-ubiquitination, in order to produce PCNA-K164 mono-ubiquitinated for further biochemical and structural studies with TLS polymerases.

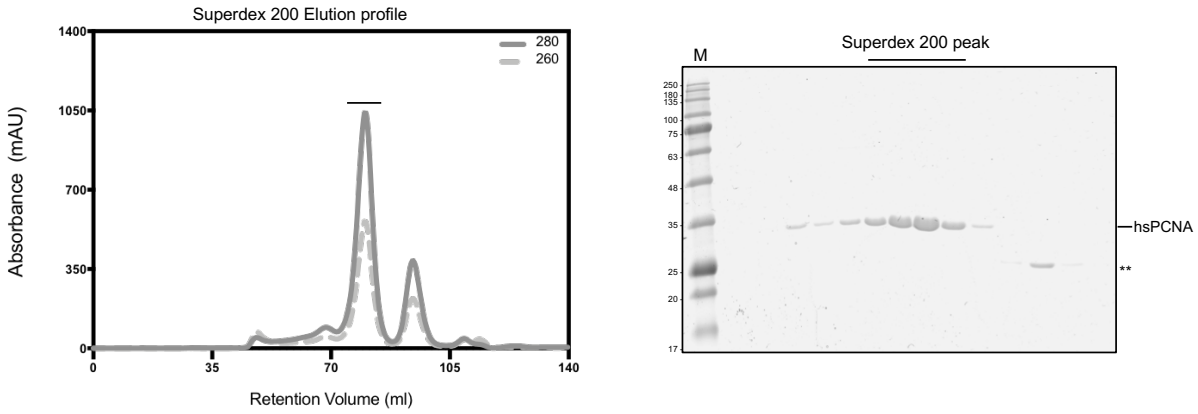


Figure2.6 Purification of hsPCNA. SEC chromatogram with absorbance at 260 nm (dotted line) and 280 nm (continuous line), shows elution of pure protein around 80 ml retention volume, in a sharp symmetric peak (indicated by black line). On the right side the corresponding SDS-PAGE, where the line represents the peak fractions, loaded respectively on the gel. A band at ~30 kDa shows the presence of pure hsPCNA (trimeric structure is denatured in SDS-PAGE), with a slight degradation at lower molecular weights, indicated by the ** in later fractions. M: molecular weight marker.

2.2.2 *In vitro* K164 mono-ubiquitination of PCNA using the promiscuous E2 UbcH5c

PCNA was previously shown to be specifically mono-ubiquitinated on K164 and that such modification played a key role in regulating TLS (Hoege et al., 2002). The addition of one ubiquitin molecule to K164 is performed by E2/E3 ubiquitin-conjugating enzymes Rad6 and Rad18 in yeast and higher eukaryotes (Hoege et al., 2002). However, low levels of Rad18-independent PCNA ubiquitination could be detected in higher eukaryotes, with the E3 ligases Rnf8 and CLR4 (Cdt2) proposed as candidates for this activity (Zhang et al., 2008) (Terai et al., 2010). A work published by Hibbert and Sixma demonstrated that is possible to obtain good amounts of K164 mono-ubiquitinated PCNA *in vitro* using the promiscous E2 enzyme UbcH5c, which is independent of a E3 ligase for specific K164 mono-ubiquitintation (Hibbert and Sixma, 2012). I reproduced and optimised the experimental set up of PCNA ubiquitination and was able to obtain preparative amounts of mono-UbPCNA (Figure 2.7 a. and b.). I introduced minor modifications to the original protocol, and the final reaction conditions are thoroughly described in Material and Methods section 4.3.1. The optimised reaction was performed in a buffer composed of 50 mM Bicine-KOH pH 9, 3 mM ATP, 2 mM TCEP, 3 mM $MgCl_2$ and a final salt concentration of 50 mM KCl. Components were added to the reaction mix in this order and amounts: 3.8 pmols of UbaI, 1.28 nmols of ubiquitin, 320 pmols of UbcH5c and 160 pmols of PCNA. As can be seen in Figure 2.7 b. almost all the PCNA present in the reaction mixture was

ubiquitinated in presence of UbcH5c. The shift, corresponding to PCNA size (+9 kDa), can be visualised on the SDS-PAGE and the modification of the protein was confirmed by mass spectrometry. The reaction was easily scaled up by a factor of 10x, yielding the conversion of 1.6 nmols of PCNA to UbPCNA. The reaction after incubation at RT was applied on a size-exclusion chromatography column Superdex 200, in order to separate UbPCNA form the rest of reaction components. As shown in Figure 2.7 a. and b. the first peak eluting contains ~90% pure UbPCNA, but it still co-elutes with a minor part of unmodified PCNA. Regardless of all the efforts in reaction optimization and purification, it was never possible to achieve a final sample composed solely of UbPCNA, as a minor part of PCNA was always present. Additionally, as suggested by native PAGE (Figure 2.7 c.) the final UbPCNA sample was composed of different trimers populations, carrying either 1, 2 or 3 ubiquitins. Native mass spectrometry (Figure 2.9) confirmed the presence all ubiquitinated PCNA species. A small population of PCNA trimers was carrying two ubiquitins for a total molecular weight of ~104 kDa. Importantly, the most represented species yielding a molecular mass of ~ 112 kDa, was corresponding to PCNA trimer with three ubiquitins (Figure 2.9).This heterogeneity was additionally displayed by cryo-EM classification and will be discussed in section 2.4.4.

Hibbert and Sixma (Hibbert and Sixma, 2012) showed that wild type (wt) UbcH5c favors the mono-ubiquitination reaction and that the S22R mutant is more stringent in generating mono-ubiquitination on K164, because of its inability of forming poly ubiquitin chains. The experimental setup presented in my work did not show any significat differences between UbcH5c wt or S22R, as both were equally good in producing K164 mono-UbPCNA. As can be observed in Figure 2.11 both UbcH5c wt or S22R mutant yielded the same amount of UbPCNA under all reaction conditions. Therefore, in my experimental set up it was irrelevant which version of UbcH5c was used.

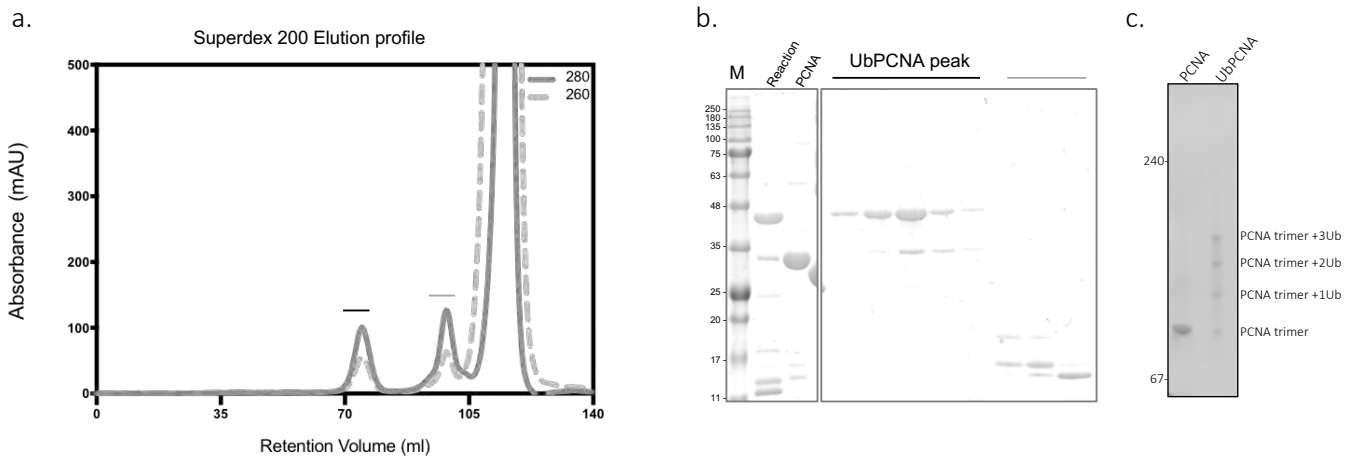


Figure 2.7 Purification of mono-UbPCNA. (a) SEC chromatogram form s200 elution of PCNA reaction products after *in vitro* ubiquitination. The first elution peak around 75 ml (black line) represents mono-UbPCNA products, the grey line corresponds to other reaction components, and the third big peak is ATP/ADP. (b) SDS-PAGE shows that almost all the UbPCNA is ubiquitinated in the reaction and the SEC elution peak (black line) is composed mostly of UbPCNA band (~35 kDa), however a small fraction of unmodified PCNA (~30 kDa) can be detected. (c) Native PAGE analysis of PCNA trimer reveals a molecular weight of ~85 kDa, UbPCNA shows multiple bands in the sample, corresponding to unmodified PCNA trimer, PCNA trimer +1 ubiquitin, +2 ubiquitins and +3 ubiquitins.

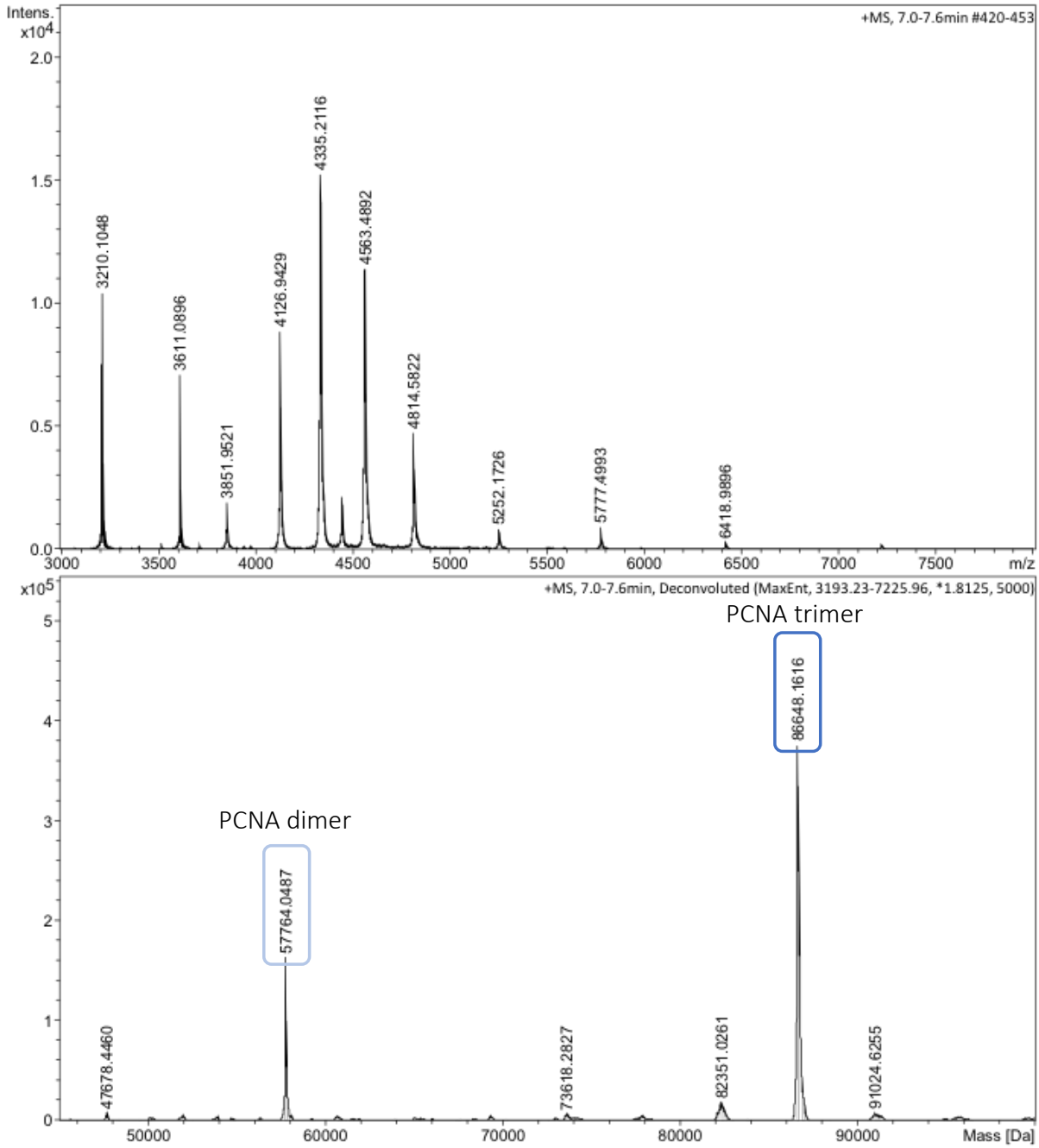


Figure 2.8 Native mass spectrometry analysis of PCNA. The native mass spectrometry peaks indicate the presence of two main populations: the peak at ~57 kDa corresponds to dimeric PCNA. The main peak at $4x10^5$ intensity is corresponding to ~86 kDa protein, corresponding to PCNA trimer.

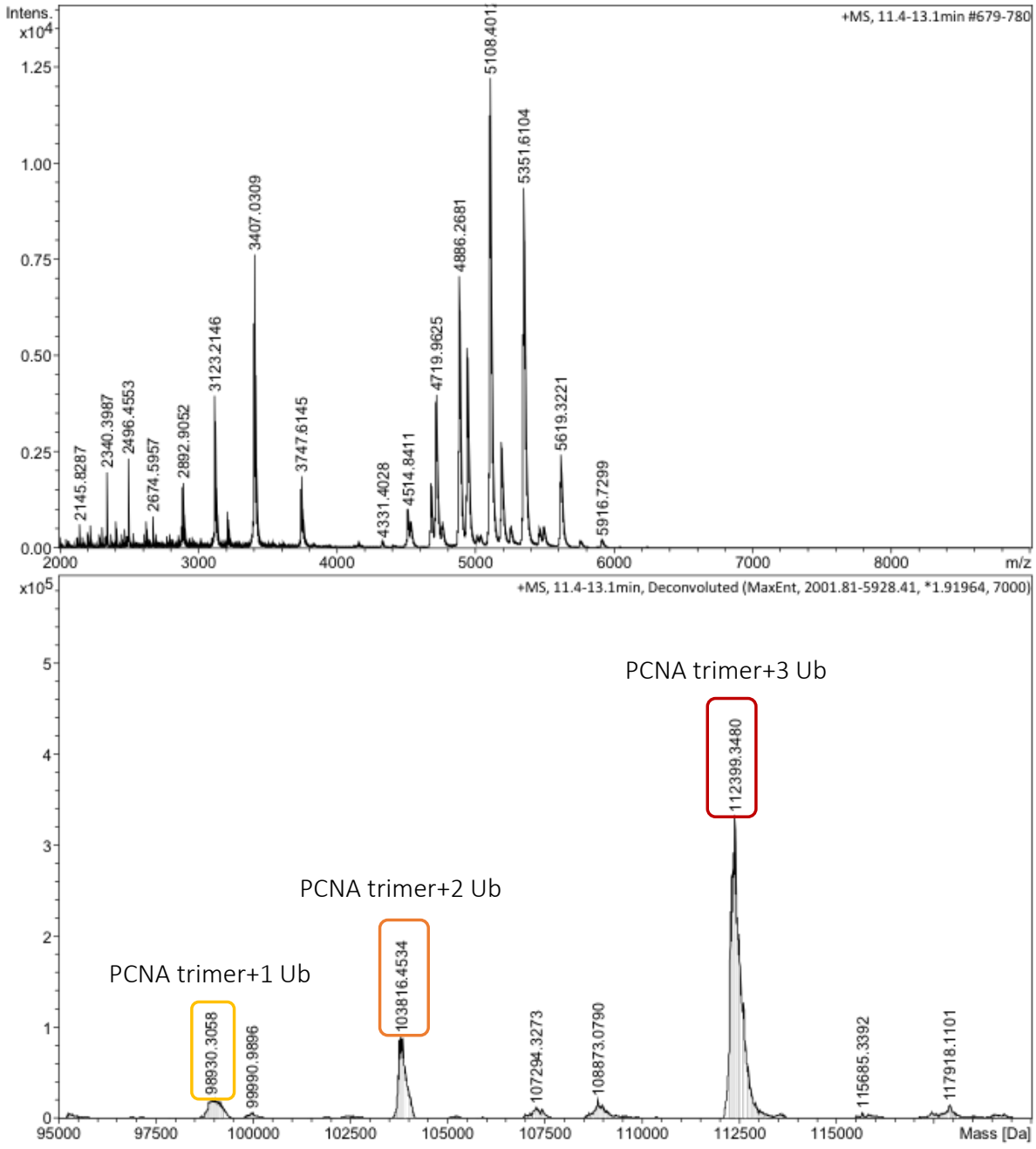


Figure 2.9 Native mass spectrometry analysis of UbPCNA. The native mass spectrometry of UbPCNA sample indicates the presence of three populations: the peak at ~98kDa corresponds to trimeric PCNA +1 ubiquitin, the peak at ~103 kDa fits to PCNA trimer +2 ubiquitins. The highest intensity peak ($3x10^5$) fits to size of PCNA trimer +3 ubiquitins (~112 kDa).

2.2.3 PCNA monomeric mutant can be purified but not efficiently ubiquitinated by UbcH5c

To further investigate different aspects of minimal TLS complexes I decided to create a PCNA mutant that would not self-associate stably and therefore remained in monomeric state in solution. To this end, I adopted the strategy published by the Lima group (Streich Jr and Lima, 2016) where a yeast monomeric from of PCNA was obtained by mutating K77, C81 and R110 to opposite charged residues. These three residues are conserved in humans respectively as K77, C81 and R110 and are located on α–helices at the interface of two adjacent monomers. K77D, C81E and K110D mutations disrupt the trimeric assembly of PCNA. The triple mutant PCNA monomer was gene-synthesised and cloned it into our bax36, with a C-terminal CPD (L204I)-8x His tag. Expression and purification were carried out as described in Material and Methods section 4.2.3. The purified protein was subjected to size-exclusion chromatography over a Superdex 75 column 16 600. As can be seen in Figure 2.10 the protein eluted at a volume of ~75 ml, indicating a molecular weight of ~30 kDa, according to standard molecular weight. From the SEC profiling it was concluded that the protein was in monomeric state. The correct mass of the protein was confirmed by ESI-TOF mass spectrometry. From a standard purification of 3 l bacterial cultures I could obtain approximately 20 milligrams of very pure protein.

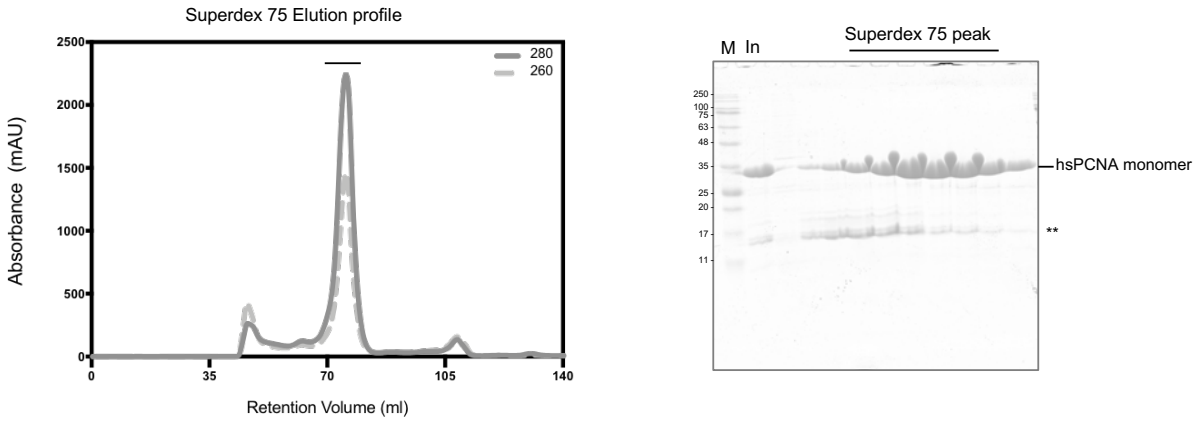


Figure 2.10 Purification of PCNA monomeric mutant. SEC chromatogram on Superdex 75 16 600, with absorbance at 260 nm (dotted line) and 280 nm (continuous line), shows elution of pure protein around 80 ml retention volume, in a sharp symmetric peak (indicated by black line). The elution volume corresponds to ~30 kDa molecular weight. On the right side the corresponding SDS-PAGE, the line represents the peak fractions, loaded respectively on the gel. A band at ~ 30 kDa shows the presence of pure PCNA monomer, with a slight degradation at lower molecular weights, indicated by the ** in later fractions. M: molecular weight marker.

Moreover, I tested whether PCNA monomeric mutant could be mono-ubiquitinated *in vitro* in a manner similar to PCNA wild type, using the established protocol with UbcH5c. Surprisingly, the reaction was much less efficient with PCNA in monomeric form. As can be appreciated in Figure 2.11 lane 5-6 (see * *), both UbcH5c wt and the S22R mutant failed to ubiquitinate monomeric PCNA efficiently. Despite exhaustive optimization of the reaction conditions, the efficiency of the reaction could not be improved. This result suggested that PCNA trimerization seems to be a prerequisite for efficient ubiquitination by promiscuous E2 UbcH5c, at least in the tested *in vitro* conditions.

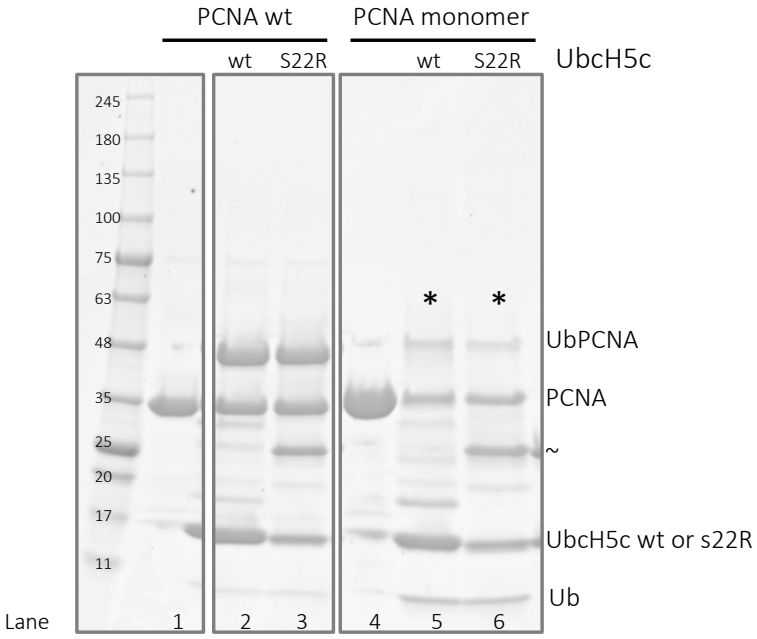


Figure 2.11 PCNA wt and monomeric mutant ubiquitination. The gel shows *in vitro* ubiquitination reactions using PCNA wt and monomeric mutant in presence of UbcH5c wt or S22R. Lane 1: PCNA wt control. Lane 2: ubiquitination reaction for 2 h at RT with PCNA wy and UbcH5c wt. Lane 3: ubiquitination reaction for 2 h at RT with PCNA wt and UbcH5c S22R. In both lanes 2 and 3 there is formation of UbPCNA product (~40 kDa) with no evident differences between the two reactions. Lane 4: PCNA monomeric mutant control. Lane : ubiquitination reaction for 2 h at RT with PCNA monomer and UbcH5c wt. Lane 3: ubiquitination reaction for 2 h at RT with PCNA monomer and UbcH5c S22R. Star indicates reactions with PCNA monomer when no or little ubiquitinated product is formed. The ~ marks a contamination present in the UbcH5c prep.

2.3 Analysis of minimal TLS complexes by size-exclusion chromatography

2.3.1 Size-exclusion chromatography analysis reveals interaction of UbPCNA with TLS polymerases

Having successfully purified full-length, stable and active polymerases, as well as PCNA, and having effectively modified PCNA by K164 mono-ubiquitination, I proceeded testing whether these proteins could form complexes *in vitro*. An efficient method to quickly detect complex formation is co-elution on size-exclusion chromatography. Therefore, I mixed proteins in a buffer suitable for complex formation (see table 4.7) and performed analytical SEC over a Superdex 200i 3.2/300 column using the Äkta micro system (see Material and Methods section 4.3.5 for details). As expected, it was observed that when mixed together, Polη and UbPCNA, as well as Polι and UbPCNA, were comigrating in a single peak compared to proteins alone. The peak corresponding to the complex was shifted toward lower retention volumes, compared to single UbPCNA or polymerase control peaks, indicating complex formation for both Polη-UbPCNA and Polι-UbPCNA (Figure 2.12 a. and c.). Presence of polymerases and UbPCNA was confirmed by SDS-PAGE analysis of peak fractions (Figure 2.12 b. and d.). An interesting observation was that complexes would co-migrate on gel filtration, only when the assembling concentration was higher than ~5 μM or 1 mg/ml. If the complexes were assembled at lower concentrations, e.g. 1 μM, the proteins would not co-migrate in a single peak, but elute separately (Appendix Figure A1). Afterwards I tested complex formation in presence of DNA. It is known that polymerase can bind to primer/template DNA substrates with good affinity, and PCNA is a DNA sliding clamp which slides across dsDNA. In order to lock the polymerase on DNA in a catalytically stable position, an excess of non-hydrolysable nucleotide dAMPNPP was added to the proteins-DNA mixture. dAMPNPP is accommodated in the polymerase catalytic site, correctly positioned between the fingers and the palm domains, and coordinated by the two divalent cations. Unlike normal deoxynucleotide, dAMPNPP has a imido group between the first and second phosphate, which does not allow the phosphodiester bond formation between the incoming nucleotide and the DNA. This trick was used to have conformationally stable DNA-bound polymerases. Then, I tested whether complexes (UbPCNA-Pols-DNA-dAMPNPP, with either Polη or Polι) were co-migrating on gel filtration. I observed stable co-migration of the DNA in the complex peak, which was monitored by absorbance at 260 nm. The complex peak in absence of DNA has a 260/280 nm absorbance

ratio around 0.5, indicating protein only, but when mixed with DNA the 260/280 nm absorbance ratio is around 1.1, indicating a DNA-protein complex (Figure 2.12 e. and g.). Both complexes of UbPCNA-Polη and UbPCNA-Polι were efficiently binding DNA in size-exclusion chromatography. Additionally, several DNA substrates with different lengths of primer/templated and different ratios of proteins-DNA were systematically screened by SEC and Theromofluor analysis (Appendix Figure A6). The best combination was shown to be a DNA length of 29/33 nt primer/template, in a slight excess compared to protein components.

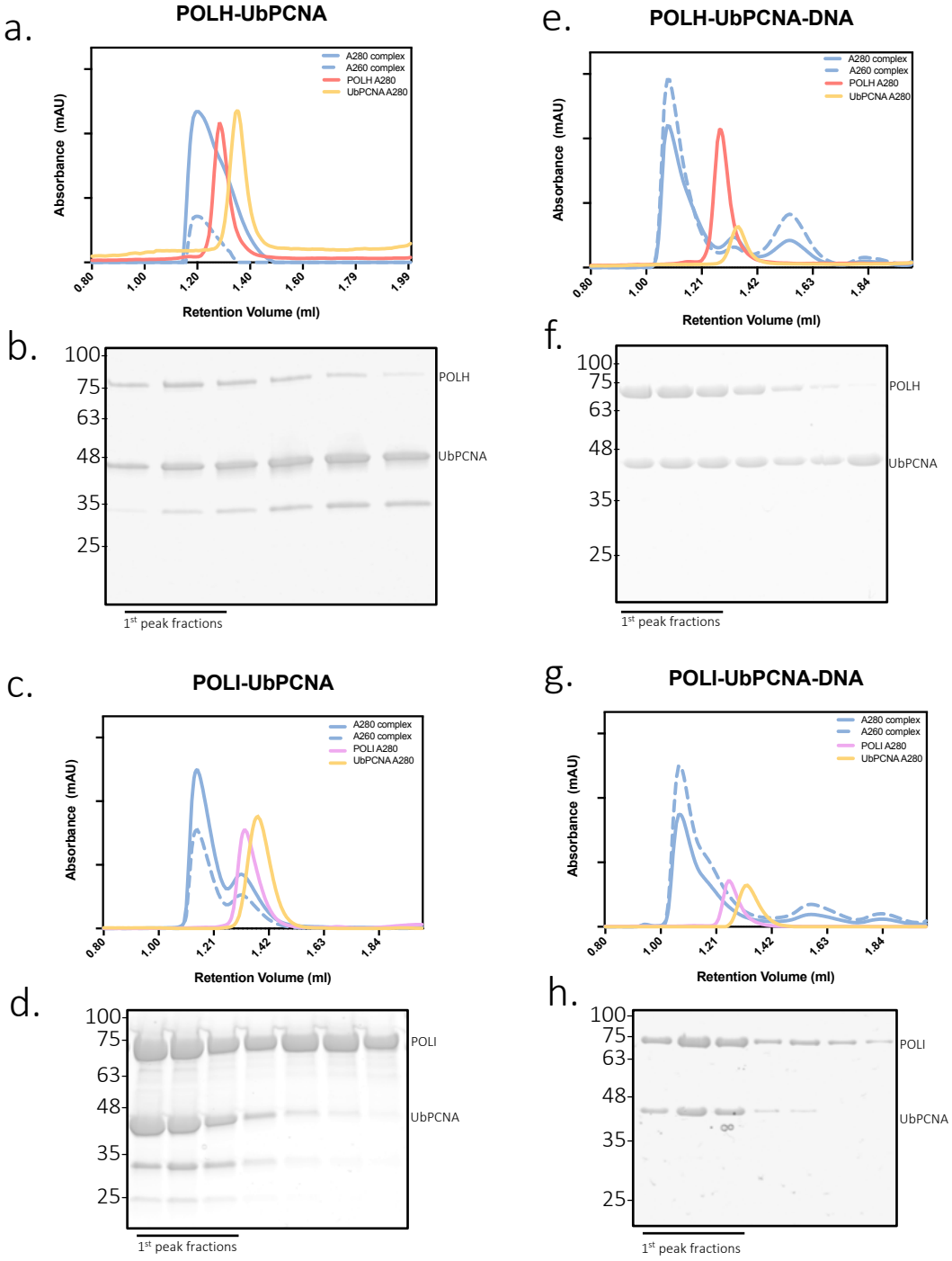


Figure 2.12 SEC analysis of UbPCNA-polymerases complexes. Complexes were reconstituted on Superdex200i 3.2/300 on Micro Äkta. (a) Elution profile with absorbance at 280 nm and 260 nm of UbPCNA and Polη complex (light blue lines) shows peak shift to the left compared to single protein control runs, receptively Polη (salmon line) and UbPCNA (yellow line). (b) UbPCNA and Polη complex corresponding fractions indicate presence of both proteins under the first elution complex peak (1st peak fractions). (c) Elution profile with absorbance at 280 nm and 260 nm of UbPCNA and Polι complex (light blue lines) shows peak shift to the left compared to single protein control runs, receptively Polι (pink line) and UbPCNA (yellow line). (d) UbPCNA and Polι complex corresponding fractions indicate presence of both proteins under the first elution complex peak (1st peak fractions). (e) Elution profile of UbPCNA-Polη-DNA-dAMPNPP complex with absorbance at 280 nm and 260 nm of UbPCNA and Polη complex (light blue lines) shows peak shift to the left compared to single protein control runs,

receptively Polη (salmon line) and UbPCNA (yellow line). The complex peak shows A260/A280 ratio of ~1, indicating a DNA-protein complex. (f) UbPCNA-Polη-DNA corresponding fractions from elution. (g) Elution profile of UbPCNA-Polι-DNA-dAMPNPP complex with absorbance at 280 nm and 260 nm of UbPCNA and Polι complex (light blue lines) shows peak shift to the left compared to single protein control runs, receptively Polι (pink line) and UbPCNA (yellow line). The complex peak shows A260/A280 ratio of ~1, indicating a DNA-protein complex. (f) UbPCNA-Polι-DNA corresponding fractions from elution show presence of both proteins.

2.3.2 Size-exclusion chromatography shows interaction of unmodified PCNA with TLS polymerases

The next step was to test binding of either Polη or Polι to unmodified PCNA. Kannouche and colleagues reported that Polη is specifically recruited to UbPCNA after DNA damage in cells, but it does not co-localize with unmodified PCNA in cells (Kannouche et al., 2002). Ever since, in the literature there has been conflicting information about the interaction between TLS polymerases and unmodified PCNA. Theoretically, specialised TLS polymerases, inaccurate and highly error-prone, should not be recruited to the replication machinery in absence of a DNA lesion, therefore their access to unmodified PCNA should be strictly regulated. It is believed that PCNA ubiquitination is the trigger point for polymerase switching and the start of translesion synthesis, therefore it was quite a surprise to find out that also unmodified PCNA could interact with Polη. I performed analytical size-exclusion chromatography, as described before for the UbPCNA containing complexes, and could observe co-migration of Polη and PCNA in a peak eluting earlier compared to proteins alone (Figure 2.13 a.). The presence of the proteins was confirmed by SDS-PAGE of fractions collected from the eluting peaks (Figure 2.13 b.). The same was observed for Polι, which similarly to Polη, formed a complex with unmodified PCNA in solution and was shown co-eluting in a left-shifted peak (Figure 2.13 c. and d.). Comparably to the UbPCNA-polymerases complexes, these assemblies were stable over gel filtration only if mixed at concentrations higher than ~1 mg/ml, indicating a concentration dependent affinity of TLS polymerases to unmodified PCNA. Finally, I tested unmodified PCNA and polymerases interaction in presence of DNA and non-hydrolysable nucleotide. As observed previously with complexes of UbPCNA-polymerases-DNA, also here complex formation for both Polη-PCNA and Polι-PCNA in presence of DNA was detected, as indicated by the 260/280 nm absorbance ratio in the SEC chromatogram. (Figure 2.13 e. and g.).

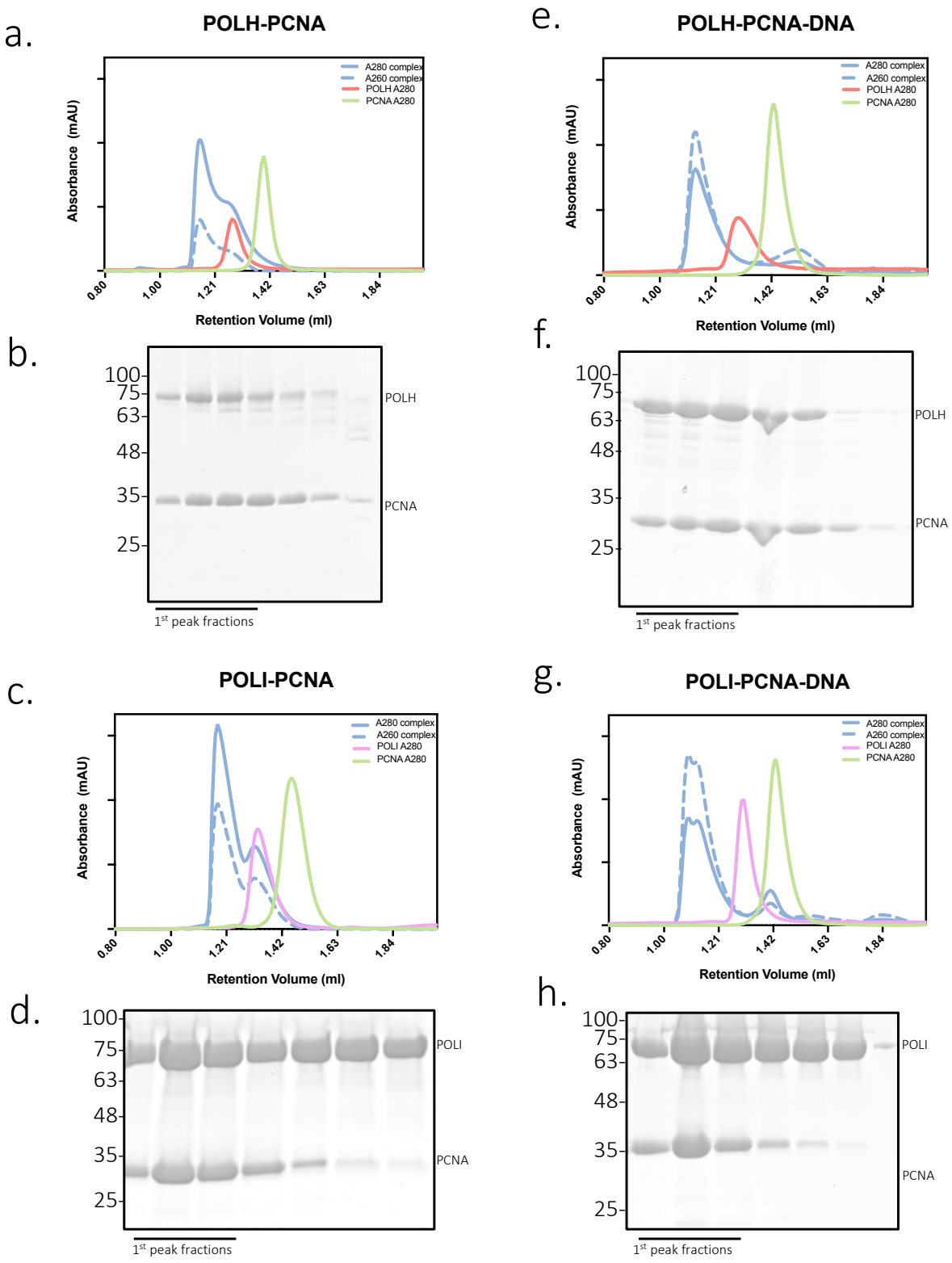


Figure 2.13 SEC analysis of PCNA-polymerases complexes. Complexes were reconstituted on Superdex200i 3.2/300 on Micro Äkta. (a) Elution profile with absorbance at 280 nm and 260 nm of PCNA and Polη complex (light blue lines) shows peak shift to the left compared to single protein control runs, receptively Polη (salmon line) and PCNA (green line). (b) PCNA and Polη complex corresponding fractions indicate presence of both proteins under the first elution complex peak (1st peak fractions). (c) Elution profile with absorbance at 280 nm

and 260 nm of PCNA and Polι complex (light blue lines) shows peak shift to the left compared to single protein control runs, receptively Polι (pink line) and PCNA (green line). (d) PCNA and Polι complex corresponding fractions indicate presence of both proteins under the first elution complex peak (1st peak fractions). (e) Elution profile of PCNA-Polη-DNA-dAMPNPP complex with absorbance at 280 nm and 260 nm of PCNA and Polη complex (light blue lines) shows peak shift to the left compared to single protein control runs, receptively Polη (salmon line) and PCNA (green line). The complex peak shows A260/A280 ratio of ~1, indicating a DNA-protein complex. (f) PCNA-Polη-DNA corresponding fractions from elution. (g) Elution profile of PCNA-Polι-DNA-dAMPNPP complex with absorbance at 280 nm and 260 nm of PCNA and Polι complex (light blue lines) shows peak shift to the left compared to single protein control runs, receptively Polι (pink line) and PCNA (green line). The complex peak shows A260/A280 ratio of ~1, indicating a DNA-protein complex. (f) PCNA-Polι-DNA corresponding fractions from elution show presence of both proteins.

In terms of stoichiometry between sliding clamp trimer and polymerases, the above-mentioned gel filtration experiments gave some initial indications, but not a well-defined picture. Despite having tested several protein-protein ratios (sliding clamp ring to polymerase or vice versa) of 0.3:1, 1:1, 1:2, 1:3 etc. the elution profiles were always quite consistent. Particularly, in case of Polη complexes, regardless of the nature of PCNA (ubiquitinated or not) and ratios tested, the polymerase was often detected in the first fractions, together with the sliding camp. When not in huge excess to sliding clamp, no unbound Polη was detected in the late eluting fraction. Dissimilar was the case of Polι containing complexes, with either PCNA or UbPCNA. Regardless of the ring : polymerase ratio tested, the late eluting fractions always contained some free unbound Polι. This preliminary observations might lead to the speculation that multiple copies of Polη could be accommodated on the same PCNA ring, whereas Polι has a minor binding occupancy on the trimeric sliding clamp.

Further experiments to understand the stoichiometry of polymerases to sliding clamp was performed using glutaraldehyde crosslinking followed by size-exclusion chromatography. Complexes with a ratio of 1 PCNA ring to 3 polymerases were assembled with either Polη or Polι respectively, and subsequently subjected to mild crosslinking with 0.05% glutaraldehyde. After quenching reaction with 100 mM Tris buffer the samples were subjected to analytical SEC onto a Superdex 200i 3.2/300 column and fractions analysed by SDS-PAGE (Figure 2.14). In the case of crosslinked PCNA-Polη complex a rather smeary band around 300 kDa could be visualised on the gel (Figure 2.14 a. start position *) indicating that the size of the crosslinked complex could correspond to one PCNA ring and three Polη molecules [86+(80x3)=326 kDa)]. When the Polι-PCNA crosslinked complex was analysed by SEC elution, a more convoluted scenario was visualised. Three different bands, also quite smeary, indicated by respectively one two or three stars (*) in Figure 2.14 b. were visible. The very faint band at ~250 kDa could correspond to PCNA trimer crosslinked with two Polι molecules, the more intense band at

roughly 160 kDa might be PCNA trimer crosslinked to one Polι. Additionally, a band around 80 kDa was visible in late eluting fractions, possibly indicating late eluting free polymerase. In conclusion, from this preliminary experiment some potential stoichiometries could be speculated, seemingly to be 1 PCNA ring to 3 Polη and 1 PCNA ring to 1 or 2 Polι.

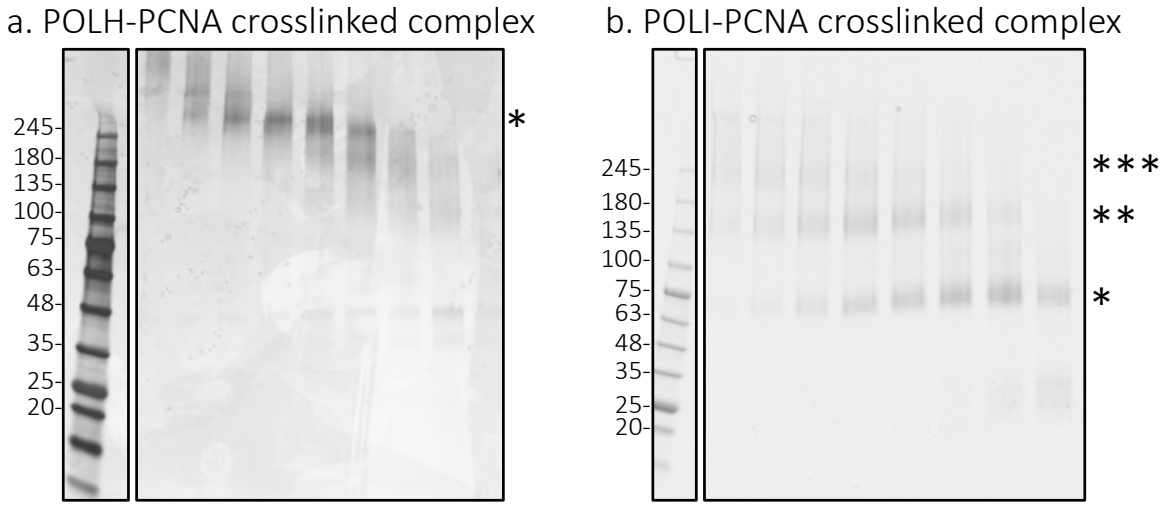


Figure 2.14 SEC analysis of crosslinked PCNA-Pols complexes. SDS-PAGE of elution fractions of glutaraldehyde crosslinked complexes. (a) PCNA- Polη crosslinked complex after SEC shows a main smeary band above the 245-marker band, marked by a star, and could correspond to PCNA trimer +3 Polη crosslinked molecules (~326 kDa). (b) PCNA- Polι crosslinked complex shows three different bands The band (*) around 75 kDa marker could be free polymerase, the above band at ~135 kDa (**) corresponds to PCNA trimer + 1 Polι (~160 kDa), and the third smeary band (***) could be PCNA trimer + 2 Polι molecules (~246 kDa).

2.3.3 PCNA monomer is not interacting with Polι in gel filtration, but slightly affects the elution of Polη

To further understand how the binding between polymerases and sliding clamp trimers is happening in solution, I performed size-exclusion chromatography of Polη and Polι respectively pre-assemble with PCNA monomer. The monomeric mutant of PCNA (K77D, C81E, K110D) was efficiently purified to homogeneity, as discussed in section 2.2.3, and consequently was subjected to complex assembly, in the same way as PCNA wild type, with both polymerases. Interestingly, no complex formation of Polι with PCNA monomer was observed (Figure 2.15 a.). From the SEC elution profile no complex peak shift was visible as the two proteins (Polι and PCNA monomer) eluted in two distinct peaks. From the analysis of peak fractions on SDS-PAGE (Figure 2.15 a.) it might appear as if PCNA monomer was co-

eluting with Polι, but that was confirmed to be an artefact due to the elution behaviour of PCNA monomer, which spread over earlier fractions also in control runs (see Appendix Figure A3). This absence of interaction of Polι with PCNA monomer is reinforcing the previous hypothesis of weak Polι interaction with PCNA. Also, it could be speculated that Polι needed a trimeric clamp interface for binding. Similarly, when I assembled complex of Polη and PCNA monomer no full peak shift was observed (Figure 2.15 b.). However, the second eluting peak presented a slight shift to the left (see star in Figure 2.15 b.), suggesting a possible transient and weak interaction of Polη with monomeric PCNA. Fractions on the gel (Figure 2.15 b.) didn't show a clear 1:1 Polη to PCNA monomer stoichiometry but a faint band around 80 kDa was detected in the late eluting fractions (Fr. 8-9-10, see star in gel of Figure 2.15 b.). This could suggest that some Polη molecules were transiently interacting with PCNA monomer in solution, even though the majority was still eluting separately. It should be emphasised that in both cases of Polι and Polη interaction with PCNA monomer, the protein concentration was very high, around 7mg/ml, a condition that would push weak protein-protein interaction. An alternative check for Polη-PCNA monomer interaction was a pulldown assay performed using N-terminally GFP tagged Polη and PCNA monomer (Appendix Figure A4). The pulldown showed a weak binding of PCNA monomer to Polη. Altogether these results would suggest that Polη does not require the trimeric PCNA quaternary structure for the interaction, but can also less efficiently bind monomeric PCNA.

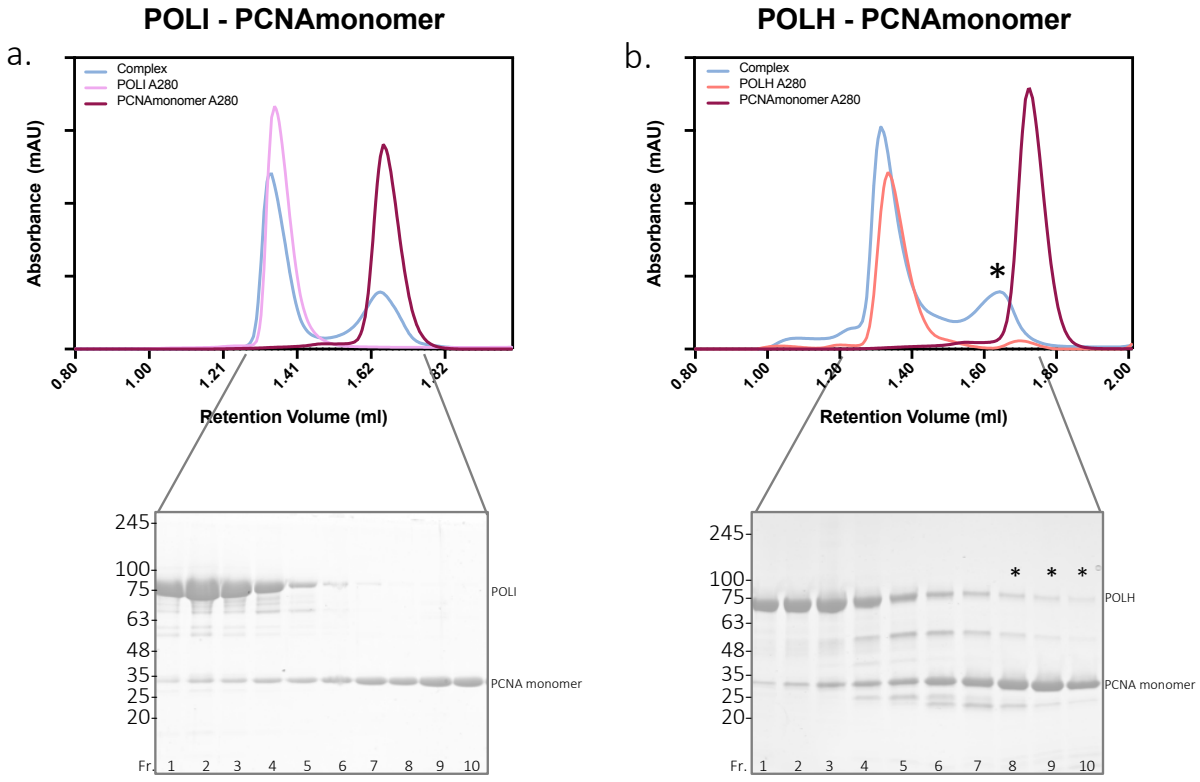


Figure 2.15 SEC analysis of PCNA monomer-polymerases. (a) SEC elution profile of Polι-PCNA monomer assembly. Blue line indicates complex run at absorbance A280 nm, pink line is Polι control (A280) and purple line is PCNA monomer control (A280). In the complex elution (blue line) there are two separate peaks eluting, and no peak shift is visible. SDS-PAGE analysis of fraction (below) shows no co-elution of Polι and PCNA monomer. The bands corresponding to PCNA monomer size in fractions 1-5 are due to a broad elution behaviour of PCNA monomer, see Appendix Figure A3 for comparison. (b) SEC elution chromatogram of Polη-PCNA monomer assembly. Blue line indicates complex run (A280 nm), salmon line is Polη control (A280) and purple line is PCNA monomer control (A280). Also no clear peak shift is observed, but the second eluting peak in the complex run (*) is slightly shifted to the left, compared to the corresponding PCNA monomer control peak (purple). In the gel (below) are fractions corresponding to the all elution profile and in fractions 8-9-10 a band around ~75 kDa is visible (marked by *) corresponding to Polη size and indicating minor recruitment of Polη to PCNA monomer, but very weak interaction results in no clear peak shift.

2.4. Structural analysis of minimal TLS complexes by cryo-EM

2.4.1 Initial negative stain observations hint that complexes disassemble on grids in absence of crosslinker

To further investigate the properties of minimal TLS complexes, I proceeded with structural analysis of reconstituted complexes using X-ray crystallography and electron microscopy. Screening for crystallisation conditions did not yield any crystals, possibly due to the high flexibility of the C-terminus of TLS polymerases, therefore this approach was abandoned. To evaluate homogeneity of samples and obtain a first low resolution model of the complex I performed initial electron microscopy analysis with negative staining. The technique of negative staining allows screening of initial conditions like particles concentrations, buffer type, protein ratios and complex assembly conditions; moreover negative staining is quite useful to quickly evaluate heterogeneity, purity, dispersion and aggregation state of proteins on grids. Polymerases complexes with UbPCNA, in presence and absence of DNA, were analysed. Firstly, grids observation followed by particle picking and 2D classification in EMAN2 (Tang et al., 2007), clearly showed that the complexes were dissociating on grids, likely due to the harsh staining procedure. The peculiar PCNA shape was predominantly visible on grids as well as in the 2D classes (Figure 2.16 a.). To tackle the problem of complex dissociation I tested samples crosslinking before application on carbon grids. The treatment of SEC preassembled Polη-UbPCNA-DNA complexes with 0.05% glutaraldehyde allowed detection of bigger particles on the grid and yielded better 2D classes (Figure 2.16 b.), which resembled crosslinked complexes. Negative staining technique is generally limited to low resolution for 3D reconstructions, but was nevertheless useful to asses initial screening conditions, particle concentration and crosslinking requirements. Having assessed that samples showed better behaviour on grids after glutaraldehyde crosslinking I proceeded to testing samples behaviour in cryogenic conditions.

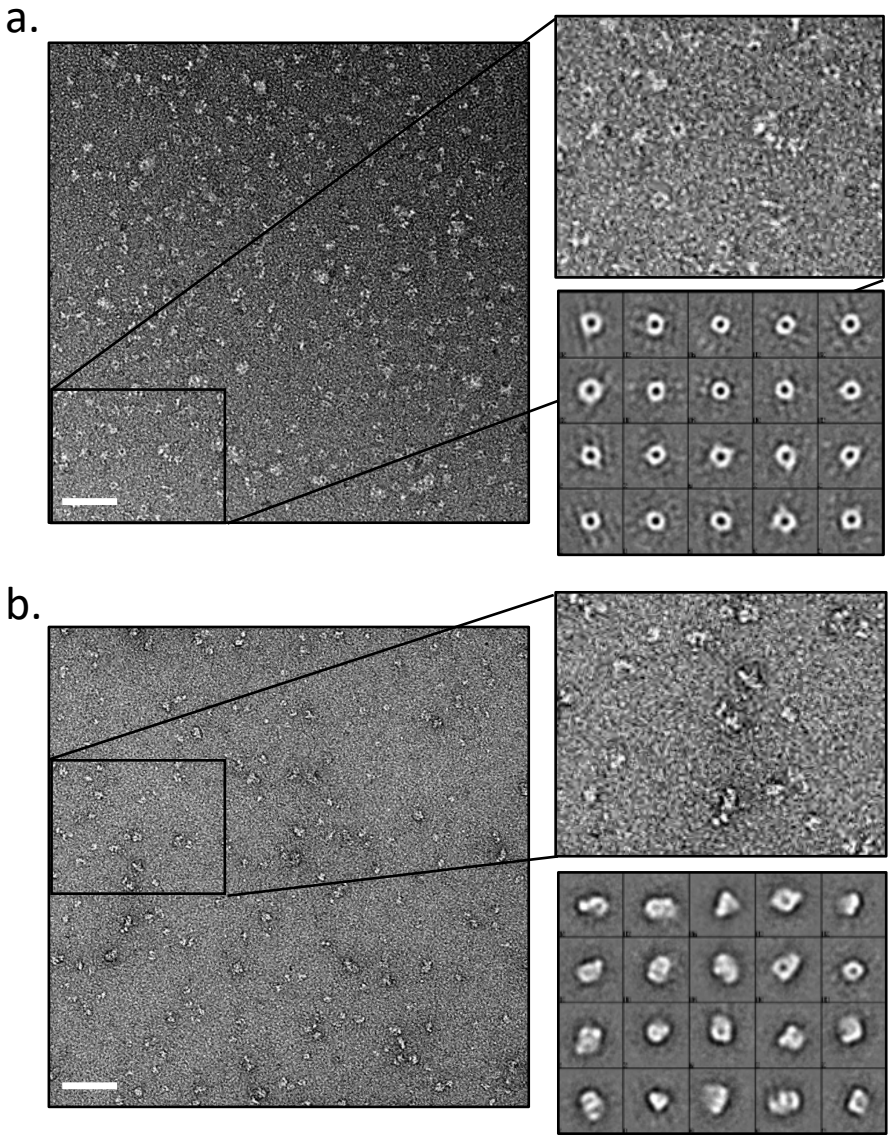


Figure 2.16 Negative stain of complexes with and without crosslinker. Representative example of negative stain EM analysis of UbPCNA-Polη complexes. (a) Micrographs of UbPCNA-Polη complex without any crosslinker shows particles (white dots) with good distribution but complex are mostly disassembled. In the zoomed in picture on the right PCNA ring is visible. The 2D classification, performed by EMAN2, show only PCNA ring features. (b) UbPCNA-Polη complex after crosslinking treatment with 0.05% glutaraldehyde for 15 minutes at RT. Micrograph (scale bar 80 nm) shows more compact particles and 2D classification depicts larger particles, where the PCNA ring is visible but additional density is also present. Scale bars 80 nm.

2.4.2 Screening of complexes in cryogenic conditions reveals sample heterogeneity

Cryo-EM sample optimisation was a long iterative process, that involved initial screening of protein concentration and ratio between complex components, as well as way of assembling the complex (either by SEC or direct mixing), DNA substrates, buffer components, crosslinking strategy, detergents etc. In order to understand which samples were the most suitable for cryo-EM high resolution data collection, several initial screening datasets for all above-mentioned conditions were collected on 200 kV Talos Arctica TEM using a Falcon 2 detector. Generally, the screening datasets were acquired overnight and were comprised of approximately 600/700 initial micrographs movies. Movies were initially corrected for beam-induced shift by MotionCor2 (Zheng et al., 2017) the contrast transfer function (CTF) correction was performed using Gctf (Zhang, 2016). Particles were initially picked either manually using EMAN2 (Tang et al., 2007) or RELION (Scheres, 2012), to get familiar with particles shapes and size, and later automatically picked by Gautomatch (Andrew Carter lab, https://www.mrc-lmb.cam.ac.uk/kzhang). Particles were extracted with appropriate box sizes and eventually binned, in RELION. After several runs of 2D classification and alignment, initial 3D model was generated either in cryoSPARC (Punjani et al., 2017a) or using RELION Stochastic Gradient Descent (SGD) algorithm (Scheres, 2016). Subsequently, particles were sorted in several rounds of 3D classification, and subjected to 3D refinement and postprocessing in RELION (Scheres, 2012) (Scheres, 2016) (Fernandez-Leiro and Scheres, 2017) (Zivanov et al., 2018).

After collecting multiple screening datasets, on different complexes, the most striking evidence from 2D classification was that complexes tended to dissociate on grids, even after crosslinking (examples of micrographs and 2D class averages of several test datasets are shown in Appendix Figure A5). In order to reduce complex dissociation on grids I systematically tested several parameters for optimization in terms of grids type, complex ratios, DNA amount, crosslinking, grids freezing strategies, detergents, etc. The choice of DNA substrate was dictated by extensive screening via Thermofluor assay (Appendix Figure A6) where higher melting temperature (T_m) indicated that complex in presence of DNA is more stable. Several DNA lengths were tested and the substrate which showed the best T_m curve was a primer/template DNA of 29/33 nt length (Appendix Figure A6). Ultimately, by using a small excess of DNA and a ratio in favour of polymerase over sliding clamp for complex assembly, as well as 0.05% glutaraldehyde crosslinking and 0.005% Tween-20 or 0.04% β-OG (Octyl-Beta-Glucosid) as detergents, I

obtained samples with a majority of intact complexes. Nonetheless, partial dissociation of the complex was always detected, likely due to the weak proteins affinity, interaction with the air-water interface, or the harsh flash freezing in liquid ethane. Conditions and level of dissociation vary from sample to sample, and are described in details for each single dataset presented in the next paragraphs.

2.4.3 Cryo-EM reconstruction of PCNA at high resolution and PCNA-Polι-DNA complex at medium resolution

After optimization of the sample behaviour in cryogenic conditions, a first dataset for high resolution reconstruction of Polι-PCNA complex was collected. The sample used for this dataset was a complex comprised of unmodified PCNA, Polι and 29/33 nt DNA in a ratio of 1 PCNA ring to 4 polymerases to 1.2 DNA. The sample was assembled in complex buffer (Table 4.7), diluted to a concentration of 0.3 mg/ml and crosslinked with 0.05 % glutaraldehyde at RT, quenched after 10 minutes and applied to cryo-EM grids, prior the application of 0.005% Tween-20, as described in Materials and Methods section 4.3.8. An overnight dataset was collected on a Titan KRIOS 300 kV TEM, equipped with energy filter and K2 camera operating in counting mode at a magnification of 105k x corresponding to 1.35 Å/pixel. Movies with 32 frames were collected, with a total dose of 42 electrons/$Å^2$ and defocus ranging from -1.5 to -3 micron. Out of 1054 movies collected, 789 were selected, based on visual evaluation of micrographs where those with excessive ice contamination or aggregation patterns were discarded. Motion correction was done by MotionCor2 and further processing was performed in RELION 2.1 (Fernandez-Leiro and Scheres, 2017). After CTF correction with Gctf, particle picking and several sorting runs in 2D classification, two main populations were observed (Figure 2.17 b. and e.). The first population corresponded to classes containing solely PCNA. The particles belonging to classes with only PCNA showed distinct secondary structure features, like the helices in the inner rim of PCNA (Figure 2.17 b.). Those particles were submitted to 3D classification and refinement. The final 3D map obtained from PCNA-only classes contained roughly 6k particles and showed a global final resolution of 6.8 Å according to FSC gold standard (Figure 2.17d.). Available crystal structure of human PCNA (PDB: 1AXC) was modelled in cryo-EM map by rigid body fitting in UCSF Chimera (Pettersen et al., 2004), showing perfect fitting of the features (Figure 2.17 c.). The only part not visible in the cryo-EM map was the intradomain connecting loop (IDCL), the flexible loop connecting the N- and C-

terminal domains of each monomer, which plays a role in polymerase binding.

Form the same dataset another population could be observe in the 2D classes (Figure 2.17 e.), corresponding to the vast majority of particle, which showed clear PCNA features bound by a more flexible and fuzzy density, corresponding to Polι. These PCNA-Polι particles were was further sorted in 2D and 3D classification and ultimately subjected to 3D auto-refine in RELION 2.1. The final set of 2D selected classes displayed high heterogeneity (Figure 2.17 e.) as 2D classification showed some dissociation of Polι form PCNA or polymerases bound with a different orientation with respect to PCNA. This heterogeneity, due to complex dissociation and different conformations, did not allow high resolution reconstruction, and a final cryo-EM map was refined to ~12 Å resolution. The medium resolution map depicted in Figure 2.17 f. showed that the PCNA moiety, resembling the donut shaped trimer, is coordinating on its side a protein of a similar size, being Polι (~80 kDa). Several contact points were visible between the PCNA and polymerase density visible in the map, but the limiting resolution did not allow any speculations. Importantly, in the dataset there was no evidence of a second polymerase binding to PCNA. Additionally, the DNA density inside PCNA's hole could also be appreciated. The relative low resolution of the complex map could be due to particles dissociation and heterogeneity, as well as different orientations and conformations of the complexes still intact in ice. In conclusion this part of the dataset, despite the low resolution, was quite ground-breaking as it was the first evidence of direct cryo-EM observation of a PCNA-Polι-DNA complex. Strikingly, PCNA alone, being a very stable and symmetric sample, showed that even a small 86 kDa protein could be visualised and refined at high resolution using cryo-EM.

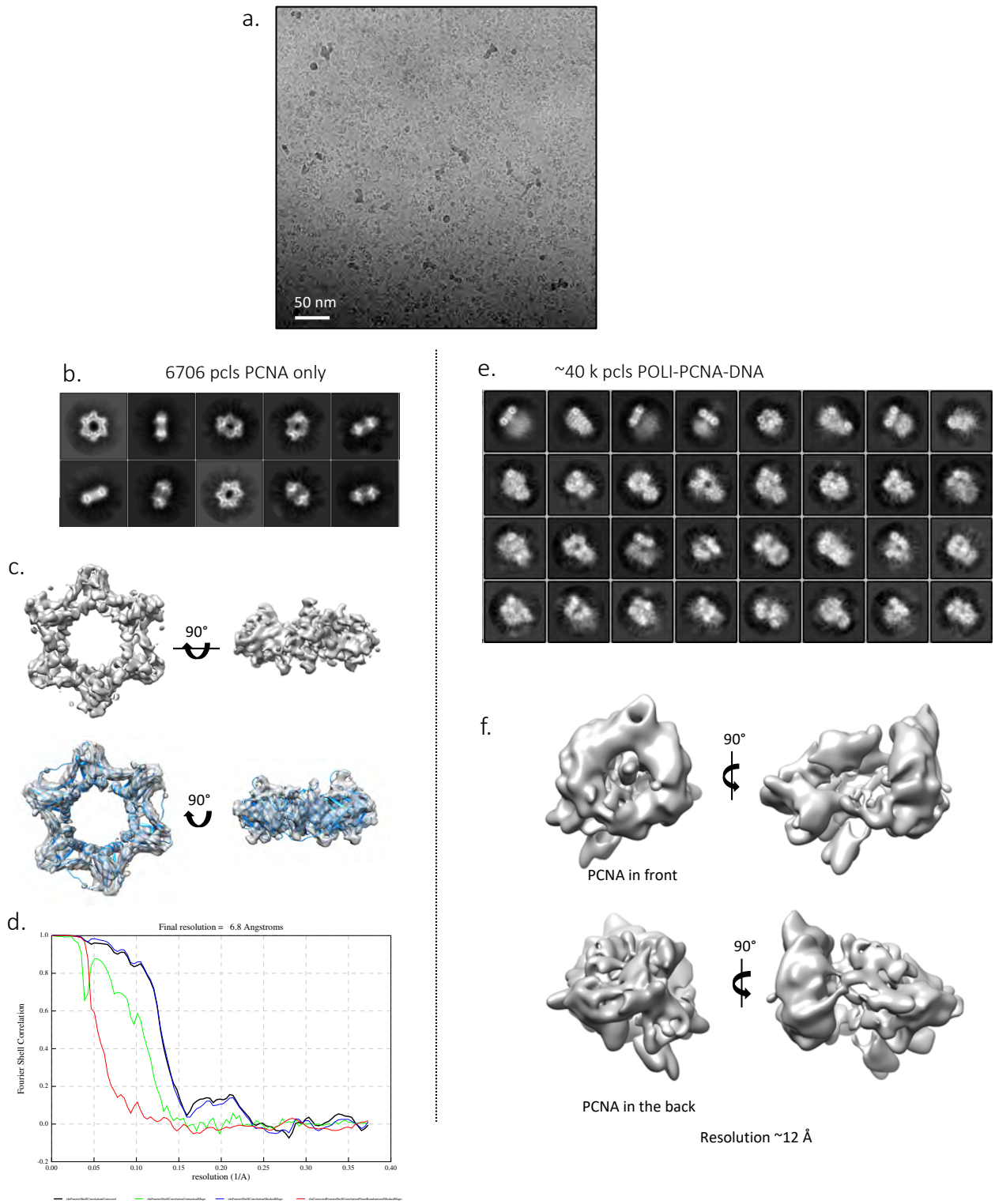


Figure 2.17 Cryo-EM analysis of Polι-PCNA-DNA complex. Cryo-EM data processing main steps of the dataset acquired on a crosslinked complex of Polι-PCNA-DNA, which showed partial dissociation on grids. (a) A representative micrograph obtained from K2 camera (scale bar 50 nm). (b) PCNA-only containing 2D classes, obtained after first sorting of the particles, comprising roughly 6k particles are a part of the whole dataset, and result from complex dissociation. High resolution features, like the helices in the inner rim or external β-sheets are visible in the 2D classes. (c) Final high resolution cryo-EM map of PCNA particles, front view and side view, showing fitting of PCNA crystal structures (PDB:1ACX). (d) Gold standard Fourier shell correlation of PCNA reconstruction, indicating a resolution of 6.8 Å at 0.143 cutoff. (e) 2D classification of most particles of the dataset,

showing complex features. The PCNA density is well refined and visible in 2D and it is bound by a fuzzy and flexible polymerase density. The 2D classes display heterogeneity in Polι positions around PCNA, indicating flexibility and different conformations. (f) Final 3D auto-refined map obtained from a subset of complex particles, showing PCNA ring features (PCNA in front). The polymerase density is occupying the orthogonal plane of PCNA and several contact points are visible between Polι and PCNA ring. Inside the ring it's visible some DNA density.

2.4.4 UbPCNA shows flexible ubiquitins in cryo-EM reconstruction

Since PCNA itself proved to be a sample suitable for cryo-EM, I subsequently tested the behavior of the more interesting UbPCNA sample on cryo grid. UbPCNA proved immediately to be a stable cryo sample, without any need of crosslinker or detergent addition prior freezing. UbPCNA at 0.5 mg/ml concentration, was applied on grids and vitrified as described in Material and Methods section 4.3.8. An overnight dataset was collected using a Talos Arctica TEM operating at 200 kV and equipped with a Falcon 2 direct electron detector. A total of 799 micrographs movies were collected at 92k x magnification (1.61 Å/pix) with a total dose of 69 electron/Å^2. The 40 frames movies were aligned using MotionCor2 with dose weighting implementation in RELION 3.0, a total of 1.8 million particles were picked by Gautomatch, extracted and subjected to several runs of sorting and classification in 2D. 2D classes of UbPCNA showed well-resolved secondary structure features, confirming the good sample quality. Strikingly, additional densities were visible (Figure 2.18) corresponding to 8.5 kDa ubiquitin moiety. The ubiquitin density could be observed at different positions in several classes indicating flexibility and movement of ubiquitin in relation to PCNA. Additionally, classes displayed considerable heterogeneity, as UbPCNA sample was partially composed of unmodified PCNA (red box, Figure 2.18 a.) and PCNA molecules with either 1, 2 or 3 ubiquitin attached (Figure 2.18 b. c. d.). This observation of several ubiquitinated species and the particle distribution in the dataset (Figure 2.18 c. d.) was in agreement with what showed previously by native mass spectrometry (see Figure 2.9). These different populations were sorted out in 2D and several initial 3D models were generated in RELION 3.0. Different 3D classes, showing PCNA with ubiquitin at variable positions were generated at a low resolution (Figure 2.18 e.). The attempt to further refine each conformation in 3D proved ineffective, regardless of the efforts in particle sorting, density subtraction and focused refinement. This was probably due to the variable ubiquitin positions in relation to PCNA and the overall small size of the particles. However, by specifically selecting the classes which showed only one ubiquitin bound to PCNA (Figure 2.18 d.), to minimize the degrees of sample complexity and heterogeneity, a better resolved 3D map was generated at a final resolution of ~7 Å, as depicted in Figure 2.18 f. This higher resolution reconstruction showed the ubiquitin moiety exposed on the side of

PCNA ring, with a slight angle towards the back face of PCNA (Figure 2.18 f.). However, this might be only one of the many conformation of ubiquitin with respect to PCNA, probably and the most favorable for particles alignment by the software.

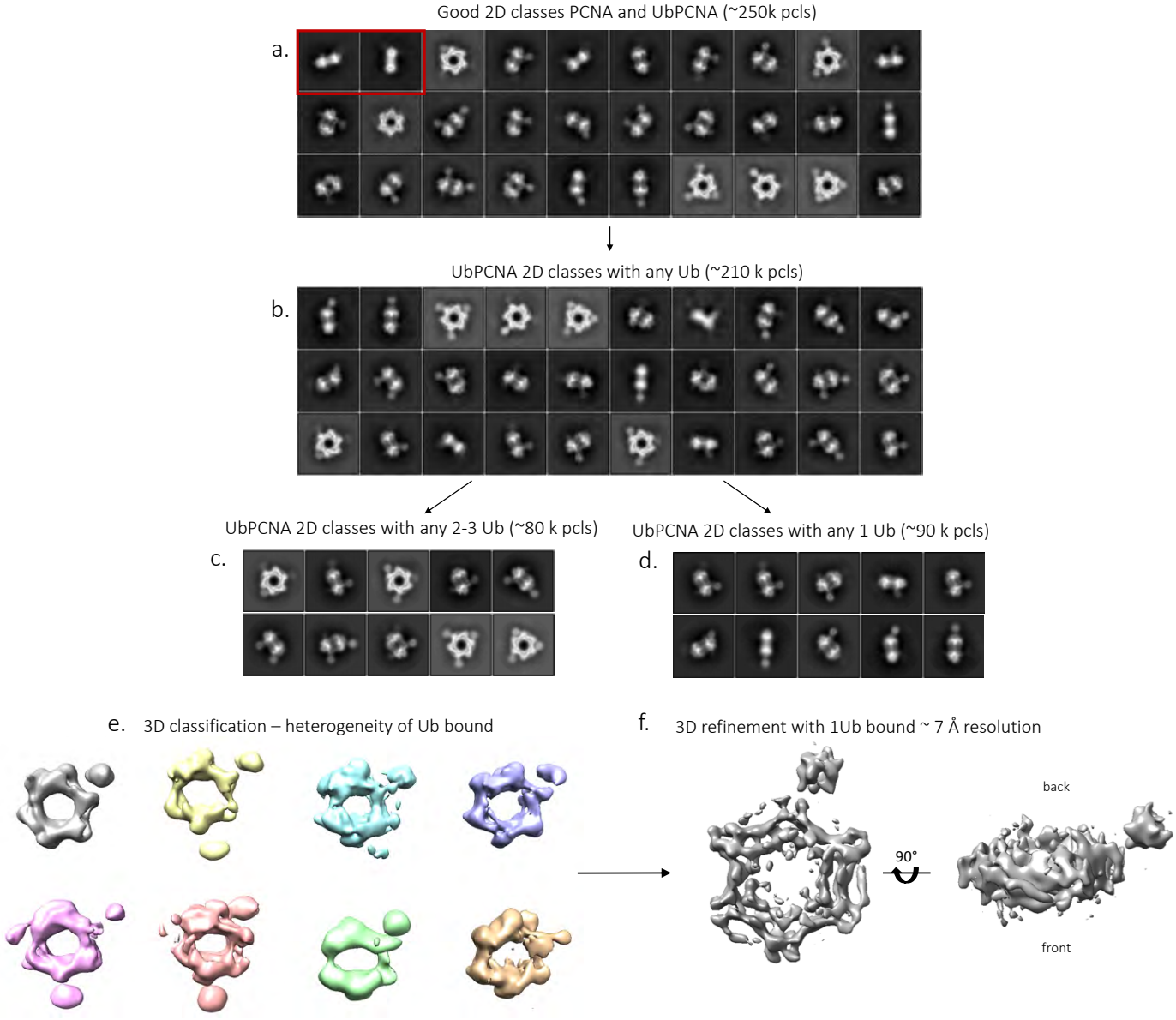


Figure 2.18 Cryo-EM of UbPCNA. 2D classes generated from processing of the dataset acquired on UbPCNA sample. (a) Initial set of 2D classes (~250k) obtained after sorting obvious junk particles (ice, carbon edges) shows well featured 2D classes, the first two (red box) are unmodified PCNA, side views. In other classes the small 8 kDa ubiquitin density is visible, occupying several positions around PCNA. (b) 2D classes which contain only UbPCNA (~210k) with either 1, 2 or 3 ubiquitins. (c) 2D classes with PCNA showing either 2 or 3 ubiquitin on. (d) 2D classes with PCNA showing only 1 ubiquitin. (e) 3D classification of particles containing any ubiquitin, the 3D classes show heterogeneity, with some classes displaying 1 blob next to PCNA (gray, violet, green) some show two blobs (yellow, blue) and some seem to have 3 blobs (pink, salmon). (f) 3D refinement of a subset of particles showing only one ubiquitin bound, displaying higher resolution features. The reconstruction resolution is ~7Å and the single ubiquitin moiety is positioned at the side of PCNA, slightly pointing towards the back face.

2.4.5 Cryo-EM reconstruction of full-length Polι bound to DNA

A second attempt to obtain high resolution data of PCNA-Polι-DNA complex was made in lights of results obtained from a Theromofluor screen, which showed a better complex stability in presence of higher amounts of DNA (Appendix Figure A6). Sample was assembled on gel filtration by mixing PCNA-Polι-DNA in a ratio of 1 PCNA ring to 3 polymerases to 6 DNA. The main peak fraction was diluted in SEC complex buffer (table 4.7) and in presence of additional 4x molar excess of DNA and dAMPNPP, crosslinked by 0.04% glutaraldehyde for 10 minutes at RT, quenched and applied to grids in the presence of 0.005% Tween-20. A dataset of 7971 micrographs movies was collected on Titan KRIOS 300 kV TEM at a magnification of 215k x, corresponding to 0.65 Å/pixel. The total dose was 98 electron/Å^2 and movies of 50 frames were recorded in counting mode with a K2 summit direct electron detector. Motion-corrected and dose-weighted micrographs by MotionCor2 were used for particle picking with Gautomatch, and all further processing steps were done using cryoSPARC. Surprisingly, 2D classes showing PCNA-Polι, that were clearly visible in previous dataset (section 2.4.3, Figure 2.17 e.), could not be visualised here. However, new 2D classes appeared (Figure 2.19 b.), showing a globular density at the extremity of a more elongated feature, which corresponded to Polι bound to DNA. A possible explanation for the unexpected disappearing of PCNA containing classes could be the diluting in presence of a large excess of DNA. By using DNA excess the Polι-PCNA complexes dissociated before crosslinking, and the higher affinity Polι-DNA complex survived intact the freezing procedure. Therefore, this unexpected yet insightful dataset showing Polι-DNA complex was further processed. The 2D class averages showed a clear Polι-DNA density (Figure 2.19 b.), and in the refined 3D map, which reached a resolution around 9 Å, the DNA density -with visible helix turns- was associated on one end with Polι density (Figure 2.19 c.). The overall resolution of Polι-DNA complex was not good enough to permit *de novo* structure modelling but it allowed a confident fit of the available Polι crystal structure (PDB: 1T3N), as shown in Figure 2.19 d. The available Polι PDB files are limited to the catalytic domain of the polymerase, since the flexible C-terminus has never been crystallised. The catalytic domain structure comprising aminoacids 26-420 (out of total 740) was manually fitted using UCSF Chimera into the EM map, showing good occupancy. Interestingly, in the cryo-EM reconstruction could be observed an additional unassigned density (Figure 2.19 d. black arrows), which could correspond to the C-terminal domain of Polι. This additional protein density was however not big enough to accommodate the missing ~300 amino acids of the C-terminus, indicating high flexibility of this region, which resulted invisible

in the EM reconstruction.

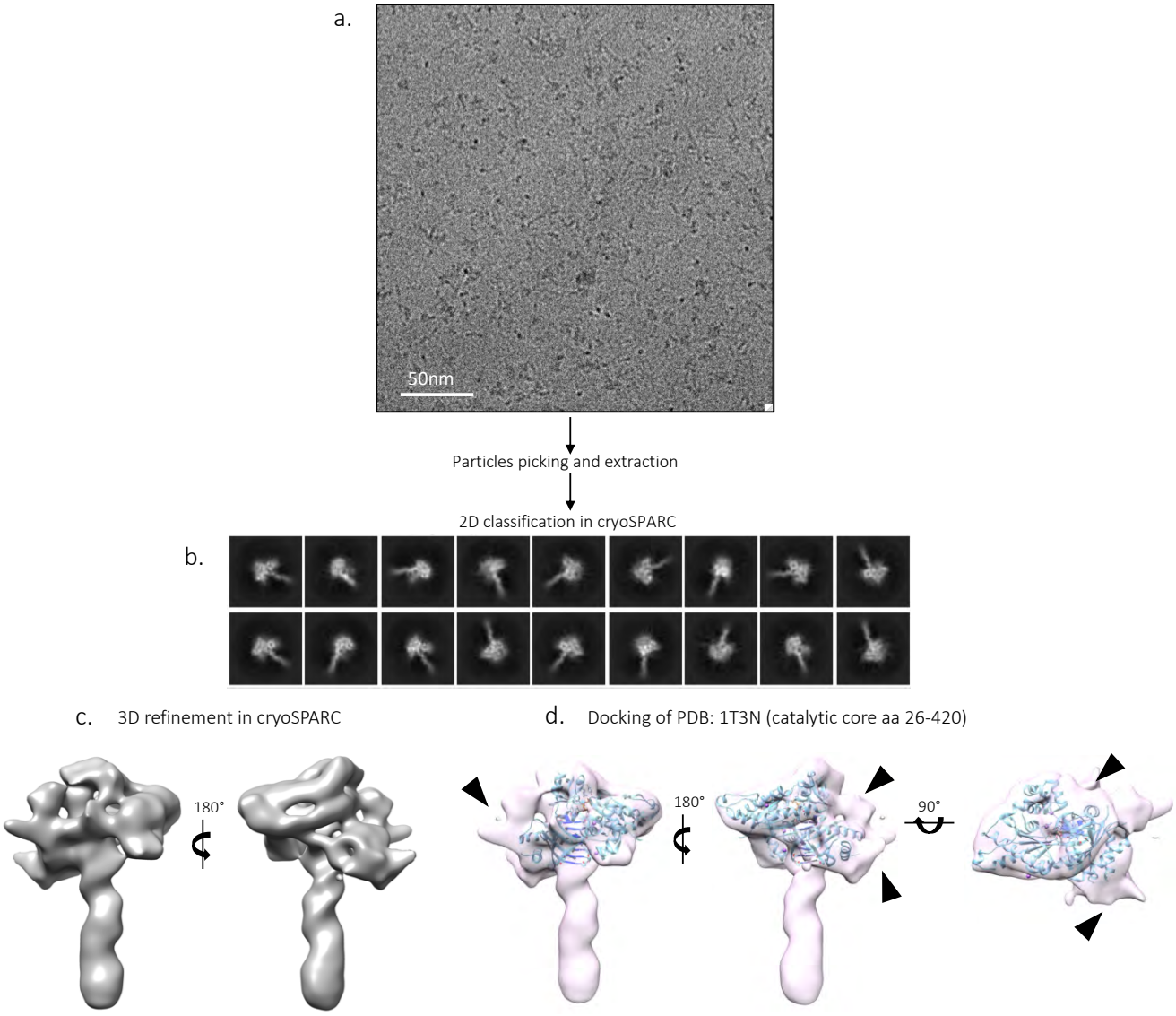


Figure 2.19 cryo-EM of Polι-DNA complex. (a) Representative micrograph (scale bar 50 nm) where the DNA is visible as black dots and rods. (b) 2D classification preformed in cryoSPARC shows clear DNA density, where the helix turns are noticeable and a globular protein density is visible on one DNA end, corresponding to Polι. (c) 3D homogeneous refinement from cryoSPARC at 9Å resolution shows good DNA density, with helix turns, and good resolution at the level of the Polι. (d) Manual fitting using UCSF Chimera of the catalytic core crystal structure of Polι (PDB: 1T3N) corresponding to 26-420 aa. The Polι the C-terminus could partially fit into the additional density visible in the cryo-EM map (indicated by black arrows). However the density is not enough to accommodate all the missing ~320 aa of the C-terminus. High flexibility of this region results in poor resolution and invisibility in the EM reconstruction.

2.4.6 High resolution cryo-EM reconstruction of Polη-UbPCNA-DNA complex proves challenging

The next attempt to obtain structural information on minimal TLS complexes by cryo-EM was endeavoured focusing on the biologically more relevant UbPCNA-Polη-DNA complex which was also the main subject of a parallel exhaustive biochemical characterisation (sections 2.5 and 2.6). A few screening datasets were collected on Arctica, in order to assess which conditions would be best for such sample in terms of proteins ratio, salt concentration, DNA amount, crosslinker and detergent but none gave any major breakthrough in terms of resolution (Appendix Figure A5). The main suspects for preventing high quality data were the glutaraldehyde crosslinking and the natural weak affinity of Polη for UbPCNA (which will be confirmed in section 2.6). Therefore, I manoeuvred some alternative ways to obtain stable UbPCNA-Polη complexes in cryo conditions by avoiding crosslinker. The following described strategies were based on parallel results obtained from biochemical characterisation (see section 2.5 and 2.6). A straightforward way to stabilise Polη interaction with UbPCNA and increase complex stability was to use the high-affinity Polη mutants pip3 and 3PIPS, specifically designed to have higher binding to sliding clamp (section 2.1.3). Furthermore, the hypothesis that a lower salt concentration in the buffer would help complex stability was tested and confirmed via pulldown assays (see section 2.6.2). Therefore, comparative cryo-EM datasets were collected for UbPCNA complexes with respectively Polη wt, pip3 and 3PIPS mutants in presence of DNA, in a low ionic strength buffer (50 mM KCl) and without any crosslinker. Besides pushing protein-protein interaction in absence of crosslinker, the lower salt buffer helped in reducing the background noise of micrographs, facilitating particles identification, although leading to higher rates of protein aggregation in thin ice (Appendix Figure A7). The three datasets (Polη wt, pip3 and 3PIPS complexes) were collected in parallel, with identical parameters in order to compare the data during processing (Appendix Figure A7). However, none of the hoped effects of lower salt or mutations were observed in the cryo-EM reconstructions. In case of Polη wt the 2D and 3D classes show only Polη and DNA features, indicating that the complex dissociated, but that the binding between Polη and DNA was maintained intact (Appendix Figure A7 a.). For both pip3 and 3PIPS mutants some PCNA-like classes were observed, both in 2D and in 3D classification, again indicating complex dissociation (Appendix Figure A7 b. and c.). The outcome of these experiment proved that, regardless of buffer ionic strength and PIP boxes mutations, crosslinking was still necessary for the complex to survive grids plunge freezing.

A final dataset, aimed at high-resolution reconstruction, was ultimately collected on UbPCNA-Polη-DNA complex, using the most optimal conditions for cryo-EM sample preparation. This condition included a ratio of 1 to 3 (UbPCNA ring to Polη) with a 20% molar excess of 29/33nt primer/template DNA, 0.05% glutaraldehyde crosslinking for 15 minutes at room temperature, and addition of 0.04% β-OG detergent right before plunge-freezing. A total of 11k micrographs movies were collected using K2 camera in counting mode on a Titan KRIOS at a magnification of 130k x, corresponding to 1.06 Å/pixel. Movies were composed of 40 frames and the total electron dose was 72 electrons/$Å^2$. The data were processed using conventional RELION 2.1 pipeline. Several iterative runs of 2D classification were necessary to remove various accidently picked particles like carbon edge and ice contamination and to isolate good particles. The data, firstly analysed via 2D classification showed heterogeneity, including dissociated complex and classes showing flexible polymerase density bound to the better resolved core of UbPCNA (Figure 2.20 b.). A subset of ~500k selected particles showing intact complex was subjected to 3D classification using as reference an initial model generated with stochastic gradient descent in RELION 2.1. Several rounds of iterative 3D classification and 3D auto-refine, as well as continuous generation of optimal initial models were carried out, in order to sort out the particles heterogeneity. Also, attempts of focused refinement by masking either PCNA or polymerase were carried out. Despite all the efforts to improve the map quality, the best result obtained was a map around 9 Å resolution (Figure 2.20 c.). The visible map details were not enough to allow *de novo* modelling. However the resolution was adequate to confidently fitting the available crystal structures of PCNA (PDB: 1AXC) and Polη catalytic domain (aa 1-432) with bound DNA (PDB: 3MR2), as well as two ubiquitin molecules (PDB: 1UBQ). The fitting was performed using UCSF Chimera. As can be observed in Figure 2.20 c., the cryo-EM map showed UbPCNA coordinating a single Polη molecule. The density corresponding to Polη located on the front side of PCNA, facing the intradomain connecting loop (IDCL). Two out of three theoretical ubiquitin molecules were manually fitted into the additional density surrounding PCNA, which was too abundant for PCNA only (Figure 2.20 d.). The additional density was visualised in the vicinity of K164 residues on two protomers (Figure 2.21 a. red stars) suggesting presence of ubiquitin in those areas. Surprisingly, the ubiquitin density didn't occupy positions at the lateral side or back face of PCNA, as described by several structural studies of UbPCNA (Freudenthal, 2010) (Zhang et al., 2012) (see Figure 2.21 b.) and as observed for the UbPCNA reconstruction in section 2.4.4. Surprisingly, the ubiquitin moiety was instead observed towards the front side of PCNA, making contacts with the polymerase density (Figure 2.21 a.). Another striking evidence was the presence of a huge unassigned

polymerase density, indicated in Figure 2.21 c. by black arrows, which could possibly correspond to the flexible Polη C-terminus making contact to PCNA and ubiquitin. Size-wise the density would fit to the unassigned 280 aminoacids of Polη C-terminus, but resolution in that area was very poor, indication high flexibility, and did not allow any modelling. Moreover, a slight angle between Polη and PCNA could be appreciated although no DNA density was visible in the cryo-EM map.

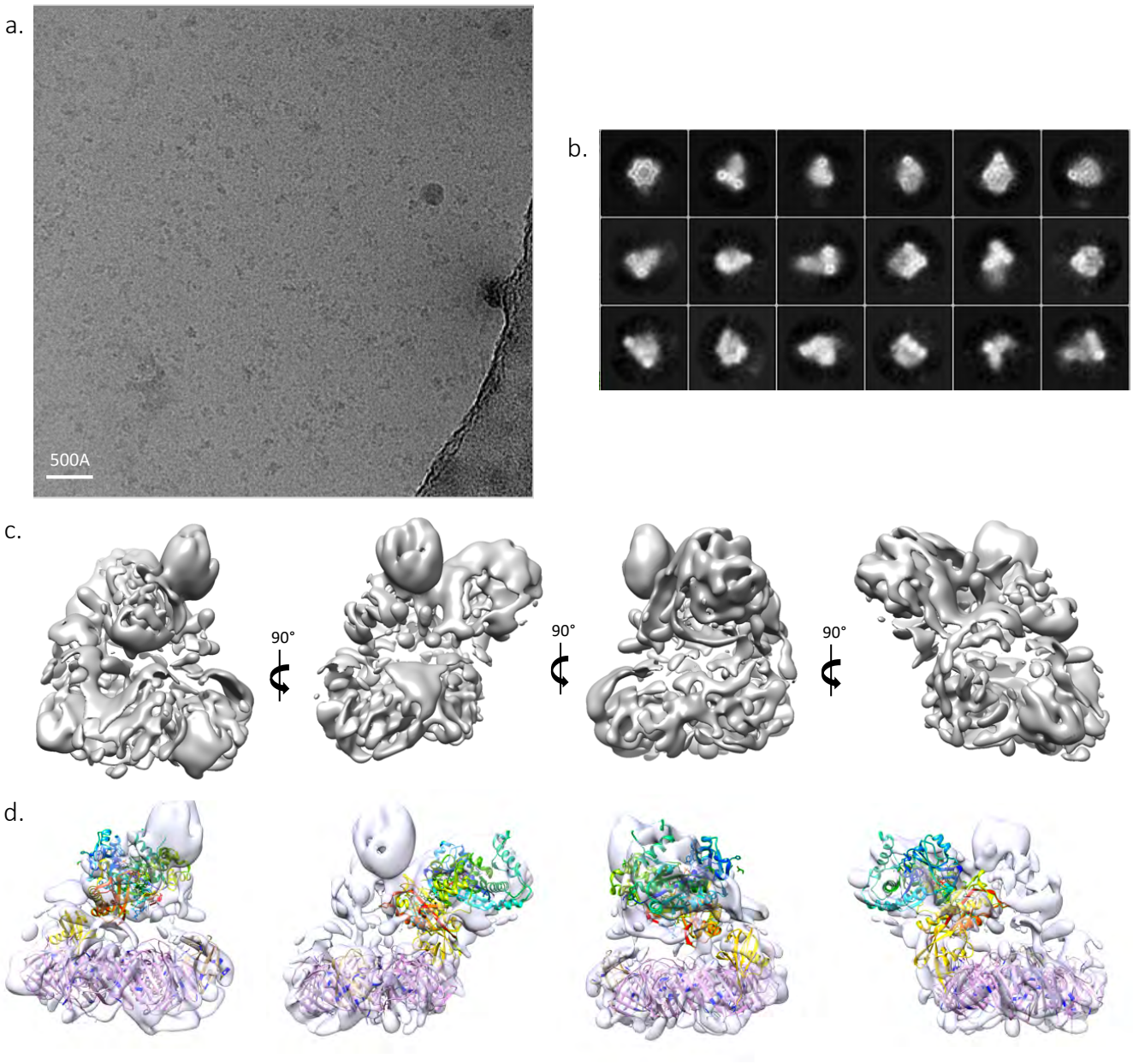


Figure 2.20 Cryo-EM of UbPCNA-Polη-DNA complex. (a) Representative micrograph of UbPCNA-Polη-DNA where particles are visible in the vitreous ice (scale bar 500Å). (b) 2D classification showing several orientations of the complex, where PCNA features are clearly visible, and the Polη density is more fuzzy and occupying different positions. (c) Final 3D auto-refinement of UbPCNA-Polη-DNA complex, where the PCNA density is

visible at the bottom (as an horizontal side view of the ring), and the polymerase is accommodated on one side. (d) Fitting of the crystal structures of Polη-DNA (PDB: 3MR2) in rainbow colour (N-terminus is blue, C-term aa 432 is red), PCNA (PDB: 1AXC) in pink(lysine residues are highlighted in blue). In the density around PCNA 2 ubiquitin molecules (PDB: 1UBQ) in yellow and brown, were manually fitted, in the proximity of K164, on two out of three PCNA protomers. A slight angle between the orthogonal plane of PCNA ring and the polymerase is visible.

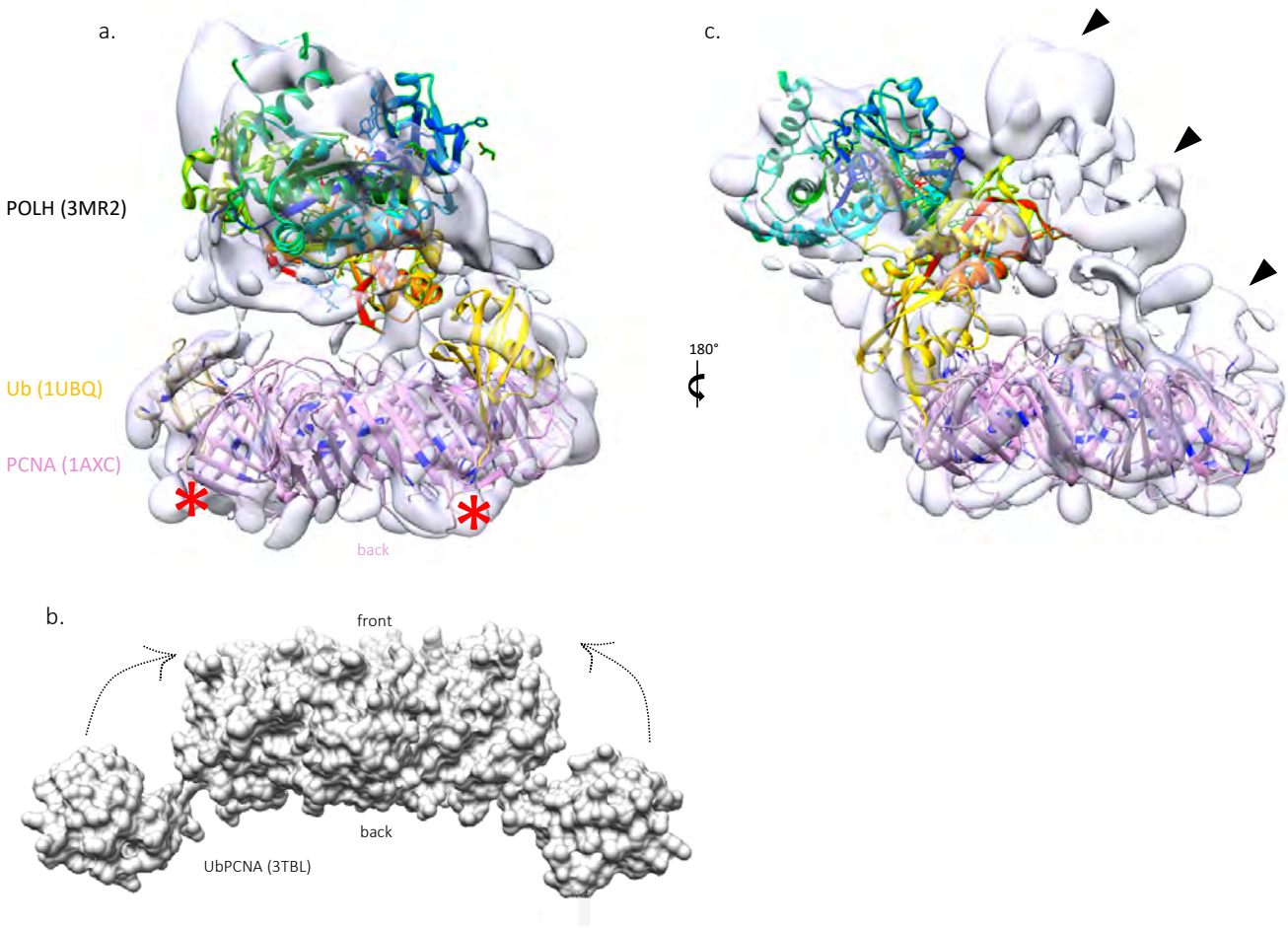


Figure 2.21 Details of UbPCNA-Polη-DNA complex. Cryo-EM reconstruction of UbPCNA-Polη-DNA with corresponding PDB files fitted in the density., zoomed in details (a) Side view of the complex where it's visible that the two ubiquitin molecules are not in the back side but flip over to the front side of PCNA, towards the polymerase density. (b) comparison model of UbPCNA crystal structure (PDB: 3TBL) where the ubiquitin is on pointing towards the back side of PCNA. (c) side view of the UbPCNA-Polη-DNA complex where a big unassigned density is visible (black arrows), which could be the flexible C-terminus of Polη making contact to the IDCL-hydrophobic pocket through PIP box and to the ubiquitin moiety through the UBZ.

2.5 Biochemical characterisation of different Polη complexes

2.5.1 Primer extension assays show Polη activity on DNA templates

As Polη is the most biologically relevant polymerase, between the two object of this study, being involved in XP-V syndrome and cisplatin related treatments, the focus was shifted to Polη for an extensive biochemical characterisation. Initially, primer extension assay was set up to verify activity of the purified polymerases. In general, primer extension allows the direct detection of polymerisation reaction *in vitro,* where the DNA primer/template substrate is elongated by the polymerase in presence of nucleotides and divalent cation. Detailed experimental procedures can be found in Material and Methods section 4.3.3. Briefly, a reaction mixture combined TDB buffer (Table 4.7), 5mM divalent cation, either $MgCl_2$ or $MnCl_2$, 500 nM DNA FAM-primer/template and 5 μM deoxynucleotide mixture. Reactions were carried out for 10 min at 37 °C and analysed by 20% TBE denaturing urea gel, which allowed to separate the FAM-labelled primer strand from the template strand and from bound proteins. The detection of the elongated primer product was performed by fluorescent imaging of the gel. Typically, a ladder product is visible as a result of primer extension, where each step represents one nucleotide addition.

Polη wild type, Polη pip3 and Polη 3PIPS mutants displayed polymerisation activity on a 12/24 nt primer/template substrate, as shown in figure (Appendix Figure A8). All proteins were active in presence of both $MgCl_2$ or $MnCl_2$ as divalent cations. In these experimental conditions Polη wt and PIP mutants did not show any difference in activity, since all could efficiently elongated the primer to full-length. The only notable difference was that the 24 nt elongated primer product (+12 nt) band was sharper and less smeary in the presence of $MnCl_2$ compared to $MgCl_2$ conditions, suggesting higher processivity of all Polη constructs in presence of manganese (Appendix Figure A8). Next, I focused on Polη wild type characterisation in presence of PCNA or UbPCNA, to test whether the presence of sliding clamp was affecting the polymerisation activity of Polη. No significant changes were detected in terms of processivity, regardless the different lengths of primer/template substrates used. A representative example is shown in Figure 2.22 where a longer substrate (9/44 nt primer/template) was used for dNTPs titration, but as can be appreciated form the gel, no difference was detected in polymerase activity in presence of PCNA or UbPCNA, compared to Polη alone.

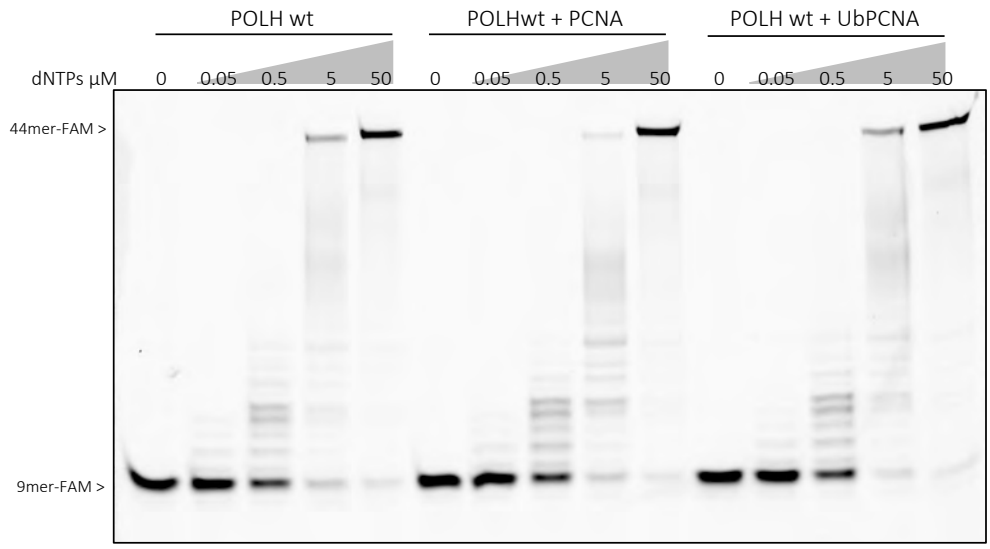


Figure 2.22 Primer extension assay, nucleotides titration. Primer extension activity of Polη on a primer/template DNA of 9/44 nt at a concentration of 500 nM with a titrations of dNTPs from 50 nM to 50 µM. Polη wt is incubated at 250 nM in the reaction mixture, either alone (first set), with 500 nM of PCNA (second set) or with 500 nM of UbPCNA (third set). No difference is visible in the reaction products comparing all three conditions.

A different strategy had to be implemented in order to detect changes in Polη activity in presence of unmodified or ubiquitinated PCNA. It is known that TLS polymerases are highly mutagenic, and even though Polη is quite faithful against cyclobutane pyrimidine dimers substrates, and efficiently incorporates correct nucleotides, the peculiar conformation of its active site makes Polη more promiscuous in nucleotide selectivity and error prone, especially against undamaged templates (Zhao et al., 2013). Aware of this promiscuous behaviour of Polη, I set up a mis-incorporation primer extension assay, where Polη was respectively incubated with all 4 different nucleotides on the same type of DNA substrate. Sets of reaction were carried out with Polη alone, in presence of PCNA and UbPCNA, respectively. As can be observed in Figure 2.23 upper panel, when the base on the template was G, the correct nucleotide C was incorporated efficiently in +1 position, and slightly extended to the +2 position. The other pyrimidine base (T) was also efficiently incorporated at +1. Interestingly, a stronger band was observed in the reaction with Polη+UbPCNA (see star in Figure 2.23) compared to reaction with Polη alone or Polη-PCNA, respectively. The purine bases A and G were also incorporated against G, confirming the error-prone nature of Polη, but there was no substantial difference between Polη alone or in presence of PCNA or UbPCNA. In Figure 2.23 lower panel, where the template base was T, the correct incoming nucleotide dATP was efficiently incorporated and a faint +2 band was visible, with no differences among Polη alone or in presence of any

sliding clamp. When the incoming nucleotide was either dGTP or dTTP, incorporation in +1 position was observed in all cases, with a slight tendency of higher dTTP incorporation in presence of PCNA or UbPCNA. Notably, when the incoming nucleotide was dCTP the +1 and +2 bands appeared stronger when Polη was working in presence of UbPCNA, compared to the other two conditions (figure * in Figure 2.23 lower panel). These results confirmed the mutagenic nature of Polη on undamaged DNA and suggested the possibility that Polη in presence UbPCNA could be more error-prone and mutagenic, compared to its normal activity without any sliding clamp.

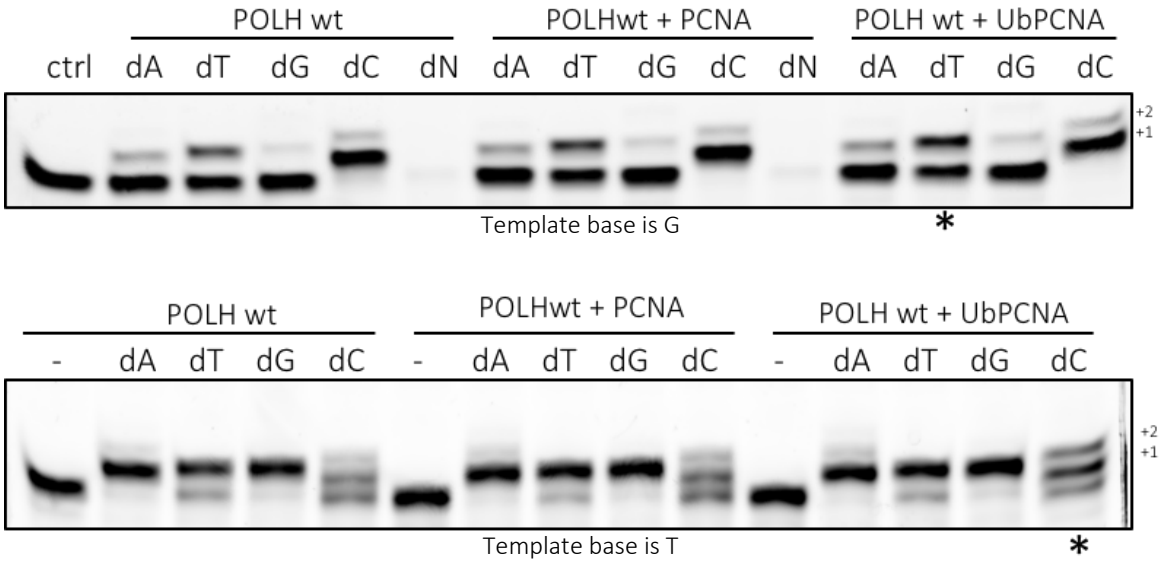


Figure 2.23 Primer extension assay, mis-incorporation. Upper panel: Primer extension activity of Polη on a primer/template DNA of with a G as template base (DNA substrates are CBs 284-201F) at 500 nM. All 4 different nucleotides (dATP, dTTP, dGTP, dCTP and dNTPs mixture) where given as substrates at 100 μM concentration. Polη wt is incubated at 250 nM in the reaction mixture, either alone (first set), with 2.5 μM of PCNA (second set) or with 2.5 μM of UbPCNA (third set). Mis-incorporation is observed in presence of all nt, stronger in cases of dT and dC. In presence of UbPCNA, Polη incorporates with higher efficiency dT against the G substrate (see *), compared to Polη alone or with PCNA. Lower panel: activity of Polη on a primer/template DNA of with a T as template base (DNA substrates are CBs 284-260F) at 500 nM. All 4 different nucleotides (dATP, dTTP, dGTP, dCTP) where given as substrates at 100 μM concentration. Polη wt is incubated at 250 nM in the reaction mixture, either alone (first set), with 2.5 μM of PCNA (second set) or with 2.5 μM of UbPCNA (third set). Mis-incorporation is stronger in the case of Polη+UbPCNA lane where dC (see *) is substrate, where the +1 and +2 bands appear stronger compared to same conditions in the Polη alone or Polη+PCNA sets.

2.5.2 DNA binding of Polη wt, pip3 and 3PIPS mutants in complex with sliding clamp

Electrophoretic mobility shift assay (EMSA) is a fast and simple method to detect protein binding to fluorescently labelled nucleic acids. Noticeably, DNA polymerases are known to interact with DNA, binding at the primer/template junction. Sliding clamp also interacts with DNA but only transiently, sliding across it. Assuming that polymerases and PCNA form a protein-protein complex, and assuming that this complex interacts with DNA, I decided to use EMSA to detect any changes in the DNA binding abilities of Polη alone or in presence of unmodified PCNA or UbPCNA. The DNA substrate used was a 24/44 nt primer/template DNA with 5'-FAM labelled primer. Proteins, either polymerases alone or in complex with PCNA or UbPCNA, respectively, were incubated at indicated concentrations (as depicted in Figure 2.24) with DNA substrate on ice for 15 minutes. Reactions were analysed on native TBE gel with 8% polyacrylamide (see material and Material and Methods section 4.3 for details). As can be observed in Figure 2.24 a., Polη wt interacted with DNA starting from a concertation of 100 nM, which would correspond to 2x the molar concertation of DNA (50 nM). At 200 mM polymerase concentration a clear shift band could be observed, indicating stable DNA-protein interaction. When the polymerase concertation was even higher (400 nM) all the DNA was shifted and the shift appeared at higher position (* Figure 2.24). Unfortunately, no difference in DNA band shift was observed when PCNA or UbPCNA were added together with Polη wt. One of the possible reasons could be that in these experimental conditions the protein-protein (polymerase-sliding clamp) interaction is transient, with the sliding clamp component dissociating from the complex and sliding off the DNA substrate. The next step was to test the DNA interaction of the Polη PIP mutants. A representative EMSA of Polη pip3 mutant with the 24/44 DNA substrate is depicted in Figure 2.24 b. Interestingly, when Polη pip3 mutant was in presence of PCNA or UbPCNA, the second shifted band became a smear (see ** Figure 2.24 b.) at higher position, suggesting that a bigger protein complex was bound to the DNA substrate. This could be explained by a stronger binding of Polη pip3 to either PCNA or UbPCNA, yielding to a more stable complex in these experimental conditions. A similar behaviour could be appreciated in the EMSA with Polη 3PIPS mutant (Figure 2.24 c.). Similar to Polη pip3, Polη 3PIPS also showed a smear (see ** Figure 2.24 c.) at higher molecular weight when the polymerase was in complex with PCNA or UbPCNA. However, in case of the 3PIPS mutant several bands at lower position were present, indicating many possible binding events of smaller molecular weight protein. This could be due to the fact that Polη 3PIPS

protein was quite unstable and showed degradation over time, as already observed during protein purification and resulting in smaller degradation products also binding to DNA.

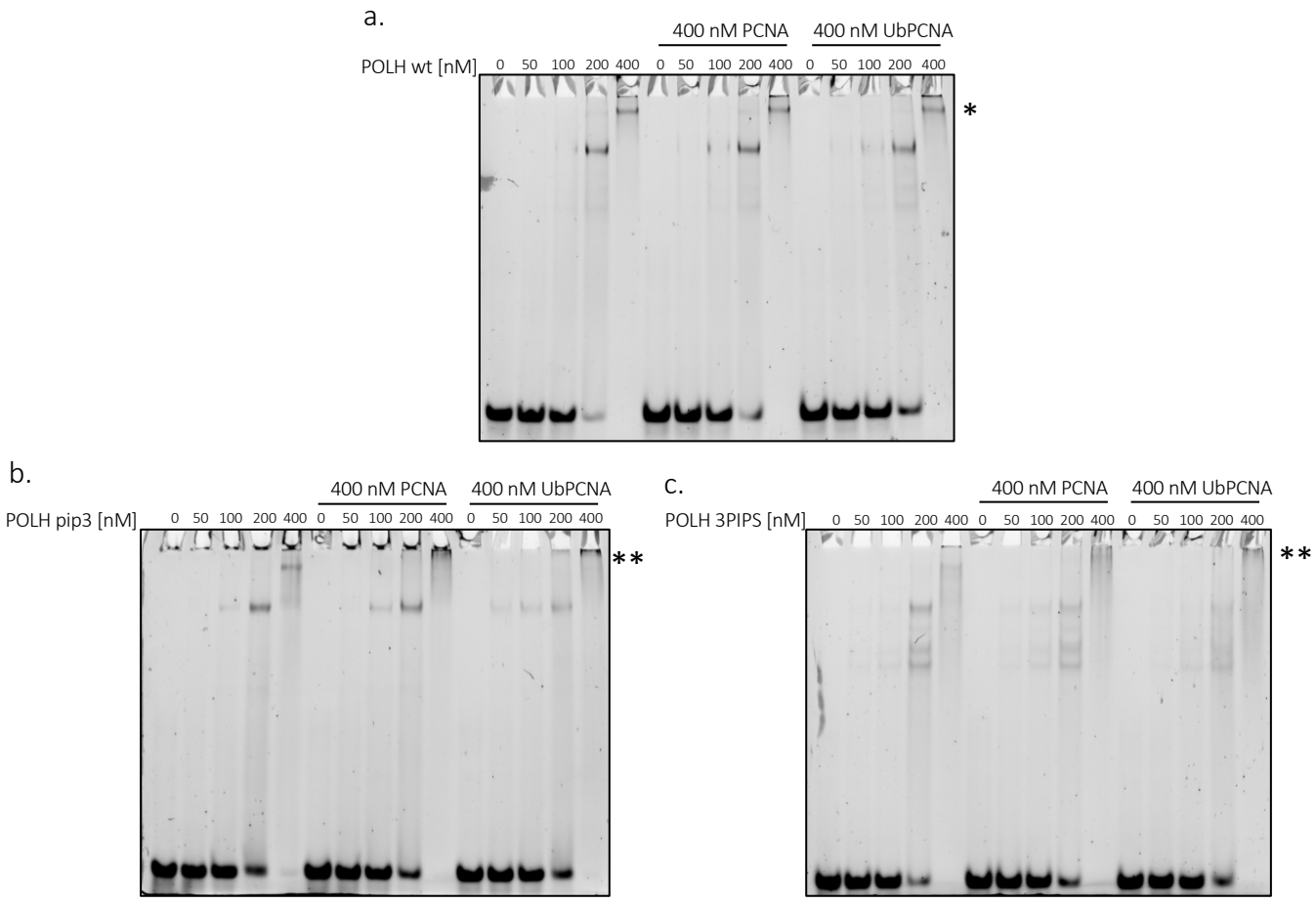


Figure 2.24 EMSA of Polη wt, pip3 and 3PIPS. (a) Electromobility shift assay of Polη wt with a primer/template (24/44 nt, substrates CBs245F-284) at 50 nM concentration. Polη is titrated as indicated in figure, and is incubated on DNA either alone (first set), in complex with 400 nM PCNA (second set) or with 400 nM UbPCNA (third set). A first shift is visible at 200 nM Polη, and a higher shift is appearing at 400 nM, but no differences are detected between Polη alone, with PCNA or with UbPCNA. (b) EMSA of Polη pip3 mutant with a the same DNA substrate at 50 nM concentration. Polη pip3 is titrated as indicated in figure, and is incubated on DNA either alone (first set), in complex with 400 nM PCNA (second set) or with 400 nM UbPCNA (third set). A first shift is visible at 200 nM Polη, and a higher smear shift is appearing at 400 nM in presence of PCNA and UbPCNA (see **), compared to Polη pip3 alone. (c) EMSA of Polη 3PIPS mutant with a the same DNA substrate at 50 nM concentration. Polη 3PIPS is titrated as indicated in figure, and is incubated on DNA either alone (first set), in complex with 400 nM PCNA (second set) or with 400 nM UbPCNA (third set). A first shift is visible at 200 nM Polη, and a higher smear shift is appearing at 400 nM in presence of PCNA and UbPCNA (see **), compared to Polη pip3 alone. Also smaller shift are visible at 200 nM, probably due to protein instability, as and these degradation products are also participating in DNA binding.

2.6 Characterising Polη binding to sliding clamp via protein pulldowns

GFP pulldowns were the method of choice to characterize the binding of each Polη construct (Figure 2.5) to PCNA and UbPCNA. The reasons for this approach were previously described in section 2.1.4. In brief, pulldowns did not require humongous amount of protein, were set up in a semi-quantitative may and allowed testing of different combinations of sliding clamp and Polη mutants in a reproducible and rapid way. N-terminal GFP tag was chosen after extensive testing of other tags, like C-terminal CPD-Twin-Strep-tag, EPEA or His tag. These tags were not optimal for experimental conditions, either leading to unspecific binding to beads or resulting in beads clumping during the procedure. The rationale was to use a N-terminal GFP tag, rather than a C-terminal tag, in order to leave the C-terminus of Polη free to capture the sliding clamp out of the mobile phase. Magnetic agarose beads GFP-Trap®_MA (Chromotek) worked effectively with all Polη constructs and did not show any unspecific protein binding, aggregation or beads clumping. I set up pulldowns in an analytical quantitative way, in order to calculate amounts of each bait and prey eluted at the end of the procedure. This allowed to estimate the amounts of bound complexes and to approximate the binding affinities of all Polη constructs to PCNA or UbPCNA. Moreover, it was possible to test protein binding in different buffers. Sequential screening of different Polη mutants allowed to investigate which parts of the regulative flexible C-terminal part were more important in recruiting, coordinating and binding either PCNA or UbPCNA.

Pulldowns assays protocol is described in details in the section 4.3.10 of Materials and Methods. In summary, all reactions were performed using LoBind tubes to minimize beads sticking or aggregation, volumes were constantly maintained to 20 µl and the full procedure was performed on ice, to avoid protein degradation. The magnetic beads were concentrated using a magnetic rack DynaMag™-2 Magnet (Thermo Fisher Scientific). Initially, 5 µl of beads slurry were washed in Pulldown buffer 1, (table 4.7). Subsequently Polη bait was incubated on beads for 1h, washed once with buffer at proper salt concentration and then beads were incubated for another hour with different preys conditions (either various protein concertation, protein type or salt concentration). Next, unbound preys were collected beads accurately washed 3x in proper buffer, according to the condition, and finally beads were dissolved in 10 µl of 1x SDS loading buffer, boiled 3 minutes at 95 °C and loaded on SDS-PAGE. Samples at different steps of the procedure like inputs, unbound baits or unbound prey and washing steps were collected

and also run on gels as controls, to confirm that protein amounts were consistent and that no material was lost during the whole procedure. Using this very accurate pulldown procedure I was able to obtain reproducible and quantitative results, confirmed by technical replicates. The first step was to establish the range of protein concentrations suitable for pulldown assay in terms of how much bait has to be bound to beads and the range of preys concertation. I determined that a concentration of 0.5 μM for Polη baits and a range between approximately between 3x and 15x molar excess for preys were working best in standard conditions. However, depending on specific purposes, concentration range varied slightly from case to case. Importantly, I also confirmed that no unspecific prey interaction was detected in all experimental conditions (Appendix Figure A9). It has to be noted that in all experimental setups, the majority of protein prey inputs was not binding efficiently to baits. If inputs and elutions bands were compared and quantified (Figure 2.25), it can be observed that only a very minor amount of preys was eluted. Specifically in case of PCNA the eluted protein amounts corresponded to ~1% of the total input, whereas in case of UbPCNA the eluted amounts where higher, in the range of 15% of total inputs (see Figure 2.25). This evidence suggested that the interaction between polymerases and sliding clamp was extremely weak, as the majority of input protein was not efficiently interacting with the baits. Nonetheless, pulldown proved a useful method in detecting binding differences of several Polη mutants to PCNA or UbPCNA, and hinted towards a higher binding efficiency of UbPCNA.

For all the pulldowns shown in the next paragraphs technical triplicates were carried out independently, quantification of bands was performed via Fiji/ImageJ (Schindelin et al., 2012) by internal normalisation of intensities values and data were plotted using GraphPad Prism version 9.0.0. (GraphPad Software, San Diego, California USA) as described in Material and Methods section 4.3.10.

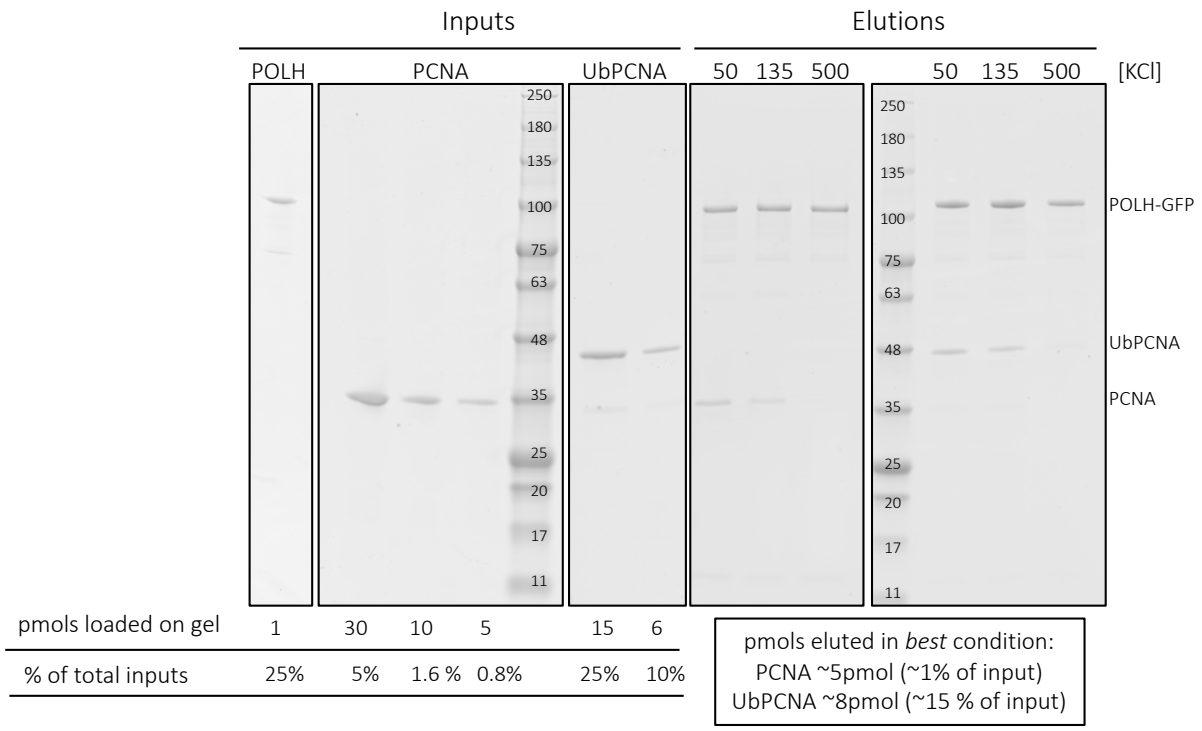


Figure 2.25 Pulldown inputs and elutions comparison. Pulldown at 3 different salt concentrations (50, 135 and 500 mM KCl) carried out with Polη as bait at 0.5 µM and PCNA or UbPCNA as preys (at 30 and 3 µM respectively) in reactions of 20 µl. On gels are loaded indicative amounts of preys (pmols loaded on gel) corresponding to a % of the total protein present in the pulldown reaction tube. The inputs band if quantified correspond to ~1% of the inputs for PCNA and to ~15% of the inputs for UbPCNA, indicating general low binding of preys to Polη. Also that a stronger binding of UbPCNA compared to PCNA is visible. Additionally the low salt conditions show higher binding compared to high salt, in all cases.

2.6.1 Polη wild type shows higher affinity for UbPCNA than PCNA

The first interesting observation was that UbPCNA showed a stronger interaction with Polη wt, compared to unmodified PCNA. As can be observed in Figure 2.26 b. lanes 5-8, the bands corresponding to UbPCNA in the elutions are more intense than the PCNA bands in the corresponding condition of PCNA pulldown (Figure 2.26 a. in lanes 6-9). Band quantification (see Figure 2.26 c.) was performed by normalising each gel internally and referring to Polη bands as maximum binding and unequivocally showed that UbPCNA is better recruited by Polη, by a factor of approximately three times, compared to unmodified PCNA.

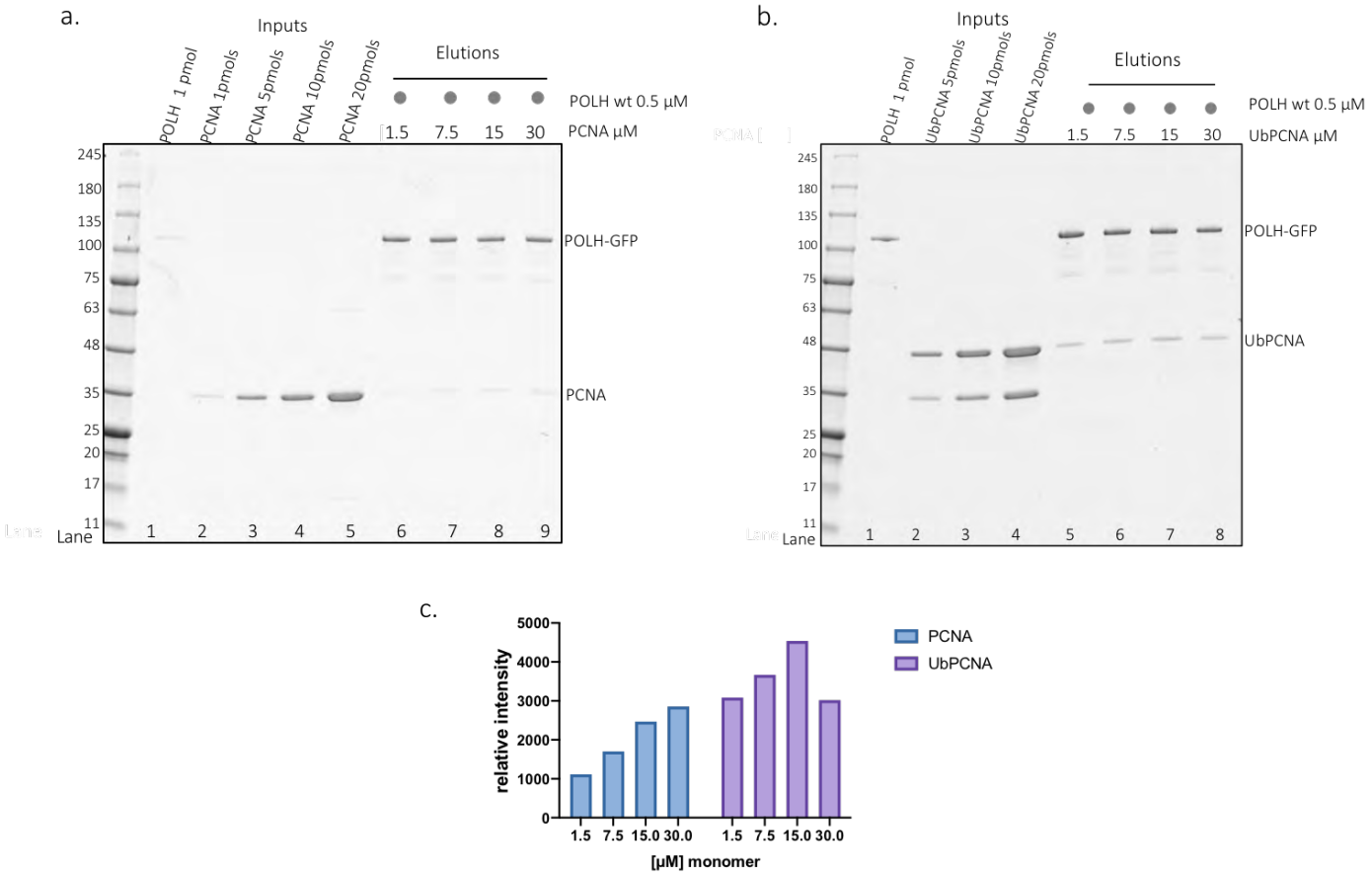


Figure 2.26 Pulldown Polη wt with PCNA and UbPCNA. (a) Pulldown gel of Polη (bait) and PCNA (prey) ranging from 1.5 to 30 µM concentration, indicated as monomer. First half (lanes 2-5) of the gel presents inputs with indicated quantities, the second half the elutions (lanes 6-9). (b) Pulldown gel of Polη (bait) and UbPCNA (prey) ranging from 1.5 to 30 µM concentration as monomer, similarly to the parallel PCNA pulldown. First half of the gel (lanes 2-4) presents inputs with indicated quantities, the second half the elutions (lanes 5-8). (c) quantification of bands, internally normalised to respective Polη, show higher binding of Polη wt to UbPCNA, compared to PCNA, in the same concentration range. The salt concentration in the buffer is 135 KCl.

Further evidence of the stronger interaction of UbPCNA came from the next set of pulldowns, where competition between PCNA and UbPCNA was tested. Briefly, such assay was performed by immobilising 0.5 µM Polη prey on beads, as described in section 2.6, and by assembling preys mixtures containing together PCNA and UbPCNA at different concentrations. As can be seen in Figure 2.27 a., when PCNA and UbPCNA were provided at the same concentration in the inputs, in a range between 1.5 and 30 µM as monomer, the elutions mainly contained UbPCNA bands. This indicated a preferential binding of UbPCNA to Polη, respect to PCNA. More interestingly, when providing imbalanced mixtures as inputs, containing a fixed amount of UbPCNA at rather low concentration (1.5 µM as monomer, corresponding to 0.5 µM trimer),

and titrating in extremely high amounts of PCNA (up to 150 μM monomeric concertation, corresponding to 50 μM trimer) the UbPCNA band were still present in the elution (Figure 2.27 b.). As can be seen in Figure 2.27 b., even when the ratio was 100 times in favour of PCNA in the inputs mixtures (see lane 4), UbPCNA was still detected. A similar observation was done when the fixed species was PCNA (kept at 15 μM monomer concentration) and UbPCNA was titrated (ranging from 1.5 to 30 μM monomer concentration) in the preys mixtures. Also in this case (Figure 2.27 c.) the most evident band in the elution was UbPCNA, again confirming a stronger preference of Polη wt for UbPCNA, compared to unmodified PCNA.

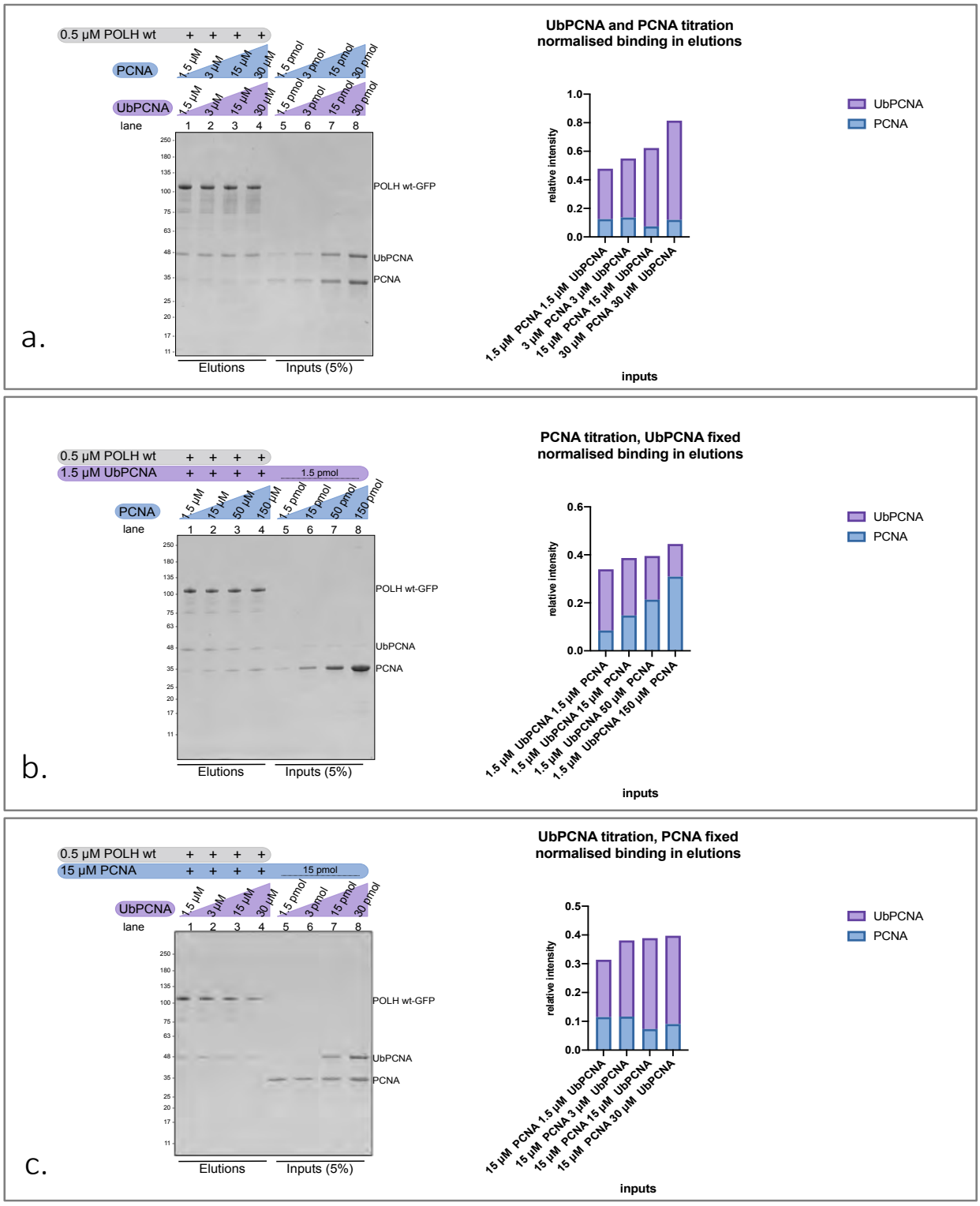


Figure 2.27 Pulldown competition of PCNA and UbPCNA. (a) Competition between same amounts of PCNA and UbPCNA, ranging from 1.5 to 30 µM (considering monomer concentration). Pulldown gel, first half shows elutions, second half shows inputs (5%). Condition of each lanes are indicated above. Quantifications of bands intensities, internally normalised to respective Pol are shown on the right side as histogram, blue represent PCNA, violet represent UbPCNA. (b) (a) Competition between same amounts UbPCNA, kept constant at 1.5 µM and PCNA titration (from 1.5 to 150µM) (all concentrations are considering monomer PCNA). Pulldown gel, first half shows elutions, second half shows inputs (5%). Condition of each lanes are indicated above. Quantifications of bands intensities, internally normalised to respective Pol are shown on the right side as histogram, blue represent PCNA, violet represent UbPCNA. (c) Competition between same amounts PCNA, kept constant at 15 µM and

UbPCNA titration (from 1.5 to 30 µM) (all concentrations are considering monomer PCNA). Pulldown gel, first half shows elutions, second half shows inputs (5%). Condition of each lanes are indicated above. Quantifications of bands intensities, internally normalised to respective Pol are shown on the right side as histogram, blue represent PCNA, violet represent UbPCNA.

2.6.2 Lower ionic strength buffer facilitates Polη binding to PCNA and UbPCNA

The following experiments aimed to find a proper concentration of both PCNA and UbPCNA, which could yield to comparable thickness of bands on gels, considering the above mentioned stronger recruitment of UbPCNA. After testing several concentrations ranges, final values were chosen and used for all the following experiments, specifically 30 µM for PCNA and 3 µM for UbPCNA (see Figure 2.28) which showed comparable band intensities in the elutions.

The next set of pulldowns was meant to test whether salt concentration played a role in protein interaction. Therefore, three different salt concentrations were tested, 50 mM and 500 mM as lowest and highest salt concentrations, and 135 mM KCl was maintained as middle concentration. The condition at 135 mM KCl was the same used in the initial pulldown assays (Figure 2.26) and PCNA-UbPCNA competition (Figure 2.27). Also, 135 mM KCl was the standard buffer salt concentration most often used in SEC complex analysis and cryo-EM (sections 2.3 and 2.4).

As can be observed in the graph in Figure 2.28, the binding of PCNA and UbPCNA to Polη wt was dependent on the buffer salt concentration, showing more intense bands at 50 mM and faint bands at 500 mM KCl. This confirmed the salt dependency of sliding clamps binding to Polη.

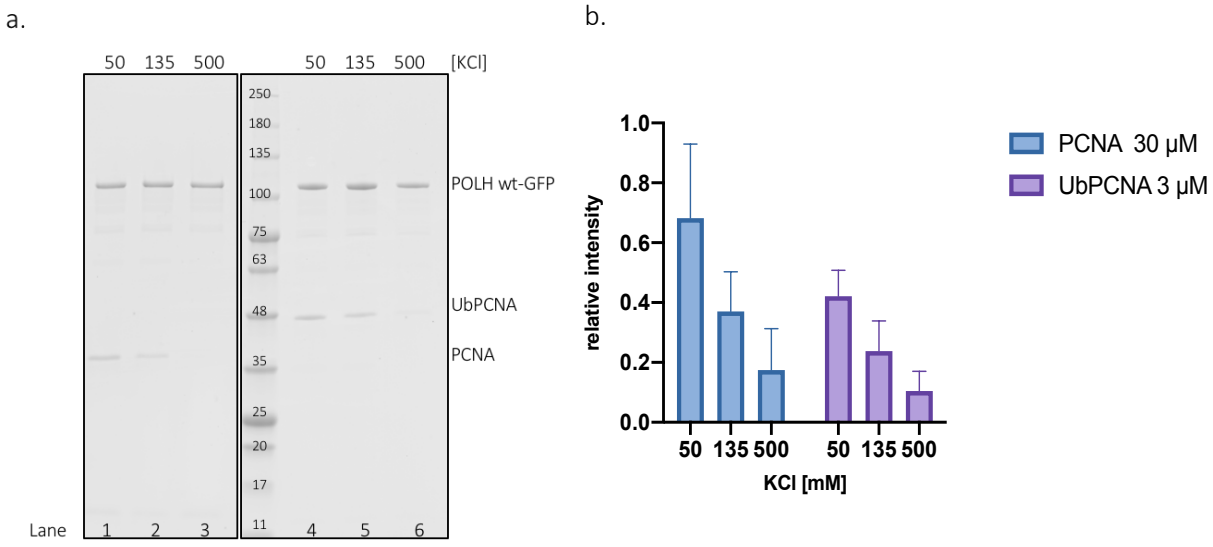


Figure 2.28 Salt dependency pulldown of Polη wt vs PCNA and UbPCNA. Elution gel of Polη pulldown with

PCNA (lanes 1-3) at 30 μM and UbPCNA (lanes 4-6) at 3 μM, at 50, 135 and 500 mM KCl respectively. (b) Gel bands relative intensities, normalised internally to Polη, are shown as histogram plot. PCNA concentration is considered as monomer.

Furthermore, all Polη constructs with C-terminal mutations either enhancing or diminishing binding to sliding clamps (Figure 2.5) were systematically tested with buffers at different salt concentrations. The salt concentrations were maintained at 50, 135 and 500 mM KCl and the Polη constructs tested were Polη wt, pip3, 3PIPS, 2PIPS, 8A, UBZ dead, pip3 dead respectively tested with PCNA and UbPCNA. As illustrated in Figure 2.29 and 2.30, all Polη mutants showed the same tendency. In every case, as can be seen from gels and quantifications, the PCNA bands (Figure 2.29) were stronger in the 50 mM KCl conditions. The same was visible in Figure 2.30, where UbPCNA bands appeared stronger on gels in the low salt conditions. This experiment confirmed that salt dependency seemed true for all Polη constructs.

This comparative experiment showed an interesting difference in behaviour of C-terminal mutants. The high-affinity mutants pip3, 3PIPS and 2PIPS showed higher recruitment of PCNA in the pulldowns, compared to Polη wild type, at all salt concentrations (Figure 2.29, compare lilac, pink and salmon bars to turquoise bars). Remarkably, the construct showing more PCNA recruitment was Polη 3PIPS, followed by pip3 and 2PIPS, possibly suggesting a potential hierarchy for PIP boxes significance. In 2PIPS mutant the absence of PIP box3 seemed to be rescued to wild type levels by the overactivation of PIP box 1 and 2. Among the C-terminal impaired mutants 8A, UBZ dead and pip3 dead, only Polη 8A showed a significantly reduced binding of PCNA, compared to wild type (Figure 2.29, compare green bars to turquoise bars). Notably, pip3 dead and UBZ dead both showed a PCNA recruitment comparable to Polη wt (Figure 2.29 - compare violet and petroleum bars to turquoise bars). In conclusion, this experiment suggested that impairment of the very C-terminal PIP box 3 could be rescued by overactivated pip box 1 and 2 and that, as expected, the inactivation of UBZ was not playing a role in recruiting unmodified PCNA. Additionally, this experiment showed that inactivation of all PIP boxes (8A construct) significantly impaired recruitment of PCNA.

Analogously to the above-mentioned experiments, when UbPCNA was given as prey (Figure 2.30) the mutants pip3 and 3PIPS and 2PIPS showed again higher recruitment of UbPCNA compared to wild type, also in a salt dependent manner (Figure 2.30 - compare lilac, pink and salmon bars to turquoise bars). Strikingly, the inactivated mutant Polη 8A proved unable to recruit UbPCNA, suggesting that a functioning UBZ and ubiquitin binding alone were not

sufficient to pulldown UbPCNA (Figure 2.30, green bars). A similar behaviour could be observed for both UBZ and pip3 dead mutants, which at all salt concentrations showed rather low ability to pulldown UbPCNA. As expected, the disruption of Polη UBZ (UBZ dead mutant, petroleum bars) affected the recruitment of UbPCNA. Quite surprisingly, also the pip3 dead Polη mutant (violet bars) showed little or no recruitment of UbPCNA, in a way comparable to the UBZ dead construct. This result could hint towards an interesting novel role of pip box 3 in the specific recruitment of UbPCNA.

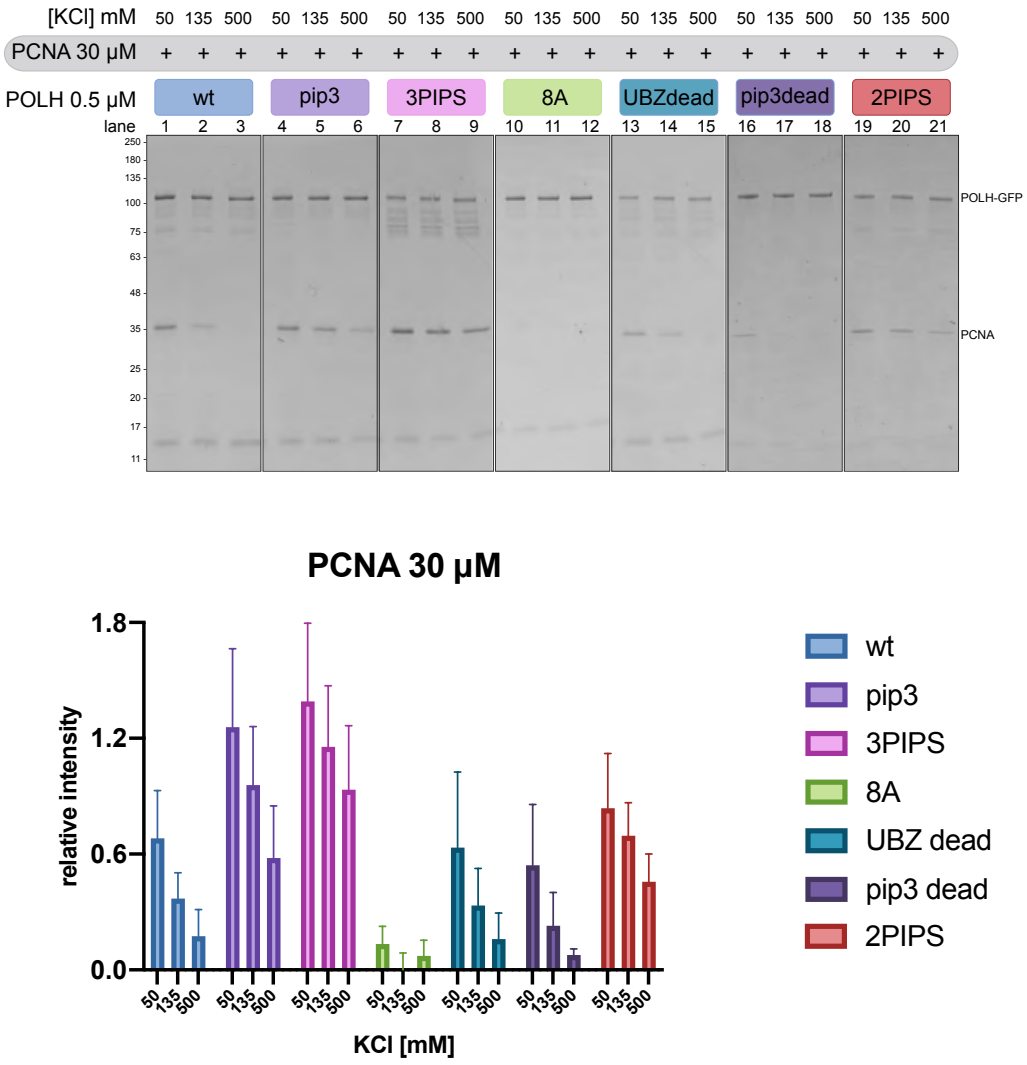


Figure 2.29 Salt dependent binding of all Polη constructs vs PCNA. Set of pulldowns with PCNA as prey at 30 µM and different salt (50, 135 and 500 mM KCl). Baits are all Polη constructs at 0.5 µM, respectively: Polη wt (turquoise), Polη pip3 (lilac), Polη 3PIPS (pink), Polη 8A (green) Polη UBZ dead (petroleum blue), Polη pip3 dead (violet), Polη 2PIPS (salmon). Representative pulldown gels elutions are shown in the upper panel, respective quantifications are shown in the lower panel, error bars are calculated as standard deviations between at least three independent experiments. Each gel is normalised independently with its internal control assuming Polη as 1.

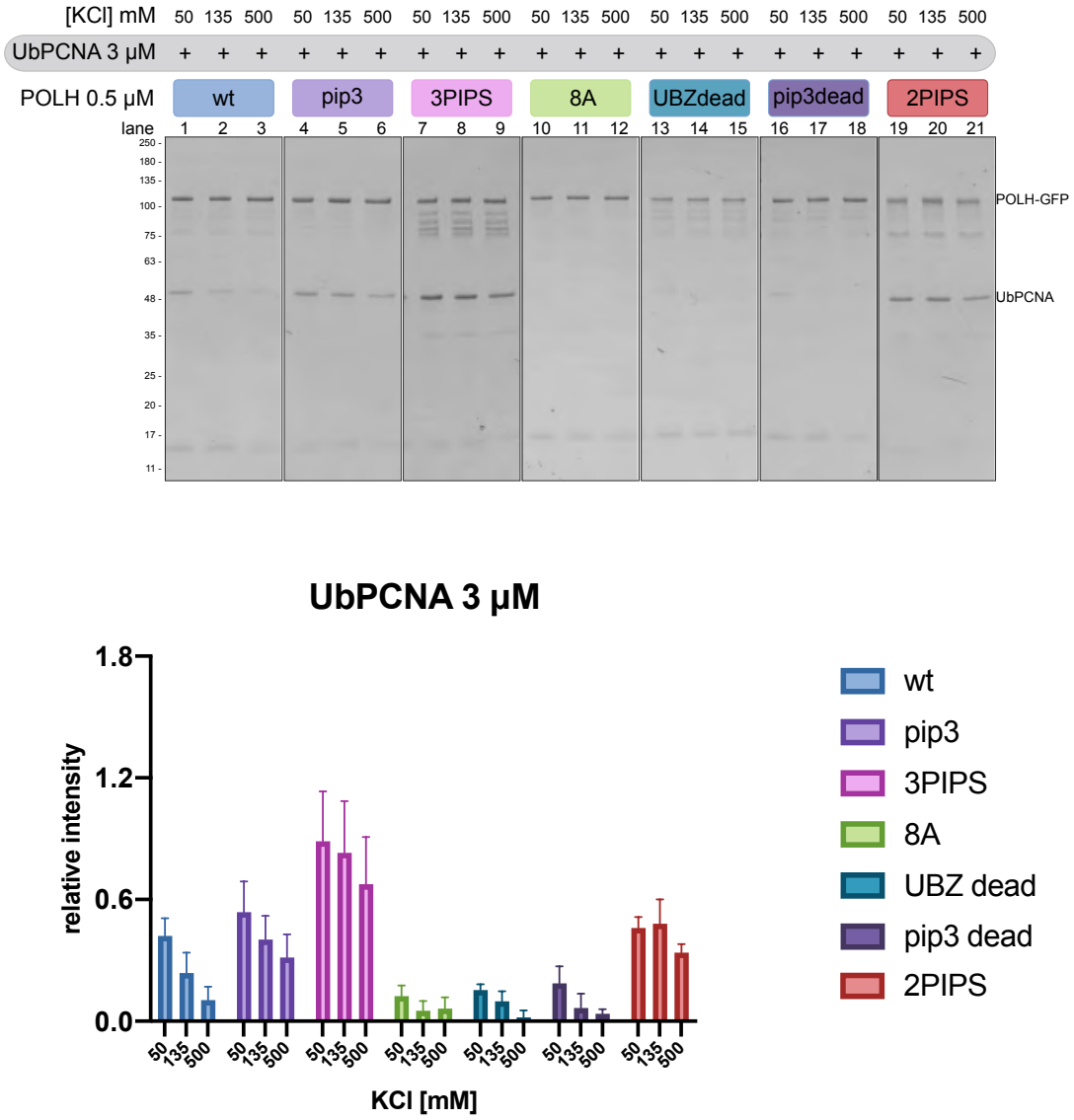


Figure 2.30 Salt dependent binding of all Polη constructs vs UbPCNA. Set of pulldowns with PCNA as prey at 3 µM and different salt (50, 135 and 500 mM KCl). Baits are all Polη constructs at 0.5 µM, respectively: Polη wt (turquoise), Polη pip3 (lilac), Polη 3PIPS (pink), Polη 8A (green) Polη UBZ dead (petroleum blue), Polη pip3 dead (violet), Polη 2PIPS (salmon). Representative pulldown gels elutions are shown in the upper panel, respective quantifications are shown in the lower panel, error bars are calculated as standard deviations between at least three independent experiments. Each gel is normalised independently with its internal control assuming Polη as 1.

2.6.3 Binding curves of all Polη constructs at low ionic strength

The next obvius step was trying to obtain binding affinities, in terms of K_d, form this optimised pulldown approach. For the following experimentsa salt concentration of 50 mM was maintained, as it proved to be most favorable for sliding clamp-Polη interaction (section 2.6.2). In parallel, as already discussed in section 2.4.6, cryo-EM analysis of Polη wt, pip3 and 3PIPS was carried out at 50 mM KCl, in the hope that such condition will lead to stronger binding in absence of crosslinker.

The protocol followed for quantification and kinetic parameters estimation from pulldown was recenlty described by Lapetina and Gil-Henn (Lapetina and Gil-Henn, 2017). However, calculating numerical values proved challenging, as non-linear regressions and application of one site specific binding model for K_d estimation did not lead to any significant value and to very unreliable R squared. Nonetheless, binding titration curves of all available constructs were carried out, since they provided good qualitative information about the general binding preferences of Polη mutants towards PCNA and UbPCNA.

Initially, Polη wt was tested at the standard concentration of 0.5 μM against PCNA and UbPCNA ranging from 0.1 to 30 μM (concentrations as monomer). The already known behaviour of Polη, preferring UbPCNA over PCNA was observed, more evident at higher clamp concentrations, as can be seen in Figure 2.31.

By comparing the binding of PCNA and UbPCNA to the high-affinity mutants pip3, 3PIPS and 2 PIPS at low salt, some intersting features came up (Figure 2.32). Polη pip3 behaved similarly to Polη wild type, showing a preferential interaction with UbPCNA (Figure 2.32 a.). The 2PIPS mutant displayed a stronger binding, if compared to wild type, but with no significant differences between PCNA and UbPCNA (Figure 2.32 b.). On the contrary and quite surprisingly, the strongest binding mutant Polη 3PIPS, showed a slightly higher affinity for unmodified PCNA than for UbPCNA (Figure 2.32 c.). These results may suggest an essential role of the three overactivated PIP boxes in recruitment of unmodified PCNA.

Additionally, sliding clamps titration curves in low salt buffer were carried out for the Polη impaired mutants 8A and UBZ dead and pip3 dead (Figure 2.33) . For Polη 8A the binding of both PCNA and UbPCNA to was completely abolished, as can be observed in Figure 2.33 a. This result confirmed what was already observed before in salt dependency pulldowns (Figure 2.29 and 2.30). This evidence suggested a crucial role of all 4 putative PIP boxes in recruitment

of both unmodified and ubiquitinated PCNA, and confirmed that the UBZ domain alone is necessary but not sufficient to pulldown UbPCNA. Unexpectedly, no or little binding was observed with either PCNA or UbPCNA to Polη UBZ dead and pip3 dead mutants, at 50 mM KCl (Figure 2.33 b. and c.). This was in contrast with what was previously observed in salt dependency pulldowns, and remains unclear. A possible explanations could be an higher insability of these proteins at low salt, resulting in worse reproducibility of binding curves.

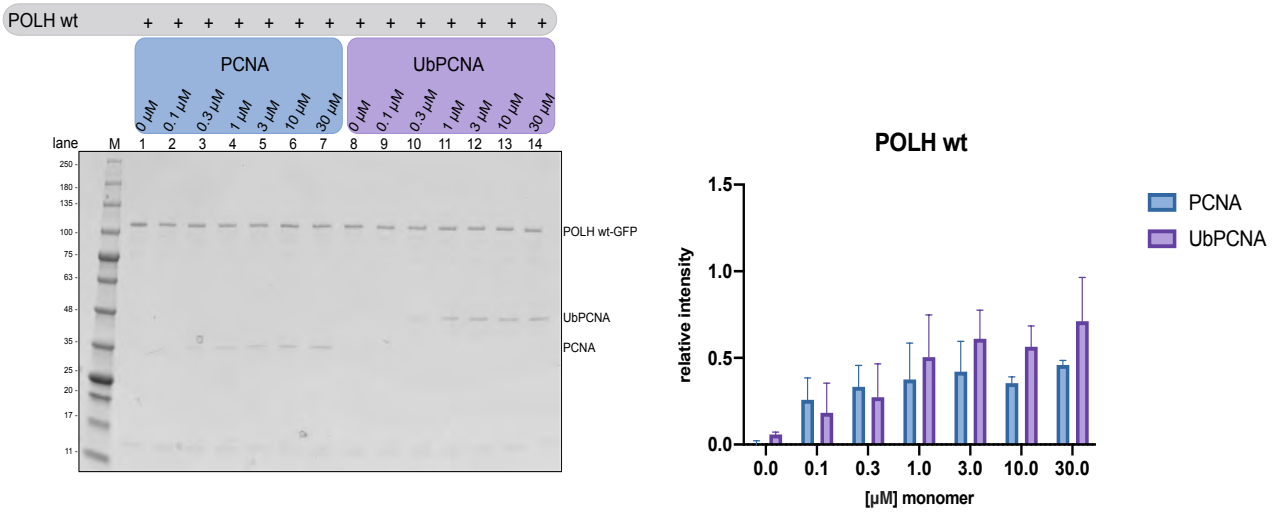


Figure 2.31 Polη wt binding at 50mM KCl. Pulldown with Polη wt as bait and PCNA and UbPCNA as preys, ranging from 0.1 to 30 µM (as monomers). The salt concentration in the pulldown buffer is 50 mM KCl. On the left side is shown a representative gel, on the right side the quantification of relative gel intensities, normalised internally to Polη band, and error bars represent standard deviation of three independent replicates.

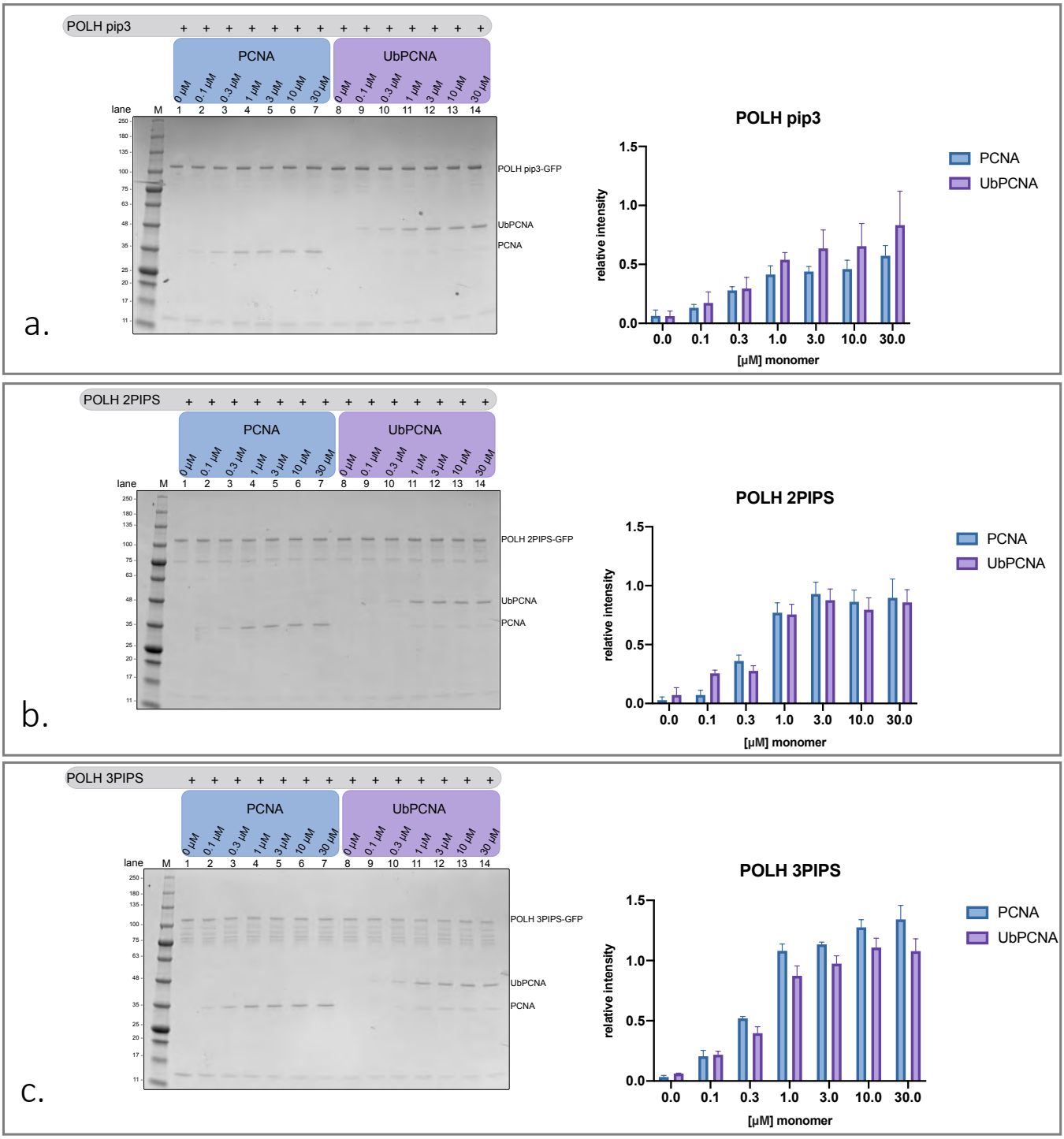


Figure 2.32 Binding curves of Polη mutants pip3, 2PIPS and 3PIPS at 50mM KCl. (a) Pulldown with Polη pip3 as bait and PCNA and UbPCNA as preys, ranging from 0.1 to 30 µM (as monomers). The salt concentration in the pulldown buffer is 50 mM KCl. On the left side is shown a representative gel, on the right side the quantification of relative gel intensities, normalised internally to Polη band, and error bars represent standard deviation of three independent replicates. (b) Pulldown with Polη 2PIPS as bait and PCNA and UbPCNA as preys, ranging from 0.1 to 30 µM (as monomers). The salt concentration in the pulldown buffer is 50 mM KCl. On the left side is shown a representative gel, on the right side the quantification of relative gel intensities, normalised internally to Polη band, and error bars represent standard deviation of three independent replicates. (c) Pulldown with Polη 3PIPS as bait and PCNA and UbPCNA as preys, ranging from 0.1 to 30 µM (as monomers). The salt concentration in the pulldown buffer is 50 mM KCl. On the left side is shown a representative gel, on the right side the quantification of relative gel intensities, normalised internally to Polη band, and error bars represent standard deviation of three independent replicates.

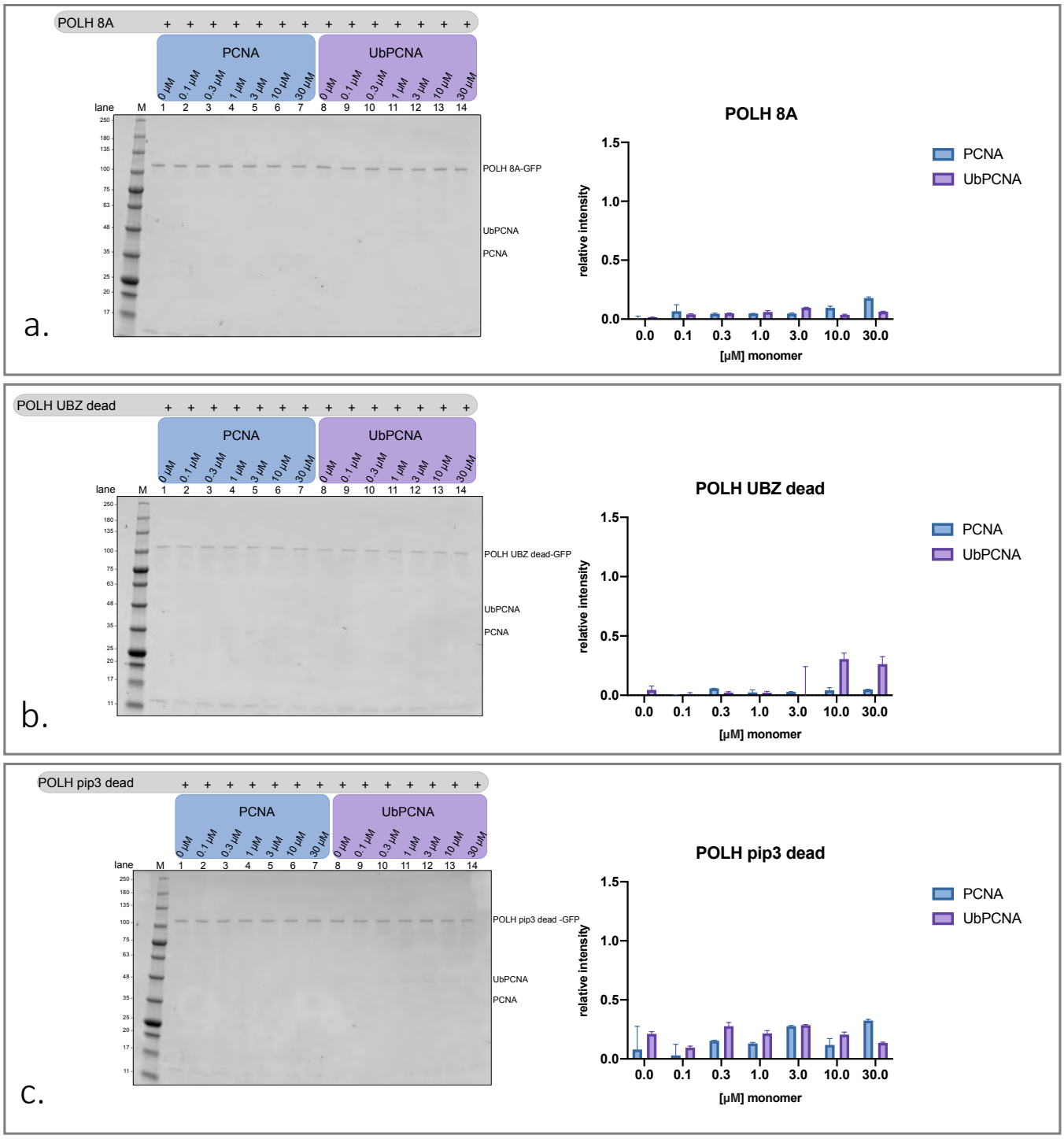


Figure 2.33 Binding curves of Polη mutants 8A, UBZ dead and pip3 dead at 50mM KCl. (a) Pulldown with Polη 8A mutant as bait and PCNA and UbPCNA as preys, ranging from 0.1 to 30 µM (as monomers). The salt concentration in the pulldown buffer is 50 mM KCl. On the left side is shown a representative gel, on the right side the quantification of relative gel intensities, normalised internally to Polη band, and error bars represent standard deviation of three independent replicates. (b) Pulldown with Polη UBZ dead mutant as bait and PCNA and UbPCNA as preys, ranging from 0.1 to 30 µM (as monomers). The salt concentration in the pulldown buffer is 50 mM KCl. On the left side is shown a representative gel, on the right side the quantification of relative gel intensities, normalised internally to Polη band, and error bars represent standard deviation of three independent replicates. (c) Pulldown with Polη pip3 dead mutant as bait and PCNA and UbPCNA as preys, ranging from 0.1 to 30 µM (as monomers). The salt concentration in the pulldown buffer is 50 mM KCl. On the left side is shown a representative gel, on the right side the quantification of relative gel intensities, normalised internally to Polη.

2.6.4 Two different polymerases recruitment by PCNA or UbPCNA

A final question was whether different types of TLS polymerases could be simultaneously recruited by either PCNA or UbPCNA. From size-exclusion experiments where both Polη and Polι were premixed with (Appendix Figure A10) the complex elution peak did not show any further shifts, compared to standard complexes, thus making any speculations on simultaneous binding of different polymerases to UbPCNA impossible. Additionally, both polymerases have the same molecular weight (~80 kDa) hence are undistinguishable on gel. Furthermore, direct interaction of Polι and Polη was tested over gel filtration (Appendix Figure A11) but no peak shift compared to controls was observed, indicating either no binding or too transient interaction to be detected by SEC. This hinted in support of the hypothesis that ubiquitination of polymerases could play a role in their mutual interaction (Vaisman and Woodgate, 2017).

Since SEC analysis did not bring definite results, I designed a double polymerase pulldown, where the immobilised bait Polη wt was exposed to either PCNA-Polι or UbPCNA-Polι preassembled complexes, as preys. As can be observed in Figure 2.34, a faint band, corresponding to Polι size, was present in the elution of lane 6, and only there. Polι band was not eluted in any of the other conditions (lanes 2-3-5). Importantly no double polymerase was detected in the elutions when the PCNA-Polι complex was preassembled and applied to Polη bait (lanes 2-3). Importantly, only when UbPCNA-Polι complex was premixed (with an excess of Polι - lane 6) the second polymerase (Polι) appeared in the elutions. This result suggested that UbPCNA, but not unmodified PCNA, is a strict condition to coordinate different types of polymerases on the same ring. Also, the fact that the recruitment was detectable only at high Polι concentrations (15 μM) points towards the already observed behaviour of Polι, which seems to have a low occupancy on sliding clamp trimers, with a ratio of one or two Polι per ring. Importantly, this experiment also highlights the highest affinity of Polη to the trimers, as Polη is effectively outcompeting Polι for the recruitment of sliding clamp.

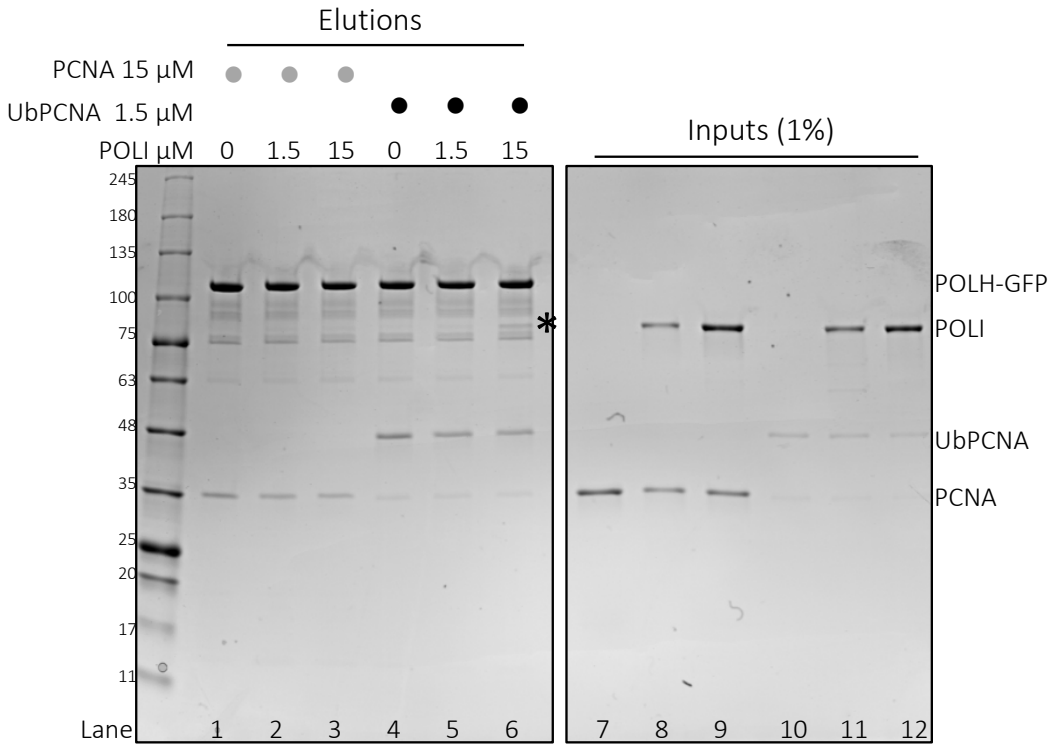


Figure 2.34 Multiple polymerases pulldown. Pulldown experiment performed with Polη wt as a bait at 0.5 µM. First gel shows elutions, second gel shows inputs for each condition. Preys are either PCNA at 15 µM as monomer (lane 1), preassembled PCNA (15 µM) + Polι (1.5 µM) (lane 2) or PCNA (15 µM)+ Polι (15 µM) (lane 3). Preys are: UbPCNA at 1.5 µM as monomer (lane 4), preassembled UbPCNA (1.5 µM) + Polι (1.5 µM) (lane 5), UbPCNA (1.5 µM) + Polι (15 µM) (lane 6). The only condition where a Polι band is visible (see *) in the elution is lane 6, which corresponds to 1.5 µM UbPCNA and 15 µM Polι.

3. Discussion

Translesion synthesis (TLS) is a DNA damage tolerance mechanism aimed to bypass DNA lesions, by exploiting specialised DNA polymerases. These polymerases, thanks to their spacious active sites and low nucleotide selectivity, are capable of incorporating nucleotides across the lesion, at the expense of accuracy. Among the 17 DNA polymerases identified in human cells, TLS polymerases mainly belong to the Y-family (Sale et al., 2012), namely Polη, Polι, Polκ and Rev1, together with a B-family enzyme Polζ (Gan et al., 2008). The recruitment of TLS polymerases to damage sites requires the central platform proliferating cell nuclear antigen (PCNA). In eukaryotes PCNA forms an homotrimer in toroidal shape which can slide on DNA and act as molecular hub for many proteins involved in DNA replication and repair. Each PCNA protomer consists of two homologous domains that are connected by the interdomain connecting loop (IDCL), which together with the underlying hydrophobic pocket constitutes the major anchor point for PCNA binding partners via their PCNA interacting peptides (PIP) boxes. The various PCNA interactions and functions in DNA metabolism are regulated via post-translational modifications. PCNA mono-ubiquitination on K164 is a key modification conserved in eukaryotes from yeast to human. PCNA-K164 ubiquitination provides a specialised platform for interaction with TLS polymerases, which have ubiquitin binding domains in the structurally disordered C-terminal regulatory region. Besides their ubiquitin binding motifs, the regulatory C-termini of TLS polymerases also contain non-canonical PIP boxes, which diverge from the consensus sequence, required for their binding to PCNA. In this thesis I presented a study on *in vitro* reconstituted minimal translesion synthesis complexes, mainly focused on Polη and Polι. With extensive biochemical and structural characterisation, this work shed light onto the protein-protein interactions and organisation within the intricated TLS network system.

3.1 TLS pols interaction with sliding clamp: relevance of PCNA trimeric state and ubiquitination

To understand the molecular interaction network of TLS polymerases and sliding clamp, the first step was to obtain the components in a near full-length state and to reconstitute stable sub-complexes. My results showed that active full-length human proteins could be successfully purified from bacterial expression system after optimizing expression and purification strategies (section 2.1). First of all, Polι was purified with a N-terminal solubility His-SUMO tag and large amounts of high purity protein were obtained. The quality of the prep was verified by primer extension assay, SEC and cryo-EM analysis (sections 2.3 and 2.4). Secondly, a novel two-step purification procedure was put in place for Polη. The codon optimized human Polη gene was expressed with a C-terminal self-cleavable solubility tag (CPD) followed by a Twin-Strep tag. Pure Polη was immediately obtained after the first affinity purification step. Primer extension assays showed that the enzyme activity was comparable to the well-characterized catalytic cores (Biertümpfel et al., 2010) confirming the good quality of the protein preparation. Furthermore, a similar purification strategy was applied to human PCNA. By exploiting the combination of C-terminal CPD-Twin-Strep tag, it was possible to obtain high amounts of very pure protein directly after affinity and size exclusion chromatography. In addition, a set of C-terminal mutants of Polη (Figure 2.5) was purified via a N-terminal GFP-Twin-Strep-tag, which was left un-cleaved to be used as affinity tag for pulldown experiments (described in section 2.6). Finally, I optimised the procedure for enzymatic ubiquitination of PCNA, using the promiscuous E2 enzyme UbcH5c to catalyse mono-ubiquitination of PCNA on K164 in the absence of a E3 ligase (Hibbert and Sixma, 2012). The reaction yielded to preparative amounts of ~ 90% pure UbPCNA. However, regardless of all the efforts in reaction optimization and purification, it was never possible to achieve a final sample composed solely of UbPCNA, as the reaction, whilst being very efficient, was still not fully converting all substrate into product. The reason for this partially incomplete PCNA ubiquitination could be due to the limiting capacity of UbcH5c to attach a single ubiquitin on each monomer of PCNA trimer in solution. In the literature there are contrasting opinions on whether PCNA ubiquitination occurs efficiently on each single monomer and which E2-E3 system is actually responsible for sequential mono-ubiquitination *in vivo* (Kanao et al., 2015). The UbPCNA sample heterogeneity was further visualised by cryo-EM analysis and will be discussed in section 3.2.

Having all players in hand - polymerases, unmodified PCNA and mono-ubiquitinated PCNA (UbPCNA) - I reconstituted minimal TLS complexes via size exclusion chromatography. As described in section 2.3, polymerase and sliding co-eluted on gel filtration, resulting in an evident peak shift when the proteins were mixed together in complex buffer, indicating complex formation. The peak shift to the left suggested formation of higher molecular weight particles, which eluted at lower retention volumes compared to single protein controls. All the combinations of polymerases (Polη and Polι) and sliding clamp (PCNA and UbPCNA) were tested, proving that complexes on SEC were forming regardless of PCNA ubiquitination state (see section 2.3.2). In addition, TLS complex formation was also tested with and without DNA. The results showed that the binding behaviour did not change in SEC, suggesting that the interaction between polymerases and PCNA can be maintained even without the presence of DNA. A first evidence from these set of experiments was that protein affinity between polymerases and sliding clamp was rather low, requiring high concentrations of proteins in order to obtain a co-migration on SEC. In fact, if the complex was assembled at low concentrations (e.g. 1 μM) the proteins would not co-migrate in a single peak, instead elute separately. Evidently, low concentrations were not ideal for protein interaction in solution, resulting in complex falling apart. A reason for this could be a weak or dynamic polymerase-sliding clamp interaction, with complexes forming and dissembling quickly in solution, a concentration dependent manner. This situation may reflect the dynamic and unstable nature of TLS complexes *in vivo*. TLS polymerases likely necessitate a quick recruitment to PCNA to incorporate across the lesion, but they also need a prompt displacement form the sliding clamp, because of their high mutagenicity.

Regarding the stoichiometry between sliding clamp trimer and polymerases, the aforementioned gel filtration experiments gave some preliminary indications, but not a clear picture. Despite having tested several protein-protein ratios (PCNA ring to polymerase of 1:1, 1:2, 1:3 etc and vice versa) the elution profiles were always coherent. Specifically, in case of Polη complexes, regardless of the nature of PCNA (ubiquitinated or not) the polymerase was often eluting in the first fractions, together with the sliding camp, suggesting a high recruitment of Polη molecules to PCNA. Dissimilar was the case of Polι complexes, with either PCNA or UbPCNA. Regardless of the ring : polymerase ratio tested, the late eluting fractions always contained some free or unbound Polι. This could indicate that PCNA trimer occupancy for Polι is less than for Polη. This preliminary observation leads to the speculation that multiple copies of Polη could be accommodated on the same PCNA (or UbPCNA) ring, whereas Polι has a

minor binding occupancy on the trimeric sliding clamp. This theory, in which only a single Polι molecule could occupy PCNA ring, was suggested very recently (Taylor, 2020) and still requires further experimental proof. In principle PCNA trimer has three identical binding interfaces, one on each protomer, and could potentially accommodate three polymerases via their PIP boxes. The interactions of PCNA hydrophobic pockets with PIP box peptides of several TLS polymerases were described by crystallography and surface plasmon resonance studies by Hishiki and co-workers. In their work they showed a ratio of 1 ring to 3 PIP peptides (Hishiki et al., 2009b), indicating independent and not cooperative binding of PIP peptides to each protomer. However, behaviour of the full-length remains elusive, and the PCNA tool belt model (described in section 1.4.7) still awaits for more experimental evidence.

To better understand the occupancy of Polη and Polι on PCNA, I performed crosslinking experiments followed by size exclusion chromatography. As described in section 2.3.2 the crosslinked PCNA-Polη complex showed a single smeary band around 300 kDa which could correspond to a complex containing one PCNA ring and three Polη molecules. Meanwhile, the Polι-PCNA crosslinked complex showed a more complicated scenario. Three different populations were detected (see Figure 2.14 b.) likely corresponding to PCNA trimer crosslinked to one or two Polι molecules, respectively. These results suggest that Polι is indeed recruited less efficiently than Polη to PCNA trimers.

To further elucidate the importance of PCNA trimerization for binding of polymerases, I performed size exclusion studies using a mutant of PCNA which cannot trimerize, instead remains monomeric in solution. Firstly, I could show that this monomeric PCNA was not efficiently mono-ubiquitinated *in vitro* by UbcH5c in contrast to the wild type PCNA (see Figure 2.11). This result suggested that PCNA trimerization appears to be an important requirement for efficient ubiquitination by UbcH5c, at least in these *in vitro* conditions. In addition, interaction of polymerases with PCNA monomer proved challenging. I tested interaction via size exclusion chromatography, as previously described, but no significant peak shift was observed (see section 2.3.3). Specifically, no interaction between Polι and PCNA monomer was detected on gel filtration (Figure 2.15 a.). This result supports the previous hypothesis of weak Polι-PCNA interaction in solution. Furthermore, it can be speculated that a binding interface created by PCNA trimerization is essential for Polι interaction. Interestingly, when the mixture of Polη and PCNA monomer was ran over SEC, no full peak shift was observed but the second eluting peak showed a slight shift to the left. This can indicate a

possible transient interaction or a minor recruitment of Polη to monomeric PCNA. This hypothesis was further confirmed by a pulldown of Polη with PCNA monomer where minor interaction was detected. This evidence indicates that Polη, in contrast to Polι, does not require a trimeric interface to interact with PCNA. However, it has to be noted that the observed interaction between PCNA monomer and Polη is just a minor recruitment, much weaker compared to trimeric PCNA-Polη interaction. To sum up, these experiments show the importance of PCNA trimeric structure for its interaction with polymerases, as for both Polη and Polι the trimeric PCNA is a best binding platform. This hypothesis is in agreement with results of Dieckman et al., who showed that PCNA trimer instability might inhibit translesion synthesis *in vitro* (Dieckman and Washington, 2013). Based on the results presented in my thesis, a prospective model of binding stoichiometry is presented in Figure 3.1. However, it has to be pointed out that size exclusion chromatography analysis is not a quantitative technique and can only give a qualitative idea of interaction, which may still be different when measured by other methods. Other quantitative methods, e.g. isothermal titration calorimetry (ITC) and microscale thermophoresis (MST), were applied as well. However, due to sample instability during the experimental procedure I couldn't obtain reliable results in a reproducible manner.

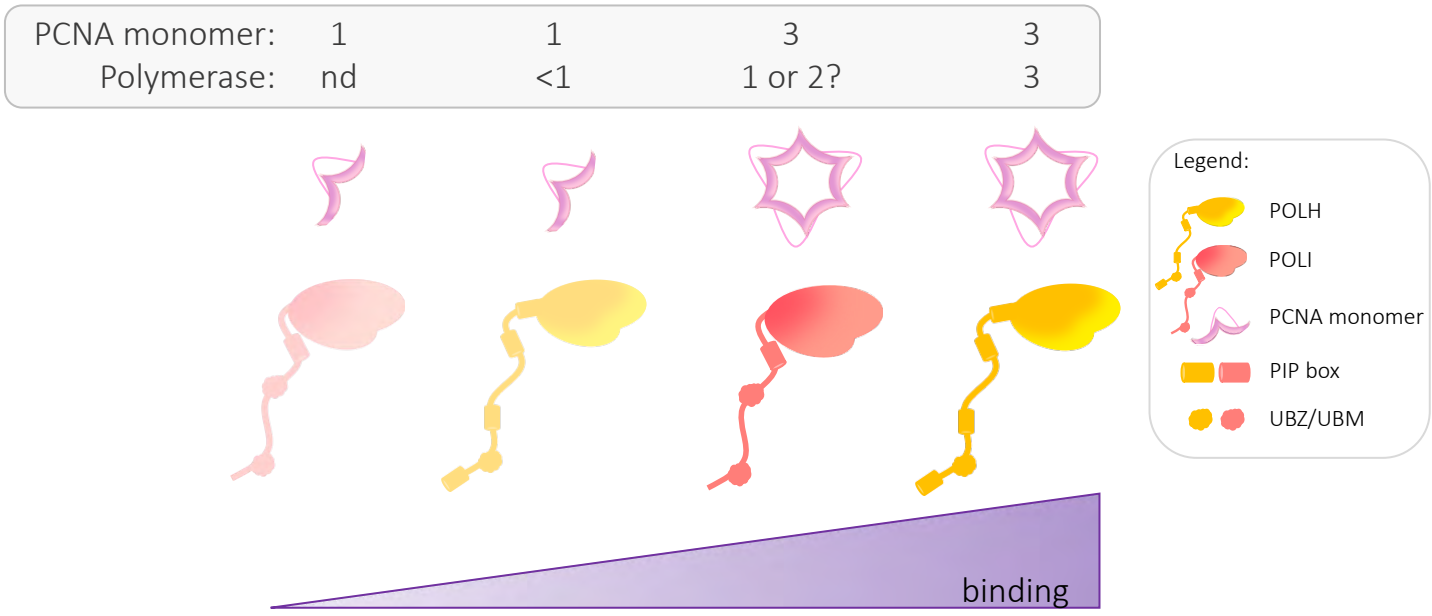


Figure 3.1 Prospective model of binding stoichiometry of polymerases to PCNA. The presented model depicts the interaction stoichiometry and relative binding "strength" of polymerases to PCNA (either monomer or trimer). See legend for accurate description of figure: PCNA is depicted in pink, polymerases are depicted as a bean (the catalytic domain) with a tail (the disordered C-terminus). Polη in yellow and Polι is in red, shades of colour are indication of the binding stoichiometry. Nd: non-detected.

3.2 Cryo-EM analysis of flexible TLS complexes reveals interesting structural features

In order to obtain structural information about binding and arrangement of minimal TLS complexes, I carried out extensive cryo-EM analysis. In the last few years the cryo-EM field witnessed the so-called "resolution revolution" (Kuhlbrandt, 2014) where thanks to technical advances in the electron microscope architecture, development of direct electron detectors (Wu et al., 2016) and the mushrooming of excellent processing algorithms and software (Scheres, 2012) (Scheres, 2016) (Punjani et al., 2017a) (Zhang, 2016) (Moriya et al., 2017), high resolution structures (< 3 Å) are solved on a daily basis. These technical breakthroughs in structural biology also benefit the DNA replication field. Several high resolution cryo-EM structures of DNA polymerase complexes were brought to light in the last years (Lau et al., 2015) (Gómez-Llorente et al., 2013) (Jain et al., 2019) (Malik et al., 2020) (Zheng et al., 2020) (Lancey et al., 2020a). Remarkably, both yeast and human Polδ structures, in presence of DNA and PCNA, were described by almost concomitant works by several groups (Jain et al., 2019) (Zheng et al., 2020) (Lancey et al., 2020a). Strikingly, few months ago the high resolution cryo-EM reconstruction of the yeast Polζ holoenzyme in presence of DNA was finally revealed (Malik et al., 2020). Moreover, a recent preprint by the De Biasio group showed for the first time an unprecedented cryo-EM reconstruction of the Y-family Polκ bound to PCNA and DNA, although in the presented map the flexible C-terminus of Polκ is invisible (Lancey et al., 2020b). Certainly, the last few years were exciting times for the cryo-EM community working on replicative or multi-subunit polymerases as Polδ or Polζ. However, the TLS polymerases object of this study don't share the luxury of having large size or several accessory regulatory subunits. Instead, Polη and Polι display an intrinsically disordered C-terminal region which makes gaining any structural information by cryo-EM quite challenging. This flexible regulatory C-terminus is a characteristic of Y-family polymerase, and plays a role in their interaction with other proteins, like PCNA, ubiquitin and other polymerases. This intrinsically disordered region it's critically contributing, thanks to its flexibility, in the prompt recruitment to, and successive displacement form, PCNA. However, due to lack of structural and biochemical data, addressing its specific functions still remains a big challenge in the field. One of the original aims of my thesis was to characterise Polη and Polι complexes with sliding clamp using cryo-EM, but the journey turned out to be rather insidious. Nonetheless, I was able to obtain interpretable structural data on both complexes as well as single proteins. Preliminary evidence from negative stain observations suggested that crosslinking was necessary to avoid complex

dissociation on grids (section 2.4.1). Several strategies, which included different crosslinking approaches, buffer salt concentration, detergents and the use of high-affinity Polη mutants (pip3 and 3PIPS, see section 2.4.6) were put in place to stabilise the complexes in cryo conditions. However, due the high intrinsic flexibility of polymerases, especially at the C-termini and their weak transient interaction with sliding clamp, structural studies were proved very challenging. In particular, high resolution features remained elusive due to overall small size of the proteins, complexes falling apart during vitrification, even after crosslinking, and intrinsic high conformational flexibility of the complexes. Nonetheless, I successfully reconstructed several maps at sub-nanometre resolution (see section 2.4).

Firstly, the behaviour of PCNA proved amenable to cryo-EM. Despite its small size (86 kDa) but thanks to its robust architecture and symmetry, PCNA reconstruction reached the highest resolution. In fact, though coming from a dataset on crosslinked PCNA-Polι-DNA complex and aimed to solve the intact complex, the reconstruction of unbound PCNA particles showed a final resolution of 6.8 Å. It is noteworthy that such result was achieved with an astonishingly low number of particles (<6000), which proved the good behaviour of the protein in cryogenic condition.

From the same dataset another population could be observed in the 2D classes, corresponding to the vast majority of particles, which showed clear PCNA features associate with a more flexible Polι density. The resolution of the complex after 3D refinement reached ~12Å and allowed some preliminary observation. Firstly, some DNA density was visible inside the PCNA ring which was bound by the polymerase density in a position that would allow the DNA passing through its catalytic core. This is the first time that the arrangement of Polι-PCNA-DNA is observed. Moreover, in this structure only one polymerase was bound to the PCNA ring and there was no multiple polymerases binding observed throughout the data processing. This would agree with the results in SEC and crosslinking analysis, that the stoichiometry of Polι and PCNA ring is 1:1. There is still a possibility that the multiple binding species exist, but the population might be too low to be detected in these experimental conditions. Finally, although the binding interface is not well resolved due to low resolution, it appears that there are several contact points between the PCNA and the polymerase density (see Figure 2.17 f.). The structure suggests that Polι possibly needs the whole trimeric PCNA structure for its recruitment. This speculation is consistent with what was described in the previous section and in Figure 3.1.

Another significant finding of this work is the visualisation, for the first time by cryo-EM, of ubiquitinated PCNA. Several studies have been using various techniques (SEC-MALS, SAXS,

NMR and crystallography) to address the mystery of ubiquitin flexibility on PCNA (Freudenthal, 2010) (Tsutakawa et al., 2011) (Zhang et al., 2012) (Hibbert and Sixma, 2012) (Tsutakawa et al., 2015). In my thesis, for the first time it was possible to clearly and directly visualise ubiquitin moieties on PCNA by cryo-EM 2D classification. Despite UbPCNA being small in size and possessing several degrees of flexibility due to the mobile ubiquitin molecules, the resulting 2D and 3D analysis was quite remarkable. Moreover, cryo-EM allowed to distinguish the coexistence of several species in the sample, which could not be separated by biochemical methods. Native mass spectrometry (section 2.2.2) showed in fact that three different populations (PCNA trimer +1Ub, +2Ub and +3Ub) were present, with the most represented population being PCNA+3Ub. The same observation was confirmed by 2D classification of UbPCNA particles during the cryo-EM dataset processing. After sorting the particles in 2D classes, it was clear that some particles showed PCNA with 1, 2 or 3 ubiquitin attached. Additionally, all the classes showed different positions of the three ubiquitins with respect to PCNA, indicating high flexibility and in agreement with what described in the literature (Tsutakawa et al., 2011) (Hibbert and Sixma, 2012) . The same flexibility could be appreciated in the 3D classification, where the resolution was limited due the small size of PCNA and the presence of three flexible ubiquitin moieties. In fact, in most of the 3D classes, the ubiquitins are visible as low resolved blobs (see Figure 2.18). However, when a particle subset containing PCNA+1Ub was selected, a high resolution reconstruction of UbPCNA was obtained and refined to ~7Å (see section 2.4.4). In this final reconstruction, the singular ubiquitin moiety was observed in a side position, slightly pointing towards the back face of PCNA. In conclusion, this cryo-EM analysis of UbPCNA confirms that UbPCNA is a flexible hub, where ubiquitin is anchored on K164 but freely moves around and these observations are in line with the previous structural studies on UbPCNA. The flexible ubiquitin moiety possibly works as fishing bait interacting with binding partners, particularly recruiting TLS polymerases.

In this work a sub-nanometre resolution map of full-length Polι in complex with DNA substrate is reported for the first time. (see section 2.4.5). The final resolution of the map was ~9 Å and allowed a confident fitting of the available crystal structure. The crystal structure of Polι (PDB: 1T3N) is limited to the catalytic domain of the polymerase, since the flexible C-terminus was never crystallised. In this case the crystal structure (PDB: 1T3N) comprising amino acids 26-420 (out of total 740) was manually fitted into the EM map, showing good agreement. From the fitting it was clear that the cryo-EM map had additional unassigned density, which could correspond to the C-terminal domain. The additional protein density in the map was however

too small to accommodate the missing ~300 amino acids of the C-terminus. This observation would suggest that the C-terminus of Polι might be extremely flexible and therefore effectively invisible in averaged cryo-EM densities, even when bound to DNA.

The only structural information of UbPCNA-Polη -DNA complex available in the literature is a low resolution map obtained by negative staining few years ago (Lau et al., 2015). Interestingly, despite the recent flourishing of cryo-EM polymerases structures, a high-resolution reconstruction of UbPCNA-Polη complex remains to be elucidated. The work presented in this thesis is a first step towards this higher goal. My reconstruction of UbPCNA-Polη-DNA complex reached a modest resolution (~9Å) which allowed confident fitting of PCNA and Polη from available crystal structures, as well as two ubiquitin molecules. As displayed in section 2.4.6 UbPCNA showed to coordinate a single Polη molecule, and no additional polymerase was visible in the map. This could be due to biochemical sample proprieties or to dissociation as a result of the vitrification procedure. The polymerase density is positioned on the front side of PCNA, facing its intradomain connecting loop (IDCL). Additionally, two out of three theoretical ubiquitin molecules were assigned to some extra density surrounding PCNA in the vicinity of K164. Unexpectedly, the ubiquitin density did not occupy positions at the lateral side, as shown in the UbPCNA map described in section 2.4.4, or back face of PCNA, as described by several structural studies on UbPCNA (Freudenthal, 2010) (Zhang et al., 2012). The two ubiquitin moieties were instead located towards the front side of PCNA, making contacts with the polymerase density. This peculiar position of ubiquitin was never experimentally observed before, but only deducted from molecular dynamic simulations (Tsutakawa et al., 2011). Additionally, the ubiquitin on the front side of PCNA is in contrast with what described in the low resolution map from Lau et al, where the ubiquitin was assigned to the back side (Lau et al., 2015). In the model presented in Figure 2.21 the ubiquitin makes more contacts with the polymerase density, suggesting a higher relevance of this structure compared to the one presented by Lau et al. This repositioning of ubiquitin to the front side of PCNA could play a pivotal role in engagement and guidance of Polη to the correct position on the DNA primer/template. The model obtained from the cryo-EM reconstruction does not, however, rule out other possible conformations of the complex. Multiple conformers of a dynamic complex in solution quite often cannot be discriminated by alignment and classification with the processing tools available so far. Another feature of this UbPCNA-Polη reconstruction was a large unassigned density on the polymerase side, which can potentially be the Polη C-terminus, making contacts to PCNA and ubiquitin. Size wise the density would fit

to the unassigned 280 amino acids of Polη C-terminal region, but the local resolution in this area, due to high flexibility, was rather poor. Interestingly, a slight angle between Polη and PCNA can be appreciated, in agreement with the descriptions in the literature (De March et al., 2017) (Lancey et al., 2020a) but no DNA density was clearly visible in the EM map.

3.3 Biochemical characterisation of Polη complexes: relevance of the flexible regulatory C-terminus and its binding to sliding clamp

Given the major biological importance of Polη in the context of UV-induced lesions bypass, its clinical relevance in Xeroderma Pigmentosum (XP-V) syndrome and the its association to chemotherapy cisplatin treatments, I decided to focus mainly on the characterisation of Polη complexes. To have a board picture of Polη interaction with PCNA and UbPCNA, I carried out extensive biochemical analysis of the complexes, in parallel to the above mentioned cryo-EM studies. A major emphasis was put on the intrinsically disordered C-terminus and its identified binding motifs, either putative PIP boxes or UBZ. In this work I identified 4 putative PIP boxes motifs at Polη C-terminus by sequence analysis, partially overlapping with motifs already presented in the literature (Masuda et al., 2015) (Boehm et al., 2016a). For convenience the identified sequences were named PIP box 0 (437-444), PIP box 1 (477-484), PIP box 2 (524-532) and PIP box 3 (701-708), respectively. Mutations aimed to reinforce or impair the binding of these PIP boxes to PCNA were carried out, resulting in the set of constructs described in section 2.1.4 (see Figure 2.5). These constructs were systematically tested in pulldown assays against PCNA and UbPCNA, in order to address interaction and binding affinities. Additionally, the high-affinity binding mutants (pip3 and 3PIPS) were more intensively characterised by primer extension assays, EMSA and subjected to preliminary cryo-EM analysis.

I showed that Polη wild type, pip3 and 3PIPS mutant possess comparable activities, using primer extension assay. Interestingly, presence of either PCNA or UbPCNA did not affect Polη processivity in primer extension, in contrast to what was observed with other experimental setups (Masuda et al., 2015) (Hedglin et al., 2016) but instead appeared to play an important role in Polη fidelity. As discussed in section 2.5.1 for Polη in complex with UbPCNA, higher mis-incorporation of wrong nucleotides was observed across undamaged template, compared to conditions with Polη alone or in complex with PCNA. Importantly, the mutagenic effect of

Polη was more significant when dTTP was provided against template G and with dCTP against template T. This could indicate a possible mutator role of Polη which is accentuated by UbPCNA presence. Polη mutagenic effect has been already described in the literature for its important role in processes like somatic hypermutation (SHM) (Zhao et al., 2013) (Zeng et al., 2001). The role of UbPCNA in this immunoglobulin production mutagenic process was already described (Arakawa et al., 2006), finding additional support in the data presented in my thesis.

Additionally, EMSA analysis was performed in parallel with Polη wild type, pip3 and 3PIPS mutants provided several interesting observations (described in section 2.5.2). In the first place the appearance of a clear DNA shift only at 4x molar excess of all Polη constructs indicates a low affinity of Polη towards DNA. This would fit to the hypothesis that Polη has low affinity to undamaged DNA but exhibits a preference for lesion-containing DNA. Presence of PCNA or UbPCNA did not affect the DNA binding behaviour of Polη wild, suggesting a weak protein-protein interaction. Interestingly, when the high affinity mutants pip3 and 3PIPS were tested in presence of PCNA and UbPCNA, a higher shift was observed, indicating the formation of more stable complexes, due to the increased protein-protein affinity.

The low affinity of Polη towards sliding clamp was further confirmed by an optimised set of pulldown assays. As described in section 2.6. the amount of prey (PCNA or UbPCNA) eluted at the end of the procedure, was only a minimal percentage compared to the actual input quantity. In case of PCNA it was estimated to be around 1%, whereas in case of UbPCNA it was much higher, in the range of ~15%. The first evident result from this set of semi-analytical pulldowns was that UbPCNA seemed to be recruited more efficiently by the bait Polη. This hypothesis was further confirmed by a set of PCNA/UbPCNA competition pulldown experiments, and proved reproducible (see section 2.6.1). From these results, it can be speculated that indeed PCNA ubiquitination plays a role in regulating binding and recruitment of Polη. In the literature there are contrasting opinions about the role of ubiquitination for Polη recruitment (Kannouche et al.) (Hedglin et al., 2016) From what shown in my work it seems that UbPCNA has a better binding affinity to Polη, compared to PCNA, although this interaction remains in the low affinity range. Another important observation was that at low salt concentrations better binding could be detected. This evidence was reproducible for all of each Polη construct tested, regardless of the mutations. At 50 mM KCl each pulldown assay worked the best, and showed utmost protein binding. Following this rationale, I tried to obtain binding curves with protein titration at low salt, but unfortunately no reliable affinity parameter

(K_d) could be extrapolated. Nonetheless, quantification of pulldown triplicates allowed to build a semi-quantitative model of Polη binding to sliding clamps, where the relative importance of PIP and UBZ motifs can be deducted (see Figure 3.2). The evidence that low salt concentration was beneficial to Polη-sliding clamp interaction also influenced the parallel ongoing optimisation of cryo-EM, as described in section 2.4.6. However, after intensive screening of complexes in low salt buffer however I did not gain any significant improvement in the EM analysis. This indicates that even biochemically optimal parameters cannot be effectively translated in satisfactory cryogenic conditions, as proteins may behave differently in the harsh vitreous ice of cryo-EM grids.

From direct comparison of all Polη mutant behaviour in pulldown assays, interesting speculations can be made. The most striking effect was that when all the PIP boxes were inactivated, no binding was detected neither of PCNA nor, surprisingly, of UbPCNA. This indicates that the sole UBZ domain of Polη is necessary but not sufficient to be recruited by UbPCNA. A fascinating behaviour was observed for both UBZ and pip3 dead mutants, which at all salt concentrations showed rather low capacity to pull down UbPCNA. As expected, the disruption of Polη UBZ clearly affected the binding of UbPCNA. Surprisingly, also the pip3 dead Polη mutant showed inability to recruit UbPCNA. This interesting observation might suggest a novel role of PIP box 3 in the recruitment of UbPCNA. Specifically, PIP box 3 seems to play a comparable role to the UBZ for the interaction with UbPCNA. However, more biochemical or structural data are needed to validate this initial hypothesis. A summary of the relative binding importance of each Polη construct to both sliding clamps is depicted in Figure 3.2.

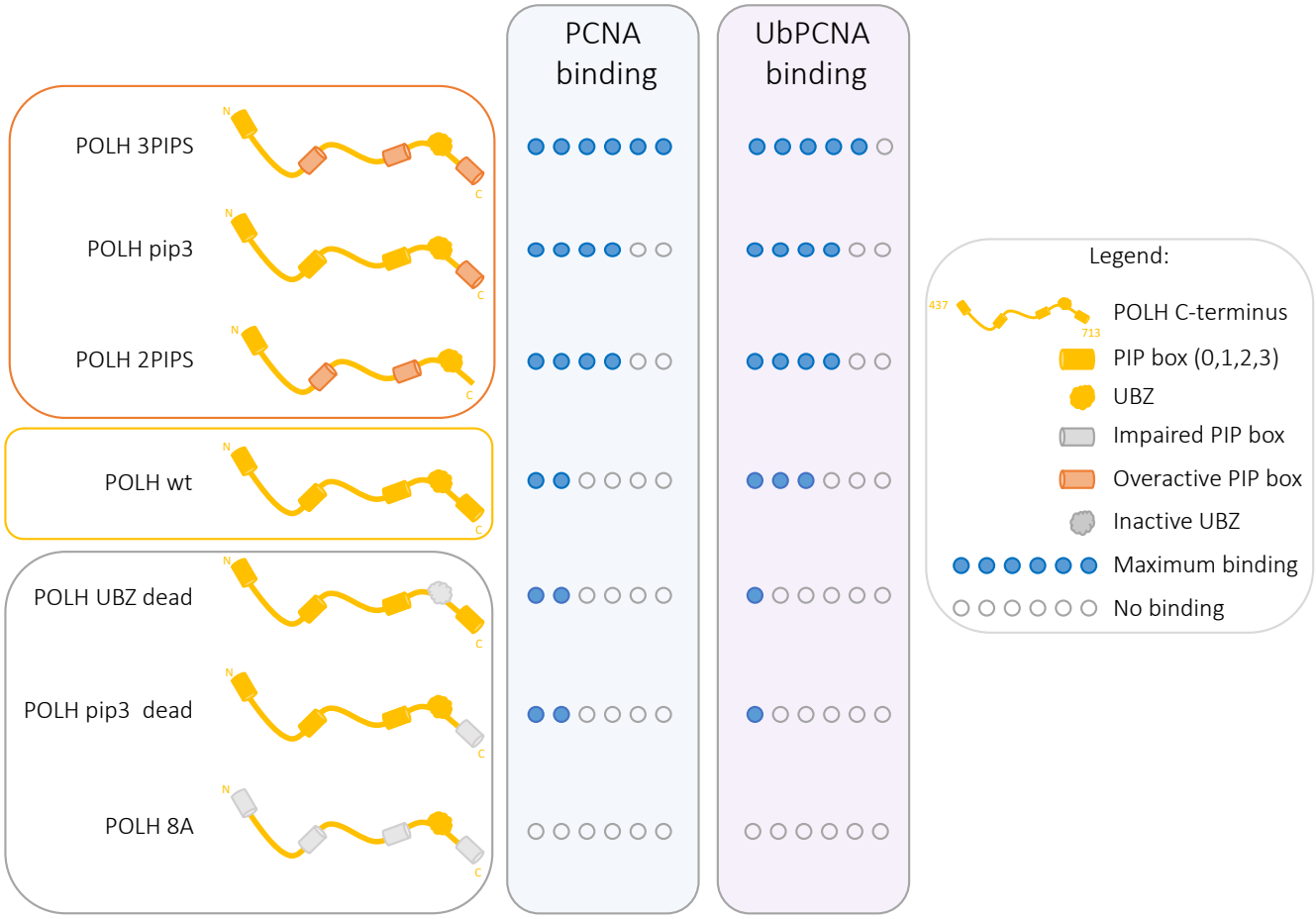


Figure 3.2 Proposed PCNA and UbPCNA interaction hierarchy of Polη C-terminal binding motifs. Here presented is a speculative model of relative binding importance of each Polη construct to PCNA and UbPCNA. Observations come from pulldown assays, the binding scale to PCNA and UbPCNA is arbitrary, but reflects the quantitative observations from pulldowns. See Legend for detail, in brief: Polη C-terminus is depicted as a yellow string, containing the 4 putative PIP boxes (PIP box 0, 1, 2, and 3, from N- to C-terminal) depicted as cylinders, and the UBZ, depicted as a cloud. Overactivated PIP box are depicted in orange, inactivated PIP boxes/UBZ are in grey.

Finally, several trial to reconstitute an *in vitro* evidence of the described tool-belt model of PCNA (Boehm et al., 2016b) were attempted by size exclusion and cryo-EM, but with little success. Exploiting a multi-polymerase pulldown approach (see section 2.6.4) I was able to show, for the first time, a possible simultaneous recruitment of two different polymerases to UbPCNA. The interaction seemed to be dependent on the ubiquitinated state of the sliding clamp, as no double polymerases binding was detected with unmodified PCNA. As described in section 2.6.4, pre-assembled UbPCNA-Polι complexes, with a large excess of polymerase, where shown to interact with the Polη bait, indicating the co-existence of two different polymerases on the same UbPCNA ring. This result is still preliminary and needs further

experimental confirmation, but is positively speaking in favour of the so-called tool belt PCNA model. Remarkably, PCNA ubiquitination seemed to play a role in multiple polymerase recruitment. A proposed model of this mechanism is depicted in Figure 3.3.

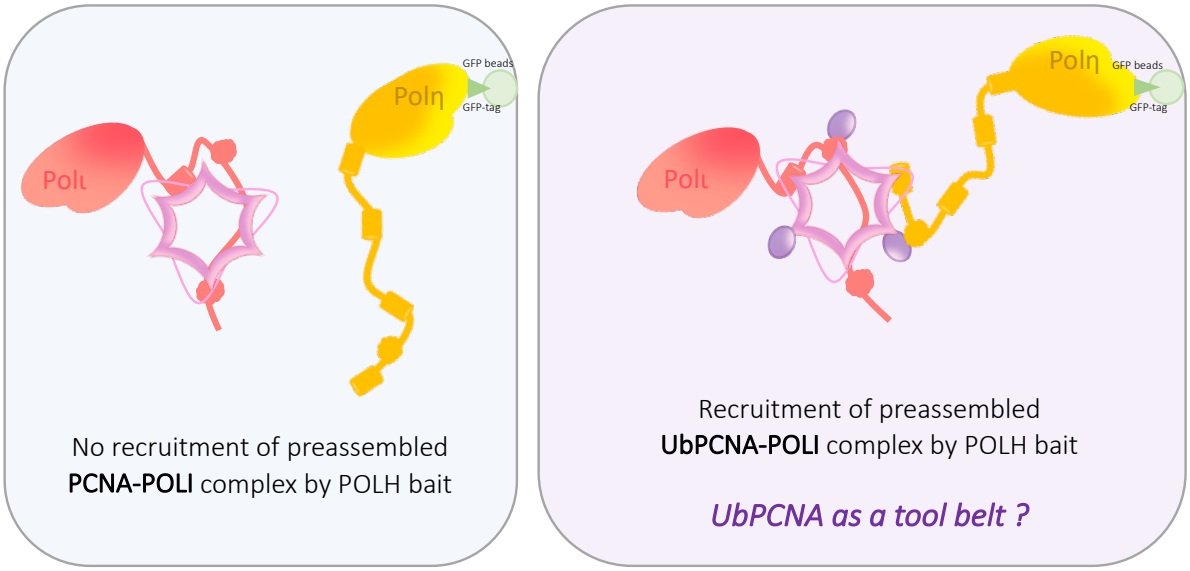


Figure 3.3 Speculative UbPCNA tool belt model. From the multi-polymerase pulldown a model of UbPCNA recruiting both polymerases can be deducted. In the left panel is depicted the case of preassembled PCNA-Polι complexes (pink ring and red polymerase) and there is no recruitment by Polη bait (yellow polymerase). In right panel the second case, when UbPCNA- Polι complex is preassembled, with an excess of Polι, there is recruitment of the complex by the Polη bait (yellow polymerase). Presumably the UbPCNA interface is capable of coordinating two different polymerases, differently from unmodified PCNA.

3.4 Conclusions and future perspectives

The combination of biochemical and structural studies presented in this thesis were aimed to shed light onto the intricate molecular mechanism of translesion synthesis. The obtained results allow a better understanding of the mechanism behind the recruitment of specialised polymerases to the sliding clamp PCNA. Importantly, I was able to obtain unprecedented cryo-EM reconstruction of single proteins as well as complexes. Given the relatively small size, high heterogeneity and flexibility of these complexes, the level of details of the structural results presented is impressive. However, the difficulty in achieving high resolution structures is an additional indication that such complexes are unstable and transient, as they are supposed to be in cellular context. Improvements in samples reconstitution, cryo-EM data collection and processing, especially utilising newly available algorithms and software could finally give an answer to the structural organisation of TLS complexes. More and more programs are being optimised to deal with sample heterogeneity and flexibility, and there is a chance that these technical improvements could help the final aim of this project. Lastly, improvements in samples biochemistry in terms of post translational modifications, DNA lesions, additional factors and additives and are essential to generate a stable and well behaved complex for effective cryo-EM reconstruction. TLS polymerases are mutagenic and therefore should be recruited only transiently and in specific context of DNA lesions. In this regard, the ubiquitination of PCNA plays an important role. The higher affinity of Polη for UbPCNA reported in this study indicates that ubiquitination positively affects the assembly of complexes, at least *in vitro*. However, the situation in cell, especially in contexts other than TLS, may be more variegate. Still the role of mutagenic polymerases is not fully understood, but the biochemical and structural studies presented here help to better understand the role of PIP boxes and ubiquitin mediated interaction with sliding clamp. However, further studies are still required, both *in vitro* and in cellular context, to fully elucidate the transient and highly regulated recruitment of TLS polymerases to the replication machinery.

4. Materials and Methods

4.1 Materials

4.1.1 Consumables and chemicals

Unless otherwise stated in the text, all chemicals were purchased from Roche Diagnostics (Mannheim, Germany), Merck (Darmstadt, Germany), Roth (Karlsruhe, Germany), Sigma-Aldrich (München, Germany) Qiagen (Hilden, Germany), and Serva (Heidelberg, Germany). Enzymes were ordered from New England Biolabs (NEB) (Frankfurt/Main, Germany), ThermoFisher Scientific (Waltham, MA, USA) or prepared by Biochemistry Core Facility of Max-Planck Institute of Biochemistry (Martinsried, Germany). The enzymes used for tag cleavage (Senp2, 3C) and for *in vitro* ubiquitination assay (hsUbaI, hsUbcH5c wt and S22R) were prepared by technical assistant Marcus Hammerl, Maren Kluegel and Sven Schkölziger. Nucleotides were ordered from Eurofins/MWG (Ebersberg, Germany). Gel and plasmid extraction kits were purchased from Macherey-Nagel (Düren, Germany). Chromatography material and columns were ordered from GE Healthcare (München, Germany), Macherey-Nagel (Düren, Germany), Qiagen (Hilden, Germany), IBA (Göttingen, Germany) and Roche (Basel, Switzerland). Crystallization Screens were obtained from Hampton Research (Aliso Viejo, CA, USA), and Qiagen (Hilden, Germany). Pre-cast gels were purchased from Bio-Rad (München, Germany). Protein concentrators were obtained from Merck Millipore (Darmstadt, Germany). EM material such as forceps and grids was ordered from Plano (Wetzlar, Germany) and Quantifoil (Großlöbichau, Germany). Negative stain carbon grids where prepared in house by Marcus Hammerl and Sven Schkölziger.

4.1.2 Antibiotic solutions

Antibiotic	Stock concentration	Final concentration
Ampicillin	100 mg/ml	100 µg/ml
Kanamycin	50 mg/ml	50 µg/ml
Chloramphenicol	34 mg/ml	34 µg/ml

Table 4.1 Antibiotic solutions

4.1.3 Media

Medium	Composition	Expression
LB	1% (w/v) Bacto Trypton, 0.5% Yeast Extract, 0.5% (w/v) NaCl, pH 7.2	*E. coli*
SOC	2% (w/v) Bacto Trypton, 0.5% Yeast Exract, 10 mM NaCl, 10 mM MgSO4, 2.5 mM KCl, 1 mM MgCl2, 0.4% glucose, pH 7.2	*E. coli*
ZY	1% (w/v) Bactor Trypton, 0.5% Yeast Extract, 100 mM PO4, 25 mM SO4, 50 mM NH4, 100 mM Na, 50 mM K, 0.5% glycerol, 0.05% glucose, 0.2% α-lactose	*E. coli*

Table 4.2 Media

4.1.4 Bacterial strains

Bacterial strain	Species	Genotype
XL1 blue	*E. coli*	*rec*A1 *end*A1 *gyr*A96 *thi*-1 *hsd*R17 *sup*E44 *rel*A1 *lac* [F' *proAB* $lacI^{q}ZΔM15$ Tn*10* (Tet^{r})]
BL21 DE3 gold	*E. coli*	B F^{-} *opm*T *hsd*S(rB^{-} mB^{-}) dcm^{+} Tet^{r} *gal* λ(DE3) *end*A Hte
BL21 DE3 gold pRARE	*E. coli*	B F^{-} *opm*T *hsd*S(rB^{-} mB^{-}) dcm^{+} Tet^{r} *gal* λ(DE3) *end*A Hte [pRARE Cam^{r}]
B834 (DE3) pLysS	*E. coli*	F^{-} *opm*T *hsd*S(rB^{-} mB^{-}) *gal dcm* met^{-} λ(DE3) pLysS (Cam^{R})

Table 4.3 Bacterial strains

4.1.5 Vectors and plasmids

Plasmids used for expression in bacterial system: bax vectors. GFP (green fluorescent protein) contained the A206K mutation to ensure monomeric state of the fluorescent dye. SUMO: small ubiquitin like modifier, CPD: cysteine protease domain. 3C: HRV-3C protease.

Name	Application	Tag	Tag cleavage
pCB-bax04	Bacterial expression of proteins	N-terminal 8x His-SUMO	Senp2
pCB-bax07	Bacterial expression of proteins	untagged	-
pCB-bax25	Bacterial expression of proteins	C-terminal CPD- Twin-Strep-tag	CPD
pCB-bax29	Bacterial expression of proteins	N-terminal GFP (Twin-Strep-tag @212)-3C	3C
pCB-bax30-CPD	Bacterial expression of proteins	N-terminal GFP (8x His@212)-3C and C-terminal CPD- Twin-Strep-tag	3C, CPD
pCB-bax36	Bacterial expression of proteins	C-terminal CPD (L204I)-8x His	CPD

Table 4.4 Plasmids for expression in bacterial system

List of cloned constructs for bacterial expression. The insertion site LIC-SUMO contains the SUMO tag always at the N-terminus. The insertion site LIC-CPD contains the CPD tag always at the C-terminus. LIC: ligation independent cloning, SUMO: small ubiquitin like modifier, CDP: cysteine protease domain. Cloning into LIC site was performed by LIC or Gibson Assembly.

Construct	Vector	Primers	Insertion site
hsPCNA	pCB-bax25	CBp576,CBp577	LIC CPD
hsPCNA monomer (K77D, C81E, K110D)	pCB-bax36	CBp576,CBp577	LIC CPD
hsPolι	pCB-bax04	CBp317, CBp318	LIC SUMO
hsUB	pCB-bax07	CBp666, CBp688	LIC
hsPolη wt	pCB-bax25	CBp450, CBp451	LIC CPD
hsPolη pip3	pCB-bax25	CBp450, CBp1232	LIC CPD
hsPolη pip3	pCB-bax29	CBp2267,CBp2269	LIC 3C
hsPolη 3PIPS	pCB-bax25	CBp450, CBp1232	LIC 3C
hsPolη 3PIPS	pCB-bax29	CBp2267, CBp2269	LIC 3C
hsPolη pip3	pCB-bax30-CPD	CBp2267, CBp1232	LIC 3C-CPD
hsPolη 3PIPS	pCB-bax30-CPD	CBp2267, CBp1232	LIC 3C-CPD
hsPolη pip box 3 dead (F707A,F708A)	pCB-bax29	CBp2267,CBp2268	LIC 3C
hsPolη 8A (F443A,L444A,F483A, F484A,F531A,F532A,F707A,F708A)	pCB-bax29	CBp2267,CBp2268	LIC 3C
hsPolη UBZ dead (D652A)	pCB-bax29	CBp2225,CBp2226	LIC 3C

Table 4.5 Cloned constructs

4.1.6 Oligonucleotides for cloning and mutagenesis

Name	Sequence (5'-3')	Application
CBp263	gcaaagcaccggccttaCCCACCTCTGAGACGGAGCAC	reverse bax/hex Cuntag-LIC primer for hsUb (end @ 76)
CBp317	accaggaacaaaccggcggccgctcgATGGAACTGGCGGACGTGGG	forward bax/hex Nsumo-LIC primer for hsPOLI (start @ 26)
CBp318	gcaaagcaccggcctcgttaTTTATGTCCAATGTGGAAATCTGATCC	reverse bax/hex Cuntag-LIC primer for hsPOLI (end @ 740)
CBp450	aagaaggagatatacatATGGCTACCGGCCAGGATCGCG	forward bax Nuntag-LIC primer for hsPOLHsyn (synthesized gene, start @ 1)
CBp451	accgcccgcgagGTGTGTCAGTGGTTTGAAAAAGGATTCCAGGG	reverse bax Ccpd(at)-LIC primer for hsPOLHsyn (synthesized, end @ 713)
CBp576	aagaaggagatatacatATGTTCGAGGCGCGCCTGGTCC	forward bax Nuntag-LIC primer for hsPCNA (start @ 1)
CBp577	accgcccgcgagAGATCCTTCTTCATCCTCGATCTTGGGAG	reverse bax Ccpd-LIC primer for hsPCNA (end @ 261)
CBp666	gttaaacaaaattattactagtTTACCCACCTCTGAGACGGAGCAC	reverse bax RBS-LIC primer for hsUB (end @ 76)
CBp688	aagaaggagatatacatATGCAGATCTTCGTGAAGACTCTG	forward bax Nuntag-LIC primer for hsUb (start @ 1)
CBp1232	agagaagatcagacgacgtttagagtggtagaagtcggtcatagaggtctgACCTTCTGGACGCGGACGTTTGTTAG	reverse mutagenesis primer for hsPOLHsyn (p21-PIP3)
CBp2225	GAACACATGGcCTATCATTTCGCCCTGGAGCTG	forward mutagenesis primer for hsPOLHsyn (D652A), UBZ
CBp2226	CAGAAAAGATTTCTGCAGCTCCAGG	reverse mutagenesis primer for hsPOLHsyn UBZ (change @ 713/cpd)
CBp2267	ggaagttctgttccaggggcccATGGCTACCGGCCAGGATCG	forward bax/hex N3c-LIC primer for hsPOLHsyn (synthetic, start @ 1)
CBp2268	cggatcctgcaaagcaccggccTTAGTGTGTCAGTGGTTTGGCTGCG	reverse bax/hex Cuntag-GA primer for hsPOLHsyn F707A, F708A inactive pip box 3 (end @ 713)
CBp2269	ggatcctgcaaagcaccggccTTAAGAGAAGATCAGACGACGTTTAGAG	reverse bax/hex Cuntag-GA primer for hsPOLHsyn pip3p21(end @ 717)

Table 4.6 Primers for cloning and mutagenesis

4.1.7 Buffers for protein purification and biochemistry

Buffer	Composition	Application
Lysis 1	50 mM Tris pH 7.5, 1M NaCl, 2 mM DTT, 1 mM EDTA 1mM AEBSF, 1µM Leupeptin, 1µM PepstatinA, 1µM Aprotinin	Lysis for PCNA
Lysis 2	20 mM HEPES pH 7.5, 1M NaCl, 5 mM DTT, 1 mM EDTA 1mM AEBSF, 1µM Leupeptin, 1µM PepstatinA, 1µM Aprotinin	Lysis for Pols

Lysis 3	50 mM Tris-HCl, 1 mM EDTA, 0.05% Triton X-100 pH 7.6 1mM AEBSF, 1µM Leupeptin, 1µM PepstatinA, 1µM Aprotinin	Lysis for Ub
Chaperone Wash Buffer	20mM Tris pH 7.5, 1M NaCl, 5mM DTT, 2mM ATP, 5 mM $MgCl_2$	Chaperone Wash for Pols
Strep 1	50 mM Tris pH 8.0, 500 mM NaCl (for pols) 200 mM NaCl (for PCNA), 1 mM EDTA	Strep Equilibration and wash
Strep 2	50 mM Tris pH 8.0, 500 mM NaCl (for pols) 200 mM NaCl (for PCNA), 10 mM IP6	Strep IP6 Elution
Strep 3	50 mM Tris pH 8.0, 500 mM NaCl (for pols) 200 mM NaCl (for PCNA), 0.1 mM EDTA, 2.5 mM Desthiobiotin	Strep Desthiobiotin Elution
Strep 4	50mMTrispH8.0,150 mM NaCl,1mMEDTA,1 mM HABA	Strep Regeneration
Strep 5	50 mM Tris pH 8.0, 150 mM NaCl, 1 mM EDTA	Strep Storage
His A	50 mM Tris pH 7.5, 500 mM NaCl (for pols) 200 mM NaCl (for PCNA), 3 mM DTT, 5% (w/v) glycerol	His Equilibration and wash
His B	50 mM Tris pH 7.5, 500 mM NaCl (for pols) 200 mM NaCl (for PCNA), 3 mM DTT, 300 mM imidazole 5% (w/v) glycerol	His Elution
Hep A	20 mM Hepes pH 7.0, 3 mM DTT, 5% (w/v) glycerol	Heparin Equilibration
Hep B	20 mM Hepes pH 7.0, 3 mM DTT, 2 M NaCl, 5% (w/v) glycerol	Heparin Elution
SP-A	50 mM $C_2H_3NaO_2$ pH 4.5	SP ubiquitin
SP-B	50 mM $C_2H_3NaO_2$ pH 4.5, 1M NaCl	SP ubiquitin
Dialysis Buffer	50 mM $C_2H_3NaO_2$ pH 4.5	Dialysis ubiquitin
SEC 1	20mM HEPES pH 7.5, 200 mM KCl, 0.5 mM TCEP, 5% w/v glycerol	SEC for PCNA
SEC 2	20mM HEPES pH 7.5, 500 mM KCl, 1 mM TCEP, 5% w/v glycerol	SEC for Pols
Ubiquitination Buffer	50 mM Bicine-KOH pH 9, 3 mM ATP, 2 mM TCEP, 3 mM $MgCl_2$, 50 mM KCl	PCNA ubiquitination
Complex Buffer	20 mM HEPES pH 7.5, 135 mM KCl, 5 mM $MgCl_2$, 5 mM DDT, 2 µM $ZnCl_2$	SEC standard complex buffer
Complex Buffer 50 KCl	20 mM HEPES pH 7.5, 50 mM KCl, 5 mM $MgCl_2$, 5 mM DDT, 2 µM $ZnCl_2$	SEC complex buffer 50 KCl
TDB Buffer	50mM Tris pH 7.5, 1 mM DTT, 0.2 mg/ml BSA 2% glycerol	Primer Extension
EMSA Buffer	20 mM Tris pH 8, 50 ug/ml BSA, 1 mM TCEP, 5% glycerol (Bromophenol blue)	EMSA
Pulldown Buffer 1	20 mM HEPES pH 7.5, 500 mM KCl, 5 mM $MgCl_2$, 5 mM DDT, 2 µM $ZnCl_2$, 0.01% NP-40	Pulldown assay high salt
Pulldown Buffer 2	20 mM HEPES pH 7.5, 135 mM KCl, 5 mM $MgCl_2$, 5 mM DDT, 2 µM $ZnCl_2$, 0.01% NP-40	Pulldown assay
Pulldown Buffer 3	20 mM HEPES pH 7.5, 135 mM KCl, 5 mM $MgCl_2$, 5 mM DDT, 2 µM $ZnCl_2$, 0.01% NP-40	Pulldown assay low salt

Table 4.7 Buffers for purifications and biochemistry

4.1.8 Oligonucleotides used in biochemical assays

Name	Sequence (5'-3')	Application
CBs284	ATCGCAGTAGCATGTGCTGAGCACTTCAGAGGTGTCGGACGCAG	Primer extension
CBs261	CACGAGTCGTGTACGATGACGCTA	EMSA
CBs201F	TAGCGTCAT	Complementary to CBs 284 and CBs261 Labelled 5'-FAM
CBs260F	TAGCGTCATCGT	Complementary to CBs 284 Labelled 5'-FAM
CBs234	TACTTCGACGCTAGTCCAGCCAT	cryoEM optimisation
CBs235	ATGGCTGGACTAGCGTCGA	cryoEM optimisation complementary to CBs234
CBs233	TGAGCATGGCTGGACTAGCGTCGA	cryoEM optimisation
CBs236	TACTTCGACGCTAGTCCAGCCATGCTCA	cryoEM optimisation complementary to CBs233
CBs230	TACTTCGACGCTAGTCCAGCCATGCTCAGGTCT	cryoEM optimisation
CBs231	AGACCTGAGCATGGCTGGACTAGCGTCGA	cryoEM optimisation complementary to CBs230
CBs276	TACTTCGACGCTAGTCCAGCCATGCTCAGGTCTTG	cryoEM optimisation
CBs277	CAAGACCTGAGCATGGCTGGACTAGCGTCGA	cryoEM optimisation complementary to CBs276
CBs278	TACTTCGACGCTAGTCCAGCCATGCTCAGGTCTTGTT	cryoEM optimisation
CBs279	AACAAGACCTGAGCATGGCTGGACTAGCGTCGA	cryoEM optimisation complementary to CBs278
CBs280	TACTTCGACGCTAGTCCAGCCATGCTCAGGTCTTGTTGG	cryoEM optimisation
CBs281	CCAACAAGACCTGAGCATGGCTGGACTAGCGTCGA	cryoEM optimisation complementary to CBs280
CBs282	TACTTCGACGCTAGTCCAGCCATGCTCAGGTCTTGTTGGGA	cryoEM optimisation
CBs283	TCCCAACAAGACCTGAGCATGGCTGGACTAGCGTCGA	cryoEM optimisation complementary to CBs282

Table 4.8 DNA substrates for primer extension, EMSA and cryo-EM

4.1.9 Equipment

Equipment	Producer
Ultrasonic homogenizer	Bandelin electronic, Berlin,Germany
Peristaltic pump	Ismatec, Cole-Parmer GmbH
ÄKTA purification systems	GE Healthcare, München, Germany
NanoDrop (ND-1000) spectrophotometer	PeqLab, Erlangen, Germany

Typhoon™ FL 7000 PhosphoImager	GE Healthcare, München, Germany
Vitrobot Mark IV	FEI (ThermoFisher Scientific)
CM200-FEG electron microscope	FEI (ThermoFisher Scientific)
Titan Krios	FEI (ThermoFisher Scientific)
Talos Arctica	FEI (ThermoFisher Scientific)
K2 Summit	Gatan, Pleasanton, USA
Falcon 2	FEI (ThermoFisher Scientific)
Eagle CCD	FEI (ThermoFisher Scientific)

Table 4.9 Equipment

4.1.10 Software

Software	Supplier/ developer
ApE	M. Wayne Davis
GraphPad Prism	GraphPad Software, La Jolla, CA, USA
Inkscape	Version 0.48.2 https://inkscape.org/en/
Protparam	ExPASy (Artimo et al., 2012)
UNICORN	GE Healthcare
Fiji/ImageJ2	(Schindelin et al., 2012)
EPU	FEI (ThermoFisher Scientific)
SerialEM	(Mastronarde, 2005)
EMAN2	(Tang et al., 2007)
Gctf	(Zhang, 2016)
Gautomatch	https://www.mrc-lmb.cam.ac.uk/kzhang
MotionCor2	(Zheng et al., 2017)
cryoSPARC	(Punjani et al., 2017b)
RELION	(Scheres, 2012)
UCSF Chimera	(Pettersen et al., 2004)

Table 4.10 Software

4.2 Methods

4.2.1 Cloning

DNA templates

All DNA templates were obtained from the cDNA library of Biochemistry Core Facility of Max Planck Institute of Biochemistry. Polη synthetic mutants or C-terminal DNA fragments (specifically Polη 8A, 3PIPS and pip3 constructs) as well as PCNA monomeric mutant, were synthetized by GeneArt String DNA fragments by ThermoFisher Scientific (Waltham, MA, USA) or by Twist Bioscience (San Francisco, CA, USA) and cloned into our vector system by LIC or Gibson Assembly.

Polymerase Chain Reaction

The constructs used in this work were amplified by Polymerase Chain Reaction (PCR) from template plasmids, cDNA or synthesized DNA. The reactions were carried out according to standard protocol and standard program (see tables 4.12 and 4.13) using Phusion High-Fidelity DNA Polymerase Master Mix (NEB). The 2x Phusion Master Mix contained already the Phusion polymerase, Phusion DNA polymerase buffer and dNTPs. The annealing temperature was adjusted to the respective primers, using the nearest neighbor method (T_m + 3°C) and ranged between 58 and 64°C. The extension time was adjusted to the size of the gene of interest with 30s / kb. The PCR product was afterwards analyzed by agarose gel electrophoresis (0.8 % w/v) in 1x TAE buffer (40mM Tris, 20mM Acetic acid, 1mM EDTA pH 8.0). The correctly sized PCR product was cut out of the gel and isolated using the NucleoSpin Gel and PCR Clean-up Kit (Macherey-Nagel, Düren, Germany).

Component	Stock concentration	Final concentration	Amount
DNA template	~ 50-200 ng/ µl	5-20 ng	1 µl
Forward primer	10 µM	0.5 µM	1 µl
Reverse primer	10 µM	0.5 µM	1 µl
Phusion master Mix	2x	1x	10 µl
DMSO	100%	5%	1 µl
ddH_2O	add to 20 µl		6 µl

Table 4.11 PCR reaction mixture

Step	Temperature °C	Duration	Cycles
Initial Denaturation	98	30 s	
Denaturation	98	20 s	
Annealing	T_m + 3°C	30 s	30
Extension	72	30 s / kb	
Final extension	72	5 min	
Cooling	8	hold	

Table 4.12 PCR program for amplification

Cloning of Mutants by Site-Directed Mutagenesis

Mutations were inserted into the gene of interest via PCR with partially overlapping primers. The forwards primer contained the mutation and is complementary with the reverse primer for 10 - 15 nucleotides at the 3' end. The mutation in the forward primer was flanked by 15 - 20 nucleotides. Both primers had a length between 30 - 40 nucleotides. For insertion of the mutation the standard PCR reaction mixture was used (see table 4.11). The PCR program for site-directed mutagenesis was applied (see table 4.13). Afterwards, the PCR product was digested with 5 U DpnI for 1 h at 37°C. For transformation, 4 µl of the digested PCR product was mixed with XL1 blue competent E. coli cells. A maximum of five colonies was subjected to amplification and plasmids were prepared with Nucleospin Plasmid Kit (Macherey-Nagel, Düren, Germany). To assess the successful insertion of the mutation the plasmids were sent to sequencing (Eurofins, Ebersberg, Germany).

Step	Temperature °C	Duration	Cycles
Initial Denaturation	98	30 s	
Denaturation	98	20 s	
Annealing	T_m + 3°C	60 s	35
Extension	72	30 s / kb	
Final extension	72	5 min	
Cooling	8	hold	

Table 4.13 PCR program for site-directed mutagenesis

Cloning into vectors by LIC or Gibson Assembly

The amplified PCR product was cloned into various pCB vectors, through BspQI digestion site, being generated by the lab of Christian Biertümpfel, by either Ligase independent cloning (LIC) or Gibson Assembly method. LIC relies on the generation of complementary single-stranded overhangs by T4 DNA polymerase, which contains a 3'-5' exonuclease activity Vector and PCR product can circularize based on their cohesive ends without ligase activity. Remaining nicks on the plasmid are repaired by *E. coli* RecA. First, the vector was linearized with restriction enzyme BspQI (NEB) for 1h at 50°C as described in 4.14. To separate between cut and uncut vector the digested vector was analyzed on a 0.8% agarose gel loading the equivalent volume of 250 ng vector per lane. The corresponding bands of linearized vector were cut and the DNA was isolated with NucleoSpin Gel and PCR Clean-up Kit (Macherey-Nagel, Düren, Germany). In order to generate the overhangs for annealing with the construct the linearized vector was processed with T4 DNA Polymerase (NEB) (see table 4.15). Likewise, the gel purified PCR product was processed with T4 DNA polymerase by mixing according to table 4.16. Insert and vector processing mixtures were incubated at room temperature for 30 min. The enzyme was heat inactivated by incubation for 20 min at 75°C. The annealing reaction was performed by mixing 4 µl T4 processed insert with 1 µl T4 processed vector and incubating 20 min at room temperature. After addition of 1µl EDTA (4.16 mM final concentration), the reaction mix was incubated for 10 min at room temperature. For transformation, chemical competent XL1-blue cells were mixed with 4 µl of annealing reaction and subjected to heat shock. A number of colonies were picked and expanded in liquid LB overnight, then plasmids were prepared with Nucleospin Plasmid Kit (Macherey-Nagel, Düren, Germany) and in order to check for correct insertion were sent to sequencing (Eurofins, Ebersberg, Germany). Gibson assembly is an isothermal, single-reaction method for assembling multiple, overlapping DNA molecules by the concerted action of a 5'-exonuclease, a DNA polymerase, and a DNA ligase. The DNA fragments are first recessed to produce ssDNA overhangs that are specifically annealed, and then they are covalently joined. Gibson master mix was provided by the Biochemistry Core Facility of Max Planck Institute of Biochemistry. The components of the 2x Gibson mix are: 5X isothermal (ISO) reaction buffer (25% PEG-8000, 500 mM Tris-HCl pH 7.5, 50 mM MgCl2, 50 mM DTT, 1 mM each of the 4 dNTPs, and 5 mM NAD). T5 exonuclease (Epicentre) Phusion DNA polymerase (New England Biolabs) Taq DNA ligase (New England Biolabs). The mixture comprises 1-5 fmol of linearized vector and 10-50 fmol

of each insert, as described in table 4.17. Reaction is performed in 20 µl final volume for 60 min at 50 °C. Subsequently 10 µl are transformed into chemically competent XL1-blue cells.

Component	Stock concentration	Amount
Vector		2 µg
Cutsmart buffer	10 x	4 µl
BspQI	10 000 U/ ml	8 U
ddH_2O		add to 40 µl

Table 4.14 Mixture for vector linearization with BspQI

Component	Stock concentration	Amount
Linearized vector		450 ng
NEB buffer 2.1	10 x	3 µl
T4 DNA polymerase	3 000 U/ ml	0.6 µl
dTTP	25 mM	3 µl
DTT	100 mM	1.5 µl
ddH_2O		add to 30 µl

Table 4.15 Mixture for T4 processing of the vector in preparation of LIC cloning

Component	Stock concentration	Amount
Linearized insert		600 ng
NEB buffer 2.1	10 x	2 µl
T4 DNA polymerase	3 000 U/ ml	0.4 µl
dATP	25 mM	2 µl
DTT	100 mM	1 µl
ddH_2O		add to 20 µl

Table 4.16 Mixture for T4 processing of the insert in preparation of LIC cloning

Component	Stock concentration	Amount
Linearized insert	-	10-50 fmol
Linearized vector	-	1-5 fmol
Gibson Assembly Mix	2x	10 µl
ddH_2O		add to 20 µl

Table 4.17 Mixture for Gibson assembly

Transformation of Competent Cells

Chemically competent XL1-blue cells were thawed from -80°C and directly put on ice. For transformation of plasmid ~10 ng and for LIC 4 µl annealing reaction were mixed respectively with 50 µl cells and incubated for 25 min on ice. The heat shock was applied at 42°C for 42 s and the cells were immediately cooled down on ice for 5 min. After adding 200 µl SOC-medium, the cells were shaken at 37°C 1000 rpm for maximum one hour. The bacteria were plated on agar-plates containing the respective antibiotic (table 4.1) and incubated over night at 37°C. The subsequent day colonies were picked, inoculated for expansion in 5ml LB media containing proper antibiotic, grown overnight at 37°C and the successive day plasmids were prepared with Nucleospin Plasmid Kit (Macherey-Nagel, Düren, Germany).

4.2.2 Protein expression

Large scale protein expression was typically carried out in 2l Erlenmeyer flasks with baffles (2l flasks were manufactured from Klimax; Glasgertebau Ochs introduced 3 baffles up to 30% of total flask height) containing 500 ml medium supplemented with appropriate antibiotics (table 4.1). The medium was inoculated with scraped colonies of freshly-transformed *E. coli* cells from LB agar plates. After inoculation the large scale bacterial cultures were raised at 37°C and 150 rpm to an optical density of OD 1.7 - 1.8. The temperature of the shaker was then reduced to 16°C and cells were cooled down at 4°C, cells were left shaking at 150 rpm for 16 h after the proper OD_{600} was reached, for protein induction. Cells were harvested by centrifugation (8000 g, 10 min) and either freshly used for protein purification or stored at -80°C.

4.2.3 Protein purification

In general, all purification steps were carried out at 4°C. Loading of the cleared lysate on the affinity and ion exchange (IEX) columns and washing steps were performed using a peristaltic pump (Reglo ICC, Ismatec). Linear gradients for elution were administered by ÄKTA Prime systems (GE Healthcare) by using mixtures of A and B buffers. Size-exclusion chromatography was carried out on ÄKTA Purifier system. Buffers used in purification procedures are listed in table 4.7

Purification of hsPCNA

Recombinantly expressed hsPCNA-CPD-Twin-Strep-tag (bax25-hsPCNA) was purified from freshly harvested or alternatively thawed pellets of BL21 gold cells. The cell pellet corresponding to 4 l bacterial culture was resuspended at 4°C with 250 ml lysis buffer (Lysis 1) additionally 1ml of lysozyme (10mg/ml) and 15 µl smDNaseI were added, cells were stirred for 20 min at 4°C and lysed by sonication (Bandelin Sonoplus, tip VS70, pulse ON/OFF 0.5/0.5 s, 40% amplitude, 15 min). The lysate was clarified by centrifugation at 75600 g for 30 min at 10°C The soluble fraction was loaded on a StrepTACTIN column (5 ml, IBA Lifescience GmbH, Göttingen, Germany) and washed with 30 column volume (CV) of buffer Strep 1. The immobilized protein complex was eluted with Strep 2 buffer containing 1mM IP_6. The self-cleavage of C-terminal tag CPD is triggered by the presence of IP_6 in the buffer, leading to elution of the untagged protein only. Elution was performed manually, loading 10 CV of buffer and collecting 1ml fractions. The remaining CPD-Twin-Strep-tag was displaced from the column by washing with Strep 3 buffer. The column was regenerated by subsequently washing with Strep 4 and Strep 5. The eluted fractions were analysed on SDS-PAGE, fractions containing pure cleaved hsPCNA were pulled an concentrated by ultra-filtration, Amicon Ultra ultrafiltration device with 10 kDa molecular weight cut-off (MWCO) on and subjected to gel filtration (Superdex 200 16/600 column, GE Healthcare) pre-equilibrated with SEC buffer (SEC 1). In the end, the purified proteins were analysed on SDS-PAGE, highly pure fractions were pooled, concentrated by ultra-filtration, aliquoted, flash frozen in liquid nitrogen and stored at -80°C. To assess protein quality samples were submitted to electrospray ionization time-of-flight (ESI-TOF) mass spectrometry (MPI Biochemistry Core Facility).

Purification of hsPCNA monomer

bax36-hsPCNA (K77D, C81E, K110D) monomeric mutant with C-terminal CPD-8x His tag was expressed in BL21 gold cells. After harvest of bacteria 3 l culture the pellet was resuspended in lysis buffer (Lys 1) 1ml of lysozyme (10mg/ml) and 15 µl smDNaseI were added, subsequently cells were lysed by sonication (Bandelin Sonoplus, tip VS70, pulse ON/OFF 0.5/1 s, 40% amplitude, 15 min). The lysate was clarified by centrifugation at 75600 g for 45 min at 10°C. After clarification and flirtation of the lysate, the soluble protein fraction was loaded on two connected Ni-columns (2x 5ml, cOmplete His-Tag purification column, Roche) and subsequently washed with 10 CV His A buffer (His A), 10 CV Chaperon Wash buffer and 10 CV His A buffer (His A). The immobilized protein complex was eluted with

buffer containing 1mM IP_6 in His A buffer, collecting 5ml fraction. The remaining CPD-8x His-tag was displaced from the column by washing with His B buffer for 10 CV. Presence of untagged hsPCNA was checked by loading fractions on SDS-PAGE. Fractions containing pure protein were concentrated by Amicon Ultra ultrafiltration device with 10 kDa molecular weight cut-off (MWCO) on and subjected to gel filtration (Superdex 75 16/600 column, GE Healthcare) pre-equilibrated with SEC buffer (SEC 1). Similar to wild type PCNA, the purified proteins were analysed on SDS-PAGE, highly pure fractions were pooled, concentrated by ultra-filtration, aliquoted, flash frozen in liquid nitrogen and stored at -80°C. To assess protein quality samples were submitted to electrospray ionization time-of-flight (ESI-TOF) mass spectrometry (MPI Biochemistry Core Facility).

Purification of hsPolι

Pellets corresponding to 6 l bacterial BL21 gold pRARE cells were thawed from -80°C storage and resuspended in 250 ml of Lysis buffer (Lysis 2) 1ml of lysozyme (10mg/ml) and 15 µl smDNaseI were added and the lysate was incubated for 20 min stirring at 4°C and subsequently sonicated (Bandelin Sonoplus, tip VS70, pulse ON/OFF 0.5/1 s, 40% amplitude, 15 min). The lysate was clarified by centrifugation at 75600 g for 45 min at 10°C, filtered through a 0.22 µm filter and loaded onto two connected Ni-columns (2x 5ml, cOmplete His-Tag purification column, Roche) and subsequently washed with 10 CV His A buffer, 10 CV Chaperon Wash buffer and 20 CV His A buffer. Elution was performed by applying His B containing 300mM imidazole and collecting 5ml fractions on the ÄKTA Prime systems (GE Healthcare) Fractions were analysed by SDS-page and the ones containing Polι band were pooled, diluted with Hep A buffer until salt concentration was 180 mM and then loaded on two connected Heparin columns (2x 5 ml, HiTrap Heparin HP, GE Healthcare) equilibrated with 9% Hep B on ÄKTA Prime systems. After short washing with 9% Hep B buffer until conductivity reached baseline, an isocratic elution gradient from 9% to 50% Hep B for 80 ml was applied, collecting 5 ml fractions. The protein eluted around 250-300 mM NaCl. Fractions were pooled, and incubated overnight with 1:100 v/v Senp2 for SUMO tag cleavage, rocking at 4°C. The day after the sample was concentrated by Amicon Ultra ultrafiltration device with 50 kDa molecular weight cut-off (MWCO) spun down and loaded on gel filtration (Superdex 200 16/600 column, GE Healthcare) pre-equilibrated with SEC buffer (SEC 2). SEC peak fractions were analysed on SDS-PAGE, pure fractions were pooled, concentrated by ultra-filtration, aliquoted, flash frozen in liquid nitrogen and stored at -80°C. To assess protein quality samples were submitted to electrospray ionization time-of-flight (ESI-TOF) mass spectrometry (MPIB-Core Facility).

Purification of hsPolη wild type, pip3 and 3PIPS mutant form bax25

The following constructs, Polη wt, pip3 and 3PIPS mutant, were cloned with a C-terminal CPD-Twin-Strep-tag (bax25) and expressed in 6 l cultures of BL21 Gold pLysS. Pellets were freshly harvested and resuspended in 250 ml Lysis buffer (Lysis 2) and 15 µl smDNaseI were added. The solution was left stirring 30 min at 4°C and lysed subsequently by sonication (Bandelin Sonoplus, tip VS70, pulse ON/OFF 0.5/0.5 s, 40% amplitude, 15 min). The lysate was clarified by centrifugation at 75600 g for 45 min at 10°C .The soluble fraction was then passed through a 0.22 µm filter and loaded on a StrepTACTIN column (5 ml, IBA Lifescience GmbH, Göttingen, Germany) and washed for 30 CV of buffer Strep1. The immobilized protein complex was eluted with buffer containing 1mM IP_6 (Strep 2 buffer). The self-cleavage of C-terminal tag CPD is triggered by the presence of IP_6 in the buffer, leading to elution of the untagged protein only. Elution was performed manually, loading 10 CV of buffer and collecting 1ml fractions. The remaining CPD-Twin-Strep-tag was displaced from the column by washing with Strep 3 buffer. The column was regenerated by subsequently washing with Strep 4 and Strep 5. The eluted fractions were analysed on SDS-PAGE and the ones containing full-length hsPolη were pooled, concentrated using Amicon Ultra ultrafiltration device with 50 kDa molecular weight cut-off (MWCO) spun down and loaded on gel filtration (Superdex 200 16/600 column, GE Healthcare) pre-equilibrated with SEC buffer (SEC2). SEC peak fractions were analysed on SDS-PAGE, highly pure fractions were pooled, concentrated by ultra-filtration, aliquoted, flash frozen in liquid nitrogen and stored at -80°C. To assess protein quality samples were submitted to electrospray ionization time-of-flight (ESI-TOF) mass spectrometry (MPI Biochemistry Core Facility).

Purification of all hsPolη constructs (wt, pip3, 3PIPS, 8A, UBZ dead and pip3 dead) with N-GFP-tag from bax29 or bax30-CPD

The following constructs, Polη wt, 8A, UBZ dead and pip3 dead were cloned into bax29 (N-terminal GFP-Twin-Strep-tag) and expressed in 6 l cultures of BL21 Gold pLysS. Pellets were freshly harvested and resuspended in 300 ml Lysis buffer (Lysis 2) 1ml of lysozyme (10mg/ml) and 15 µl smDNaseI were added. The solution was left stirring 15 min at 4°C and lysed subsequently by sonication (Bandelin Sonoplus, tip VS70, pulse ON/OFF 0.5/0.5 s, 40% amplitude, 15 min). The lysate was clarified by centrifugation at 75600 g for 45 min at 10°C. The soluble fraction was then passed through a 0.22 µm filter and loaded on a StrepTACTIN column (5 ml, IBA Lifescience GmbH, Göttingen, Germany) and washed for 30 column

volume of buffer Strep1. The immobilized protein complex was eluted with buffer containing 7.5 mM destiobiotin (Strep 3 buffer). Elution was performed manually, loading 10 CV of buffer and collecting 1ml fractions, which were analysed on SDS-PAGE using a GFP compatible-loading dye and without boiling the samples, to preserve GFP fluorescence. Gels were scanned by 650 nm laser with a Typhoon FL 7000 PhosphoImager (GE Healthcare, München, Germany). Afterwards gels were stained with Coomassie blue. Fractions containing full-length GFP tagged proteins were pooled, concentrated using Amicon Ultra ultrafiltration device with 50 kDa molecular weight cut-off (MWCO) spun down and loaded on gel filtration (Superdex 200 16/600 column, GE Healthcare) pre-equilibrated with SEC buffer (SEC2). SEC peak fractions were analysed on SDS-PAGE using a GFP compatible-loading dye and checked for fluorescent signal at 650 nm. Pure fractions were pooled, concentrated by ultra-filtration, aliquoted, flash frozen in liquid nitrogen and stored at -80°C. To assess protein quality samples were submitted to electrospray ionization time-of-flight (ESI-TOF) mass spectrometry (MPI Biochemistry Core Facility).

Constructs pip3 and 3PIPS, cloned into bax29, showed specific degradation during the above mentioned purification procedure. Mass spectrometry analysis revealed that both constructs were losing the last C-terminal PIP box 3. One degradation product, coming from the 3PIPS attempted purification, showed to have intact PIP box 1 and PIP box 2 p21 substitutions, but was missing the PIP box 3. This construct was kept and refer to as 2PIPS and used in pulldown assays. Polη pip3 and 3PIPS constructs were then cloned into a bax30-CPD vector, with a N-terminal GFP (8x His@212)-3C tag and C-terminal CPD- Twin-Strep-tag. The presence of the C-CPD-Twin-Strep tag prevented the C-terminal degradation at the level of the PIP box 3. These two constructs were expressed in 6 l cultures of BL21 Gold pLysS and purified as described for bax25 constructs, as described before. Briefly, cells were harvested and pellets resuspended in 300 ml Lysis Buffer (Lysis 2) and 15 µl smDNaseI were added. The solution was left stirring 20 min at 4°C and lysed subsequently by sonication (Bandelin Sonoplus, tip VS70, pulse ON/OFF 0.5/0.5 s, 40% amplitude, 15 min). The lysate was clarified by centrifugation at 75600 g for 45 min at 10°C .The soluble fraction was then passed through a 0.22 µm filter and loaded on a StrepTACTIN column (5 ml, IBA Lifescience GmbH, Göttingen, Germany) and washed for 30 column volume (CV) of buffer Strep 1. The immobilized protein complex was eluted with buffer containing 1mM IP_6 (Strep 2 buffer). Elution was performed manually, loading 10 CV of buffer and collecting 1ml fractions. The eluted fractions were analysed on SDS-PAGE using a GFP compatible-loading dye and checked for fluorescent signal at 650 nm. Fractions containing pure proteins were pooled, concentrated using Amicon

Ultra ultrafiltration device with 50 kDa molecular weight cut-off (MWCO) spun down and loaded on gel filtration (Superdex 200 16/600 column, GE Healthcare) pre-equilibrated with SEC buffer (SEC 2). SEC peak fractions were analysed on SDS-PAGE followed by GFP detection, pure fractions were pooled, concentrated by ultra-filtration, aliquoted, flash frozen in liquid nitrogen and stored at -80°C. To confirm the correct molecular weight proteins were submitted to electrospray ionization time-of-flight (ESI-TOF) mass spectrometry (MPI Biochemistry Core Facility).

Purification of hsUB

The gene for human ubiquitin was cloned in a vector without tag, and ubiquitin purification was carried out by $HClO_4^-$ precipitation. Briefly, BL21 gold cell pellets corresponding to 2 l culture were resuspend in 50 ml Lysis buffer 3 and 1ml of lysozyme (10mg/ml) and 15 µl smDNaseI were added. The solution was left stirring 30 min at 4°C and lysed subsequently by sonication (Bandelin Sonoplus, tip VS70, pulse ON/OFF 0.5/0.5 s, 40% amplitude, 15 min). lysate was spun down at 75600 g for 45 min, transferred to a glass beaker (on ice) with a stir bar in it. Slowly, over a period of 3 min 350 µl of 70% $HClO_4^-$ were added, in small drops, until pH drops to 4.5. Fundamental, in order to avoid local pH dropping too fast, was to slowly add small drops of the acid, and keep the solution stirring on ice. The "milky" solution was kept stirring on ice for few more minutes, until the pH stabilised around 4.5, pH was measured using pH paper. The solution was subsequently spun at 75600 g for 30 min to separate precipitate proteins from soluble ubiquitin. The supernatant was dialyzed against 2 l of Dialysis buffer overnight at 4°C. The day after a second run of dialysis was carried out, as previously described, for 4h, followed by loading of the sample on on SP cation exchange column (5 ml, HiTrap SP HP, GE Healthcare) equilibrated with SP-A buffer. The column was moved to ÄKTA Prime for washing with 4 CV of SP-A until conductivity reached baseline. The protein was eluted with a linear gradient from 0% to 50% SP-B buffer, in 10 CV, and 5 ml fractions were collected. Fractions were analysed on a 18%gel for SDS-PAGE, and the ones containing pure ubiquitin were concentrated by Amicon Ultra ultrafiltration device with 3 kDa MWKO, spun down and loaded on Superdex 75 16/600 column, GE Healthcare) pre-equilibrated with SEC buffer (SEC1). Peak fractions were then analyzed on gel and fractions containing 95% pure ubiquitin were pooled, concentrated by Amicon Ultra ultrafiltration device with 3 kDa MWKO to a final concentration of 214 µM , aliquoted, flash frozen in liquid nitrogen and stored at -80°C. To assess protein quality a small aliquot of diluted sample was submitted to electrospray ionization time-of-flight (ESI-TOF) mass spectrometry (MPI Biochemistry Core Facility).

4.3 Biochemical assays

Determination of protein concentration and purity

Protein purity was assessed by SDS-PAGE, by monitoring the absorption ratio A260/A280 during size-exclusion chromatography and by Nanodrop measurements. Protein concentrations were measured by NanoDrop spectrophotometer (ThermoFisher Scientific (Waltham, MA, USA) at absorbance 280nm and 260nm. The extinction coefficients were calculated using the ProtParam tool (ExPASy, SIB bioinformatics resource portal).

SDS-PAGE

Purity of protein was assessed by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) according to Laemmli (Laemmli, 1970) using homemade 12.5, 15 or 18% gels. For analytical purposes like pulldown assay, pre-cast 4-20% gradient Tris-Glycine gels were purchased from Bio-Rad (München, Germany). SDS-Gels were stained with Coomassie brilliant blue or in case of very faint signal with Oriole Fluorescent gel Staining Bio-Rad (München, Germany).

Native PAGE and denaturing urea PAGE for DNA and protein analysis

To analyse tetrameric structure of PCNA and UbPCNA Native PAGE was carried out using 6% Tris-glycine gels (1x Tris-glycine, 6% 19:1 Acrylamide:Bisacrylamide, 0.2% TEMED, 0.08% APS) were 150 pmols of protein was loaded . Gels were stained with Coomassie blue for protein detection. To analyse the DNA-binding properties of Polη mutants, electro mobility shift assay (EMSA) was carried out with self- made Tris-Borate-EDTA gels (1x Tris-Borate-EDTA, 8% 19:1 Acrylamide:Bisacrylamide, 1% glycerol 0.2% TEMED, 0.08% APS,) left polymerize for 30 min at RT. After loading the samples, the gels were ran into an ice container for 120 min at 90 V in 1x TBE. The gels were analysed by fluorescence imaging at 650 nm with a Typhoon FL 7000 PhosphoImager (GE Healthcare, München, Germany). To resolve polymerase activity reactions, carried out with *in vitro* primer extension assay, products were loaded onto denaturing urea PAGE. Briefly the gels, composed of 1x Tris-Borate-EDTA, 20% 19:1 Acrylamide:Bisacrylamide, 8M urea, 0.2% TEMED, 0.08% APS, were casted in house and left polymerise for 45 min at RT. After loading of the samples the gels were ran in 1x TBE buffer for 60 min at 5W at RT. Gels were analysed by fluorescence imaging (excitation 473 nm, filter 520 nm) via a Typhoon FL 7000 PhosphoImager, GE Healthcare).

4.3.1 Enzymatic PCNA-K164 mono-ubiquitination

PCNA ubiquitination reaction was performed according to Hibber and Sixma (Hibbert and Sixma, 2012) with slight modifications. Briefly, reactions were performed in ubiquitination buffer (table 4.7) by mixing of 160 pmols of PCNA, 3.8 pmols of Uba I, 320 pmols of UbcH5c S22R or wild type, 1.28 nmols of ubiquitin, adding H_2O to reach a final volume of 40 µl. UbaI and UbcH5c (wt or S22R) were enzymes already present in the Biertümpfel lab. In some cases, the UbcH5c ubiquitin ligase was pre incubated with 1:200 w/w Senp2, 30 min at RT, for SUMO tag cleavage. Ubiquitination reactions were incubated at 37 °C for 2 hours, or alternatively 2h at RT followed by ON at 4°C. 10x scaled up reactions were set up for large scale production of UbPCNA, often running 2 or more reactions in parallel. Mono-ubiquitinated PCNA was separated by the rest of reaction components via size-exclusion chromatography on Superdex 200 10/300 column, GE Healthcare) pre-equilibrated with SEC buffer (SEC1). Fractions were analysed by SDS-PAGE and the ones containing UbPCNA species were pooled, concentrated by Amicon Ultra ultrafiltration device with 10 kDa MWKO to a suitable working concertation, flash frozen in liquid nitrogen and stored at -80°C.

4.3.2 Native mass spectrometry

Samples of PCNA and UbPCNA were buffer exchanged in 100 mM ammonium acetate buffer, a volatile salt compatible with the technique. Both samples, PCNA and UbPCNA were concentrated and simultaneously buffer exchanged in Amicon Ultra ultrafiltration device with 10 kDa MWKO until a final concentration of 0.1 mg/ml and were subjected to native state analysis using maXis II™ ETD (Bruker). The whole experimental procedure and data analysis was carried out by Dr. Nagarjuna Nagaraj in the Biochemistry Core Facility at the Max-Planck Institute of Biochemistry (Martinsried, Germany).

4.3.3 Primer extension assay

To test the activity of TLS polymerases an *in vitro* primer extension assay was set up. Briefly, DNA substrates CBs284, CBs261F, CBs201F (table 4.8) were dissolve in annealing buffer (50 mM Tris-HCl pH 8.0, 50mM NaCl, 0.1 mM EDTA). Equimolar amounts of 5'FAM labelled primer and unlabeled template were incubated. Annealing reaction was performed by heating to 85°C for 5 min, followed by slow-cooling to room temperature. Annealed primer template DNA substrates were then diluted in water to working concentrations for primer extension assay.

To test polymerase activity through primer extension assay a reaction mixture was prepared by combining 1x TDB buffer (table 4.7), 5mM divalent cation (MgCl2 or MnCl2), 500 nM DNA 5'FAM-primer/template and 5 μM deoxynucleotide mixture. 9 μl of mix was aliquoted in tubes and 1μl of enzyme (500 nM final concentration) was added, to start the reaction in a total of 10 μl volume. Reactions were incubated for 10 min at 37 °C and terminated by adding 5 μl of stop solution (90% formamide, 1x TBE, 1mg/ml bromophenol blue) heated to 95°C for 5 min and immediately cooled down. Samples were analysed by running 20% TBE denaturing urea gel as described before. For mis-incorporation or nucleotide titrations the reaction was set up by combining 1x TDB buffer (table 4.7) 5mM divalent cation ($MgCl_2$ or $MnCl_2$), 500 nM DNA 5'FAM-primer/template, 250 nM of polymerase and different amounts of PCNA/UbPCNA. The reaction was started by the addition of 1 μl of different nucleotides (dATP, dTTP, dCTP, dGTP or dNTPs mixture) at various concentrations, in a final volume of 10 μl and incubated at 10 min at 37 °C and terminated as previously described. Samples were analysed by running 20% TBE denaturing urea gel as described before.

4.3.4 Electromobility shift assay (EMSA)

For EMSA analysis, proteins were incubated with DNA substrates for 15 min on ice in EMSA buffer (table 4.7). Proteins were titrated at different concentration form 0 to 400 nM, and DNA was at a final concentration of 50 nM. PCNA and UbPCNA were kept constant at 400 nM in the conditions where required. The formed complexes were analysed by native PAGE. For first analysis DNA binding was analysed in a protein concentration range from 0.25 to 2.5 or 5 μM. Gels were analysed by fluorescence imaging (excitation 473 nm, filter 520 nm, Typhoon FL 7000 PhosphoImager, GE Healthcare).

4.3.5 Complex reconstitution by size-exclusion chromatography

For analytical size-exclusion chromatography, the protein-DNA and protein-protein complexes as well as single proteins, were buffer exchanged and concentrated using Amicon Ultra ultrafiltration device with 10 kDa molecular weight cut-off (MWCO) to Complex buffer (table 4.7), variation on the salt in the buffer where applied as described in main text. Samples were spun down and injected on analytical Superdex 200i 3.2/300 (ÄKTAmicro system (GE

Healthcare). Fractions were collected, mixed with 6x SDS sample buffer and analysed on SDS-PAGE. Fractions containing proteins of interest were further submitted to biochemical characterisation and structural analysis by negative stain or cryo electron microscopy as described in main text.

4.3.6 Thermal shift assay (Thermofluor)

For Thermofluor analysis different protein-DNA complexes, assembled as described in section 4.3.5, were mixed with 1µ of 1:10 Sypro Orange dye (5000x, Invitrogen) in a 96-well PCR plate (Eppendorf). The plate was sealed and heated in a real-time PCR system (Eppendorf) from 20°C to 80°C applying increments of 0.5°C. Changes in fluorescence were observed simultaneously with excitation and emission wavelengths of 470 nm and 550 nm, respectively. To obtain the binding constant, a Boltzmann model was used to fit the fluorescence data:

$$y = Top + \frac{Bottom - Top}{1 + e^{\frac{x - Kd}{dx}}}$$

The entire set of Thermofluor measurements and data analysis were carried out in quadruplets by Dr. Fabian Bonneau and Dr. Claire Basquin.

4.3.7 Negative stain EM sample preparation and observation

The technique of negative staining allows the absorption of the protein sample to a thin layer of ~10 nm continuous carbon support film, which is rendered hydrophilic by glow discharging. Afterwards the sample is coated by a layer of a heavy metal salt, usually uranyl acetate, which will strongly scatter the electrons hitting the sample once inside the EM. This will produce a background or silhouette image of the protein (hence a 'negative stain' image). This technique has the advantages to be fast and easy to perform at RT and allows screening of initial conditions like particles concentration in order to evaluate heterogeneity, purity, dispersion and aggregation state.

For negative-stain EM, homemade carbon-coated grids were prepared and glow discharged for 30 seconds using a GloQube Plus Glow discharge system (Quantum Design GmbH, Darmstadt, Germany) right before use. 5 µl of sample was applied and incubated for 1 min, blotted, washed three times in water, blotted and stained in 2% uranyl acetate for 1 min. The prepared specimens

were visualized with a FEI CM200 with an operating voltage of 160 kV equipped with an Eagle CCD camera with a pixel size of 2.16 Å/pix. For 2D classifications, particles were manually or automatically picked using RELION or EMAN2. Particle extraction and reference free 2D classification was performed in RELION.

4.3.8 Cryo-EM sample preparation

For preparation of cryo-EM samples, complexes or single proteins were collected from analytical SEC or directly after mixing and concentration, according to specific cases. In some cases samples were subjected to crosslinking with 0.05% Glutaraldehyde for 10 min at RT, quenched with 100mM Tris-HCl pH 7.5, for other samples (single UbPCNA) no crosslinker was applied. Samples were then supplemented with detergent, either 0.005% Tween-20 or 0.04% β-OG, which was added just before plunging. 3 µl of sample (at different concentrations, ranging from 0.2-0.8 mg/ml) were placed on a glow discharged R 1.2/1.3 Cu200 mesh grid (Quantifoil), blotted for 3.5 sec at force 4, and plunge-frozen immediately using a Vitrobot Mark IV (FEI/ThermoFisher). Grids were screened for particle concentration on Talos Arctica (FEI/ThermoFisher) and initial dataset were collected with Falcon 2 direct electron detector (FEI) using the EPU software (FEI). Final datasets were collected on a Titan Krios (FEI/ThermoFisher) 300 kV TEM, equipped with a K2 direct electron detector (Gatan). Each different cryo-EM dataset was collected with different parameters for specific samples and aims.

4.3.9 Cryo-EM data collection and analysis

The dataset yielding to PCNA reconstruction was collected on sample composed of PCNA-Polι and DNA that dissociated after freezing. The dataset was collected overnight on a Titan KRIOS equipped and energy filter and with K2 summit camera (Gatan). 1054 movies were collected at 105 000 x magnification (1.35 Å/pixel) using SerialEM at a total dose of 42 $e^-/Å^2$. The 32 frames of each micrograph were imported in RELION 2.1 and aligned using its MotionCor2 own implementation. CTF estimation was performed using Gctf and particles were picked using Gautomatch,. Coordinates were imported in RELION and particles were extracted with a box size of 200 pix corresponding to 270 Å. Several runs of classification were carried out in 2D. Classes containing PCNA only classes were identified and separated from complex-bound classes. PCNA 2D classes were further sorted in 2D and a final pool of ~6k particles was used

to generate an initial model using SGD algorithm in RELION. 3D classification with 4 classes yielded to equal classes distribution so all particles were subjected to 3D auto-refine and postprocessing using a tight mask created in RELION. Final PCNA resolution reached 6.8 Å according to FSC gold standard. The complex bound particles, were classified in 2D, yielding to a final set of ~40k particles. Those particles were sorted in 3D classification using a *de novo* generated initial model in RELION. Several steps of 3D classification and refinement, with different K and T parameters were performed, but the final resolution of the complex-bound map was limited.

The dataset of UbPCNA (sample at 0.5 mg/ml) was collected on a Talos Artica 200 kV scope, equipped with a Falcon 2 direct detector (FEI). Data collection was performed by EPU. 799 total micrographs movies were collected at 92k x magnification (1.61 Å/pix) with a total dose of 69 electron/Å^2. The 40 frames movies were aligned using MotionCor2 with dose weighting implementation in RELION 3.0, a total of 1.8 million particles were picked by Gautomatch, extracted with a box size of 156 pixels (corresponding to 250Å) and subjected to several runs of sorting and classification in 2D. Many 3D initial models were generated in RELION 3.0, according to the three different populations of UbPCNA observed. Extensive 3D classification was performed but due to sample heterogeneity no high resolution feature was obtained. Classes showing only one ubiquitin bound to PCNA were selected and refined separately in 3D, yielding to a final map of ~ 7Å resolution.

The dataset which lead to the Polι-DNA map was collected on a sample composed PCNA-Polι-DNA diluted with a 4x excess of DNA after size-exclusion elution, and before crosslinking with 0.05% glutaraldehyde. The sample dissociated and only Polι-DNA complexes survived the vitrification. The dataset of 7971 micrographs movies was collected on Titan KRIOS 300 kV TEM with K2 summit detector (Gatan) at a magnification of 215k x, corresponding to 0.65 Å/pixel using SerialEM. The total dose was 98 electron/Å^2 and movies of 50 frames were recorded in counting mode. Motion-corrected and dose-weighted micrographs by MotionCor2 were used for particle picking and 220k initial particles were picked with Gautomatch, and extracted in RELION with a box size of 336 pixels (220Å), CTF correction was performed by Gctf in RELION and all further processing steps were done using cryoSPARC (Punjani et al., 2017b). Several rounds of 2D classification were performed, then *ab intio* 3D classification was performed followed by heterogenous refinement and a final set of 23k particles was submitted to homogeneous refinement, which reached a final resolution of 9 Å. Masking and subtracting the DNA density and focusing on the polymerase part didn't bring any improvements in

processing outcomes. Fitting of the PDB: 1T3N was performed in Chimera (UCSF) by rigid body fitting.

The final dataset collected on UbPCNA-Polη-DNA complex was collected over 3 days on a Titan KRIOS equipped and energy filter and with K2 summit camera (Gatan). A total of 11k micrographs movies were collected in counting mode at a magnification of 130k x, corresponding to 1.06 Å/pixel. Movies were composed of 40 frames and the total electron dose was 72 electrons/$Å^2$. The data were processed using RELION 3.0 pipeline, MotionCor2 RELION own implementation was used to correct for induced beam motion and align movies with Dose Weighting option, CTF correction was done by Gctf. Template free particle picking was performed by Gautomatch. Particles were extracted with a box size of 232 pixels and binned 1x for the 2D classification runs. 3D initial model was generated in RELION from a subset of ~60k selected particles and several runs of 3D classification with different K and T parameters were performed, to sort out heterogeneity. Finally, a set of 175k particles was subjected to iterative runs of 3D classification, initial model generation and 3D auto-refine, particles were unbinned and a final 3D map was obtained at a nominal resolution of 7Å. PDB files of Polη catalytic domain (3MR2), PCNA (1AXC) and ubiquitin (1UBQ) were manually fitted into the density map using Chimera fit into map option.

4.3.10 Pulldown assays

For pulldown assays all reactions and were performed in 1.5 ml LoBind tubes (Eppendorf) to minimize beads sticking. Volumes for binding and washing for all the step of the procedure were of 20 µl, the final elution volume was of 10 µl 1x SDS buffer. The full procedure was performed on ice, to avoid protein degradation. The magnetic beads (GFP-Trap®_MA from Chromotek) were concentrated using a magnetic rack DynaMag™-2 Magnet (Thermo Fisher Scientific). Initially, 5 µl of beads slurry were washed 3x in 20 µl Pulldown buffer 1 (table 4.7). Then 20 µl of 0.5 µM Polη bait (total of 10 pmols) diluted in Pulldown buffer 1was incubated on beads for 1h on ice, for each experimental condition, with gentle manual twirling every ~10 min. A sample of bait unbound fraction (bait FT) was collected for SDS-PAGE analysis. Beads were washed once with Pulldown Buffer at proper salt concentration for the experimental setup (see table 4.7) and beads were incubated for another hour with different preys conditions. Respectively different preys conditions included various protein concertation, protein type or salt concentration in buffer. Generally 20 µl of prey were incubated for 1h on ice, with gentle

manual mixing. Next unbound preys were collected, and samples kept for eventual SDS-PAGE analysis of preys unbound fraction (prey FT). Beads were accurately washed 3x in proper buffer, according to the condition, and finally beads were dissolved in 10 µl of 1x SDS loading buffer, heated 3 min at 95 °C and loaded on SDS-PAGE without spinning the tubes. Samples collected at different steps of the procedure like inputs, unbound baits or unbound prey and washing were also occasionally run on gels, to confirm that protein amounts were consisted and that no material was lost during the whole procedure. Samples were run on 4-20% gradient Tris-Glycine gels (Bio-Rad) and stained by Coomassie brilliant blue. Technical triplicates were performed for each pulldown.

The protocol followed for quantification and kinetic parameters estimation was adapted from Lapetina and Gil-Henn (Lapetina and Gil-Henn, 2017). Briefly, gels were scanned, and bands were quantified using Fiji/ImageJ (Schindelin et al., 2012) by drawing a tight box around the band and subtracting background signal for each single lane. Intensity of each band was normalised internally to each gel, to a Polη band as. Relative intensity values were analysed using GraphPad Prism version 9.0.0. (GraphPad Software, San Diego, California USA) using XY tables with three replicates values. Data were plotted using grouped graph type summary data histogram. Additionally attempts of calculating K_d values were attempted in pulldowns with sliding clamp titrations. Prism fit with non-linear regressions and one site specific binding model was performed according to Lapetina and Gil-Henn protocol (Lapetina and Gil-Henn, 2017). R squared parameters were bad (<0.5) and numerical K_d values were not reasonable, therefore weren't considered as correct. Final gel images and graphs were assembled and modified using Inkscape version 0.48.2.

Appendix

A.1 No complex formation is observed on SEC at low concentrations

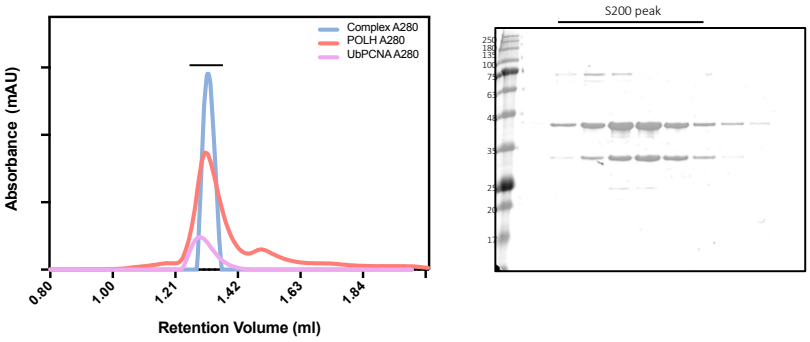
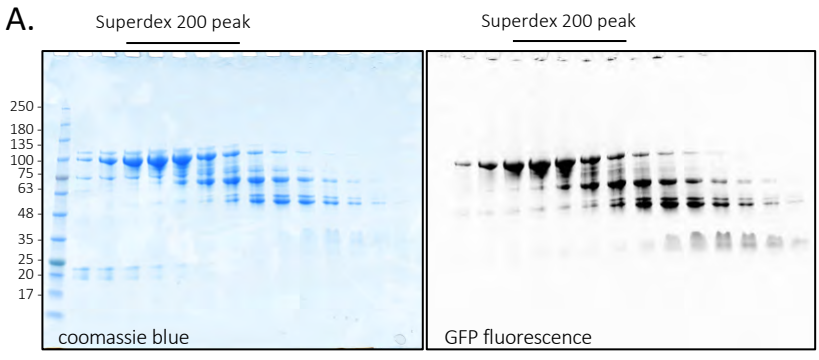


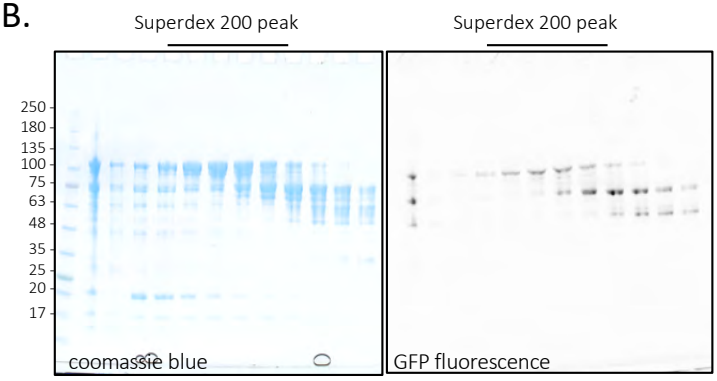
Figure A1. Size exclusion chromatography of UbPCNA-Polη complex at 1 μM. SEC chromatogram from S200i analysis of premixed UbPCNA-Polη at 1 μM in complex buffer. Blue line represents complex elution (absorbance atA280 nm), salmon line is Polη controls (A280) and pink line is UbPCNA control (A280). As can be seen, there is no shift when the proteins are mixed together at low concentrations (1 μM), indicating complex disassembling. On the right side SDS-PAGE confirms no co-elution of UbPCNA and Polη.

A.2 Purification of all Pol constructs with a N-GFP tag



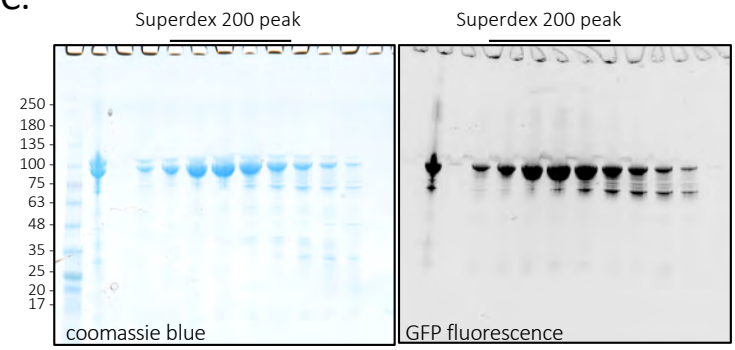
Construct: bax29- Polη wt [N-GFP (Twin-Strep-tag @212)-3C]
Theoretical mass: 109497.530 Da
Experimental mass: 109493.277 Da

B.



Construct: bax29- Polη pip3 [N-GFP (Twin-Strep-tag @212)-3C]
Theoretical mass: 108420.551 Da
Experimental mass: 106797.155 Da Missing the last 25 aa

C.



Construct: bax30-CPD Polη pip3 [N-GFP (His@212)-3C and C-CPD-Twin-Strep-tag]

After CPD- cleavage -> Theoretical mass: 108420.551 Da
Experimental mass: 108416.133 Da

D.

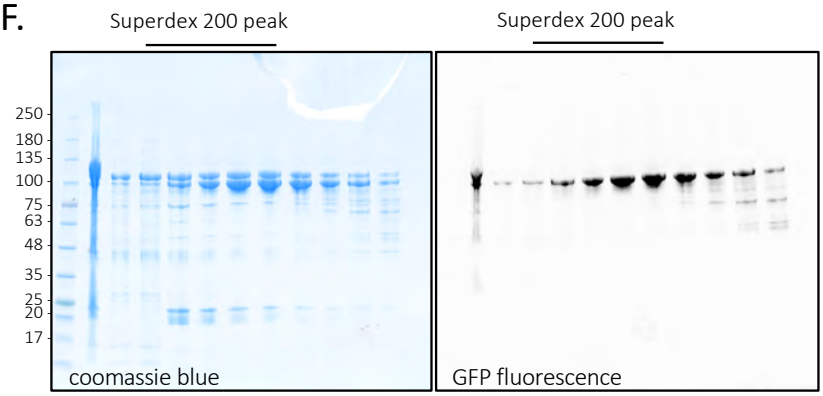
Construct: bax29- Polη 3PIPS [N-GFP (Twin-Strep-tag @212)-3C]
Theoretical mass: 108805.239 Da
Experimental mass: 107182.539 Da Missing the last 25 aa -> kept as Polη 2PIPS

E.

Construct: bax30-CPD Polη 3PIPS [N-GFP (His@212)-3C and C-CPD-Twin-Strep-tag]

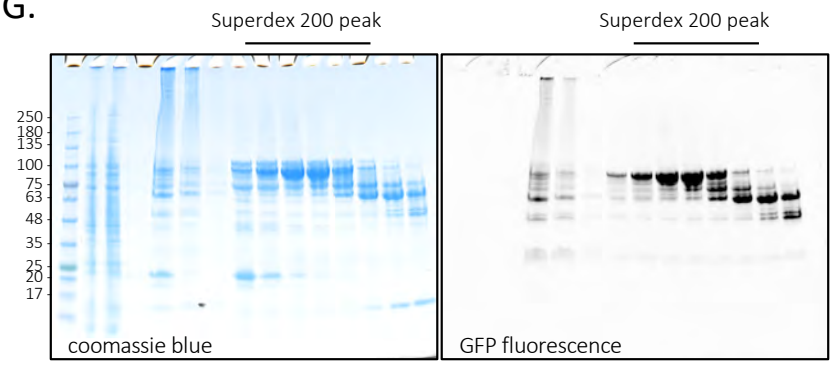
After CPD- cleavage -> Theoretical mass: 108805.239 Da
Experimental mass: 108801.561 Da

F.



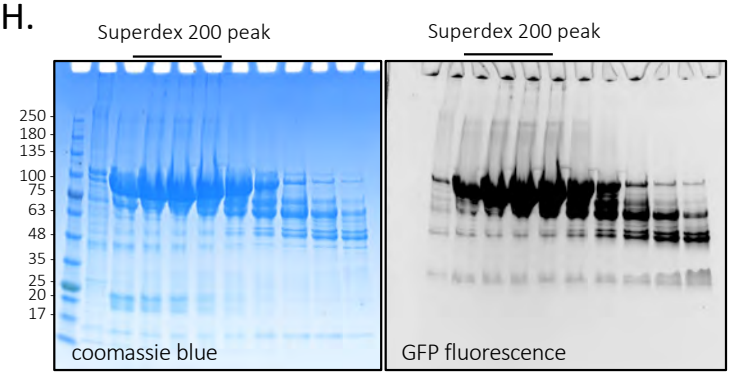
Construct: bax29- Polη 8A [N-GFP (Twin-Strep-tag @212)-3C]
Theoretical mass: 108923.167 Da
Experimental mass: 108918.512 Da

G.



Construct: bax29- Polη UBZ dead [N-GFP (Twin-Strep-tag @212)-3C]
Theoretical mass: 109454.022 Da
Experimental mass: 109449.267 Da

H.



Construct: bax29- Polη pip3 dead [N-GFP (Twin-Strep-tag @212)-3C]
Theoretical mass: 109344.903 Da
Experimental mass: 109341.082 Da

Figure A2. Purification of all Polη constructs with N-GFP tag. (a) Purification of bax29-Pol η wt, SDS page Coomassie stained and GFP fluorescence detection. Theoretical and experimental MS values are indicated. (b) Purification of bax29-Pol η pip3, SDS page Coomassie stained and GFP fluorescence detection. Theoretical and experimental MS values are indicated and show missing of the last 25 aa a the C-terminus. (c) Purification of bax30-CPD-Pol η pip3, SDS page Coomassie stained and GFP fluorescence detection. Theoretical and experimental MS values are indicated and show correct value of FL N-GFP-Pol η pip3. (d) Purification of bax29-Pol η 3PIPS, SDS page Coomassie stained and GFP fluorescence detection. Theoretical and experimental MS values are indicated and show missing of the last 25 aa a the C-terminus, the construct was kept as N-GFP-Polη 2PIPS. (e) Purification of bax30-CPD-Pol η 3PIPS, SDS page Coomassie stained and GFP fluorescence detection. Theoretical and experimental MS values are indicated and show correct value of FL N-GFP-Pol η 3PIPS. (f) Purification of bax29-Pol η 8A, SDS page Coomassie stained and GFP fluorescence detection. Theoretical and experimental MS values are indicated. (g) Purification of bax29-Pol η UBZ dead, SDS page Coomassie stained and GFP fluorescence detection. Theoretical and experimental MS values are indicated. (h) Purification of bax29-Pol η pip3 dead, SDS page Coomassie stained and GFP fluorescence detection. Theoretical and experimental MS values are indicated.

A.3 PCNA monomer elution profile from SEC

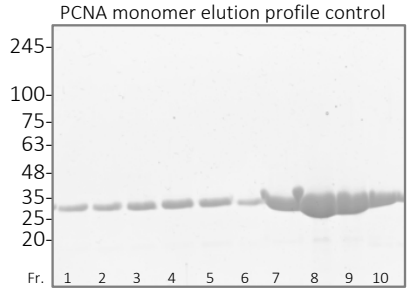


Figure A3. PCNA monomer elutions profile from s200i. SDS-PAGE of fractions corresponding to the elution profile of PCNA monomer on a S200i 2.3/200 column. The fraction number is indicated and corresponds to the same numbers showed in main text (Figure 2.15) and indicates a broad elution of PCNA monomer also in earlier fractions (1-6).

A.4 Pulldown of Polη with PCNA monomer

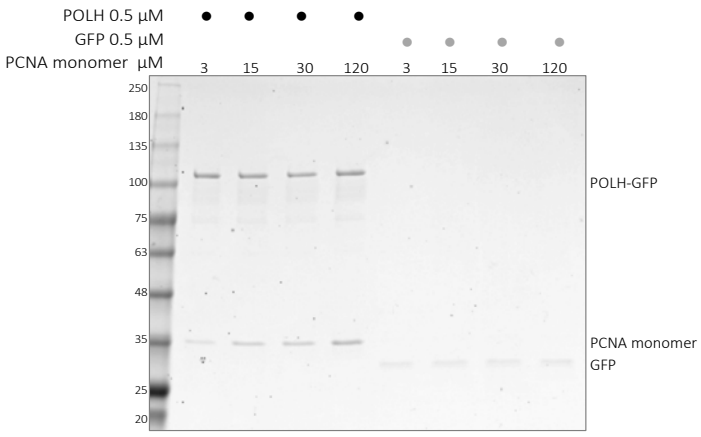


Figure A4. Pulldown of Polη and PCNA monomer. Pulldown assay performed by immobilising 0.5 µM of Polη as bait and by titrating different amounts of PCNA monomer as preys (amount indicated in figure). Elutions show appearance of PCNA band, indicating recruitment of PCNA monomer to Polη. The second half of the gel shows no unspecific interaction of PCNA monomer to GFP.

A.5 Cryo-EM dataset heterogeneity and particles disassembly in different conditions

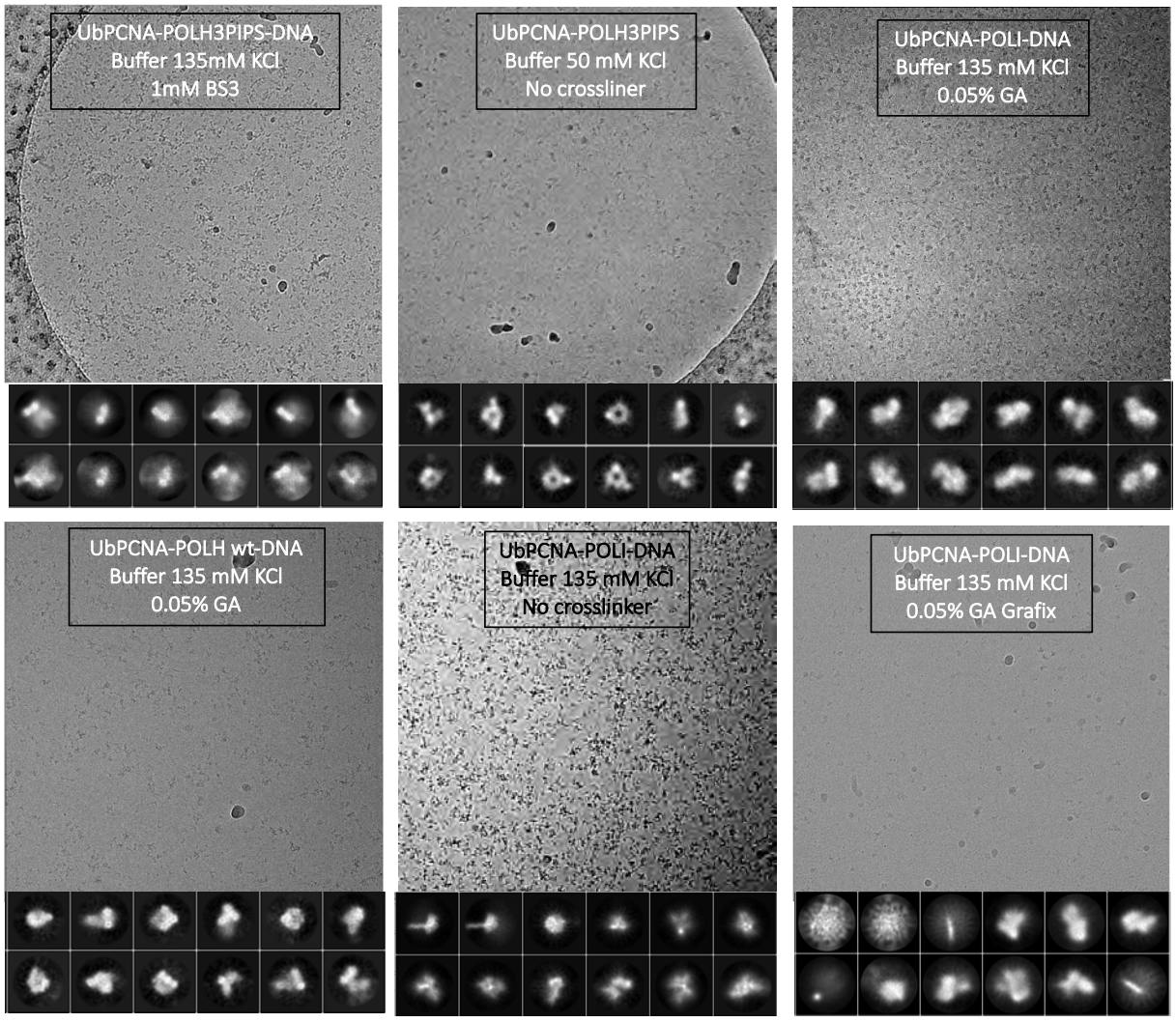


Figure A5. Screening of different complexes in cryo-EM with different conditions. Summary of few representative cryo-EM datasets collected on Arctica. As can be appreciated from micrographs and 2D classes, regardless of sample type and reconstitution (+/- DNA, presence UbPCNA or PCNA, or polymerase type) buffer salt concentration (135 or 50 mM KCl), crosslinker type and approach (glutaraldehyde GA, or BS3, batch or GraFix) all dataset showed particles disassembly and heterogeneity.

A.6 Thermofluor analysis of complex

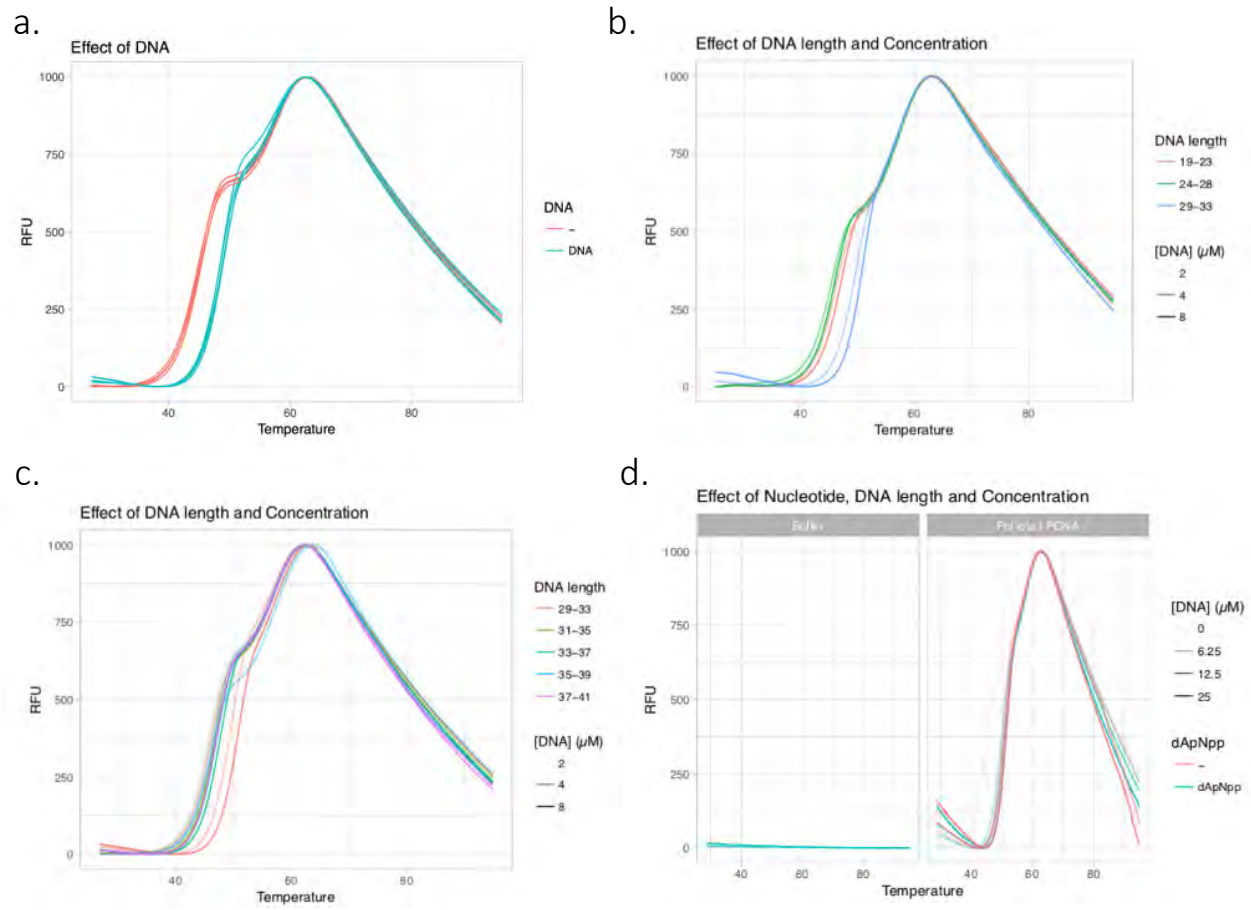


Figure A6. Theromofluor analysis of the effect of DNA on complexes. The Thermofluor analysis was conducted using PCNA-Polι complex at 2 μM concentration, the all experimental setup and data analysis was performed by Dr. Fabien Bonneau. (a) Effect of presence of a 29/23 nt DNA. (b) effect of DNA length and concentration, with short DNA lengths (see Legend). (c) effect of DNA length and concentration, with long DNA lengths (see Legend). (d) General effect of presence of higher concentrations of 29/33 nt DNA and dAMPNPP. In general, fluorescent curves shifted to the right (higher Tm) indicate more stable complexes

A.7 Comparative cryo-EM analysis of Polη wt and PIP mutants without crosslinker at 50 mM KCl

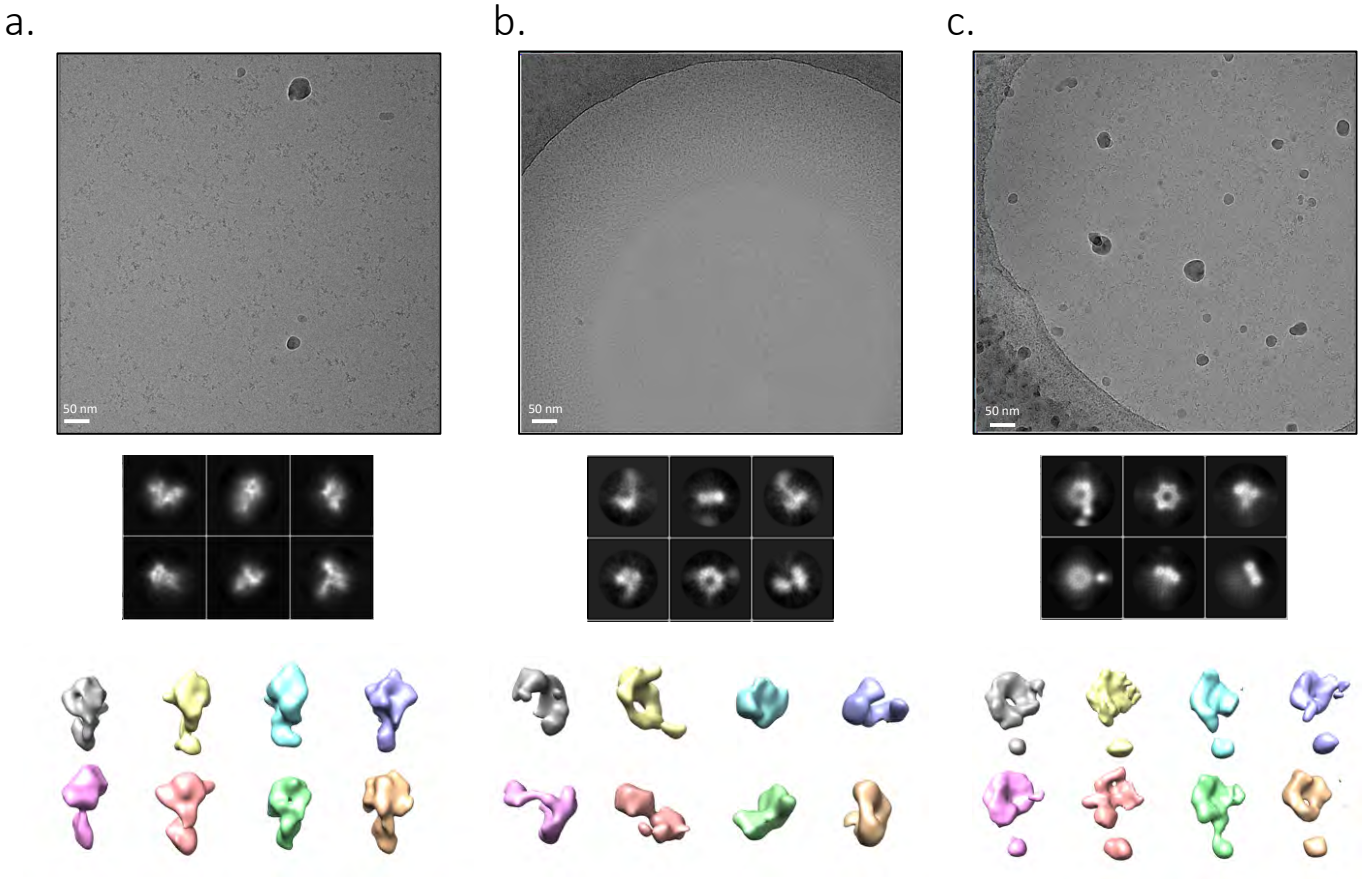


Figure A7. Comparative cryo-EM analysis of Polη wt and PIP mutants without crosslinker at 50 mM KCl. (a) Shows representative micrograph (scale bar 50 nm), 2D class averages and 3D classification of a dataset collected in low salt buffer (50 mM KCl) in absence of crosslinker on the Polη wt -UbPCNA-DNA complex. (b) Shows representative micrograph (scale bar 50 nm), 2D class averages and 3D classification of a dataset collected in low salt buffer (50 mM KCl) in absence of crosslinker on the Polη pip3-UbPCNA-DNA complex. (c) Shows representative micrograph (scale bar 50 nm), 2D class averages and 3D classification of a dataset collected in low salt buffer (50 mM KCl) in absence of crosslinker on the Polη 3PIPS-UbPCNA-DNA complex. In all cases there is are no clear complex-looking 2D or 3D classes, indicating disassembling of proteins regarding of low salt or PIP mutations. All dataset were collected on Arctica and processed in parallel with similar parameters.

A.8 Activity assays via primer extension of Polη wt and PIP mutants

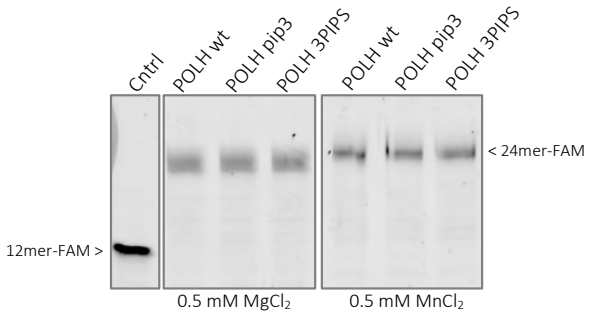


Figure A8. Primer extension of Polη wt and PIP mutants. Primer extension activity of Polη construct on a 12/24 primer/template DNA substrate with a 5'-FAM labelled primer. Activity is evaluated in presence of $MgCl_2$ and $MnCl_2$ as divalent cation, showing slight more processivity to full-length (24mer) in presence of manganese, for all Polη constructs.

A.9 Pulldown control for unspecific preys binding

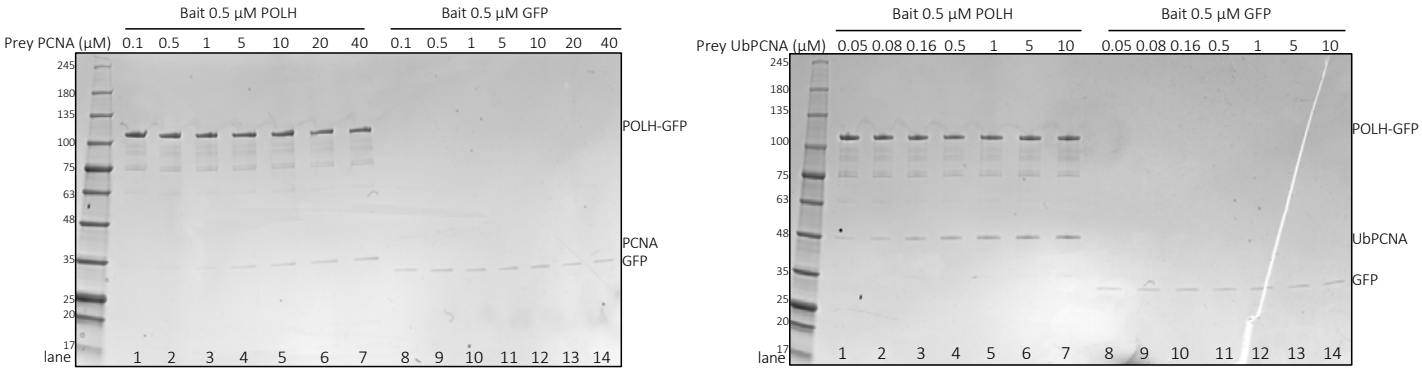


Figure A9. Pulldown controls with binding to GFP baits of PCNA and UbPCNA preys. No unspecific binding of preys (PCNA and UbPCNA) is detected at any concentration, only specific pulldown in presence of polymerase is detected in the elutions.

A.10 Multi-polymerase binding to UbPCNA via SEC

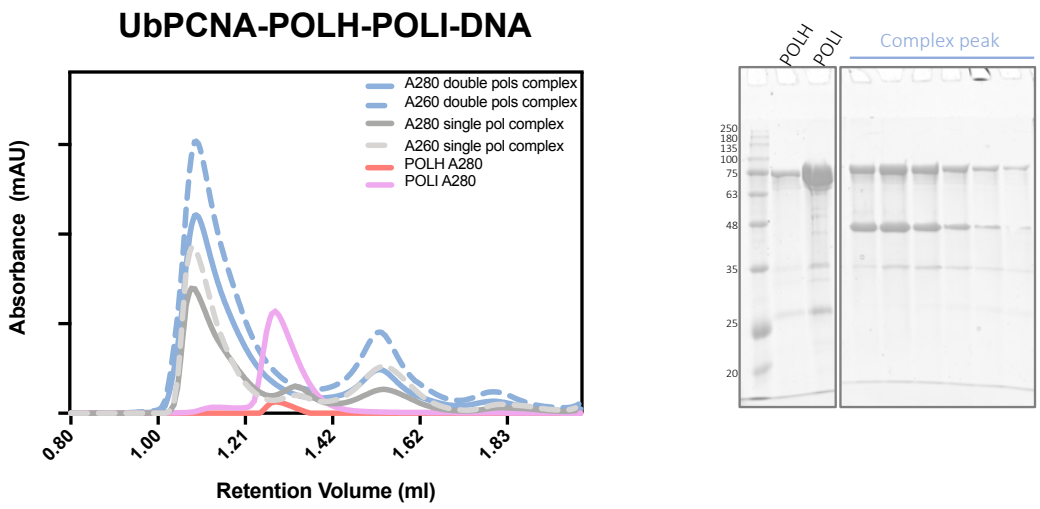


Figure A10. Multi-polymerase binding to UbPCNA via SEC. S200i elution profile of UbPCNA-Polη-Polι-DNA complex (light blue lines) does not show any additional peak shift compared to classical single polymerase complex (grey lines). Continuous lines indicate A280 nm and dotted lines A260 nm. Corresponding SDS-PAGE analysis of peak fractions shows indistinguishable Polη or Polι bands due to nearly identical size.

A.11 Direct Polη-Polι binding attempt via SEC

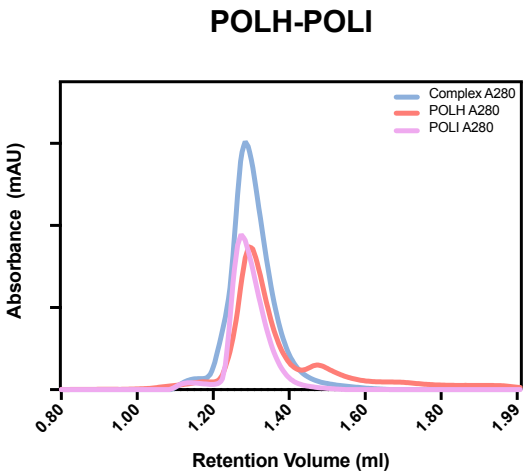


Figure A11. Direct Polη-Polι- binding attempt via SEC analysis. No interaction between Polη (salmon line A280) and Polι (pink line A280) is detected over SEC as corresponding to protein mixture (blue line) is showing any peak shift.

Abbreviations

°C	Degree Celsius
6-4PPs	6-4 pyrimidine photoproducts
2D	Two-dimensional
3D	Three-dimensional
8-oxoG	8-oxoguanine
Å	Angstrom
A	Alanine
aa	Amino acid
AEBSF	4-(2-aminoethyl) benzenesulfonyl fluroride hydrochlororide
APIM	AlkB homologue PCNA interacting motif
ATP	Adenosine triphosphate
β-OG	Octyl-Beta-Glucosid
BER	Base excision repair
bp	base pair
BRCA	Areast cancer susceptibility protein 1
BSA	Bovine serum albumin
CBP	CREB-binding protein
CCD	Charge-coupled device
CPD	Cysteine protease domain
CPD	*cis–syn* cyclobutane pyrimidine dimer
CRL4(Cdt2)	Cullin-4-RING ligase (CRL4)-Ddb1-Cdt2
cryo-EM	Cryogenic electron microscopy
CTD	C-terminal domain
DDR	DNA damage response
DDT	DNA damage tolerance
DNA	Deoxyribonucleic acid
DSB	Double-stranded break
dsDNA	Double-stranded DNA
dATP	Deoxyadenosine triphosphate
dAMPNPP	2'-Deoxyadenosine-5'-[(α,β)-imido]triphosphate
dCTP	Deoxycytidine triphosphate
dGTP	Deoxyguanosine triphosphate
dNTPs	Deoxynucleotid(s)e triphosphate
dTTP	Deoxythymidine triphosphate
DTT	Dithiothreitol
dRP-lyase	5′-deoxyribose phosphate-lyase activity
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
EM	Electron microscopy
EMSA	Electrophoretic mobility shift assay
ESI-TOF MS	Electron spray ionization - time of flight mass spectrometry
FAM	Fluorescein amidite
FEN1	Flap endonuclease 1
FL	Full length
FT	Flow-through
GA	Glutaraldehyde
GFP	Green fluorescent protein
Hepes	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

HLTF	Helicase-like transcription factor
HR	Homologus Recombination
I	Isolucine
ICL	Intrastrand-crosslink repair
IDCL	Intradomain connecting loop
IDPs	Intrinsically disordered proteins
IDRs	Intrinsically disordered region
IPTG	Isopropyl-b-D-thiogalactopyranoside
ISG15	Interferon-stimulated gene 15
ITC	Isothermal titration calorimetry
K	Lysine
K_d	Dissociation constant
KCl	Potassium chloride
L	Leucine
LIC	Ligase independent cloning
LIG1	DNA ligase I
LF	Little Finger domain
M	Methionine
M	Molar (concentration)
MAD2L2	Mitotic Arrest Deficient 2 Like 2
MIP	Mlh1 interacting proteins
$MgCl_2$	Magnesium chloride
$MnCl_2$	Manganese chloride
MMR	Mismatch repair
MS	Mass spectrometry
MST	Microscale thermophoresis
NaCl	Sodium chloride
NEDD8	Neural-precursor-cell-expressed developmentally down-regulated 8
NER	Nucleotide excision repair
NHEJ	Non-homologous end joining
NLS	Nuclear localization signal
NMR	Nuclear magnetic resonance
NP-40	Nonidet P-40, octylphenoxypolyethoxyethanol
nt	Nucleotide(s)
ON	Overnight
$p21^{WAF1/CIP1}$	Cyclin-dependent kinase inhibitor 1, p21
PAD	Polymerase associate domain
PAGE	Polyacrylamide gel electrophoresis
PDB	Protein data bank
PCNA	Proliferating cell nuclear antigen
PIP	PCNA interacting peptide
pol	Polymerase
POLI	Human DNA polymerase iota gene
POLH	Human DNA polymerase eta gene
Polα	DNA polymerase alpha- Primase
Polβ	DNA polymerase beta
Polδ	DNA polymerase delta
Polε	DNA polymerase epsilon
Polη	DNA polymerase eta
Polι	DNA polymerase iota
Polκ	DNA polymerase kappa

Polλ	DNA polymerase lambda
Polζ	DNA polymerase zeta
POLD1	DNA Polymerase delta 1, catalytic subunit, p125
POLD2	DNA Polymerase delta subunit 2, p50
POLD3	DNA Polymerase delta subunit 3, p66
POLD4	DNA Polymerase delta subunit 4, p12
PTM	Post-translational modification(s)
R	Arginine
Rad6-Rad18	E2-E3 Ubiquitin ligase complex involved in DNA repair
RAD30	*Saccharomyces cerevisiae* DNA polymerase eta
REV1	Reversionless 1 (DNA polymerase)
REV7	Reversionless 7, subunit of DNA polymerase zeta
REV3L	Reversionless 3-like, human DNA polymerase zeta catalytic subunit
RFC	Replication fator C
RING	Really interesting new gene
RIR	REV1 interacting region
RNF8	RING finger protein 8
RPA	Replication protein A
RT	Room temperature
RT-	Reverse transcriptase (polymerase family)
SAXS	Small angle X-ray scattering
SDS	Sodium dodecyl sulphate
SEC	Size exclusion chromatography
SEC-MALS	Size exclusion chromatography- multi angle light scattering
SENP2	Sentrin-specific protease 2
sGRS	Short gap repair synthesis
SHPRH	SNF2 histone linker PHD RING helicase
SHM	Somatic hypermutation
SLiM	Short Linear Motifs
SPR	Surface plasmon resonance
Srs2	ATP-dependent DNA helicase Srs2
SUMO	Small ubiquitin modifier
TBE	Tris-borate-EDTA
TdT	Terminal deoxynucleotidyl transferase
TLS	Translesion synthesis
TMEJ	Theta-mediated end joining
TS	Template switch
Ub	Ubiquitin
UbaI	Ubiquitin-like modifier activating enzyme 1
Ubc13–Mms2	E2 ubiquitin-conjugating enzyme complex
UbcH5c	Human Ubiquitin-conjugating enzyme E2-D3
UBM	Ubiquitin binding motif
UbPCNA	PCNA mono-ubiquitinated on lysine 164
UBZ	Ubiquitin binding zinc finger
UV	Ultraviolet
Q	Glutamine
wt	Wild type
XP	Xeroderma pigmentosum
XP-V	Xeroderma pigmentosum variant
X-ray	Röntgen radiation

List of Figures

List of Tables

Acknowledgments

Firstly, I would like to express my sincerest gratitude to my supervisor, Dr. Christian Biertümpfel, for his expertise, enthusiasm and motivation. His guidance supported me through all the time of my PhD project and in these final steps of thesis writing. Thanks a lot for offering me the chance to work on the TLS project, for your support, supervision and trust throughout the years. Furthermore, I would like to particularly thank Dr. Naoko Mizuno, for her guidance and support during the last period of my doctorate, as well as for the valuable inputs and scientific discussion during our lab meetings. I am very grateful you always showed interested in my project, and your useful advices, EM-related and not, really helped guiding my project forward.

A special thanks goes to my doctoral thesis supervisor, Prof. Elena Conti, for giving me the opportunity to perform my research in her department, for the support during my all PhD time, for her trust throughout the years as part of my TAC committee and for providing an excellent research atmosphere in the department. Additionally, I would like to express my gratitude to my other TAC members, Prof. Karl Duderstadt and Prof. Andreas Ladurner, for their informative and helpful comments and encouragement.

A big thanks to the MPIB Crystallization Facility, for all the crystallization attempts and trials. Additionally I would like to thanks MPIB the Biochemistry Core Facility for providing a vibrant infrastructure, excellent technical setups and scientific support. Especially, I express my gratitude to Dr. Stephan Uebel, for the many scientific advices and to Dr. Nagarjuna Nagaraj, for sample analysis via native mass spectrometry. A big thanks goes to the amazing MPIB cryo-EM facility, thanks to Dr. Mike Strauss for his help and advice on my initial steps in cryo-EM and for the final data collection. Additionally, a huge thanks goes to the actual cryo-EM “crew”, although I didn’t get to collect any big dataset since you joined, I am very great of the help, technical guidance and relentless presence of Dr. Daniel Bollschweiler and Dr. Tillman Schäfer. Thanks teaching me how not to burn the microscopes down! In terms of microscopy I should particularity thank Prof. Baumeister and the whole Baumeister department for the great opportunities of TITAN data collection and use of CM200. Additionally, huge thank goes to Dr. Rajan Prabu, for his support in my last TITAN dataset, for maintaining an impeccable computing environment, and for his constant assistance with processing issues.

My thanks extent to all present and past members of the Conti Department for the great scientific support and the enjoyable lab atmosphere. Big thanks to Claire Basquin for the help with all the biophysics and to Fabien Bonneau for running the thermofluor analysis. Thanks to Jerôme for the synchrotron data collections, and to Karina and Sabine for the patients in my delayed crystallization submission. Thanks to Peter for maintaining the Aekta systems perfectly running and for being always available every time I messed something up. Thanks to Christian Benda for your constant support in computing and for the help in fixing, multiple times and in big hurry, my laptop issues. Thanks to Walter and Ivan for the general IT technical support and to Ulli and Petra for the help in the administrative part. Thanks to Ariane for the help in initial expression test, to Judith and Joerg for invisibly keeping everything smooth and in order in the department. Furthermore, I was very happy to be part of the International Max-Planck Research School for Molecular Life Sciences (IMPRS-LS). I appreciate the support I received from the coordination office and the great scientific and friendly atmosphere.

A huge thanks to all the Conti's Postdocs and PhD student, past and present, for the invaluable inputs and suggestions during group meetings, the general help in the lab and in the EM bunker. Thank you all for maintaining the best atmosphere possible, between group meeting, scientific talks, Get Togethers and coffee breaks. Also I believe that the incredible Conti volleyball team members deserves a big acknowledgment for making the dream of beating the Pfander's finally true. A special thanks goes to Jana, for being a good friend and believing in me more than myself, and for all the good life advices you gave me drinking tea or having ice-cream on our sunny balcony. An enormous thank goes to Mahesh, for your constant support and listening and for being a valuable friend. Also, thanks for proofreading my thesis, I wouldn't (would not) be here if it wasn't (was not) for you. I really hope all your inputs to ameliorate my thesis writing skills were of some use.

The list of people who helped me during my PhD is endless, but I guess my lab mates need a special treatment. Thanks Shun for being available form my very first day, supporting me with your great knowledge and really thanks a lot for the reading of my thesis, from Amsterdam, you really helped me a lot. Infinite thanks go to Iuliia, with your positive attitude and charisma you really helped me through difficult times, and especially thanks for your valuable help in reading and proofreading my thesis. A big thanks goes to Carina, for your help in the lab and for your moral support. I want to thank Ying for being the best master student I could wish for, as well as Markus, Maren, Sven and Wolfgang for technical support thought-out the years. The Biertümpfel and Mizuno labs always were like one big lab, and outside the institute we are also

good friends, so I want to thank all the present and past members of both labs for all the good times. Especially thanks to Qianmin and Dirk for guiding me into my initial steps approaching cryo-EM and processing. Big thanks to Dirk for reprocessing some of my data in the attempts to make some "magic" and for being the best grumpy lab mate and, above all, a great friend. Thanks to Chris, for your positivity, constant trust and for partially sharing the blame of "speaking Italian". Thanks to Nirakar, for sharing your geeky processing skills and your contagious laughs. Thank you all guys for the great time we had together in the lab, outside, and recently on our beer zooms. Finally, thank you Steffi for being not only the best lab-mate, the best Get Together-organizer, the best late-night-complains-listener, the best emergency-couch-provider, the best concert-pal, the best car-driver, the best party-beast and the best volleyball player I know. Thank you Steffi for being the best friend I could wish for during these insidious PhD years, I am really lucky to have you by my side.

Outside the lab, but not too far, a gigantic thanks goes to Lorenzo, thanks for always being available for a siga, a chat and for reading my results. Hope can pay back the favour soon. We shared a lot, from Pavia to Franz, and I'm grateful to have you as a friend.
I wouldn't have survived this foolish year without the daily (or hourly) constant love and support of my special friends. Thank you Dali, Giorgia and Stefania, for always being there, with our messages on spam and 'biological' problems. You were the best thing of this year, and I wouldn't have finished this thesis without your continuous motivation, love and our endless therapeutic calls. We are *Seta.* Lastly, but not less importantly, I would like NOT to thank Franzburger, for his constant distractions and for continuously jumping on the keyboard. Even though with your paws you tried to write some part of this thesis, thanks for nothing, stupid annoying adorable fat cat!

Finally the biggest thanks goes to my family, who always believed in me. Without your endless love and encouragement I wouldn't have reached this point. Thanks Mamma for loving me unconditionally and thank you Papi for being my rock and my example: I hope one day I could be just half the person you are. The deepest thanks goes to the two halves of my heart, my two siblings. Thank you Diego for making me realise that some problems in life are relative, this thesis is dedicated to you and is a small demonstration that with strength and commitment every battle can be won. I believe in you Ghigo! Thank you Luci for being the strongest person I know, for being my mental support, my laugh, my love, my happiness and the best *Sore* ever.

References

A, P., Xu, X., Wang, C., Yang, J., Wang, S., Dai, J., and Ye, L. (2018). EZH2 promotes DNA replication by stabilizing interaction of POLδ and PCNA via methylation-mediated PCNA trimerization. Epigenetics Chromatin *11*, 44.

Acharya, N., Yoon, J.-H., Gali, H., Unk, I., Haracska, L., Johnson, R.E., Hurwitz, J., Prakash, L., and Prakash, S. (2008). Roles of PCNA-binding and ubiquitin-binding domains in human DNA polymerase in translesion DNA synthesis. Proc. Natl. Acad. Sci. *105*, 17724–17729.

Alt, A., Lammens, K., Chiocchini, C., Lammens, A., Pieck, J.C., Kuch, D., Hopfner, K.-P., and Carell, T. (2007). Bypass of DNA Lesions Generated During Anticancer Treatment with Cisplatin by DNA Polymerase. Science *318*, 967–970.

Andersen, P.L., Xu, F., and Xiao, W. (2008). Eukaryotic DNA damage tolerance and translesion synthesis through covalent modifications of PCNA. Cell Res. *18*, 162–173.

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

Armstrong, A.A., Mohideen, F., and Lima, C.D. (2012). Recognition of SUMO-modified PCNA requires tandem receptor motifs in Srs2. Nature *483*, 59–63.

Artimo, P., Jonnalagedda, M., Arnold, K., Baratin, D., Csardi, G., de Castro, E., Duvaud, S., Flegel, V., Fortier, A., Gasteiger, E., et al. (2012). ExPASy: SIB bioinformatics resource portal. Nucleic Acids Res. *40*, W597–W603.

Avkin, S., Adar, S., Blander, G., and Livneh, Z. (2002). Quantitative measurement of translesion replication in human cells: Evidence for bypass of abasic sites by a replicative DNA polymerase. Proc. Natl. Acad. Sci. *99*, 3764–3769.

Baldeck, N., Janel-Bintz, R., Wagner, J., Tissier, A., Fuchs, R.P., Burkovics, P., Haracska, L., Despras, E., Bichara, M., Chatton, B., et al. (2015). FF483–484 motif of human Polη mediates its interaction with the POLD2 subunit of Polδ and contributes to DNA damage tolerance. Nucleic Acids Res. *43*, 2116–2125.

Bauer, G.A., and Burgers, P.M.J. The yeast analog of mammalian cyclin/proliferating-cell nuclear antigen interacts with mammalian DNA polymerase 6. 5.

Beard, W.A., and Wilson, S.H. (2003). Structural Insights into the Origins of DNA Polymerase Fidelity. Structure *11*, 489–496.

Beard, W.A., Shock, D.D., Vande Berg, B.J., and Wilson, S.H. (2002). Efficiency of Correct Nucleotide Insertion Governs DNA Polymerase Fidelity. J. Biol. Chem. *277*, 47393–47398.

Bi, X. (2015). Mechanism of DNA damage tolerance. World J. Biol. Chem. *6*, 48.

Bienko, M. (2005). Ubiquitin-Binding Domains in Y-Family Polymerases Regulate Translesion Synthesis. Science *310*, 1821–1824.

Biertümpfel, C., Zhao, Y., Kondo, Y., Ramón-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 η. Nature *465*, 1044–1048.

Blackburn, E.H., Greider, C.W., and Szostak, J.W. (2006). Telomeres and telomerase: the path from maize, Tetrahymena and yeast to human cancer and aging. Nat. Med. *12*, 1133–1138.

Boehm, E.M. (2016). The regulation of translesion synthesis through the binding and activation of polymerases by PCNA. Doctor of Philosophy. University of Iowa.

Boehm, E.M., and Washington, M.T. (2016). 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.

Boehm, E.M., Gildenberg, M.S., and Washington, M.T. (2016a). The Many Roles of PCNA in Eukaryotic DNA Replication. In The Enzymes, (Elsevier), pp. 231–254.

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

Bomar, M.G., D'Souza, S., Bienko, M., Dikic, I., Walker, G.C., and Zhou, P. (2010). Unconventional Ubiquitin Recognition by the Ubiquitin-Binding Motif within the Y Family DNA Polymerases ι and Rev1. Mol. Cell *37*, 408–417.

Boudsocq, F., Kokoska, R.J., Plosky, B.S., Vaisman, A., Ling, H., Kunkel, T.A., Yang, W., and Woodgate, R. (2004). Investigating the Role of the Little Finger Domain of Y-family DNA Polymerases in Low Fidelity Synthesis and Translesion Replication. J. Biol. Chem. *279*, 32932–32940.

Branzei, D., Seki, M., and Enomoto, T. (2004). Rad18/Rad5/Mms2-mediated polyubiquitination of PCNA is implicated in replication completion during replication stress. Genes Cells *9*, 1031–1042.

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

Bruck, I., Goodman, M.F., and O'Donnell, M. (2003). The Essential C Family DnaE Polymerase Is Error-prone and Efficient at Lesion Bypass. J. Biol. Chem. *278*, 44361–44368.

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

Brzovic, P.S., Lissounov, A., Christensen, D.E., Hoyt, D.W., and Klevit, R.E. (2006). A UbcH5/Ubiquitin Noncovalent Complex Is Required for Processive BRCA1-Directed Ubiquitination. Mol. Cell *21*, 873–880.

Cazzalini, O., Sommatis, S., Tillhon, M., Dutto, I., Bachi, A., Rapp, A., Nardo, T., Scovassi, A.I., Necchi, D., Cardoso, M.C., et al. (2014). CBP and p300 acetylate PCNA to link its degradation with nucleotide excision repair synthesis. Nucleic Acids Res. *42*, 8433–8448.

Chatterjee, N., and Walker, G.C. (2017). Mechanisms of DNA damage, repair, and mutagenesis: DNA Damage and Repair. Environ. Mol. Mutagen. *58*, 235–263.

Chen, J., Chen, S., Saha, P., and Dutta, A. (1996). p21CiPl/Wafl disrupts the recruitment of human Fenl by proliferating-cell nuclear antigen into the DNA replication complex. Proc Natl Acad Sci USA 6.

Cleaver, J.E. (1972). Xeroderma Pigmentosum: Variants with Normal Dna Repair and Normal Sensitivity to Ultraviolet Light. J. Invest. Dermatol. *58*, 124–128.

Cordeiro, T.N., Chen, P., De Biasio, A., Sibille, N., Blanco, F.J., Hub, J.S., Crehuet, R., and Bernadó, P. (2017). Disentangling polydispersity in the PCNA−p15PAF complex, a disordered, transient and multivalent macromolecular assembly. Nucleic Acids Res. *45*, 1501–1515.

Davies, A.A., Huttner, D., Daigaku, Y., Chen, S., and Ulrich, H.D. (2008). Activation of Ubiquitin-Dependent DNA Damage Bypass Is Mediated by Replication Protein A. Mol. Cell *29*, 625–636.

De Biasio, A., Sánchez, 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.

De Bont, R. (2004). Endogenous DNA damage in humans: a review of quantitative data. Mutagenesis *19*, 169–185.

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.

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.

Despras, E., Delrieu, N., Garandeau, C., Ahmed-Seghir, S., and Kannouche, P.L. (2012). Regulation of the specialized DNA polymerase eta: Revisiting the biological relevance of its PCNA- and ubiquitin-binding motifs. Environ. Mol. Mutagen. *53*, 752–765.

Dieckman, L.M., and Washington, M.T. (2013). PCNA trimer instability inhibits translesion synthesis by DNA polymerase η and by DNA polymerase δ. DNA Repair *12*, 367–376.

Dovrat, D., Stodola, J.L., Burgers, P.M.J., and Aharoni, A. (2014). Sequential switching of binding partners on PCNA during in vitro Okazaki fragment maturation. Proc. Natl. Acad. Sci. *111*, 14118–14123.

Elia, A.E.H., Boardman, A.P., Wang, D.C., Huttlin, E.L., Everley, R.A., Dephoure, N., Zhou, C., Koren, I., Gygi, S.P., and Elledge, S.J. (2015). Quantitative Proteomic Atlas of Ubiquitination and Acetylation in the DNA Damage Response. Mol. Cell *59*, 867–881.

Fernandez-Leiro, R., and Scheres, S.H.W. (2017). A pipeline approach to single-particle processing in *RELION*. Acta Crystallogr. Sect. Struct. Biol. *73*, 496–502.

Frank, E.G. (2017). DNA polymerase ι_ The long and the short of it! DNA Repair 5.

Freudenthal, B.D. (2010). Structure of monoubiquitinated PCNA and implications for translesion synthesis and DNA polymerase exchange. Mol. Biol. *17*, 7.

Freudenthal, B.D., Ramaswamy, S., Hingorani, M.M., and Washington, M.T. (2008). Structure of a Mutant Form of Proliferating Cell Nuclear Antigen That Blocks Translesion DNA Synthesis † ‡. Biochemistry *47*, 13354–13361.

Freudenthal, B.D., Gakhar, L., Ramaswamy, S., and Washington, M.T. (2009). A charged residue at the subunit interface of PCNA promotes trimer formation by destabilizing alternate subunit interactions. Acta Crystallogr. D Biol. Crystallogr. *65*, 560–566.

Gali, H., Juhasz, S., Morocz, M., Hajdu, I., Fatyol, K., Szukacsov, V., Burkovics, P., and Haracska, L. (2012). Role of SUMO modification of human PCNA at stalled replication fork. Nucleic Acids Res. *40*, 6049–6059.

Gan, G.N., Wittschieben, J.P., Wittschieben, B.Ø., and Wood, R.D. (2008). DNA polymerase zeta (pol ζ) in higher eukaryotes. Cell Res. *18*, 174–183.

Ghosal, G., and Chen, J. (2013). DNA damage tolerance: a double-edged sword guarding the genome. 28.

Gibbs, P.E.M., McDonald, J., Woodgate, R., and Lawrence, C.W. (2005). The Relative Roles *in Vivo* of *Saccharomyces cerevisiae* Pol η, Pol ζ, Rev1 Protein and Pol32 in the Bypass and Mutation Induction of an Abasic Site, T-T (6-4) Photoadduct and T-T *cis-syn* Cyclobutane Dimer. Genetics *169*, 575–582.

Giglia-Mari, G., Zotter, A., and Vermeulen, W. (2011). DNA Damage Response. Cold Spring Harb. Perspect. Biol. *3*, a000745–a000745.

Gómez-Llorente, Y., Malik, R., Jain, R., Choudhury, J.R., Johnson, R.E., Prakash, L., Prakash, S., Ubarretxena-Belandia, I., and Aggarwal, A.K. (2013). The Architecture of Yeast DNA Polymerase ζ. Cell Rep. *5*, 79–86.

González-Magaña, A., and Blanco, F.J. (2020). Human PCNA Structure, Function and Interactions. Biomolecules *10*, 570.

Gonzalez-Magaña, A., Ibáñez de Opakua, A., Romano-Moreno, M., Murciano-Calles, J., Merino, N., Luque, I., Rojas, A.L., Onesti, S., Blanco, F.J., and De Biasio, A. (2019). The p12 subunit of human polymerase δ uses an atypical PIP box for molecular recognition of proliferating cell nuclear antigen (PCNA). J. Biol. Chem. *294*, 3947–3956.

Guan, J., Yu, S., and Zheng, X. (2017). NEDDylation antagonizes ubiquitination of proliferating cell nuclear antigen and regulates the recruitment of polymerase η in response to oxidative DNA damage. Protein Cell.

Gueranger, Q., Stary, A., Aoufouchi, S., Faili, A., Sarasin, A., Reynaud, C.-A., and Weill, J.-C. (2008). Role of DNA polymerases η, ι and ζ in UV resistance and UV-induced mutagenesis in a human cell line. DNA Repair *7*, 1551–1562.

Gulbis, J.M., Kelman, Z., Hurwitz, J., O'Donnell, M., and Kuriyan, J. Structure of the C-Terminal Region of p21WAF1/CIP1 Complexed with Human PCNA. 10.

Guo, C., Tang, T.-S., Bienko, M., Parker, J.L., Bielen, A.B., Sonoda, E., Takeda, S., Ulrich, H.D., Dikic, I., and Friedberg, E.C. (2006). Ubiquitin-Binding Motifs in REV1 Protein Are Required for Its Role in the Tolerance of DNA Damage. Mol. Cell. Biol. *26*, 8892–8900.

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

Haracska, L., Yu, S.-L., Johnson, R.E., Prakash, L., and Prakash, S. (2000). Efficient and accurate replication in the presence of 7,8-dihydro-8-oxoguanine by DNA polymerase η. Nat. Genet. *25*, 458–461.

Haracska, L., Acharya, N., Unk, I., Johnson, R.E., Hurwitz, J., Prakash, L., and Prakash, S. (2005). A Single Domain in Human DNA Polymerase Mediates Interaction with PCNA: Implications for Translesion DNA Synthesis. MOL CELL BIOL *25*, 8.

Haracska, L., Kondratick, C.M., Unk, I., Prakash, S., and Prakash, L. Interaction with PCNA Is Essential for Yeast DNA Polymerase Function. 9.

Havens, C.G., and Walter, J.C. (2009). Docking of a Specialized PIP Box onto Chromatin-Bound PCNA Creates a Degron for the Ubiquitin Ligase CRL4Cdt2. Mol. Cell *35*, 93–104.

Hedglin, M., Pandey, B., and Benkovic, S.J. (2016). Characterization of human translesion DNA synthesis across a UV-induced DNA lesion. ELife *5*, e19788.

Hedglin, M., Aitha, M., Pedley, A., and Benkovic, S.J. (2019). Replication protein A dynamically regulates monoubiquitination of proliferating cell nuclear antigen. J. Biol. Chem. *294*, 5157–5168.

Hendel, A., Krijger, P.H.L., Diamant, N., Goren, Z., Langerak, P., Kim, J., Reißner, T., Lee, K., Geacintov, N.E., Carell, T., et al. (2011). PCNA Ubiquitination Is Important, But Not Essential for Translesion DNA Synthesis in Mammalian Cells. PLoS Genet. *7*, e1002262.

Hibbert, R.G., and Sixma, T.K. (2012). Intrinsic Flexibility of Ubiquitin on Proliferating Cell Nuclear Antigen (PCNA) in Translesion Synthesis. J. Biol. Chem. *287*, 39216–39223.

Hishiki, A., Hashimoto, H., Hanafusa, T., Kamei, K., Ohashi, E., Shimizu, T., Ohmori, H., and Sato, M. (2009a). Structural Basis for Novel Interactions between Human Translesion Synthesis Polymerases and Proliferating Cell Nuclear Antigen. J. Biol. Chem. *284*, 10552–10560.

Hishiki, A., Hashimoto, H., Hanafusa, T., Kamei, K., Ohashi, E., Shimizu, T., Ohmori, H., and Sato, M. (2009b). Structural Basis for Novel Interactions between Human Translesion Synthesis Polymerases and Proliferating Cell Nuclear Antigen. J. Biol. Chem. *284*, 10552–10560.

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.

Huang, T.T., Nijman, S.M.B., Mirchandani, K.D., Galardy, P.J., Cohn, M.A., Haas, W., Gygi, S.P., Ploegh, H.L., Bernards, R., and D'Andrea, A.D. (2006). Regulation of monoubiquitinated PCNA by DUB autocleavage. Nat. Cell Biol. *8*, 341–347.

Iakoucheva, L.M., Brown, C.J., Lawson, J.D., Obradović, Z., and Dunker, A.K. (2002). Intrinsic Disorder in Cell-signaling and Cancer-associated Proteins. J. Mol. Biol. *323*, 573–584.

Ito, J., and Braithwaite, D.K. (1991). Compilation and alignment of DNA polymerase sequences. Nucleic Acids Res. *19*, 4045–4057.

Jackson, A.L., and Loeb, L.A. (2001). The contribution of endogenous sources of DNA damage to the multiple mutations in cancer. Mutat. Res. Mol. Mech. Mutagen. *477*, 7–21.

Jackson, S.P., and Bartek, J. (2009). The DNA-damage response in human biology and disease. Nature *461*, 1071–1078.

Jain, R., Rice, W.J., Malik, R., Johnson, R.E., Prakash, L., Prakash, S., Ubarretxena-Belandia, I., and Aggarwal, A.K. (2019). Cryo-EM structure and dynamics of eukaryotic DNA polymerase δ holoenzyme. Nat. Struct. Mol. Biol. *26*, 955–962.

Jansen, J.G., Langerak, P., Tsaalbi-Shtylik, A., van den Berk, P., Jacobs, H., and de Wind, N. (2006). Strand-biased defect in C/G transversions in hypermutating immunoglobulin genes in Rev1-deficient mice. J. Exp. Med. *203*, 319–323.

Johnson, K.A. (2010). The kinetic and chemical mechanism of high-fidelity DNA polymerases. Biochim. Biophys. Acta BBA - Proteins Proteomics *1804*, 1041–1048.

Johnson, R.E. (1999). hRAD30 Mutations in the Variant Form of Xeroderma Pigmentosum. Science *285*, 263–265.

Johnson, R.E., Prakash, S., and Prakash, L. (1999). Efficient Bypass of a Thymine-Thymine Dimer by Yeast DNA Polymerase, Pol. Science *283*, 1001–1004.

Johnson, R.E., Washington, M.T., Haracska, L., Prakash, S., and Prakash, L. (2000). Eukaryotic polymerases i and z act sequentially to bypass DNA lesions. *406*, 5.

Johnson, R.E., Prakash, L., and Prakash, S. (2005). Biochemical evidence for the requirement of Hoogsteen base pairing for replication by human DNA polymerase. Proc. Natl. Acad. Sci. *102*, 10466–10471.

Kanao, R., Masuda, Y., Deguchi, S., Yumoto-Sugimoto, M., Hanaoka, F., and Masutani, C. (2015). Relevance of Simultaneous Mono-Ubiquitinations of Multiple Units of PCNA Homo-Trimers in DNA Damage Tolerance. PLOS ONE *10*, e0118775.

Kannouche, P., Fernández de Henestrosa, A.R., Coull, B., Vidal, A.E., Gray, C., Zicha, D., Woodgate, R., and Lehmann, A.R. (2002). Localization of DNA polymerases η and ι to the replication machinery is tightly co-ordinated in human cells. EMBO J. *21*, 6246–6256.

Kannouche, P.L., Wing, J., and Lehmann, A.R. Interaction of Human DNA Polymerase □ with Monoubiquitinated PCNA: A Possible Mechanism for the Polymerase Switch in Response to DNA Damage. Mol. Cell 10.

Kannouche, P.L., Wing, J., and Lehmann, A.R. Interaction of Human DNA Polymerase □ with Monoubiquitinated PCNA: A Possible Mechanism for the Polymerase Switch in Response to DNA Damage. Mol. Cell 10.

Kelch, B.A. (2016). Review: The lord of the rings: Structure and mechanism of the sliding clamp loader. Biopolymers *105*, 532–546.

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

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.

Krijger, P.H., Tsaalbi-Shtylik, A., Wit, N., van den Berk, P.C.M., de Wind, N., and Jacobs, H. (2013). Rev1 is essential in generating G to C transversions downstream of the Ung2 pathway but not the Msh2+Ung2 hybrid pathway: Molecular immunology. Eur. J. Immunol. *43*, 2765–2770.

Krishna, S.R., and Gary, S. Crystal Structure of the Eukaryotic DNA Polymerase Processivity Factor PCNA. 11.

Kuhlbrandt, W. (2014). The Resolution Revolution. Science *343*, 1443–1444.

Kunkel, T.A. (2009). Evolving Views of DNA Replication (In)Fidelity. Cold Spring Harb. Symp. Quant. Biol. *74*, 91–101.

Laemmli, U.K. (1970). Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. Nature *227*, 680–685.

Lancey, C., Tehseen, M., Raducanu, V.-S., Rashid, F., Merino, N., Ragan, T.J., Savva, C.G., Zaher, M.S., Shirbini, A., Blanco, F.J., et al. (2020a). Structure of the processive human Pol δ holoenzyme. Nat. Commun. *11*, 1109.

Lancey, C., Tehseen, M., Takahashi, M., Sobhy, M.A., Ragan, T.J., Crehuet, R., Hamdan, S.M., and De Biasio, A. (2020b). Cryo-EM structure of Pol κ−DNA−PCNA holoenzyme and implications for polymerase switching in DNA lesion bypass (Biochemistry).

Lapetina, S., and Gil-Henn, H. (2017). A guide to simple, direct, and quantitative in vitro binding assays. J. Biol. Methods *4*, 62.

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

Lehman, A.R., Kirk-Bell, S., Arlett, C.F., Paterson, M.C., Lohman, P.H., de Weerd-Kastelein, E.A., and Bootsma, D. (1975). Xeroderma pigmentosum cells with normal levels of excision repair have a defect in DNA synthesis after UV-irradiation. Proc. Natl. Acad. Sci. *72*, 219–223.

Leung, W., Baxley, R., Moldovan, G.-L., and Bielinsky, A.-K. (2018). Mechanisms of DNA Damage Tolerance: Post-Translational Regulation of PCNA. Genes *10*, 10.

Liang, Z., Diamond, M., Smith, J.A., Schnell, M., and Daniel, R. (2011). Proliferating cell nuclear antigen is required for loading of the SMCX/KMD5C histone demethylase onto chromatin. Epigenetics Chromatin *4*, 18.

Lin, W., Xin, H., Zhang, Y., Wu, X., Yuan, F., and Wang, Z. (1999). The human REV1 gene codes for a DNA template-dependent dCMP transferase. Nucleic Acids Res. *27*, 4468–4475.

Lin, Y.-C., Li, L., Makarova, A.V., Burgers, P.M., Stone, M.P., and Lloyd, R.S. (2014). Error-prone Replication Bypass of the Primary Aflatoxin B $_1$ DNA Adduct, AFB $_1$ -N7-Gua. J. Biol. Chem. *289*, 18497–18506.

Lindahl, T. (1993). Instability and decay of the primary structure of DNA. Nature *362*, 709–715.

Ling, H., Woodgate, R., and Yang, W. Crystal Structure of a Y-Family DNA Polymerase in Action: A Mechanism for Error-Prone and Lesion-Bypass Replication. 12.

Lior-Hoffmann, L., Wang, L., Wang, S., Geacintov, N.E., Broyde, S., and Zhang, Y. (2012). Preferred WMSA catalytic mechanism of the nucleotidyl transfer reaction in human DNA polymerase κ elucidates error-free bypass of a bulky DNA lesion. Nucleic Acids Res. *40*, 9193–9205.

Livneh, Z., Z, O., and Shachar, S. (2010). Multiple two-polymerase mechanisms in mammalian translesion DNA synthesis. Cell Cycle *9*, 729–735.

Makarova, A.V., and Kulbachinskiy, A.V. (2012). Structure of human DNA polymerase iota and the mechanism of DNA synthesis. Biochem. Mosc. *77*, 547–561.

Makarova, A.V., Grabow, C., Gening, L.V., Tarantul, V.Z., Tahirov, T.H., Bessho, T., and Pavlov, Y.I. (2011). Inaccurate DNA Synthesis in Cell Extracts of Yeast Producing Active Human DNA Polymerase Iota. PLoS ONE *6*, e16612.

Malakhov, M.P., Mattern, M.R., Malakhova, O.A., Drinker, M., Weeks, S.D., and Butt, T.R. (2004). SUMO fusions and SUMO-specific protease for efficient expression and purification of proteins. J. Struct. Funct. Genomics *5*, 75–86.

Malik, R., Kopylov, M., Gomez-Llorente, Y., Jain, R., Johnson, R.E., Prakash, L., Prakash, S., Ubarretxena-Belandia, I., and Aggarwal, A.K. (2020). Structure and mechanism of B-family DNA polymerase ζ specialized for translesion DNA synthesis. Nat. Struct. Mol. Biol. *27*, 913–924.

Mastronarde, D.N. (2005). Automated electron microscope tomography using robust prediction of specimen movements. J. Struct. Biol. *152*, 36–51.

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

Masutani, C., Kusumoto, R., Yamada, A., Dohmae, N., Yokoi, M., Yuasa, M., Araki, M., Iwai, S., and Hanaoka, F. (1999). The XPV (xeroderma pigmentosum variant) gene encodes human DNA polymerase □. *399*, 5.

Masutani, C., Kusumoto, R., Iwai, S., and Hanaoka, F. Mechanisms of accurate translesion synthesis by human DNA polymerase h. 10.

Matsui, I., Matsui, E., Yamasaki, K., and Yokoyama, H. (2013). Domain Structures and Inter-Domain Interactions Defining the Holoenzyme Architecture of Archaeal D-Family DNA Polymerase. Life *3*, 375–385.

McCulloch, S.D., and Kunkel, T.A. (2008). The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases. Cell Res. *18*, 148–161.

McCulloch, S.D., Kokoska, R.J., Masutani, C., Iwai, S., Hanaoka, F., and Kunkel, T.A. (2004). Preferential cis–syn thymine dimer bypass by DNA polymerase h occurs with biased fidelity. *428*, 4.

McDonald, J.P., Rapic-Otrin, V., Epstein, J.A., Broughton, B.C., Wang, X., Lehmann, A.R., Wolgemuth, D.J., and Woodgate, R. Novel Human and Mouse Homologs of Saccharomyces cerevisiae DNA Polymerase □. 11.

McIntyre, J. (2020). Polymerase iota - an odd sibling among Y family polymerases. DNA Repair 11.

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.

Mehta, A., and Haber, J.E. (2014). Sources of DNA Double-Strand Breaks and Models of Recombinational DNA Repair. Cold Spring Harb. Perspect. Biol. *6*, a016428–a016428.

Miyachi, K.; Fritzler, M.J.; Tan, E.M. Autoantibody to a Nuclear Antigen in Proliferating Cells. J. Immunol.1978, 121, 2228.

Mizukami, S., Kim, T.W., Helquist, S.A., and Kool, E.T. (2006). Varying DNA Base-Pair Size in Subangstrom Increments: Evidence for a Loose, Not Large, Active Site in Low-Fidelity Dpo4 Polymerase. Biochemistry *45*, 2772–2778.

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

Moon, A.F., Garcia-Diaz, M., Batra, V.K., Beard, W.A., Bebenek, K., Kunkel, T.A., Wilson, S.H., and Pedersen, L.C. (2007). The X family portrait: Structural insights into biological functions of X family polymerases. DNA Repair *6*, 1709–1725.

Moriya, T., Saur, M., Stabrin, M., Merino, F., Voicu, H., Huang, Z., Penczek, P.A., Raunser, S., and Gatsogiannis, C. (2017). High-resolution Single Particle Analysis from Electron Cryo-microscopy Images Using SPHIRE. J. Vis. Exp. 55448.

Motegi, A., Sood, R., Moinova, H., Markowitz, S.D., Liu, P.P., and Myung, K. (2006). Human SHPRH suppresses genomic instability through proliferating cell nuclear antigen polyubiquitination. J. Cell Biol. *175*, 703–708.

Nair, D.T., Johnson, R.E., Prakash, S., Prakash, L., and Aggarwal, A.K. (2004). Replication by human DNA polymerase-i occurs by Hoogsteen base-pairing. *430*, 4.

Nair, D.T., Johnson, R.E., Prakash, L., Prakash, S., and Aggarwal, A.K. (2006). An Incoming Nucleotide Imposes an anti to syn Conformational Change on the Templating Purine in the Human DNA Polymerase-ι Active Site. Structure *14*, 749–755.

Nair, D.T., Johnson, R.E., Prakash, L., Prakash, S., and Aggarwal, A.K. (2009). DNA Synthesis across an Abasic Lesion by Human DNA Polymerase ι. Structure *17*, 530–537.

Nelson, J.R., Lawrence, C.W., and Hinkle, D.C. (1996a). Deoxycytidyl transferase activity of yeast REV1 protein. Nature *382*, 729–731.

Nelson, J.R., Lawrence, C.W., and Hinkle, D.C. (1996b). Thymine-Thymine Dimer Bypass by Yeast DNA Polymerase zeta. Science *272*, 1646–1649.

Nick McElhinny, S.A., Havener, J.M., Garcia-Diaz, M., Juárez, R., Bebenek, K., Kee, B.L., Blanco, L., Kunkel, T.A., and Ramsden, D.A. (2005). A Gradient of Template Dependence Defines Distinct Biological Roles for Family X Polymerases in Nonhomologous End Joining. Mol. Cell *19*, 357–366.

Nick McElhinny, S.A., Gordenin, D.A., Stith, C.M., Burgers, P.M.J., and Kunkel, T.A. (2008). Division of Labor at the Eukaryotic Replication Fork. Mol. Cell *30*, 137–144.

Ohashi, E., Hanafusa, T., Kamei, K., Song, I., Tomida, J., Hashimoto, H., Vaziri, C., and Ohmori, H. (2009). Identification of a novel REV1-interacting motif necessary for DNA polymerase κ function. Genes Cells *14*, 101–111.

Ohmori, H., Hanafusa, T., Ohashi, E., and Vaziri, C. (2009). Separate Roles of Structured and Unstructured Regions of Y-Family DNA Polymerases. In Advances in Protein Chemistry and Structural Biology, (Elsevier), pp. 99–146.

Ollis, D.L., Brick, P., Hamlint, R., Xuong, N.G., and Steitz, T.A. (1985). Structure of large fragment of Escherichia coli DNA polymerase I complexed with dTMP. 5.

Ortega, J., Li, J.Y., Lee, S., Tong, D., Gu, L., and Li, G.-M. (2015). Phosphorylation of PCNA by EGFR inhibits mismatch repair and promotes misincorporation during DNA synthesis. Proc. Natl. Acad. Sci. *112*, 5667–5672.

Park, J.M., Yang, S.W., Yu, K.R., Ka, S.H., Lee, S.W., Seol, J.H., Jeon, Y.J., and Chung, C.H. (2014). Modification of PCNA by ISG15 Plays a Crucial Role in Termination of Error-Prone Translesion DNA Synthesis. Mol. Cell *54*, 626–638.

Peng, B., Ortega, J., Gu, L., Chang, Z., and Li, G.-M. (2019). Phosphorylation of proliferating cell nuclear antigen promotes cancer progression by activating the ATM/Akt/GSK3β/Snail signaling pathway. J. Biol. Chem. *294*, 7037–7045.

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

Plosky, B.S., Vidal, A.E., de Henestrosa, A.R.F., McLenigan, M.P., McDonald, J.P., Mead, S., and Woodgate, R. (2006). Controlling the subcellular localization of DNA polymerases ι and η via interactions with ubiquitin. EMBO J. *25*, 2847–2855.

Prestel, A., Wichmann, N., Martins, J.M., Marabini, R., Kassem, N., Broendum, S.S., Otterlei, M., Nielsen, O., Willemoës, M., Ploug, M., et al. (2019). The PCNA interaction motifs revisited: thinking outside the PIP-box. Cell. Mol. Life Sci. *76*, 4923–4943.

Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017a). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat. Methods *14*, 290–296.

Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017b). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat. Methods *14*, 290–296.

Rothwell, P.J., Mitaksov, V., and Waksman, G. (2005). Motions of the Fingers Subdomain of Klentaq1 Are Fast and Not Rate Limiting: Implications for the Molecular Basis of Fidelity in DNA Polymerases. Mol. Cell *19*, 345–355.

Rudd, S.G., Bianchi, J., and Doherty, A.J. (2014). PrimPol—A new polymerase on the block. Mol. Cell. Oncol. *1*, e960754.

Saintigny, Y. (2001). Characterization of homologous recombination induced by replication inhibition in mammalian cells. EMBO J. *20*, 3861–3870.

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.

Sale, J.E. (2013). Translesion DNA Synthesis and Mutagenesis in Eukaryotes. Cold Spring Harb. Perspect. Biol. *5*, a012708–a012708.

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

Sánchez, 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.

Scheres, S.H.W. (2012). RELION: Implementation of a Bayesian approach to cryo-EM structure determination. J. Struct. Biol. *180*, 519–530.

Scheres, S.H.W. (2016). Processing of Structurally Heterogeneous Cryo-EM Data in RELION. In Methods in Enzymology, (Elsevier), pp. 125–157.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *9*, 676–682.

Shachar, S., Ziv, O., Avkin, S., Adar, S., Wittschieben, J., Reißner, T., Chaney, S., Friedberg, E.C., Wang, Z., Carell, T., et al. (2009). Two-polymerase mechanisms dictate error-free and error-prone translesion DNA synthesis in mammals. EMBO J. *28*, 383–393.

Shen, A., Lupardus, P.J., Morell, M., Ponder, E.L., Sadaghiani, A.M., Garcia, K.C., and Bogyo, M. (2009). Simplified, Enhanced Protein Purification Using an Inducible, Autoprocessing Enzyme Tag. PLoS ONE *4*, e8119.

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

Stary, A., Kannouche, P., Lehmann, A.R., and Sarasin, A. (2003). Role of DNA Polymerase η in the UV Mutation Spectrum in Human Cells. J. Biol. Chem. *278*, 18767–18775.

Stelter, P., and Ulrich, H.D. (2003). Control of spontaneous and damage-induced mutagenesis by SUMO and ubiquitin conjugation. Nature *425*, 188–191.

Stodola, J.L., and Burgers, P.M. (2016). Resolving individual steps of Okazaki-fragment maturation at a millisecond timescale. Nat. Struct. Mol. Biol. *23*, 402–408.

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

Streich Jr, F.C., and Lima, C.D. (2016). Capturing a substrate in an activated RING E3/E2–SUMO complex. Nature *536*, 304–308.

Takawa, M., Cho, H.-S., Hayami, S., Toyokawa, G., Kogure, M., Yamane, Y., Iwai, Y., Maejima, K., Ueda, K., Masuda, A., et al. (2012). Histone Lysine Methyltransferase SETD8 Promotes Carcinogenesis by Deregulating PCNA Expression. Cancer Res. *72*, 3217–3227.

Tang, G., Peng, L., Baldwin, P.R., Mann, D.S., Jiang, W., Rees, I., and Ludtke, S.J. (2007). EMAN2: An extensible image processing suite for electron microscopy. J. Struct. Biol. 9.

Taylor, H. Structural Study of the Complex Between DNA Polymerase Iota and Ub-PCNA. 84.

Terai, K., Abbas, T., Jazaeri, A.A., and Dutta, A. (2010). CRL4Cdt2 E3 Ubiquitin Ligase Monoubiquitinates PCNA to Promote Translesion DNA Synthesis. Mol. Cell *37*, 143–149.

Timinskas, K., Balvočiūtė, M., Timinskas, A., and Venclovas, Č. (2014). Comprehensive analysis of DNA polymerase III α subunits and their homologs in bacterial genomes. Nucleic Acids Res. *42*, 1393–1413.

Tissier, A., Frank, E.G., McDonald, J.P., Iwai, S., Hanaoka, F., and Woodgate, R. (2000). Misinsertion and bypass of thymine–thymine dimers by human DNA polymerase ι. EMBO J. *19*, 5259–5266.

Tissier, A., Janel-Bintz, R., Coulon, S., Klaile, E., Kannouche, P., Fuchs, R.P., and Cordonnier, A.M. (2010). Crosstalk between replicative and translesional DNA polymerases: PDIP38 interacts directly with Polη. DNA Repair *9*, 922–928.

Trincao, J., Johnson, R.E., Escalante, C.R., Prakash, S., Prakash, L., and Aggarwal, A.K. Structure of the Catalytic Core of S. cerevisiae DNA Polymerase □: Implications for Translesion DNA Synthesis. Mol. Cell 10.

Tsutakawa, S.E., Van Wynsberghe, A.W., Freudenthal, B.D., Weinacht, C.P., Gakhar, L., Washington, M.T., Zhuang, Z., Tainer, J.A., and Ivanov, I. (2011). Solution X-ray scattering combined with computational modeling reveals multiple conformations of covalently bound ubiquitin on PCNA. Proc. Natl. Acad. Sci. *108*, 17672–17677.

Tsutakawa, S.E., Yan, C., Xu, X., Weinacht, C.P., Freudenthal, B.D., Yang, K., Zhuang, Z., Washington, M.T., Tainer, J.A., and Ivanov, I. (2015). Structurally Distinct Ubiquitin- and Sumo-Modified PCNA: Implications for Their Distinct Roles in the DNA Damage Response. Structure *23*, 724–733.

Vaisman, A., and Woodgate, R. (2017). Translesion DNA polymerases in eukaryotes: what makes them tick? Crit. Rev. Biochem. Mol. Biol. *52*, 274–303.

Vaisman, A., Frank, E.G., McDonald, J.P., Tissier, A., and Woodgate, R. (2002). pol□-dependent lesion bypass in vitro. Mutat. Res. 14.

Vidal, A.E., Kannouche, P., Podust, V.N., Yang, W., Lehmann, A.R., and Woodgate, R. (2004). Proliferating Cell Nuclear Antigen-dependent Coordination of the Biological Functions of Human DNA Polymerase ι. J. Biol. Chem. *279*, 48360–48368.

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

Wang, Y., Woodgate, R., McManus, T.P., Mead, S., McCormick, J.J., and Maher, V.M. (2007). Evidence that in Xeroderma Pigmentosum Variant Cells, which Lack DNA Polymerase η, DNA Polymerase ι Causes the Very High Frequency and Unique Spectrum of UV-Induced Mutations. Cancer Res. *67*, 3018–3026.

Washington, M.T., Johnson, R.E., Prakash, L., and Prakash, S. (2002). Human DINB1-encoded DNA polymerase is a promiscuous extender of mispaired primer termini. Proc. Natl. Acad. Sci. *99*, 1910–1914.

Washington, M.T., Minko, I.G., Johnson, R.E., Wolfle, W.T., Harris, T.M., Lloyd, R.S., Prakash, S., and Prakash, L. (2004). Efficient and Error-Free Replication Past a Minor-Groove DNA Adduct by the Sequential Action of Human DNA Polymerases and . MOL CELL BIOL *24*, 7.

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.

Wilson, R.C., Jackson, M.A., and Pata, J.D. (2013). Y-Family Polymerase Conformation Is a Major Determinant of Fidelity and Translesion Specificity. Structure *21*, 20–31.

Woodgate, R. A plethora of lesion-replicating DNA polymerases. 6.

Wu, S., Armache, J.-P., and Cheng, Y. (2016). Single-particle cryo-EM data acquisition by using direct electron detection camera. Microscopy *65*, 35–41.

Xu, H., Zhang, P., and Liu, L. A Novel PCNA-Binding Motif Identified by the Panning of a Random Peptide Display Library. 9.

Yang, W. (2003). Damage repair DNA polymerases Y. Curr. Opin. Struct. Biol. *13*, 23–30.

Yang, W. (2005). Portraits of a Y-family DNA polymerase. FEBS Lett. *579*, 868–872.

Yang, W. (2014). An Overview of Y-Family DNA Polymerases and a Case Study of Human DNA Polymerase η. Biochemistry *53*, 2793–2803.

Yang, W., and Gao, Y. (2018). Translesion and Repair DNA Polymerases: Diverse Structure and Mechanism. Annu. Rev. Biochem. *87*, 239–261.

Yang, W., and Woodgate, R. (2007). What a difference a decade makes: Insights into translesion DNA synthesis. Proc. Natl. Acad. Sci. *104*, 15591–15598.

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

Zeng, X., Winter, D.B., Kasmer, C., Kraemer, K.H., Lehmann, A.R., and Gearhart, P.J. (2001). DNA polymerase η is an A-T mutator in somatic hypermutation of immunoglobulin variable genes. Nat. Immunol. *2*, 537–541.

Zeng, X., Negrete, G.A., Kasmer, C., Yang, W.W., and Gearhart, P.J. (2004). Absence of DNA Polymerase η Reveals Targeting of C Mutations on the Nontranscribed Strand in Immunoglobulin Switch Regions. J. Exp. Med. *199*, 917–924.

Zhang, K. (2016). Gctf: Real-time CTF determination and correction. J. Struct. Biol. *193*, 1–12.

Zhang, H., Gibbs, P.E.M., and Lawrence, C.W. (2006). The *Saccharomyces cerevisiae rev6-1* Mutation, Which Inhibits Both the Lesion Bypass and the Recombination Mode of DNA Damage Tolerance, Is an Allele of *POL30* , Encoding Proliferating Cell Nuclear Antigen. Genetics *173*, 1983–1989.

Zhang, S., Chea, J., Meng, X., Zhou, Y., Lee, E.Y.C., and Lee, M.Y.W.T. (2008). PCNA is ubiquitinated by RNF8. Cell Cycle *7*, 3399–3404.

Zhang, Y., Yuan, F., Xin, H., Wu, X., Rajpal, D.K., Yang, D., and Wang, Z. Human DNA polymerase κ synthesizes DNA with extraordinarily low fidelity. 10.

Zhang, Y., Yuan, F., Wu, X., Taylor, J.-S., and Wang, Z. Response of human DNA polymerase ι to DNA lesions. 8.

Zhang, Z., Zhang, S., Lin, S.H.S., Wang, X., Wu, L., Lee, E.Y.C., and Lee, M.Y.W.T. (2012). Structure of monoubiquitinated PCNA: Implications for DNA polymerase switching and Okazaki fragment maturation. Cell Cycle *11*, 2128–2136.

Zhao, Y., Biertumpfel, C., Gregory, M.T., Hua, Y.-J., Hanaoka, F., and Yang, W. (2012). Structural basis of human DNA polymerase -mediated chemoresistance to cisplatin. Proc. Natl. Acad. Sci. *109*, 7269–7274.

Zhao, Y., Gregory, M.T., Biertumpfel, C., Hua, Y.-J., Hanaoka, F., and Yang, W. (2013). Mechanism of somatic hypermutation at the WA motif by human DNA polymerase. Proc. Natl. Acad. Sci. *110*, 8146–8151.

Zheng, S.Q., Palovcak, E., Armache, J.-P., Verba, K.A., Cheng, Y., and Agard, D.A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat. Methods *14*, 331–332.

Zhou, B.-L., Pata, J.D., and Steitz, T.A. (2001). Crystal Structure of a DinB Lesion Bypass DNA Polymerase Catalytic Fragment Reveals a Classic Polymerase Catalytic Domain. Mol. Cell *8*, 427–437.

Zivanov, J., Nakane, T., Forsberg, B.O., Kimanius, D., Hagen, W.J., Lindahl, E., and Scheres, S.H. (2018). New tools for automated high-resolution cryo-EM structure determination in RELION-3. ELife *7*, e42166.