2016-07-15

Structural Mechanisms of the Sliding Clamp and Sliding Clamp Loader: Insights into Disease and Function: A Dissertation

Caroline M. Duffy
University of Massachusetts Medical School

STRUCTURAL MECHANISMS OF THE SLIDING CLAMP AND SLIDING CLAMP LOADER: INSIGHTS INTO DISEASE AND FUNCTION

A Dissertation Presented

By

CAROLINE MARY DUFFY

Submitted to the Faculty of the
University of Massachusetts Graduate School of Biomedical Sciences, Worcester
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

July 15, 2016

Biochemistry and Molecular Pharmacology

STRUCTURAL MECHANISMS OF THE SLIDING CLAMP AND SLIDING CLAMP LOADER: INSIGHTS INTO DISEASE AND FUNCTION

A Dissertation Presented
By

CAROLINE MARY DUFFY

This work was undertaken in the Graduate School of Biomedical Sciences

Biochemistry and Molecular Pharmacology

The signature of the Thesis Advisor signifies
validation of Dissertation content

Brian A. Kelch, Ph.D., Thesis Advisor

The signatures of the Dissertation Defense Committee signify
completion and approval as to style and content of the Dissertation

Daniel Bolon, Ph.D., Member of Committee

Kendall Knight, Ph.D., Member of Committee

Nicholas Rhind, Ph.D., Member of Committee

Peter Chien, Ph.D., External Member of Committee

The signature of the Chair of the Committee signifies that the written dissertation meets
the requirements of the Dissertation Committee

William Royer, Ph.D., Chair of Committee

The signature of the Dean of the Graduate School of Biomedical Sciences signifies
that the student has met all graduation requirements of the school.

Anthony Carruthers, Ph.D.,
Dean of the Graduate School of Biomedical Sciences

July 15, 2016

ACKNOWLEDGMENTS

It’s truly been a privilege working in Dr. Brian Kelch’s lab. I greatly admire his tenacity, his insatiable love for all science, and his well-documented inability to let a chalk talk go uninterrupted with insightful questions. I appreciate his forthright opinions, support, and deep commitment to his students’ success. Most of all, I will always be grateful he took me into his lab. Despite the challenges grad school threw at me, I feel incredibly lucky to have trained with him and I fully expect he will do great things.

I’m thankful to say I’ve genuinely enjoyed working with my lab mates. I could not have done the bulk of my thesis work without the help of Dr. Brendan Hilbert. Not only do I appreciate our friendship, but I admire his meticulous scientific mind. He has made me a better scientist and I’m indebted to him for his expertise, especially for help with all things crystallographic and ITC. Nicholas Stone and Janelle Hayes joined the lab not long after I did and I’m grateful we have been close both in and out of lab. Our small group has always been incredibly supportive of one another and I will miss working with all of them.

Working in the Biochemistry and Molecular Pharmacology Department has also been a privilege. I’ve loved having the opportunity to chat science in as many offices as I have over the years, and I have enjoyed the department’s collaborative atmosphere. I appreciate the feedback I’ve received from my awesome TRAC/DEC committee members, past and present: Dr. Bill Royer (my

fantastic chair), Dr. Sean Ryder, Dr. Nick Rhind, Dr. Ken Knight, Dr. Dan Bolon, and my external member, Dr. Peter Chien. In addition, I received invaluable training, technical support, and/or advice from Dr. Mary Munson, Luca Leone, Ellen Nalivaika, Dr. Djade Soumana, Melissa Greven, Livio Dukaj, Dr. Maggie Heider, Dr. Dave Mofford, and Josh Gardner. It takes a village!

I've made some life-long friends in grad school, and whether we were commiserating or celebrating, it was always great to have people who knew exactly the frustration and joy you can (sometimes simultaneously) have while working towards your PhD. To my immediate family, Mom, Dad, Meg, and Tim, you have always encouraged me blindly and loudly, and most times I just needed to hear that I could do this.

Finally, to my partner-in-crime, Todd Monroe: You have supported me since going to graduate school was a mere thought. Thank you for patiently accepting without complaint the crazy schedule that working in a lab brings, understanding phrases like "running a gel" meant I would probably be late, and "concentrating protein" *definitely* meant I would be late. For better or worse, many of our relationship milestones coincided with graduate school milestones—I'm sorry again for going to the synchrotron the weekend you were going to propose. I didn't know. But I got a structure and a husband out of it regardless. Thank you for your unwavering love and patience, your constant support, and thank you for doing most of the dishes and laundry during the time it took to write this thesis. I could not have done this without you.

ABSTRACT

Chromosomal replication is an essential process in all life. This dissertation highlights regulatory roles for two critical protein complexes at the heart of the replication fork: 1) the sliding clamp, the major polymerase processivity factor, and 2) the sliding clamp loader, a spiral-shaped AAA+ ATPase, which loads the clamp onto DNA.

The clamp is a promiscuous binding protein that interacts with at least 100 binding partners to orchestrate many processes on DNA, but spatiotemporal regulation of these binding interactions is unknown. Remarkably, a recent disease-causing mutant of the sliding clamp showed specific defects in DNA repair pathways. We aimed to use this mutant as a tool to understand the binding specificity of clamp interactions, and investigate the disease further. We solved three structures of the mutant, and biochemically showed perturbation of partner-binding for some, but not all, ligands. Using a fission yeast model, we showed that mutant cells are sensitive to select DNA damaging agents. These data revealed significant flexibility within the binding site, which likely regulates partner binding.

Before the clamp can act on DNA, the sliding clamp loader places the clamp onto DNA at primer/template (p/t) junctions. The clamp loader reaction couples p/t binding and subsequent ATP hydrolysis to clamp closure. Here we show that composition (RNA vs. DNA) of the primer strand affects clamp loader

binding, and that the order of ATP hydrolysis around the spiral is likely sequential. These studies highlight additional details into the clamp loader mechanism, which further elucidate general mechanisms of AAA+ machinery.

TABLE OF CONTENTS

List of Tables

List of Figures

List of Third Party Copyrighted Material

Figure Number	Publisher	Licensing Number
Figure 1.1	Elsevier	3886000117594
Figure 1.5	AAAS	3873220605050
Figure 1.6	AAAS	3873220605050

Preface

Publications derived from work contained within this thesis:

In CHAPTER II: **Duffy, C. M.**, Hilbert, B. J., & Kelch, B. A. (2015). A Disease-Causing Variant in PCNA Disrupts a Promiscuous Protein Binding Site. *Journal of Molecular Biology*, *428*(6), 1023–1040.

In APPENDIX 2: Heider, M. R., Gu, M., **Duffy, C. M.**, Mirza, A. M., Marcotte, L. L., Walls, A. C., et al. (2016). Subunit connectivity, assembly determinants and architecture of the yeast exocyst complex. *Nature Structural & Molecular Biology*, *23*(1), 59–66.

CHAPTER I

Introduction

DNA replication and repair:

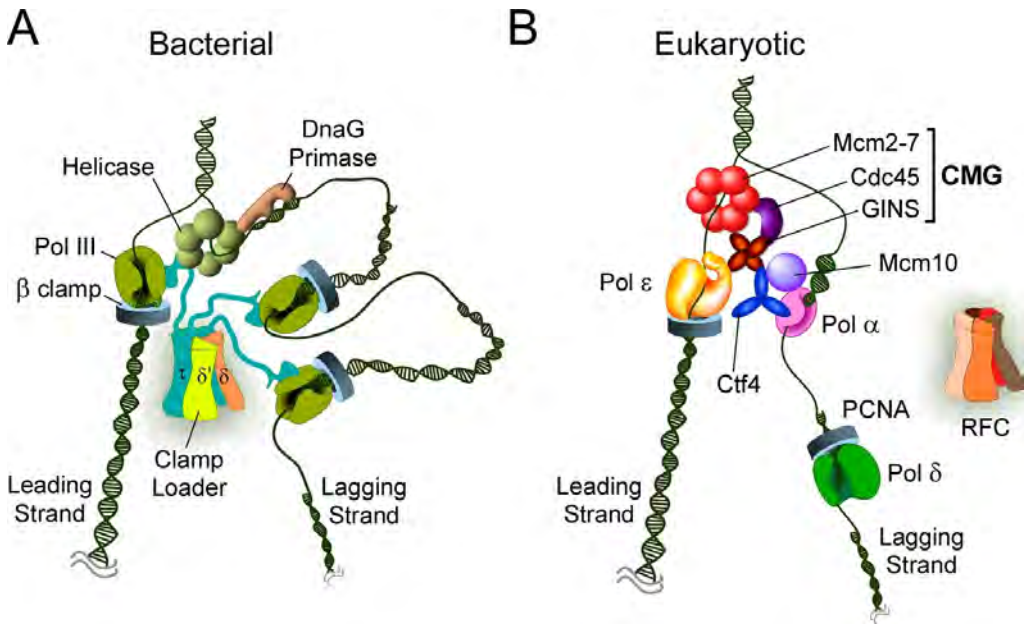
Accurate and efficient replication of DNA is an essential process required for all cellular life. There is a vast host of players at the replication fork that ensure DNA is unwound at the right place and time, and meticulously copied. The replication fork moves at remarkably high speeds (~1000 nucleotides per second in bacteria) (1), and requires the coordinated efforts of up to 31 different polypeptides (as recently discovered in eukaryotic replication) (2). A fascinating analogy from Tania Baker and Stephen Bell sets this information to a more useful scale: Assuming DNA was 1 m in diameter, the bacterial genome would be replicated at roughly 600km/hour using replication machinery that is the size of a FedEx truck (3). This analogy also offers a striking example of the accuracy by the replication machinery—even given the relatively large size and speed of these machines, mistakes would only be made every 170 km (3). However, mistakes that are not corrected can compromise genomic integrity, and lead to disease.

Although DNA polymerase is the major enzyme that carries out DNA synthesis, it is joined by a diverse array of molecular machines that regulate and coordinate the complex reactions required for DNA replication (**Table 1.1**). A dynamic protein complex known as the replisome carries out this coordination (**Figure 1.1**)(4). A series of regulated replication initiation steps lead to replisome assembly, and subsequent replication. This thesis will discuss two proteins at the heart of the replisome: the sliding clamp, and the sliding clamp loader.

Table 1.1 The minimal machinery used in DNA replication

		Bacterial	Eukaryotic	
Initation	**Origin recognition**	DnaA	ORC	Pre-RC (inactive)
			Cdc6	
	Helicase loading	DnaC	Cdt1	
			MCM	
	Helicase	DnaB	CMG (Cdc45 + MCM + GINS)	
	Primase	DnaG	Polα	
Replisome	**Helicase**	DnaB	CMG (Cdc45 + MCM + GINS)	
	Primase	DnaG	Polα	
	Polymerase	Pol III	Pol ε (leading)	
			Pol δ (lagging)	
	Sliding clamp	β	PCNA	
	Sliding clamp loader	γ/τ complex	RFC	
	Single-stranded binding proteins	SSB	RPA	

Figure 1.1: Replisomes of bacteria and eukaryotes.
(A) The bacterial replisome is organized by the clamp loader, which contains a tau subunit homotrimer with extensions that bind three C-family DNA polymerases and connect to the helicase. The helicase encircles the lagging strand. Primase is a single subunit that acts to prime synthesis. DNA loops form during Okazaki fragment synthesis as a consequence of the connection between the leading and lagging strand polymerases via the clamp loader.
(B) The eukaryotic replisome is organized by the 11-subunit CMG helicase (composed of Cdc45, the Mcm2-7 hexamer which encircles the leading strand and the 4-subunit GINS heterotetramer). GINS binds to the leading strand polymerase Pol ε. The 4-subunit Pol α-Primase interacts with CMG through the Ctf4 homotrimer, which also binds a GINS subunit of CMG. Pol α also extends RNA primers to form hybrid RNA-DNA primers. Primers are further extended into Okazaki fragments by Pol δ that functions with the PCNA clamp. Direct connections of Pol δ and the RFC clamp loader to other replisome components are currently unknown, and thus the lagging strand DNA may not form loops. The bacterial SSB tetramer and eukaryotic RPA ssDNA-binding proteins protect lagging strand ssDNA from nucleases and are not shown for clarity.
Reprinted with permission from Elsevier, License number 3886000117594. (5)



Replication initiation

DNA replication initiates at sites in the genome known as origins. Initiator proteins bind DNA at origins in order to start the process of replication. All initiator proteins are members of the AAA+ ATPase family (ATPase associated with diverse cellular activities) (6). Origin recognition and initiator binding follow similar principles, with some important differences. *Escherichia coli* (*E. coli*) contain only a single origin, *oriC*, which contains several sequence-specific regions (7). Multiple copies of the bacterial initiator protein, DnaA, form a helical filament and bind at this single origin (8). The DnaA filament uses ATP to unwind a single-strand DNA (ssDNA) bubble from which two replications forks form (9). Eukaryotic cells contain multiple origins per chromosome; multiple origins are necessary in order to replicate larger genomes in a timely manner (10). With the exception of the budding yeast *Saccharomyces cerevisiae* (*S. cerevisiae*) (11), eukaryotic origins do not have as well-defined DNA sequence specificity and instead might be defined by DNA nucleosome positioning that allows access to the origin (12, 13). The eukaryotic initiator protein is the origin recognition complex, ORC (14), which, together with Cdc6, binds to origins (15) The primary role of these origin-binding initiator ATPases is to load DNA helicases onto DNA. (**Table 1.1**).

Helicase loading and activation is a critical step because it is the limiting step for replication; the activated helicase recruits the rest of the replication machinery to the replication fork (16). Additionally, once DNA is replicated,

subsequent helicase reloading is tightly suppressed to prevent reinitiation in order to maintain genome stability (17). Helicases couple ATP hydrolysis to strand separation in order to create the leading and lagging strands. In bacteria, helicase activation is relatively simple compared to eukaryotes: the AAA+ ATPase DnaC helps load the helicase (DnaB) (7) onto the ssDNA region unwound by the initiator protein (DnaA) (18). Assembling an active helicase in eukaryotic cells is much more complicated than in bacteria. Because eukaryotes contain multiple origins per chromosome, replication re-initiation can lead to genomic instability (17). Therefore, eukaryotic cells split helicase loading and subsequent activation into two defined phases of the cell cycle (13, 16, 19). These phases are tightly regulated by additional proteins: during G1 phase, two factors, the MCM2-7 complex and Cdt1, bind to the initiator complex ORC:Cdc6. This constitutes the pre-replication (pre-RC) complex and its assembly is known as origin licensing. The pre-RC complex remains inactive until passage into S-phase, where the Cdc45 and GINS complex bind to the assembly. This binding activates the major replicative helicase complex which is known as CMG (Cdc45, the GINS complex, and MCM2-7) (19-21). The active helicase then unwinds the double stranded DNA (dsDNA), exposing two ssDNA fragments for bidirectional replication. After unwinding, the subsequent replication events are unregulated and occur spontaneously.

The last major initiation step requires primer synthesis at the replication fork. No known polymerase initiates DNA synthesis *de novo* and all cellular

polymerases require an initial RNA or DNA primer complementary to the template strand. The bacterial and eukaryotic primases, DnaG and Polα respectively, synthesize an initial RNA primer as well as primers for the lagging strand Okazaki fragments (22). The eukaryotic primase synthesizes a RNA/DNA hybrid primer, and acts at Okazaki fragments that form every 100-200bp (23).

The replisome

All dividing cells contain machinery referred to as a replisome (**Figure 1.1**) to replicate DNA simultaneously on both leading and lagging strands. Replisome organization varies between all three forms of life but contains at least six different core components: a replicative DNA polymerase, a helicase, a primase, a circular sliding clamp, a sliding clamp loader, and single-stranded DNA binding proteins (SSB) (**Table 1.1**) (4, 13). The role of the helicase and primase were discussed above. SSBs prevent re-annealing, as well as protect against nuclease digestion. DNA polymerases require the sliding clamp for high speed DNA synthesis. The clamp is a ring-shaped protein that tethers the polymerase to DNA, thus increasing its processivity. However, because the clamp is a closed ring, it must be placed onto DNA by the sliding clamp loader (24).

The bacterial replisome is organized to couple leading and lagging strand replication, known as the trombone model. In the trombone model, the lagging strand of DNA is looped through the lagging strand polymerase (akin to the shape of a trombone), requiring loading and unloading of the sliding clamp, along

with primase activity repeatedly along the lagging strand (25, 26). Along with helicase DnaB, and primase DnaG, replicative polymerase Pol III is recruited to the bacterial replication fork, including accessory proteins β-sliding clamp, and the clamp loader τ/γ complex, and SSB (13). The bacterial replisome is organized by the clamp loader complex; extensions from the τ subunits contact both Pol III at leading and lagging strands, and also contact the DnaB helicase (27). The clamp loader complex loads the β-clamp to the primer/template junction, formed by the DnaG primase (**Figure 1.1a**).

Organization of the eukaryotic replisome differs significantly from bacterial replisomes (**Figure 1.1b**). Three separate polymerases act at the eukaryotic replisome: pol ε and pol δ which act on the leading and lagging strands, respectively, as well as the primase/polymerase pol α, which initiates primers at each Okazaki fragment for the lagging strand (22). Analogous to bacterial SSB, the eukaryotic replication protein A (RPA) binds to ssDNA regions on the lagging strand after helicase unwinding. The eukaryotic replicative helicase CMG encircles the leading strand and enhances leading and lagging strand polymerase activity (2). Unlike the bacterial replisome, the eukaryotic clamp loader complex, replication-factor C (RFC) does not does directly bind to the replicative DNA polymerases or to the replicative helicase (2, 28). Instead, the eukaryotic clamp loader serves to open and load PCNA, the eukaryotic sliding clamp, onto primed sites of DNA.

Replicative polymerases contain proofreading activity to ensure accuracy, mistakes occur approximately once every 10^5 to 10^6 nucleotides added, depending on the fidelity of the polymerase (29). Additionally, replicative polymerases often encounter lesions of various sizes on DNA that act like roadblocks, causing the polymerase to stall. If the polymerase cannot replicate past these lesions, the replication fork can collapse, leading to DNA damage in the form of double strand breaks (DSBs). These, and additional types of DNA damage require DNA repair pathways, discussed briefly below.

DNA damage and repair

DNA repair pathways are critical to correct both endogenous and exogenous types of DNA damage in order to maintain genome integrity. Endogenous damage occurs in the case of replication errors and damage caused by reactive oxygen species from metabolic byproducts (30). Exogenous damage includes damage caused by lesion-inducing radiation or mutagenic chemicals (30). There are a variety of damage response pathways, such as direct removal of the DNA damage, damage checkpoint activation, transcriptional response, or apoptosis (31). Activating these responses depend on the type and severity of DNA damage. In both prokaryotes and eukaryotes, examples of pathways used to remove DNA damage are base excision repair (BER), mismatch repair (MMR), nucleotide excision repair (NER), and recombinational repair. Because DNA replication is tightly linked to the cell cycle, most DNA

damage activates checkpoint pathways that prevent premature progression into the next phase of the cell cycle (30). Depending on the type of damage, these pathways ensure that even a simple remedy of a single damaged base, or excision and re-synthesis of long stretch of DNA, can be repaired before cellular division. These pathways employ the use of a diverse array of enzymes, including, but not limited to nucleases, helicases, polymerases, recombinases, ligases, glycosylases, kinases, and phosphatases to maintain genome integrity (31).

Failure of DNA repair often leads to disease. In particular, there are several DNA repair disorders, such as ataxia telangiectasia (AT), xeroderma pigmentosum (XP), and Cockayne syndrome (CS) that are caused by failure of various DNA repair pathways. AT is caused by a mutation in one of the master checkpoint kinases, ATM (ataxia telangiectasia mutated) (32). XP is caused by mutations in enzymes responsible for NER and patients have a heightened sensitivity to UV light (33). Mutations in CSB or CSA, proteins involved in transcription-coupled NER repair cause Cockayne syndrome (34). Recently, a novel human disorder, ataxia telangiectasia-like disorder-2 (ATLD-2), was discovered to share similar clinical features with AT, XP, and CS (**Table 1.2**). ATLD-2 is caused by a single point mutation in the human sliding clamp (35). As discussed further below, the sliding clamp facilitates the actions of enzymes that function in DNA replication, but it also coordinates multiple cellular pathways through its abundant protein interactions.

Table 1.2: Symptoms of ALTD-2 and similar DNA repair diseases.

Symptoms of PCNA disorder	Ataxia Telangiectasia	Xeroderma Pigmentosum	Cockayne Sydrome
Small stature	+	++	+
Microcephaly	-	++	+
Telangiectasia	++	-	-
Extreme Photosensitivity	-/+	++	++
Intellectual disability	-/+	++	++
Ataxia	++	++	+
Neurodegeneration	++	++	++
Hearing loss	-	++	++
Premature aging	+	++	++

The sliding clamp

Sliding clamps are essential for processive DNA replication

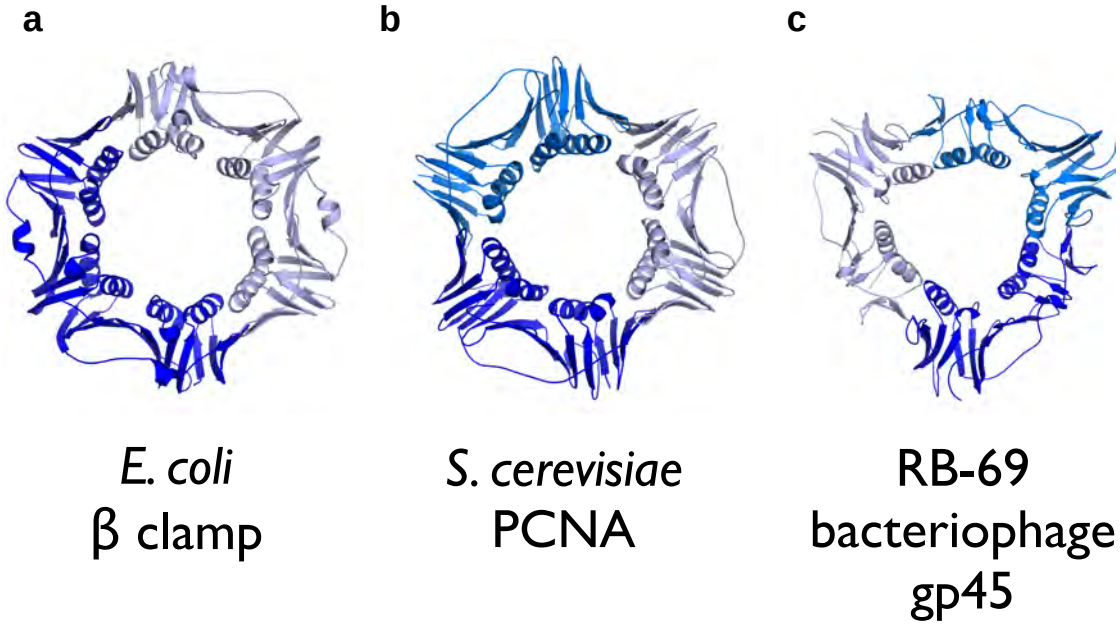
Sliding clamps are evolutionarily conserved accessory proteins that tether polymerases to DNA. This tethering prevents the polymerase from dissociating from DNA, dramatically increasing processivity. For example, bacterial polymerase Pol III dissociates from the template after approximately 10 bp in the absence of the clamp. In the presence of the clamp, this distance extends to over 80 kb (1, 36). In addition to the polymerase staying bound to DNA, the clamp also increases the rate of the replicative polymerase synthesis 40 to 100-fold in bacterial and phage systems, respectively (37-40). In eukaryotes, the presence of clamp and its clamp loader stimulates polymerase activity ~25-30 fold, thus increasing processive DNA synthesis from ~100bp to ~4kb (41). Therefore, without the tethering activity of the clamp, the efficiency of replication is severely diminished.

Sliding clamps share a widely conserved structure

How does the sliding clamp act as a molecular tether? Initial biochemical experiments using the *E. coli* β-clamp elegantly predicted the sliding clamp would be ring-shaped and topologically linked to DNA (42). The β-clamp showed tight association with circular dsDNA, but easily dissociated after DNA was linearized using restriction endonucleases, suggesting the β-clamp “slides” off the end of DNA (42).

Remarkably, when the first structure of the β-clamp from *E. coli* was published, it was indeed a toroid shaped protein, consisting of a homodimeric ring approximately 80Å in diameter (43, 44) (**Figure 1.2a**). While the structure was solved in the absence of DNA, several features suggested how DNA associates with the clamp. Despite the fact that clamps are overall negatively charged, the β-clamp structure revealed positively charged residues lining the central cavity of the clamp, which measures approximately 35Å—wide enough to accommodate A- or B- form dsDNA (29). B-form dsDNA modeled into the structure had no steric clashes, and molecular simulations predicted DNA threaded through the pore between the phosphate-backbone of DNA and positively charged side chains lining the pore. Molecular dynamics further predicted DNA threaded through the central pore of the clamp (45). This hypothesis was later confirmed by a crystal structure of the bacterial β-clamp bound to primed DNA (46). This structure revealed that the residues that line the clamp's pore make specific contacts with both ssDNA and dsDNA. However, these residues were dispensable for tight DNA-binding, and instead play a primary role in directing the clamp loading process (47).

Figure 1.2: Clamps from distinct forms of life share a similar architecture despite limited sequence conservation. All identical protomers are shown in shades of blue. *Left*: The β-clamp from *E.coli* is a homodimer (PDB: 2POL) *Middle* and *right*: The sliding clamps from yeast and bacteriophage are homotrimers. (PDBs: 1PLQ and 1B77) (43, 48, 49)

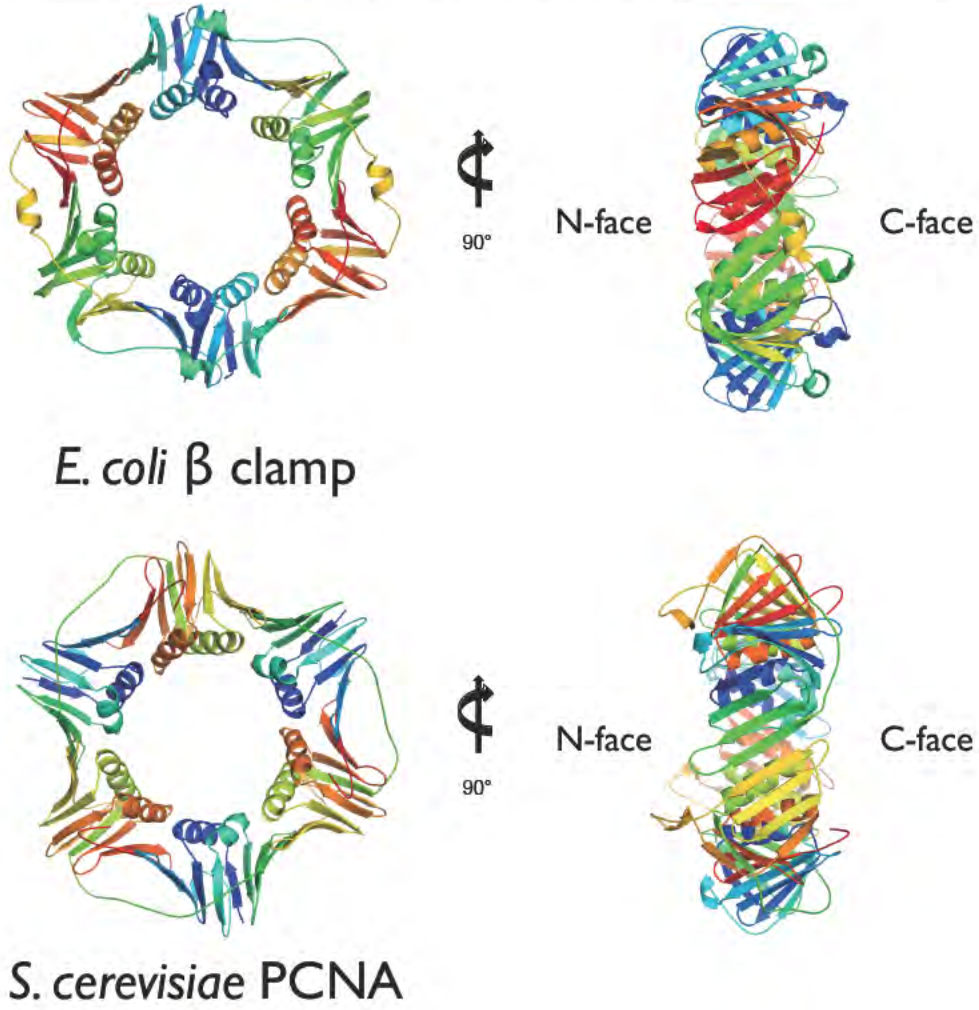


Despite low sequence conservation, the general architecture of sliding clamps is highly conserved throughout all domains of life (5, 49, 50). The clamp from *S. cerevisiae*, called proliferating cell nuclear antigen (PCNA) (49), and the clamp from bacteriophage RB-69, called gp45 (48) both share the same overall architecture as the *E. coli* β-clamp despite limited sequence similarity. Despite only ~15% sequence identity between PCNA and β-clamp, the RMSD after alignment on alpha carbons is less than 1Å (49). The major difference between these structures is in oligomeric assembly: bacterial clamps are homodimers, eukaryotic and viral clamps are homotrimers. Different classes of archaeal

clamps use either hetero- and homotrimers (43, 49, 51-53) (**Figure 1.2**). Each protomer contains two subdomains (as in trimeric clamps) or three subdomains (as in dimeric clamps). Regardless of the sequence conservation or oligomerization all clamps share similar pseudo-hexameric symmetry (48, 49, 54, 55) reflecting structural and functional conservation.

A head-to-tail arrangement within the pseudo-hexameric symmetry has two important implications for binding partners interacting with the clamp. First, the arrangement of the monomers within the ring produces two asymmetrical faces of the clamp, termed the N-face and C-face, each possessing distinct roles (**Figure 1.3**). This asymmetry is proposed to orient the polymerase, which binds to the C-face, toward the growing end of DNA (44). The N-face serves as the site for post-translational modifications for eukaryotic clamps (56, 57). Second, the placement of the subdomains within each monomer in a head-to-tail arrangement produces a long connection between both subdomains, termed the interdomain-connecting linker, or IDCL. Dimeric clamps possess two IDCLs while trimeric clamps have three. The IDCL protrudes out into solvent on the C-face side of the clamp and plays a critical role in PCNA function as the major binding site for sliding clamp-interacting proteins.

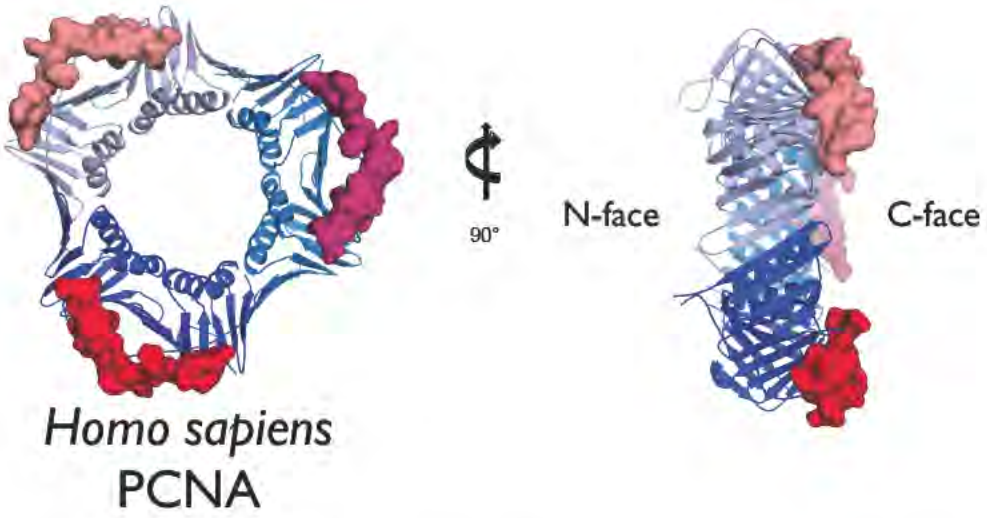
Figure 1.3: Clamps are organized in a head-to-tail arrangement. Protomers within the *E. coli* β-clamp and *S. cerevisiae* PCNA clamp are colored from N- to C-terminus, blue to red. This arrangement produces two distinct faces of the clamp, termed the N- and C-face. Additionally, the subdomains formed by each monomer are also arranged head-to-tail, necessitating a linker between N- and C-terminal domains. Most clamp interacting partners bind on the C-face, at the linker between subdomains. (PDBs: 2POL, 1PLQ) (43, 49)



The IDCL is a conserved binding site in all sliding clamps

The co-crystal structure of human PCNA (hPCNA) bound to a peptide from the cell-cycle inhibitor, $p21^{CIP1}$ identified the IDCL as the primary site for protein binding (58) (**Figure 1.4**)**.** The peptide binds each protomer in the ring at the IDCL, with an overall 1:1 stoichiometry, and without significant conformational changes in hPCNA. The peptide contains a 3_{10} helix, which "plugs" into a hydrophobic binding pocket on hPCNA formed primarily by the IDCL and the C-terminal subdomain (58). The peptide forms additional contacts, including salt-bridges and hydrogen bonds, along the length of the IDCL. This structure also highlighted the importance of the IDCL location; binding to this solvent exposed loop does not occlude the pore, which would prevent DNA binding. Importantly, binding at the IDCL allows the sliding clamp to be a platform for enzymes that then act on DNA.

Figure 1.4: Human PCNA binds to cell cycle regulator p21^{CIP1} peptide with 1:1 stoichiometry. PCNA protomers are shown in shades of blue, and p21^{CIP1} peptide is shown as surface representation in shades of red. p21^{CIP1} binds to the IDCL site on the C-face side of PCNA. The IDCL connects the N- and C-terminal subdomains of PCNA. (PDB: 1AXC). (58)



Subsequent studies have illustrated that the IDCL is a protein-binding platform for sliding clamp interacting proteins. More recent structures of hPCNA bound to interaction partners have revealed universal binding at the IDCL (58-62) (63-65). Crystal structures of sliding clamps from bacteriophage RB-69 clamp (66) and E. coli clamp (48) illustrate the IDCL is a conserved binding site. For example, a co-crystal structure of bacterial β-clamp bound to a peptide from Pol II showed the peptide bound to the IDCL (66). A co-crystal structure of the gp45 clamp from bacteriophage RB-69 bound to a peptide from DNA polymerase also shows bacteriophage clamp interacting partners bind to the IDCL (48). Co-crystal

structures also indicate that partner binding generally does not significantly change the conformation of the sliding clamp, although a recent study of an archaeal protein from *Thermococcus kodakarensis* discovered a novel inhibitor protein, TIP (*Thermococcales* inhibitor of PCNA) and reveals it distorts PCNA and prevents trimer formation (65).

Because sliding clamps have multiple partner binding sites, they can bind more than one partner simultaneously. Therefore, sliding clamps can act as a "toolbelt"; multiple ligands can bind simultaneously to then coordinate their actions (67). For example, the *E. coli* β-clamp can simultaneously bind two different polymerases, the high-fidelity Pol III, and an error-prone bypass polymerase, Pol IV (68). When Pol III stalls on DNA at a lesion, Pol IV can take over to replicate past the lesion. Demonstrations of the clamp from the archaeon *Sulfolobus solfataricus*, which is a heterotrimer, coordinates the actions of DNA polymerase, the flap endonuclease FEN1 and DNA ligase for Okazaki fragment maturation (69, 70). The toolbelt model is less clear in the case of the eukaryotic clamp, PCNA, where the clamp might mediate sequential switching of binding partners to help coordinate cellular processes (71). Regardless of the exact mechanism, the sliding clamp interacts with a variety of proteins at the IDCL, and as such, the clamp acts as a platform for these proteins to act on DNA.

PCNA interacting partners contain conserved binding motifs

The majority of the clamp's binding partners are known as PCNA-interacting partners, or PIPs. PIPs contain a conserved motif, called the PIP-box, defined as a eight-amino acid consensus sequence, QxxΨxxΩΩ, where ψ is a hydrophobic residue, L, M, or I, and Ω is either of the aromatic residues, F or Y (72). PIPs bind to the C-face side of PCNA at the IDCL. The molecular contacts between PCNA and PIPs have been extensively characterized: to date, there are 13 co-crystal structures of WT human PCNA bound to various PIPs (58-62) (63-65). The conserved glutamine residue in the PIP-box inserts into a small binding pocket on the C-terminal end of PCNA (58). The hydrophobic and aromatic residues in the PIP-motif form a 3_{10} helix that binds a larger hydrophobic pocket on PCNA. In line with the clamp's role as a processivity factor, most polymerases and many replication proteins contain a PIP-box, but PIPs are also found in proteins involved in a diverse array of cellular activities, including DNA repair, cell cycle and apoptosis regulation, chromatin remodeling, and epigenetic inheritance (**Table 1.3**) (57).

Table 1.3: A subset of PCNA-interacting partners (PIPs)

DNA replication		DNA repair	
Pol δ	Rfc1	Rad18	MPG
Pol ε	Rfc2	Rad5	NTH1
Pol η	Rfc4	Msh3	hMYH
Pol κ	Fen1	Msh6	APE1
Pol λ	Lig1	Mih1	APE2
Pol β	Topo IIα	EXO1	XRCC1
Rev1	Cdt1	UNG2	XPG
Chromatin remodeling		**Cell cycle and apoptosis**	
CAF-1	p300	CDK2	Cyclin D1
WSTF	HDAC1	EGFR	Gadd45
Epigenetic inheritance		p21	ING1b
DNMT1		p57	p53

A second, and lesser-characterized PCNA interaction motif is the AlkB homolog 2 PCNA interacting motif, or APIM. The APIM is defined by a five amino acid consensus sequence [KR]-[FYW]-[LIVA]-[LIVA]-[KR] (73). APIM interaction with PCNA is cell-cycle dependent, as well as partially dependent on post-translational modifications on PCNA. PIPs and APIMs appear to share a common binding site within the hydrophobic pocket created by the IDCL on the PCNA C-face, based on site directed mutagenesis within APIMs and competition studies (74). To date, there are no co-crystal structures of any APIM bound to PCNA, so the exact molecular contacts are yet unknown.

The textbook view of the clamp's major function is that of a processivity factor, a recent proposal suggests an important role for PCNA as a marker for

newly replicated DNA (5). PCNA interactions during replication are critical, but PCNA interactions facilitate a diverse array of cellular activities besides DNA replication (**Table 1.3**). PCNA's numerous partners (>100) imply that its role exceeds processivity. Indeed, not all polymerases require the clamp; as previously mentioned, pol ε is highly active in absence of PCNA (2), and replicative DNA polymerases from bacteriophages phi29, T5 and T7 do not require a sliding clamp for processive activity (75-78). These additional roles, as well as how the clamp interacts with so many additional proteins, will be discussed further below.

Additional PCNA roles beyond processivity

PCNA interactions facilitate a diverse range of cellular activities beyond replication by recruiting enzymes to newly replicated DNA (**Table 1.3**). As previously mentioned, these interactions are mediated by conserved motifs (the PIP-box or the APIM-motif) on PCNA partner proteins Additionally, post-translational PCNA modifications also modulate its interactions with other proteins, particularly with proteins involved in DNA damage responses (56). Post-translational modifications occur on the N-face of PCNA and structural studies indicate the modifications do not significantly alter PCNA conformations (79, 80). I discuss several examples of PCNA interacting proteins and their diverse functions below.

PCNA in DNA repair

PCNA plays an essential role in all types of DNA repair mechanisms through its many interactions with DNA repair proteins. PCNA is required for mismatch repair (MMR), and associates with mismatch repair proteins within the MutSα and MutSβ sensor complexes (81). MMR is strand-specific and the process is dependent on having a marker for newly replicated DNA. PCNA may assist in this marking newly synthesized DNA because its C-face is oriented along the newly synthesized strand.

Base excision repair (BER) relies on PCNA at nearly every step. PCNA interacts with glycosylases, such as UNG2, that remove damaged bases from the backbone (82). After glycosylation, APE1 nicks the phosphodiester bond in the abasic site, which requires an APE1-PCNA interaction. In short-patch BER, Polβ uses PCNA to synthesize the single gap, and the XRCC1-DNA ligase complex operates with PCNA to seal the nick. Interestingly, XRCC1 is an example of a PIP that does not contain the PIP-box motif. In long-patch repair, Pol δ or ε are recruited to PCNA to perform repair synthesis, and both FEN1 and DNA ligase I use PCNA to seal the nick, akin to the Okazaki fragment maturation process.

PCNA interacts with several XP proteins implicated in nucleotide excision repair (NER). The XPA protein contains an APIM-motif and is thought to be directed to sites of damage through its interaction with PCNA (83). The XPG protein contains a PIP-box to facilitate nuclease activity required for NER (84).

Mutations in XP proteins result in the DNA repair disorder xeroderma pigmentosum.

Ribonucleotide excision repair (RER) is also dependent on PCNA interactions. The RNaseH2 complex is responsible for excising the majority of misincorporated rNTPs from DNA (85). The catalytic subunit of the complex, RNaseH2B, contains a PIP-box (61) and PCNA localizes RNaseH2B to sites of damage, as well as enhances its enzymatic activity (86).

PCNA also plays important roles in processes preventing double strand breaks (DSBs). If the replicative polymerases cannot replicate past a lesion on the template strand, monoubiquitinated-PCNA recruits translesion synthesis polymerases η, κ or ι to DNA for bypass replication to prevent fork collapse (87, 88). It is then thought that ISG15 modification (Interferon-Stimulated Gene 15 or ISGylation) of PCNA is used to terminate error prone TLS synthesis (89). Finally, SUMOylated-PCNA on Lys164 by Ubc9 may play an uncharacterized role in preventing recombination by recruiting the anti-recombinogenic Srs2 helicase (87) (90, 91).

Cell cycle and apoptosis

Cell cycle control in eukaryotes is critical for preventing re-replication. PCNA interacts with cell cycle regulating proteins, Cdt1 and p21. Cdt1 helps load the MCM2-7 complex at sites of origins, and this interaction is critical for replication-dependent proteolysis of Cdt1 (92). The PCNA-p21^{CIP1} interaction

inhibits replication by blocking the polymerase-binding site on PCNA; $p21^{CIP1}$ is the tightest-known PIP-box (59, 93). PCNA bound $p21^{CIP1}$ inhibits binding to additional factors, barring any PCNA-mediated processes.

PCNA both positively and negatively regulates apoptosis. For example, PCNA interacts with members of the Gadd45 family via a PIP-box to suppress apoptosis (94). In contrast, PCNA interactions with ING1, a tumor suppressor, and this interaction increases apoptosis in response to UV light (95).

Epigenetic inheritance, chromatin assembly and remodeling

DNA is commonly modified via methylation in order to encode epigenetic information. A key contributor to epigenetic methylation markers is the methyltransferase DNMT1. DNMT1 contains a PIP-box and localizes to replication foci, as well as to sites of DNA repair, via its interactions with PCNA (96, 97). Nucleosome assembly pathways are carried out by the histone chaperones CAF-1 or the HIR complex. PCNA helps recruit CAF-1 to sites of replication so CAF-1 can promote nucleosome assembly (98).

Open questions and scope of thesis

Given the large number of PCNA binding partners and functions, PCNA must employ mechanisms for spatiotemporal regulation of partner binding. Not all PIPs bind to PCNA with the same affinity, so the sequence variation within the binding motif may serve as one discriminator for what PCNA can bind and when.

Posttranslational modification of PCNA (i.e., ubiquitination or SUMOylation) also regulates specificity. Although these methods regulate partner binding, additional mechanisms must exist given the large number of binding partners. A clue towards PCNA binding regulation comes from a recent study describing a PCNA mutant variant.

Ataxia telangiectasia-like syndrome 2 (ATLD-2), is caused by the Ser228Ile mutation in human PCNA (35). ATLD-2 patients present some clinical features that overlap with known DNA repair disorders, such as xeroderma pigmentosum, Cockayne syndrome, and ataxia telangiectasia. Patients exhibit a combination of symptoms such as microcephaly, hearing loss, abnormal blood vessel formation (telangiectasia), heightened sensitivity to sunlight, and premature aging (**Table 1.2**). Initial cell-based studies demonstrated NER is disrupted while DNA replication and several other PCNA-dependent functions are surprisingly normal (35). For example, $PCNA^{S228I}$ protein levels, localization, and DNA replication function are comparable to WT. DNA replication occurs to the same extent and same speed as with WT PCNA, as indicated by BrdU incorporation and DNA fiber analysis, respectively. However, cell-lines derived from affected patients suggested significant defects in NER. Pull-down assays and a mass spectrometry study indicated loss of binding partners that act in DNA repair pathways. However, the proteins whose PCNA-interactions were disrupted *in vivo* (as measured by mass spec) do not function in NER. Regardless, this study revealed a PCNA mutant that affects specific PCNA-regulated pathways

while leaving other functions intact (35).

This study was surprising given a PCNA mutant caused a very specific defect in DNA repair. Given the many PCNA-mediated cellular pathways that could be disrupted (DNA replication, nucleosome assembly, chromatin remodeling, cell cycle control, epigenetic inheritance), why was DNA repair specifically disrupted? Are DNA repair pathways intrinsically physiologically fragile, or would disruption of any of the other of PCNA-mediated pathways also lead to disease? We sought to use the $PCNA^{S228I}$ mutation as a tool to determine how particular pathways are mediated by PCNA function. Specifically, we aimed to understand the structural, and biochemical effects of this mutant, which I discuss in Chapter II, as well as the cellular effects of this mutant using an *in vivo* model system, which I discuss in Chapter III.

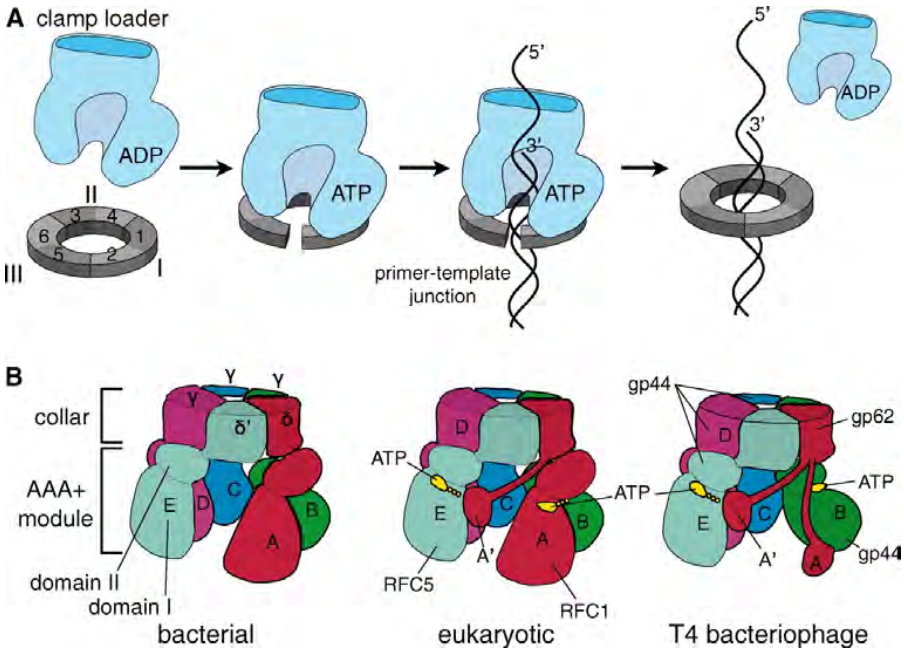
The sliding clamp loader

Because the clamp is a closed-ring, it must be actively loaded onto DNA. A multi-subunit ATPase, known as the sliding clamp loader, is responsible for opening the clamp and placing it at primed sites of DNA (24, 99, 100). The clamp loader recognizes primed sites containing a recessed 3' end, formed by a short primer synthesized by primase, and a 5' template overhang, known as the primer/template junction. The clamp loader couples ATP binding and hydrolysis to clamp closure around primer/template junctions, followed by clamp loader ejection from both the clamp and DNA (100) (**Figure 1.5**). A combination of

biochemical studies, as well as elegant structures of the clamp loader bound to its substrates helped reveal the overall clamp loader binding mechanism.

Figure 1.5: Clamp loaders and sliding clamps.

(A) Clamp-loading reaction. The clamp loader has low affinity for both clamp and primer/template DNA in the absence of ATP. Upon binding ATP, the clamp loader can bind the clamp and open it. The binding of primer/template DNA activates ATP hydrolysis, leading to ejection of the clamp loader. (B) Three classes of clamp loaders. Bacterial clamp loaders are pentamers consisting of three proteins: δ (A position), γ (B, C, and D positions), and δ' (E position). Eukaryotic clamp loaders (RFCs) consist of five different proteins, with the A subunit containing an A' domain that bridges the gap between the A and E AAA+ modules. The T4 bacteriophage clamp loader consists of two proteins: gp44 (the B, C, D, and E subunits) and gp62 (the A subunit).



Clamp loader architecture and subunit composition

Structural studies of clamp loaders from multiple organisms have revealed a conserved architecture of the clamp loader subunits (101-104). All clamp loaders are members of the AAA+ family of ATPases, which typically link ATP binding and hydrolysis to mechanical motion. (6, 105). AAA+ proteins are a subfamily of the additional strand conserved glutamate (ASCE) family of ATPases, contained within the superfamily of P-loop NTPases (106). Clamp loaders are unique in that they are pentamers that do not form a closed ring, unlike the typical ringed hexamers within the AAA+ family (105). Each subunit contains three minimal domains; the N-terminal of domain I contains the P-loop responsible for ATP binding (**Figure 1.5b**) (106). Domain II is made up of a short alpha-helical domain termed the "lid" (107). The ATP binding site is at the interface of domains I and II, and together, these domains make up the AAA+ module (102, 107). Finally, domain III is C-terminal to the AAA+ module and oligomerizes to form a disk-like structure referred to as the "collar" (101, 102), holding the pentamer together.

All clamp loader complexes share a similar subunit organization (**Figure 1.5**). The five subunits of the clamp loader assemble into a spiral shape with a gap between the first and the fifth subunit. Because the nomenclature for different clamp loader subunits has been confusing, a universal naming scheme has been employed to describe the identity of five subunits as letters A-E. The A subunit is positioned to the right of the gap between the first and the fifth subunit.

The remaining B,C,D, and E subunits are arranged counter-clockwise from the gap. (102). Non-bacterial clamp loaders also contain an additional subdomain off the A subunit called the A' domain, and this primarily serves as a linker for the gap between the A and E subunit, as well as providing an additional contact site for the clamp (102-104). Bacterial clamp loaders lack the A' domain (101), but also contain two χ and ψ subunits, which are dispensable for clamp loading activity, and tether to SSB within the bacterial replisome to couple clamp loader activity to the lagging strand (108).

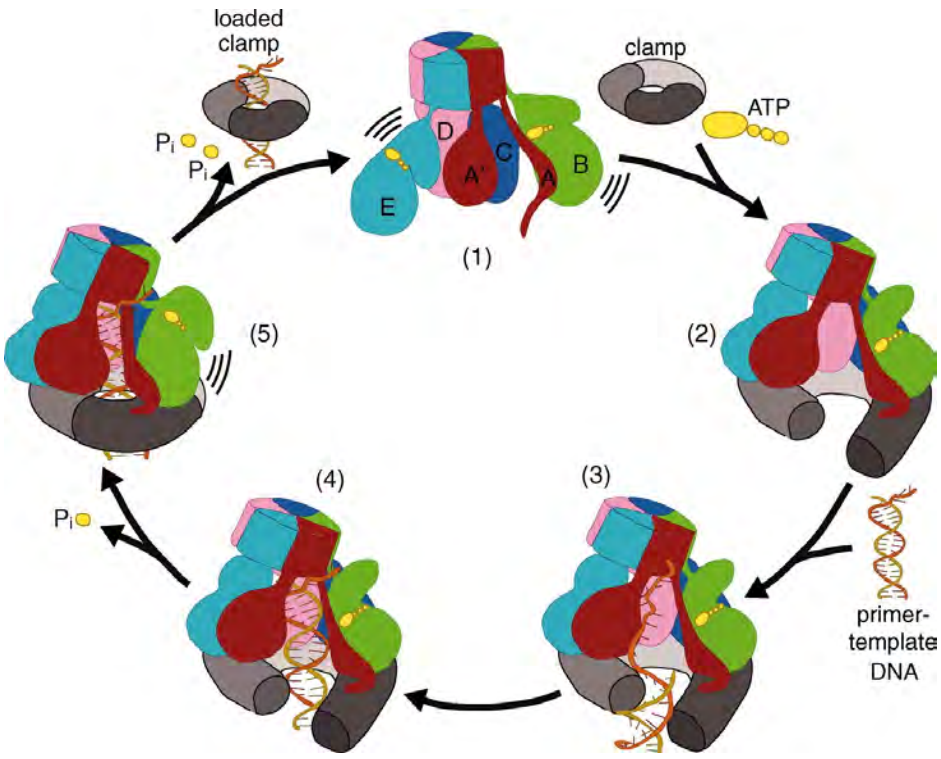
The clamp loader subunit composition and stoichiometry varies across life forms. Viral and archaeal clamp loaders consist of two proteins in a 4:1 stoichiometry, representing the simplest known composition (109). In the T4 bacteriophage, gp62 encodes for the unique subunit at position A, and gp44 encodes for the four identical ATPase subunits residing in positions B-E (103, 110). Bacterial clamp loaders are comprised of three different proteins, γ (or τ), δ and δ', creating an overall 3:1:1 stoichiometry. The non-ATPase subunits δ and δ' occupy positions A and E, respectively (101). In *E. coli*, the three ATPase subunits are encoded by the *dnaX* gene, which produces a full length, τ subunit, or a truncated product, γ, and occupy positions B, C, and D (101, 111, 112). Finally, the eukaryotic clamp loader, RFC, consists of RFC1 through RFC5. The subunits are arranged 1-4-3-2-5 for positions A-E, respectively (102). All five subunits can bind nucleotide (102), but only the ATPase activity of subunits B, C, and D are necessary for clamp loading activity (113). Despite these differences in

subunit composition, all clamp loader share a similar architecture and overall conserved mechanism.

The clamp loader reaction

The clamp loader coordinates activity through its three substrates: ATP, the sliding clamp, and primer/template DNA (**Figure 1.6**) Early biochemical studies revealed the clamp loader opens the clamp and loads it onto primer/template DNA in an ATP-dependent manner, raising significant questions about the order of substrate binding, and how hydrolysis is coupled to clamp closure (114, 115). Several important structural studies (102-104, 116) helped elucidate these questions and build a model for the conserved clamp loader reaction. I will discuss the order and implications of binding events in the clamp loader reaction below.

Figure 1.6. A detailed mechanism for the clamp loading reaction.
The reaction cycle for the T4 clamp loader is shown as a schematic diagram. (1) In the absence of ATP, the clamp loader AAA+ modules cannot organize into a spiral shape. (2) Upon ATP binding, the AAA+ modules form a spiral that can bind and open the clamp. (3) Primer/template DNA must thread through the gaps between the clamp subunits I and III and the clamp loader A and A´ domains. (4) Upon DNA binding in the interior chamber of the clamp loader, ATP hydrolysis is activated, most likely through flipping of the switch residue and release of the Walker B glutamate. (5) ATP hydrolysis at the B subunit breaks the interface at the AAA+ modules of the B and C subunits and allows closure of the clamp around primer/template DNA. Further ATP hydrolyses at the C and D subunits dissolve the symmetric spiral of AAA+ modules, thus ejecting the clamp loader because the recognition of DNA and the clamp is broken. The clamp is now loaded onto primer/template DNA, and the clamp loader is free to recycle for another round of clamp loading. (103)



ATP binding

Prior to binding the clamp and loading it onto DNA, the clamp loader must first bind ATP (**Figure 1.6**). ATP binding drives a conformational change in the clamp loader, which brings the subunit interfaces together using an “arginine finger” (101, 114, 117, 118). The arginine finger is a residue that reaches into the active site of a neighboring subunit to stabilize the negative charge on the γ-phosphate, and is a conserved residue within the AAA+ family (106). ATP binding therefore causes a conformational change that primes clamp loader for clamp binding.

Clamp binding

After ATP-binding, the clamp loader can bind and open the sliding clamp. How the clamp loader opens the sliding clamp differs slightly across life forms, due to the differences in stability of the sliding clamp ring across species. The *E. coli* clamp forms a very stable ring that has a dimer dissociation constant in the pM range (119, 120). Thus, the clamp loader must somehow force the dimer interface open. The structure of the *E. coli* δ clamp loader subunit bound to a β-clamp monomer demonstrates δ distorts the curvature of the monomer, likely opening the interface between the β-clamp dimer (121). This study suggested the clamp loader acts as a “wrench” to force the clamp open. This structure lacks the rest of the clamp loader subunits γ and δ’, but δ is capable of binding to the β-clamp alone in the absence of ATP. In the *E. coli* system, ATP is then required

for DNA binding and clamp closure (115).

Bacteriophage T4 clamp loader likely binds to an open clamp. Bacteriophage T4 clamp is the least stable clamp in solution (119), and is thought to frequently sample an open conformation. Therefore, the T4 clamp loader likely traps a transient open conformation (103, 122) instead of actively prying it open, as was seen with the *E. coli* clamp loader. One T4 crystal structure that contains two complexes in the asymmetric unit shows the clamp open in one clamp loader:clamp complex and closed in the other complex (103). In the crystal structure there are extensive contacts between the clamp loader subunits and the clamp. The clamp changes conformation from planar ring to a right-handed lockwasher (54, 103). This conformational change adopts a spiral that matches the helical symmetry of DNA, which has important implications for primer/template DNA binding, discussed below.

It is not clear whether the eukaryotic RFC clamp loader actively opens PCNA (as in the *E. coli* β-clamp) or traps an open conformation (as in the T4 gp45 clamp). Compared to the sliding clamps from *E. coli* and T4 bacteriophage, PCNA has intermediate stability (119, 120). The field currently lacks a structure of the clamp loader bound to an open clamp in the absence of DNA, but a crystallographic study revealed the structure of the eukaryotic clamp loader RFC bound to a closed PCNA clamp in the presence of the slowly hydrolyzable nucleotide analog ATPγS (102). This structure likely represents an encounter complex; nucleotide and the clamp have bound, but the clamp remains closed.

The next putative step would be opening of the clamp, followed by primer/template binding.

Primer/template binding

The structure of RFC bound to closed PCNA provided insight into how primer/template DNA might thread through the clamp pore (102). Modeling dsDNA into the complex illustrates the interior clamp loader spiral matches the major grove of dsDNA. The collar domains physically block the modeled recessed 3’ primer strand, whereas the 5’ template strand threads the gap created between the A and E subunits. These features lead to what is known as the “notched screw cap model” for how dsDNA screws up into the inner clamp loader chamber, and how the collar domains provide specificity for recognizing the 5’ template ssDNA overhang (123). It is worth noting that the arginine fingers in each of the clamp loader subunits were mutated to glutamine to prevent ATP hydrolysis. As a result, the pitch between subunits is not uniform, distorting the spiral symmetry (102). However, several subsequent studies support this hypothesis (103, 116).

The crystal structure of the *E. coli* clamp loader bound to primer/template DNA revealed the clamp loader AAA+ modules adopt a highly symmetrical right-handed spiral tracking the minor groove of DNA, and unexpectedly, the clamp loader contacts the template strand alone (116). The clamp loader does not contact the primer strand, which explains how the clamp loader accommodates

both DNA and RNA primers. A study investigating the effects of mutations in the clamp loader was also consistent with the interactions of the template strand and residues lining the clamp loader central cavity (124). The AAA+ modules also induce conformational changes of the dsDNA from B-form to nearly A-form, close to the physiologically relevant A-form of a RNA/DNA hybrid (125). Modeling a RNA/DNA primer/template into *E. coli* clamp loader structure suggested the clamp loader can accommodate an RNA primer. Thus, the clamp loader spiral can accommodate both DNA and RNA primers.

The crystal structure of the T4 clamp loader bound to an open clamp and DNA supports several previous studies explaining how primer/template binds (103). The gap at the open clamp interface is ~9Å wide, which is not wide enough to accommodate dsDNA but wide enough to allow ssDNA to pass, suggesting that the clamp likely binds ssDNA first. The structure of *E. coli* β-clamp bound to the δ-subunit of the clamp loader complex (121) supports this mechanism: the δ-subunit “wrench” relaxes the curvature of a β-monomer such that a superposition with the closed *E. coli* clamp (43) shows a ~15Å gap between the dimer interface. This is wide enough for ssDNA passage, but not wide enough for dsDNA (121). Similar to the *E. coli* clamp loader structure bound to primer/template DNA (116), the T4 complex structure also revealed primer/template DNA adopted an A-form conformation. Importantly, the T4 structures also suggested how DNA binding is coupled to ATP hydrolysis using an allosteric “switch” residue discussed below.

ATP hydrolysis and clamp loader ejection

DNA binding is directly coupled to ATP hydrolysis, and subsequent clamp loader ejection from the closed clamp bound to primer/template DNA (118, 126, 127). The switch residue (Lys80 in T4 gp44 subunits; Lys100 in *E. coli* γ subunits) is a conserved basic residue that interacts with the catalytic glutamate in the Walker B motif preventing it from catalyzing ATP hydrolysis (103, 116, 124). Upon DNA binding, the phosphate-backbone of the template strand directly contacts the switch residue, which releases the catalytic glutamate to assume its active conformation for ATP hydrolysis.

ATP hydrolysis triggers a rigid-body conformational change in AAA+ module of the clamp loader. One of the crystal structures of the T4 clamp loader bound to clamp, nucleotide, and primer/template DNA shows one subunit, the B subunit, has hydrolyzed ATP (103). In the ADP bound state, the AAA+ module from the B-subunit moves ~7Å away from the C-subunit. This conformational change allows clamp closure. The B-subunit movement breaks the spiral symmetry within the clamp loader subunits, and also breaks clamp interactions, explaining why the clamp is closed in the hydrolyzed structure. Continued hydrolysis of the other subunits would therefore eject the clamp loader from the clamp:primer/template complex.

Open questions and Scope of thesis

While the structural and biochemical data provide a detailed view of the

clamp loader reaction, there are several open questions. A major question is how the clamp loader recognizes and binds primer/template DNA that uses a RNA primer. All structures and most biochemical data have used a DNA/DNA primer/template, which partially satisfy our understanding of eukaryotic clamp loaders, but not prokaryotic or viral clamp loaders. Eukaryotes use the dual primase/polymerase, Pol α to synthesize a hybrid primer strand containing both RNA and DNA (128, 129). However, bacteria and bacteriophage both use primers comprised entirely of RNA (22). Given the sheer abundance of these life forms on earth (130, 131), the most common substrate for the clamp loader would be a RNA/DNA primer/template. The second major gap in knowledge concerns how ATP hydrolysis proceeds around the spiral. The clamp loader structure showing one subunit has hydrolyzed ATP raises the question whether ATP hydrolysis around the ring is random, or ordered. Other AAA+ and ASCE members hydrolyze ATP sequentially, such as E1 helicase, Rho helicase, and the phi29 motor (132-134). We wanted to understand the mechanism by which the clamp loader would recognize an RNA primer, as well as investigate the order of ATP hydrolysis between clamp loader subunits, which I will discuss in Chapter IV.

CHAPTER II

A disease-causing variant in PCNA disrupts a promiscuous protein binding site

Introduction

The replication of chromosomal DNA requires a highly conserved protein complex, known as the sliding clamp, for replicase function. Sliding clamps are found in all cellular life, as well as many viruses (4), and are required for cellular proliferation (57, 135). All sliding clamps share a similar multi-lobed ring structure, despite low sequence homology and different oligomerization states between the sliding clamps of eubacteria and those found in archaebacteria and eukaryotes (24, 43, 49, 53, 54, 58). In eukaryotes, the sliding clamp is known as Proliferating Cell Nuclear Antigen, or PCNA, and forms a homotrimeric assembly (58).

The sliding clamp surrounds DNA and acts as a processivity factor for DNA polymerases and other enzymes that act on DNA. The increase in processivity is striking because isolated bacterial replicative polymerases can only synthesize a handful of base pairs (~10 bps) before dissociating from the DNA (136), but in the presence of the clamp these polymerases catalyze ~80 kbps before dissociation (1). Moreover, sliding clamps are critical for the rapid rate of DNA replication, increasing the rate of nucleotide incorporation by up to ~100-fold (37-39). In addition to its role in facilitating DNA replication, PCNA plays a critical role in coordinating various cellular networks to ensure for successful copying and repair of DNA (5). When damage is too severe for the basal repair machinery, PCNA acts to assemble proteins that slow or halt cell cycle progression until the damage is treated. PCNA is necessary not only for the faithful transmission of the

genetic material, but also for proper inheritance of epigenetic information in the form of nucleosome deposition and DNA methylation (96, 137-140). Finally, PCNA functions not just within the confines of the nucleus but also in the cytosol of neutrophils, where it plays an important but nebulous role in regulating apoptosis (141-143).

The sliding clamp accomplishes the difficult task of coordinating multiple pathways by binding various proteins that then carry out their function on DNA. Scores of proteins are capable of binding PCNA but, because PCNA is a homotrimer, it can bind up to three ligands simultaneously (68, 70, 144). Thus PCNA can act as a 'toolbelt' to assemble multiple enzymes that cooperate to complete a given task. For example, heterotrimeric archaeal PCNA simultaneously binds the replicative DNA polymerase PolB1, Flap Endonuclease I (FEN1), and DNA Ligase I to cooperatively catalyze Okazaki fragment maturation (70), although this role is less clear in eukaryotes (71). Because of PCNA's complex role in coordinating various activities and enzymes both in series and in parallel, PCNA is a 'hub' protein that connects multiple different pathways and activities within the network of the cell (145).

PCNA primarily uses a single site to bind a majority of its interaction partners, which are known as PCNA-interacting proteins or PIPs (57, 58, 145). Each subunit contains one binding site, with three identical sites lining one face of the ring. Each binding site is comprised of a groove formed by the C-terminal domain and the interdomain connecting loop or IDCL, with some partners binding

to the N-terminal domain as well (58). Most known PIPs use a simple sequence motif known as the PIP-box to bind PCNA. The PIP-box motif has a signature sequence QxxØxx$\vartheta\vartheta$ (where is Ø a hydrophobic residue I, L, or M and ϑ is either of the aromatic residues F or Y) and forms a 3_{10} helix upon binding within the groove formed by the C-terminal domain and the IDCL (58, 59). A second motif, known as the AlkB homolog 2 PCNA interacting motif (or APIM), has been identified to bind PCNA (73). However, binding of APIM sequences to PCNA is largely uncharacterized. Because a large and diverse array of PIPs must bind PCNA in a spatiotemporally regulated manner, the factors that control PIP binding are critical for proper cellular physiology. As an elegant example of this molecular balancing act, yeast harboring a mutated PCNA with increased PIP-box affinity exhibit severe DNA replication and repair defects (146). Thus, interactions with PCNA have evolved to have precise dynamic exchange between interaction partners. However, given the complexity of the vast PCNA interaction network, the factors that govern PIP binding have been difficult to elucidate.

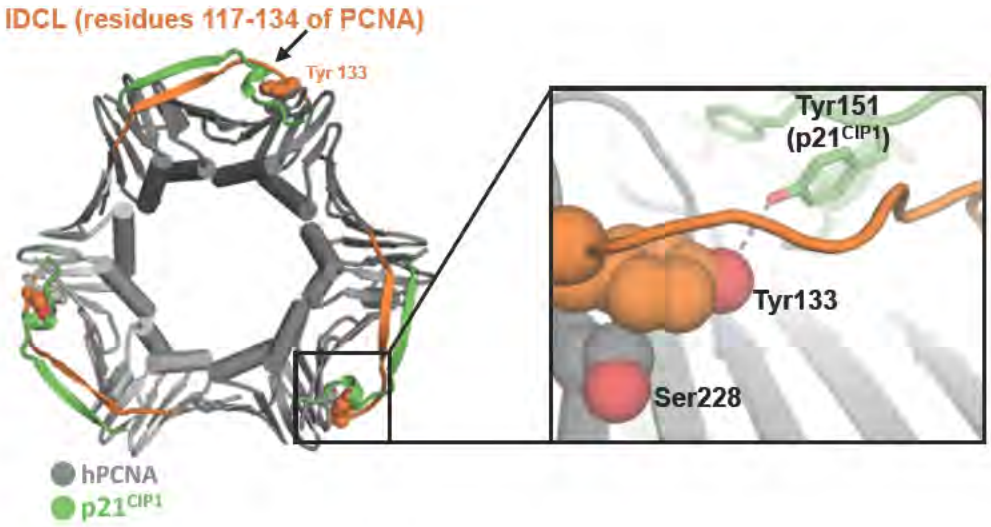
A hypomorphic mutation in the human PCNA (hPCNA) gene was recently identified to result in an autosomal recessive disorder with symptoms similar to DNA repair diseases such as Xeroderma Pigmentosum, Cockayne Syndrome, and Ataxia Telangiectasia (35). In patient-derived cells, DNA replication seems to proceed normally but nucleotide excision repair (NER) is disrupted. Thus the major cause of the disease was ascribed to a defect in NER (35). The specific

effect on NER is remarkable because the mutation, conversion of Ser228 to isoleucine (S228I), is at a buried site proximal to the promiscuous PIP-binding pocket (**Figure 2.1**). The mutation is hypothesized to alter the binding site, and preliminary Stable Isotope Labeling by Amino Acids in Culture (SILAC) mass spectrometry experiments suggested that the mutation alters binding of some PIPs (35). However, how the mutation affects hPCNA structure and its effect on PIP binding affinity are still unknown.

Here we report the crystal structure of hPCNA-S228I and identify a major disruption of the IDCL and the PIP binding groove. This disruption results in idiosyncratic binding defects in several PIP-box peptides. We further show that binding of PIP-box peptides to hPCNA-S228I occurs through distinct binding modes, with certain inter-molecular interactions overcoming the disruption of the PIP binding groove. Our results illustrate the complexity of interactions between PCNA and PIPs and highlight novel cellular pathways that may contribute to disease. Finally, our results have implications for understanding the plasticity of PIP binding.

Figure 2.1: Structure of hPCNA with p21^{CIP1} PIP-box peptide.

The PCNA homotrimeric ring is shown in shades of gray, with the p21^{CIP1} PIP-box peptide shown in green, and the IDCL region is shown in orange (residues 117-134). The site of the disease mutation Ser228 as well as the key IDCL residue Tyr133 are shown as spheres. Inset shows detailed view of the p21^{CIP1} PIP-box peptide and its proximity to the two key residues Ser228 and Tyr133. (PDB code 1AXC (58)).



Materials and Methods

Mutagenesis

Site-directed mutagenesis was performed using a protocol similar to QuikChange (147) using untagged hPCNA-WT in a pET3c vector (kindly provided by J. Kuriyan) and primers with the sequence 5'-ggtgacactcattatgtctgcagatgtacccttgttgtagag-3', and 5'-ctgcagacataatgagtgtcaccgttgaagagagtggagtggc-3'. Enzymes were purchased from New England BioLabs. Oligonucleotides for mutagenesis were purchased from IDT.

Expression and purification:

All proteins were expressed in E. coli BLR-DE3 cells containing either the pET3c-WT- or hPCNA-S228I construct. Cultures were grown in 2x-YT media containing 100 µg/mL of ampicillin antibiotic at 37°C. After reaching an OD600 between 0.6-0.8, cultures were incubated at 4°C for 20 minutes before being induced with 1 mM IPTG to and placed at 18°C for overnight expression. Cells were pelleted by centrifugation, resuspended in buffer A (25 mM Tris pH 7.5, 10% vol/vol glycerol, 2 mM DTT) and flash frozen in liquid nitrogen.

Cell pellets were thawed and lysed via cell disruptor (Microfluidics Inc.) four times to ensure sufficient DNA shearing. Lysate was centrifuged at 14,000 rpm, for 40 minutes at 4°C. Cleared lysates were filtered and loaded onto sequential S and Q columns (GE Healthcare) pre-equilibrated in buffer A. After loading

cleared lysates, the S column was removed and Q column was washed with 2 column volumes of buffer A. PCNA was eluted using a gradient of buffer B (25mM Tris pH 7.5, 10% vol/vol glycerol, 2mM DTT, 1M NaCl). PCNA-containing fractions eluted between 29%-33% buffer B were pooled and concentrated to load onto a Sephacryl-200 gel filtration column (GE Healthcare) pre-equilibrated in GF buffer (20mM Tris pH 7.5, 5% vol/vol glycerol, 2mM DTT). hPCNA-containing fractions were pooled, concentrated to >30mg/mL, flash frozen in liquid nitrogen and stored at -80°C.

Crystallization, structure determination and refinement

All proteins were crystallized using the hanging drop vapor diffusion method. Apo hPCNA-S228I drops were set with final protein concentration between 5-20mg/mL and crystals formed in 100mM sodium acetate pH 4.5, 2.0 M ammonium sulfate. Cryogenic buffer contained elevated amounts of ammonium sulfate precipitant supplemented with 25% (v/v) glycerol. Diffraction data was collected at Advanced Photon Source beamline 23-ID-B at a wavelength of 1.033 Å. For co-crystallization of hPCNA-S288I to $p21^{CIP}$ peptide was pre-incubated at 1.5 molar excess of protein (final drop concentration between 5-20 mg/mL). Co-crystals formed in 100mM HEPES pH 7.5, 25% PEG 3350, which also served as the cryogenic buffer supplemented with excess $p21^{CIP}$ peptide. Diffraction data was collected at SIBYLS 12.3.1 beam line at the Advanced Light Source with a wavelength of 1.000 Å. For co-crystallization of

hPCNA-S228I to FEN1pep, peptide was pre-incubated at 1.25 molar excess of protein (final drop concentration between 5-20mg/mL). Co-crystals formed in 2.35 M ammonium sulfate, 100 mM sodium acetate pH 4.6, 40% (v/v) (±)-1,3-Butanediol. Cryogenic buffer contained elevated levels of ammonium sulfate precipitant supplemented with FEN1pep and 20% (v/v) glycerol. Diffraction data were collected at Advanced Photon Source beamline 23-ID-B at a wavelength of 1.033 Å. All datasets were processed using HKL-3000 (148). Structures were solved by molecular replacement (MR) using wild-type hPCNA as a search model (serine 228 was removed for MR) using Phaser (149). Model building and refinement were performed with Coot (150) and PHENIX (151). Analysis in phenix xtriage revealed twinning for co-crystals of hPCNA-S228I-p21 and hPCNA-S228I-FEN1pep, as has been commonly seen for other hPCNA crystals (58, 59). Data for both the $p21^{CIP}$ peptide and FEN1pep could be processed in point group P6, but all molecular replacement solutions resulted in unstable refinement with high R_{work} and R_{free}. Therefore, symmetry was reduced to P3, and solutions yielded stable refinement with acceptable R values. The twin laws [-h,-k,l] and [h,-h-k,-l] were applied throughout refinement for the $p21^{CIP}$ and FEN1pep structures, respectively.

Isothermal Titration Calorimetry

Thermodynamics of peptide binding to WT and mutant hPCNA were determined by isothermal titration calorimetry using a Micro-Cal ITC200 system

(Malvern). Peptides (FEN1pep=STQGRLDDFFKVTGSL-OH; RNaseH2B= DKSGMKSIDTFFGVKNKKKIGKV-OH) were purchased from 21st Century Biochemicals and synthesized to >95% purity. The $p21^{CIP}$ peptide (CGRKRRQTSMTDFYHSKRRLIFS) was a gift from David King, UC Berkeley/HHMI. Peptides were solubilized in water and diluted in reducing PBS (137 mM NaCl, 2.7 mM KCl, 1 mM TCEP, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.5). hPCNA was buffer exchanged into reducing PBS using PD-10 desalting columns (GE Healthcare). In a typical experiment, the peptide ligand (400 μM for $p21^{CIP}$; 750 μM for FEN1pep and RNaseH2B in the syringe) was titrated over 19 injections of 2 μL into 20-25 μM of hPCNA, thermostated to 30°C in the sample cell. The quantity of heat released by peptide binding to hPCNA was measured by integrating the area under the curve of each injection peak. The heat of dilution of peptide injected into buffer was subtracted from the raw data before analysis. Data were analyzed and fit to a one-site binding model using ORIGIN software (MicroCal). Errors reported are weighted averages from multiple independent experiments (n=4 for hPCNA-S228I with $p21^{CIP}$; all others n=3).

Structural and Sequence analysis

Superposition performed with SuperPose (152) or PyMOL. Buried surface area was calculated with 'Protein interfaces, surfaces and assemblies' (PISA) service at the European Bioinformatics Institute (http://www.ebi.ac.uk/pdbe/prot_int/pistart.html(153). Hydrogen bonds calculated

using the VADAR server (154).

Results

Structure of hPCNA-S228I

To elucidate how the S228I variant alters hPCNA conformation, we crystallized the mutant protein and determined its structure to 2.7 Å resolution (**Table 2.1**). The overall structure of hPCNA-S228I is largely similar to that of wild-type (64) (**Figure 2.2a**), with an global RMSD of 0.7 Å for simultaneous superposition of all three chains (152). Moreover, the electrostatic surface within the pore of hPCNA-S228I is nearly identical to that of wild-type (**Figure 2.3**), suggesting that the interaction of the sliding clamp with DNA (45-47, 155) is unaffected by the S228I mutation.

Although the global structure of hPCNA-S228I is similar to wild-type, there are substantial changes near and extending away from the mutation site, including significant changes in the IDCL. Notably, a residue that contacts the mutation site, Tyr133, rotates outwards by nearly 90° to prevent clashing with the Ile228 sidechain gamma methyl group (**Figure 2.2b**). Tyr133 is a highly conserved residue (**Figure 2.2c**) within the IDCL that lines the PIP-binding pocket in wild-type hPCNA (**Figure 2.1**). This rotation of Tyr133 is

Table 2.1: Crystallographic Statistics

	hPCNA-S228I Apo	hPCNA-S228I:p21^{CIP1}	hPCNA-S228I:FEN1pep
Data Collection			
Wavelength (Å)	1.033	1.000	1.033
Resolution range (Å)	42.85 - 2 67 (2.76 - 2.67)	34.44 - 1.93 (2.0 - 1.93)	30 95 - 2.07 (2.148 - 2.074)
Space group	P 43 21 2	P 3	P 3
Unit cell a,b,c	162.324, 162 324, 139 837	143.295, 143.295, 41.402	80.721, 80.721, 66 556
Unit cell α,β,γ	90, 90, 90	90, 90, 120	90, 90, 120
Total reflections	431888 (40717)	553500 (51330)	113471 (10985)
Unique reflections	53660 (5213)	71743 (7159)	29297 (2881)
Multiplicity	8.0 (7.7)	7.7 (6.8)	3.9 (3.8)
Completeness (%)	99.79 (98.34)	99.78 (98.65)	99.70 (98.33)
Mean /sigma(I)	12.3 (3 32)	14 (2.96)	9.7 (2.1)
Wilson B-factor	47 65	28.1	36.6
Twin fraction	-	0.49	0.38
R-merge	0.124 (0 837)	0.090 (0 568)	0.078 (0.636)
R-meas	0.132 (0 899)	0.097 (0 612)	0.089 (0.737)
R-pim	0.046 (0 322)	0.035 (0 227)	0.044 (0.369)
CC1/2	0.997 (0.711)	0.997 (0 904)	0.997 (0.685)
CC*	0.999 (0 912)	0.999 (0 974)	0.999 (0.902)
Refinement			
R-work	0.1888 (0.2547)	0.1623 (0.2902)	0 202 (0.3319)
R-free	0.2263 (0.2961)	0.2211 (0 2613)	0.2292 (0.3247)
Twin law	-	-h,-k,l	h,-h-k,-l
Number of non-hydrogen atoms	6040	7085	4129
macromolecules	5804	6384	4030
water	236	701	99
Protein residues	753	820	523
RMS(bonds)	0.014	0.016	0.004
RMS(angles)	1 63	1.59	0.73
Ramachandran favored (%)	99	99	98
Ramachandran outliers (%)	0.4	0	0 2
Clashscore	6.44	17 02	4.95
Average B-factor	50 90	30 80	49 90
macromolecules	50 90	30.10	50.10
solvent	48 80	37 00	42.40

Figure 2.2: S228I mutation perturbs the IDCL conformation.
A) Overlay of the structure of hPCNA-S228I with the hPCNA-WT (PDB code 1VYM (64)). The overall C_α RMSD is 0.7 Å. Images on bottom-left and on bottom-right show zoomed-in views of the WT- and hPCNA-S228I comparison, from the top and side of the ring, respectively. Any residue that differs in C_α position by >2 Å is displayed with red arrows. Most of the arrows are in the IDCL and/or PIP binding site. Ile228 is shown as spheres.

B) The S228I mutation induces a large rotation in Tyr133. hPCNA-S228I (blue) is superposed onto the hPCNA-WT structure(64) (gray). The larger steric bulk of the isoleucine sidechain forces Tyr133 into an alternate conformation.

C) Sequence alignment of hPCNA. Top: Residues within the IDCL are highly conserved, including tyrosine at position 133 (purple triangle). Bottom: Serine is conserved at position 228 (purple triangle), with the exception of *S. cerevisiae*.

D) & E) Relative B-factors suggest that the hPCNA-S228I IDCL is less dynamic than WT. hPCNA-WT is shown in panel (D) and hPCNA-S228I in panel (E). The thickness and color of the chain are scaled by the normalized B factor (low to high B factor = thin to thick, blue to red). The highest flexibility is seen in the IDCL and in a few surface loops.

Figure 2.2: S228I mutation perturbs the IDCL conformation.

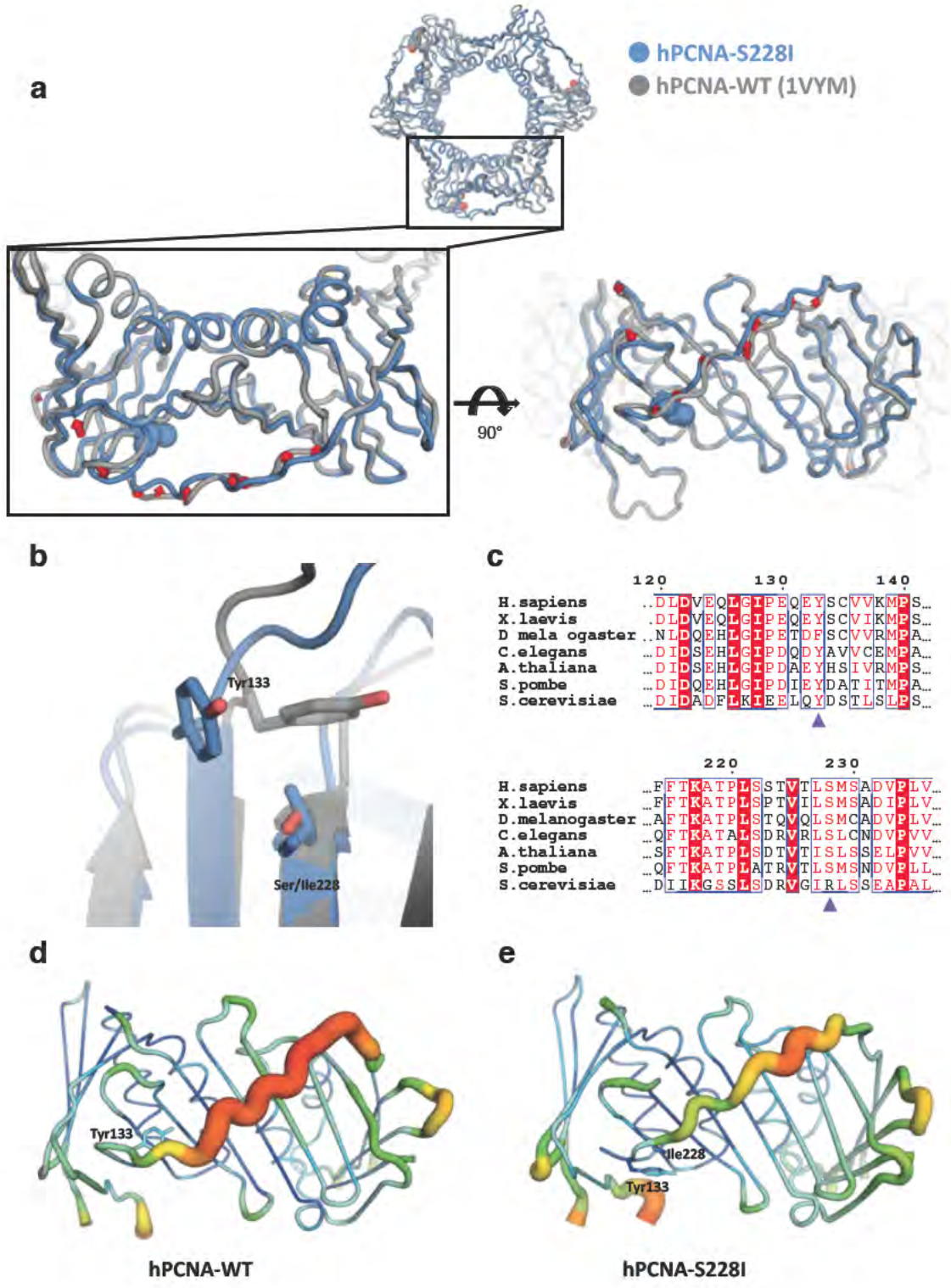
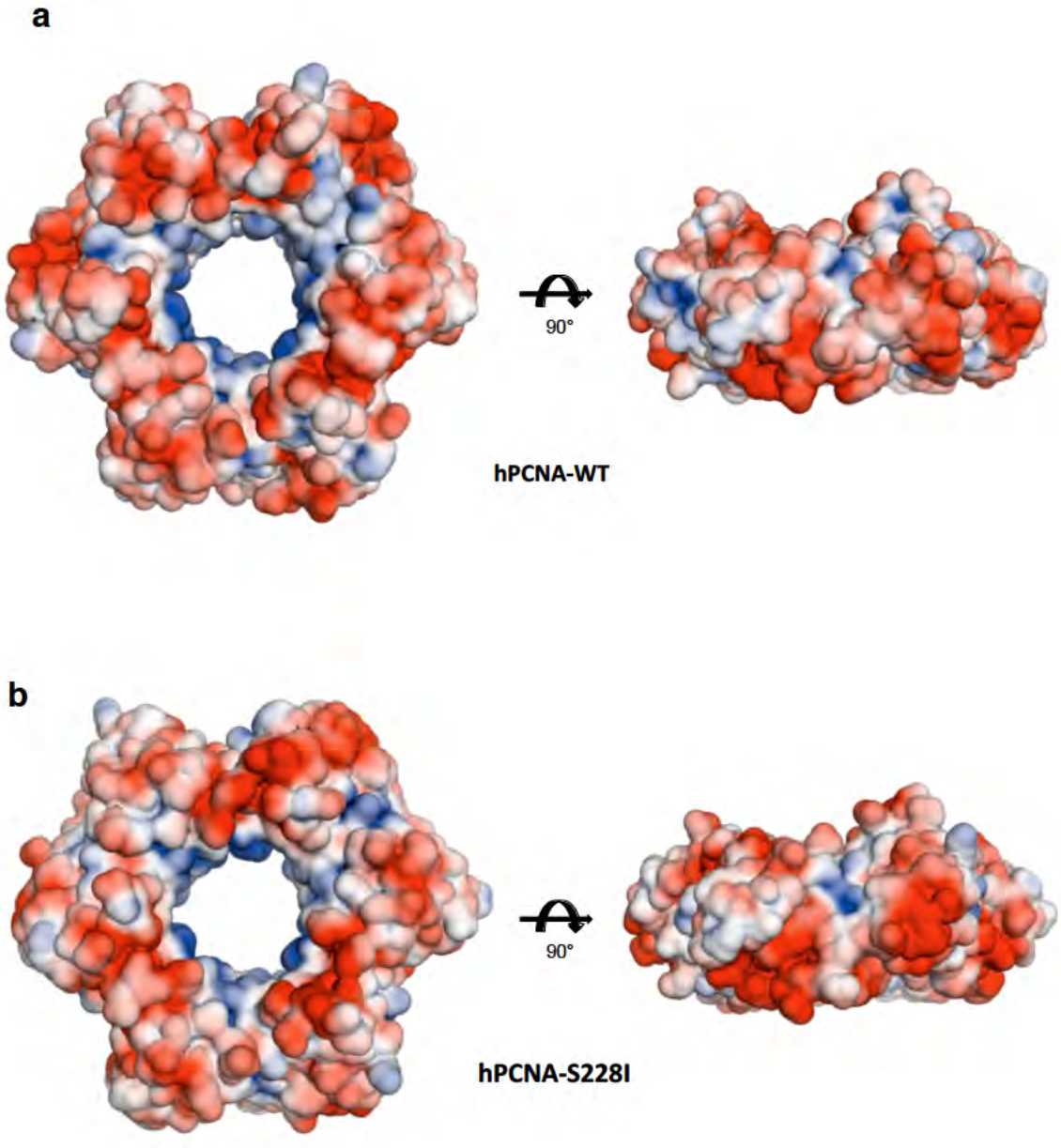


Figure 2.3: **PCNA surface electrostatics.**

The surface electrostatics of hPCNA-WT (1VYM (64)) and hPCNA-S228I (A & B, respectively). Electrostatics were generated in PyMOL using the APBS plugin (156). Figure is colored by electrostatic potential, red to blue (negative to positive).



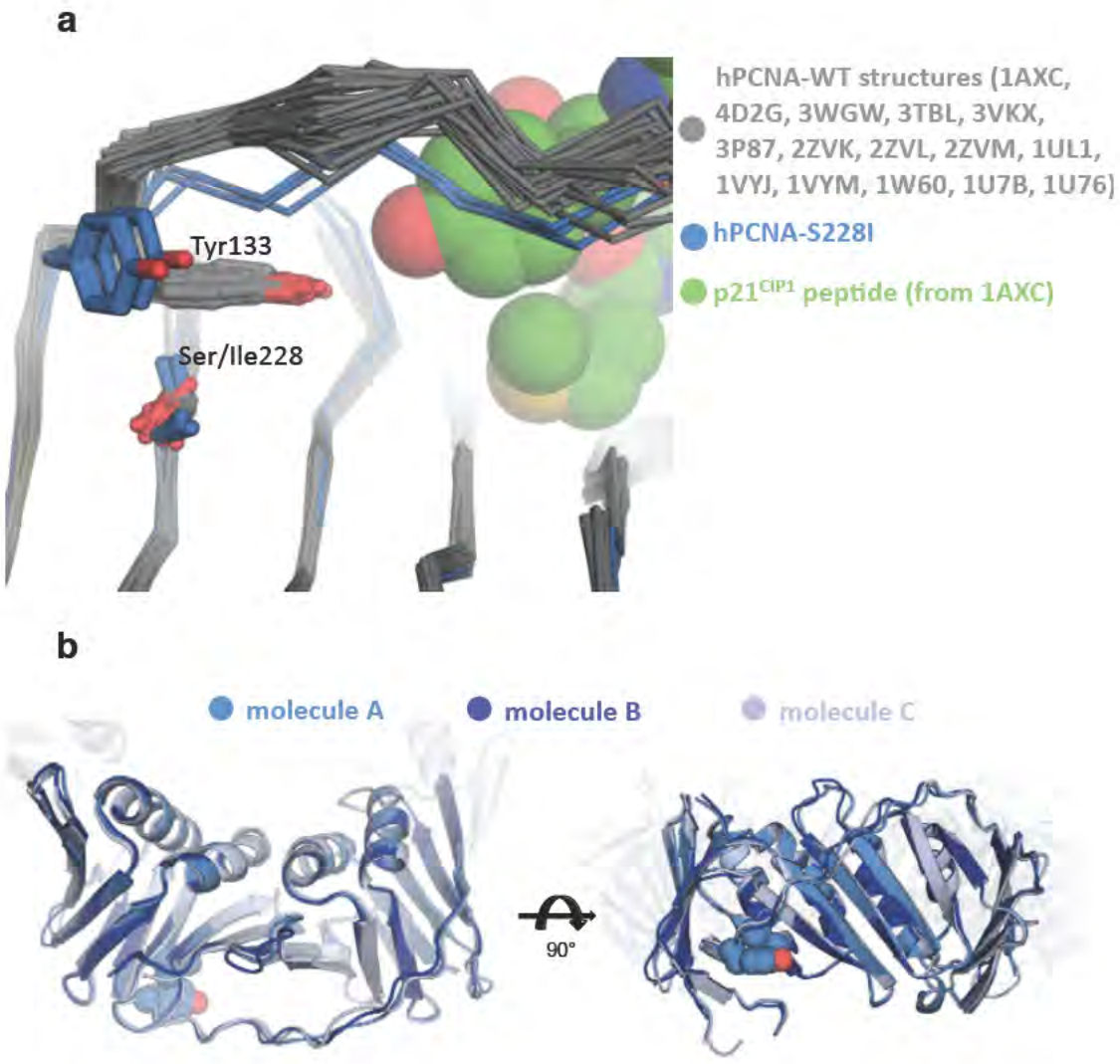
unprecedented in hPCNA; in the 15 different crystallographic structures of hPCNA (with ligands and without) (58-62, 64, 157-159), none show a significant change in the conformation of Tyr133 (**Figure 2.4a**). The 'flipped out' orientation of Tyr133 is the predominant conformation but is likely not the sole conformation populated, because F_o-F_c maps reveal significant positive density adjacent to the modeled conformation of Tyr133. This flat-shaped density is found in all three subunits and can be strengthened by averaging the maps over non-crystallographically related PCNA protomers (**Figure 2.5a**). Superposition of the hPCNA-WT structure places the canonical conformation of Tyr133 into this density. However, our attempts to model this alternative conformation, even at low occupancy, lead to an increase in R_{free} and decreased map quality. Thus our results suggest that Tyr133 accesses the canonical conformation but with an occupancy too low to accurately model.

The rotation of Tyr133 significantly perturbs the entire IDCL, with substantial conformational changes observed in IDCL residues far from the site of the mutation. In all three copies of hPCNA in the asymmetric unit, we observe a similar conformation of the IDCL (**Figure 2.4b**). This result suggests that the altered IDCL conformation is not due to crystal packing forces because each hPCNA protomer experiences a different crystalline environment. The conformational changes in the IDCL are mainly composed of twisting motions of

Figure 2.4: Conformation of the IDCL in hPCNA-S228I.

(A) A novel conformation of the IDCL in apo hPCNA-S228I. All chains of hPCNA from many different structures of hPCNA-WT shown in gray (PDB codes: 1AXC, 3WGW, 3TBL, 3VKX, 3P87, 2ZVL, 2ZVM, 1UL1, 1VYJ, 1VYM, 1W60, 1U7B, 1U76) are superposed with the three copies of the hPCNA-S228I structure (shown in blue). Only hPCNA-S228I shows major deviation of the IDCL due to rotation outward of Tyr133. The $p21^{CIP}$ PIP-box peptide bound to hPCNA is shown to indicate the binding site.

(B) All three copies of hPCNA-S228I show an altered conformation of the IDCL.



the backbone, such that each residue's sidechain is oriented in a different direction than in the canonical IDCL configuration. The twisting motion is prevalent even out to Asp120, which is thirteen residues and ~39 Å away from Tyr133. Thus, the local change in the orientation of Tyr133 triggers long-range reorganization of the IDCL.

Despite the substantial disruption of the canonical IDCL conformation, the B-factors for the IDCL of hPCNA-S228I are relatively low as compared to those of the wild-type IDCL. To quantify these changes, we normalize the B-factors of the IDCL residues to the average of the entire hPCNA protein. We observe that the normalized B-factors for the IDCL of wild-type hPCNA are ~40% higher than those from hPCNA-S228I (**Figure 2.2d&e**). This result suggests that the S228I mutation results in a perturbation of the IDCL dynamics, in addition to the more obvious changes in conformation.

S228I mutation disrupts PIP binding

Because the IDCL constitutes the majority of the PIP-box interaction surface, the conformation of PIP-box binding pocket is substantially disrupted in S228I. The molecular surface near the PIP-box binding groove is significantly different than in hPCNA-WT (**Figure 2.6a&b**). The volume and depth of the binding groove appears much smaller and shallower than the hPCNA-WT. To quantify this change, we measured the size and shape of the PIP-box binding cavity using the program CAVER (160). This analysis revealed

Figure 2.5: Stereo diagram of NCS averaged difference maps.

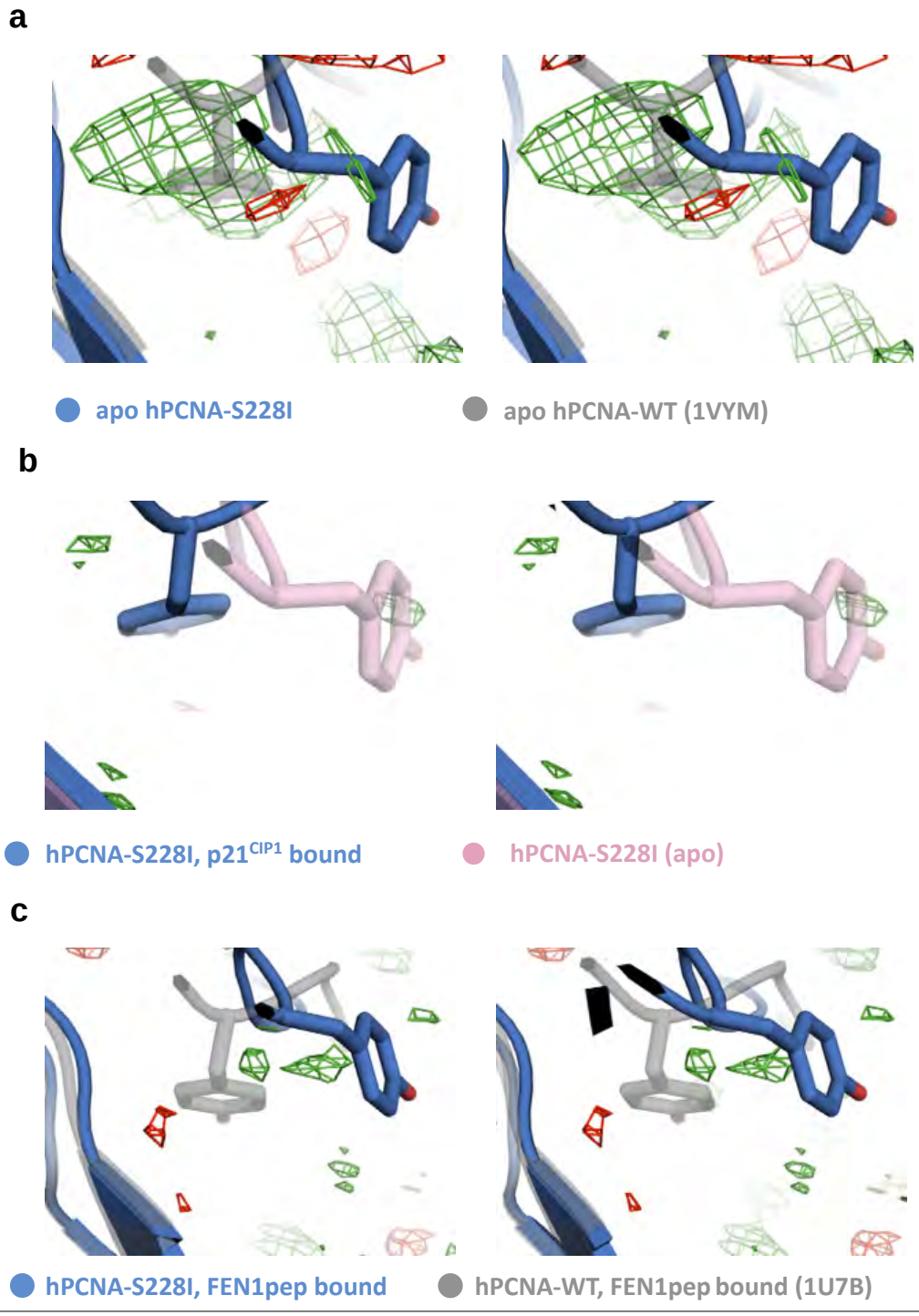
Green and red mesh represent positive and negative density, respectively.

A) NCS-averaged difference (F_o-F_c) map of apo hPCNA-S228I contoured at 3.5σ. The structure of WT-hPCNA (translucent gray) is superposed onto that of hPCNA-S228I. Note that the 'flipped-in' conformation of Tyr133 fits the positive density adjacent to the 'flipped-out' Tyr133 conformation that is modeled (blue).

B) NCS-averaged difference (F_o-F_c) map of hPCNA-S228I bound to the $p21^{CIP}$ PIP-box peptide contoured at 2.25σ. The structure of apo hPCNA-S228I (translucent purple) is superposed onto that of hPCNA-S228I:$p21^{CIP}$ (blue). Note that no significant positive density corresponds to the 'flipped-out' conformation.

C) NCS-averaged difference (F_o-F_c) map of hPCNA-S228I bound to FEN1pep contoured at 2.25σ. The structure of hPCNA-WT bound to FEN1pep (translucent gray) is superposed onto that of hPCNA-S228I:FEN1pep (blue). Note that no significant positive density corresponds to the 'flipped-in' conformation.

Figure 2.5: Stereo diagram of NCS averaged difference maps.



that the cavity in hPCNA-S228I is only ~33% of the volume observed in wild-type and the shape of the binding pocket is greatly altered (**Figure 2.6c&d**). Moreover, superposition of the PIP-box motif from RNaseH2B (61) into the hPCNA-S228I binding pocket reveals that the observed conformation is incompetent for PIP-box binding (**Figure 2.6e**). In fact, superposition of all known PIP-box motifs (58, 59, 61, 62, 64, 157) results in steric clashes. Given these results, we hypothesize that the S228I mutation disrupts PIP binding.

To test our hypothesis, we focused on the PCNA binding regions of three different PIPs: (1) $p21^{CIP}$, a cell cycle regulation protein(161), (2) the DNA repair protein RNaseH2B, a subunit of the RNaseH2 complex that degrades R-loops and single rNMPs incorporated into DNA during replication(162), and (3) Flap Endonuclease I (FEN1), which functions in both DNA replication (Okazaki fragment maturation) and DNA repair (base excision repair) (163). We investigated the binding of FEN1 and RNaseH2B because the previous SILAC experiments(35) suggested that these two PIPs had differential binding in response to the S228I mutation: FEN1 showed decreased binding to the disease variant, and RNaseH2B showed increased binding. Although the SILAC experiments suggest relative changes in binding, this method does not provide a quantitative measurement of binding affinity and is prone to artifacts(164). $p21^{CIP}$ was chosen for study because it is the tightest binding PIP-box sequence known(165) and

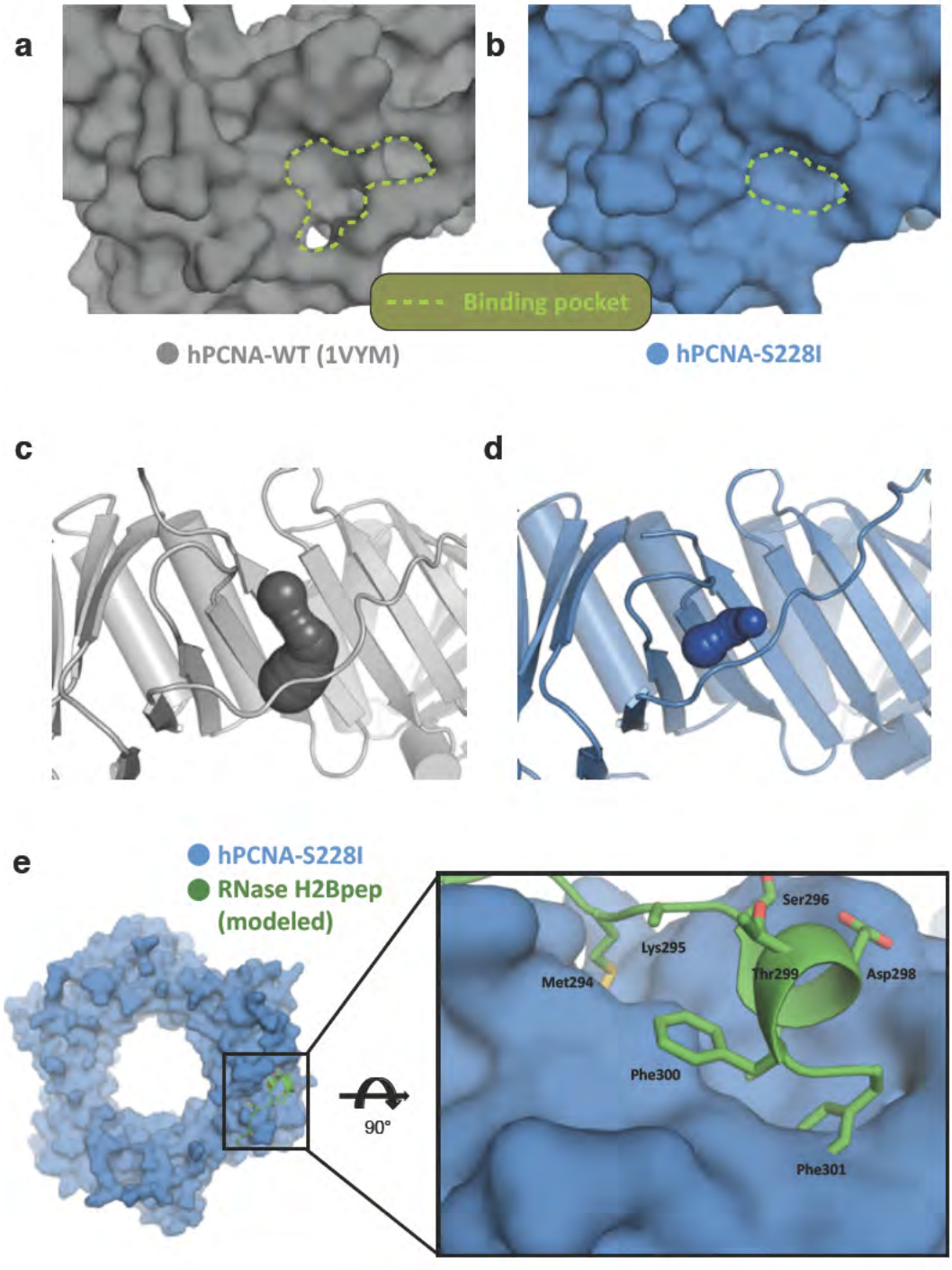
Figure 2.6: Disruption of the binding site in hPCNA-S228I.

A & B) The binding site surface of hPCNA-S228I is altered. (A) The surface of apo hPCNA-WT is shown from the top of the ring with the binding site for the second aromatic residue of the PIP-box motif outlined in green. (B) The surface of hPCNA-S228I is shown with the same binding site outlined in green.

C) & D) Constricted binding site volume in hPCNA-S228I. The volume of the binding pocket is shown as space-filling volume as calculated by Caver(160). The volume of the hPCNA-WT (panel C; gray) binding pocket is substantially larger than hPCNA-S228I (panel D; blue). View is shown from the side of the ring.

E) Binding pocket is incompatible with PIP-box binding. The structure of hPCNA bound to the PIP-box motif of RNase H2B is superposed onto the apo hPCNA-S228I structure (only RNase H2B peptide is shown). The constricted binding site would clash with the PIP-box, most notably at Phe301 of RNase H2B, which is the second aromatic residue of the PIP-box motif.

Figure 2.6: Disruption of the binding site in hPCNA-S228I.



has been useful for crystallographic (58) and binding studies (59, 93, 166).

We measured the binding affinity of these three peptides by isothermal titration calorimetry (ITC). We found that all three peptides lose affinity for S228I relative to hPCNA-WT (**Table 2.2**), contrary to what was predicted by SILAC(35). However, the magnitude of the effect is dependent on the particular PIP-box sequence. hPCNA-S228I exhibits only a subtle (~0.2 kcal/mol) loss in binding energy for the tight-binding $p21^{CIP}$ PIP-box peptide ($K_{d,wt}$ = 560 ± 140 nM, $K_{d,S228I}$ = 810 ± 20 nM) (**Figure 2.7a**). In contrast hPCNA-S228I has a much larger loss in affinity for the PIP-box peptides derived from RNaseH2B or FEN1. The FEN1 PIP-box peptide binds hPCNA-WT with a K_d of 9.7 ± 0.4 μM but has no measurable affinity by ITC for hPCNA-S228I (**Figure 2.7b**). Some heat evolves upon titration of the FEN1 PIP-box with hPCNA-S228I, illustrating that the enthalpy of binding of FEN1 peptide is greatly reduced for hPCNA-S228I. However, the small heat of binding is not sufficient to fit the data to obtain reliable estimates of the K_d or thermodynamics parameters of binding. Likewise, the PIP-box peptide from RNaseH2B binds hPCNA-WT with a K_d of 35 ± 1 μM, but there is no measurable binding with hPCNA-S228I (**Figure 2.7c**). In this case, there is negligible heat evolved upon titration of the RNaseH2B peptide with hPCNA-S228I above that of the heat of peptide dilution, suggesting significant impairment of binding. Thus, the S228I mutation has idiosyncratic effects on the binding of different PIP-box motifs.

Table 2.2: ITC Statistics

Protein	Ligand	K_d (μM)	ΔG (kcal mol^{-1})	ΔH (kcal mol^{-1})	$-T\Delta S$ (kcal mol^{-1})
hPCNA-WT	p21^{CIP1‡}	0.56 ± 0.14	-8.7 ± 0.2	-12.3 ± 2.0	3.7
hPCNA-S228I	p21$^{CIP1◇}$	0.87 ± 0.07	-8.4 ± 0.1	-11.4 ± 0.4	3.0
hPCNA-WT	FEN1pep‡	9.7 ± 0.4	-6.90 ± 0.02	-2.0 ± 0.2	-4.9
hPCNA-S228I	FEN1pep$^{◇}$	No fit	No fit	No fit	No fit
hPCNA-WT	RNase H2B‡	35 ± 1	-6.20 ± 0.02	-4.0 ± 0.5	-2.1
hPCNA-S228I	RNase H2B$^{◇}$	No fit	No fit	No fit	No fit

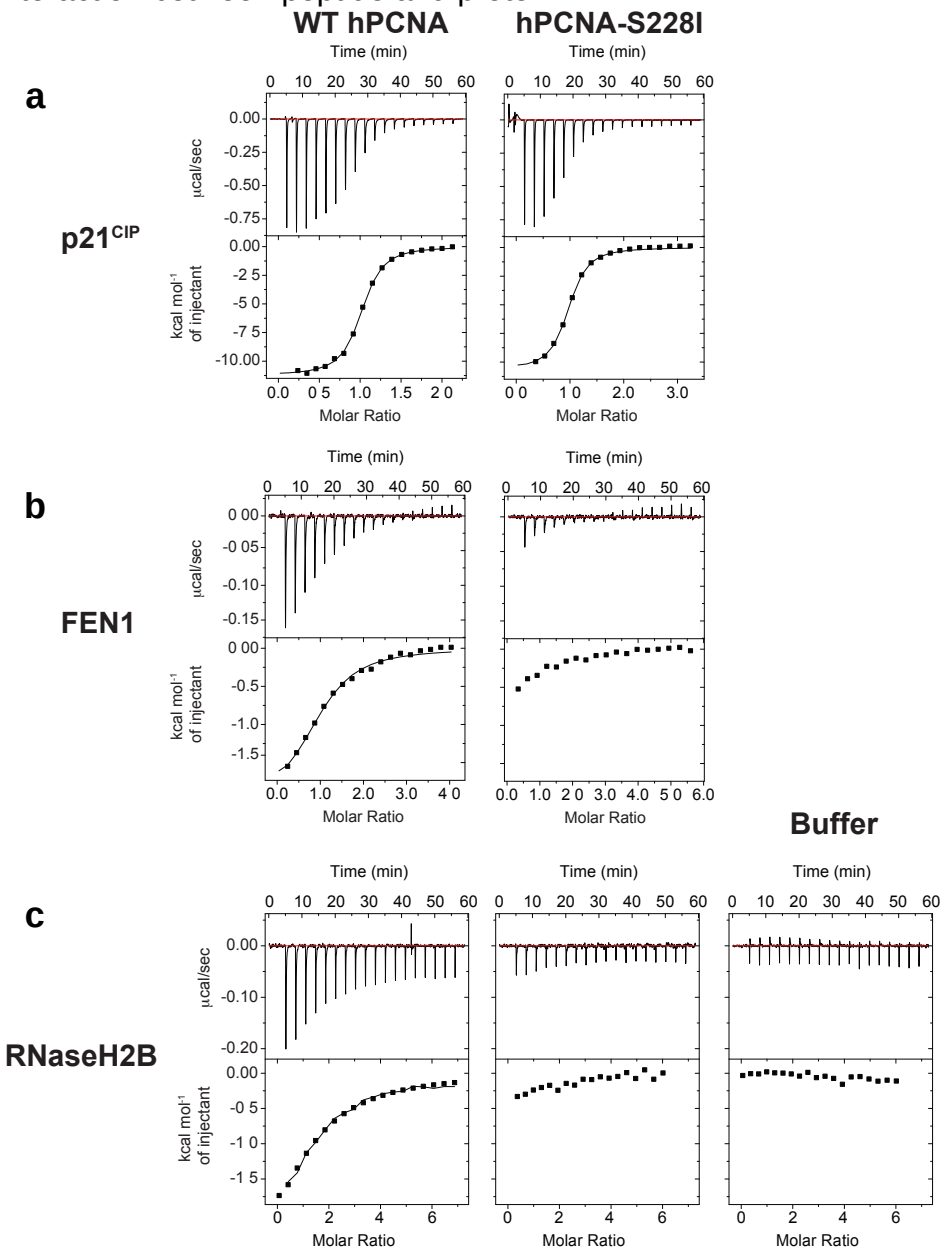
‡Three measurements $^{◇}$Four Measurements

Our ITC measurements allow us to investigate the enthalpic and entropic components of the PIP-box binding energy. We limit our comparison of the thermodynamic parameters of binding (WT compared to S228I) to just the p21^{CIP} peptide because the S228I mutant reduces the enthalpy of binding for FEN1 too much for reliable fitting and we observe no measurable evidence of binding for RNaseH2B to hPCNA-S228I. For both WT- and hPCNA-S228I, p21^{CIP} peptide binding is driven by enthalpy and entropy disfavors binding. Although within the error of our measurements, the enthalpy of binding is more negative (favoring binding) for hPCNA-WT than hPCNA-S228I ($\Delta\Delta H_{S228I\text{-}WT}$ = 1.5 kcal/mol). Likewise the entropy of binding is more positive (disfavoring binding) for hPCNA-WT than for hPCNA-S228I ($T\Delta\Delta S$ = 1.3 kcal/mol, for a net $\Delta\Delta G$ of 0.2 kcal/mol).

p21^{CIP} PIP-box binding reverts IDCL conformation to that seen for wild-type

Our binding studies indicate that the effects of the S228I mutation are complex and will depend on individual atomic interactions and forces. To identify the atomic interactions that allow for PIP-box binding to hPCNA-S228I, we have solved the 1.9-Å crystal structure of hPCNA-S228I bound to the p21^{CIP} PIP-box peptide (**Table 2.1**). The co-crystals of hPCNA-S228I and p21^{CIP} PIP-box peptide

Figure 2.7. Affects of the S228I mutation on binding energetics of select PIP-box motifs. ITC data for WT- and hPCNA-S228I with various PIP-box peptides. See Table 2.2 for results.
A) ITC measurements for WT- and hPCNA-S228I with $p21^{CIP}$ peptide.
B) ITC measurements for WT- and hPCNA-S228I with FEN1 peptide.
C) ITC measurements for WT- and hPCNA-S228I with RNaseH2B peptide. The third column shows titration of the RNaseH2B peptide into buffer, which gives the heat of peptide dilution. Note that the heat evolved upon titration with hPCNA-S228I is of the same scale as the heat of dilution, suggesting no significant interaction between peptide and protein.



contain three separate chains per asymmetric unit which are organized into an unusual arrangement in which each of the three hPCNA subunits in the asymmetric unit forms a separate three-fold symmetric ring within the context of the overall crystal (**Figure 2.8**). Thus there are three independent rings of hPCNA arranged in the crystal. Each ring is similar to the others with obvious density for the $p21^{CIP}$ peptide bound (**Figure 2.9a**). The IDCL conformation in hPCNA-S228I is substantially altered upon $p21^{CIP}$ peptide binding (**Figure 2.9b**). Even though isolated $p21^{CIP}$ is an intrinsically disordered protein(167), the core PIP-box motif forms a 3_{10}-helix upon binding PCNA, while the region from Ser153 to Phe159 of $p21^{CIP}$ forms an anti-parallel sheet with residues 120 through 126 of the IDCL. This conformation is essentially the same conformation as observed for hPCNA-WT (58) (**Figure 2.9c**). Importantly, Tyr133 of hPCNA reverts to the canonical conformation seen in all structures of hPCNA-WT, with the exception that the entire residue is shifted upwards by ~1Å away form Ile228 (**Figure 2.9d**). Moreover, $p21^{CIP}$ buries the same amount of area for both hPCNA-WT and hPCNA-S228I (2150 ± 10 $Å^2$ and 2147 ± 80 $Å^2$, respectively). The IDCL also becomes more ordered upon $p21^{CIP}$ binding with the IDCL B-factor higher by only ~10% than the overall protein B-factor for both wild-type hPCNA and the S228I mutant (**Figure 2.10a&b**).

To investigate why $p21^{CIP}$ PIP-box binding affinity is largely unaffected by the S228I mutation, we used the crystal structure to examine key interactions. A possible explanation for the small effect on $p21^{CIP}$ PIP-box affinity is that a

Figure 2.8: Packing in the hPCNA-S228I:p21^{CIP1} crystals.

The proteins in one asymmetric unit are shown as solid outlined cartoons, with each hPCNA chain displayed in a different color and p21^{CIP} shown in green. Symmetry-related hPCNA copies are shown as translucent 'ghosts' to illustrate how the rings are arranged within the context of the crystal.

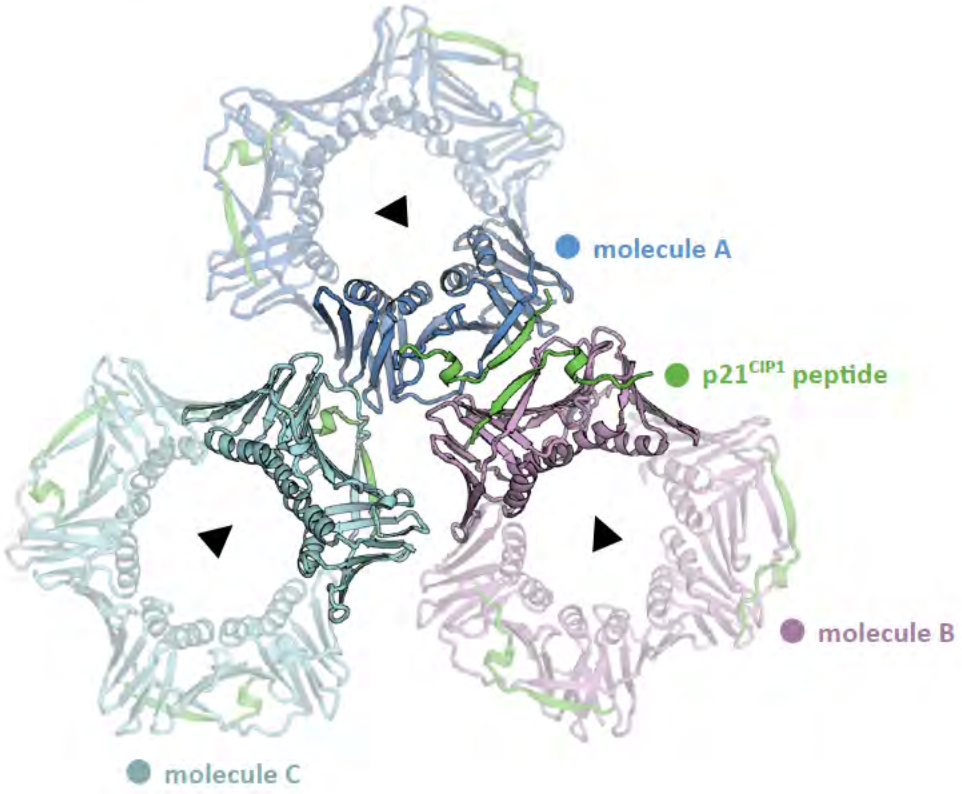
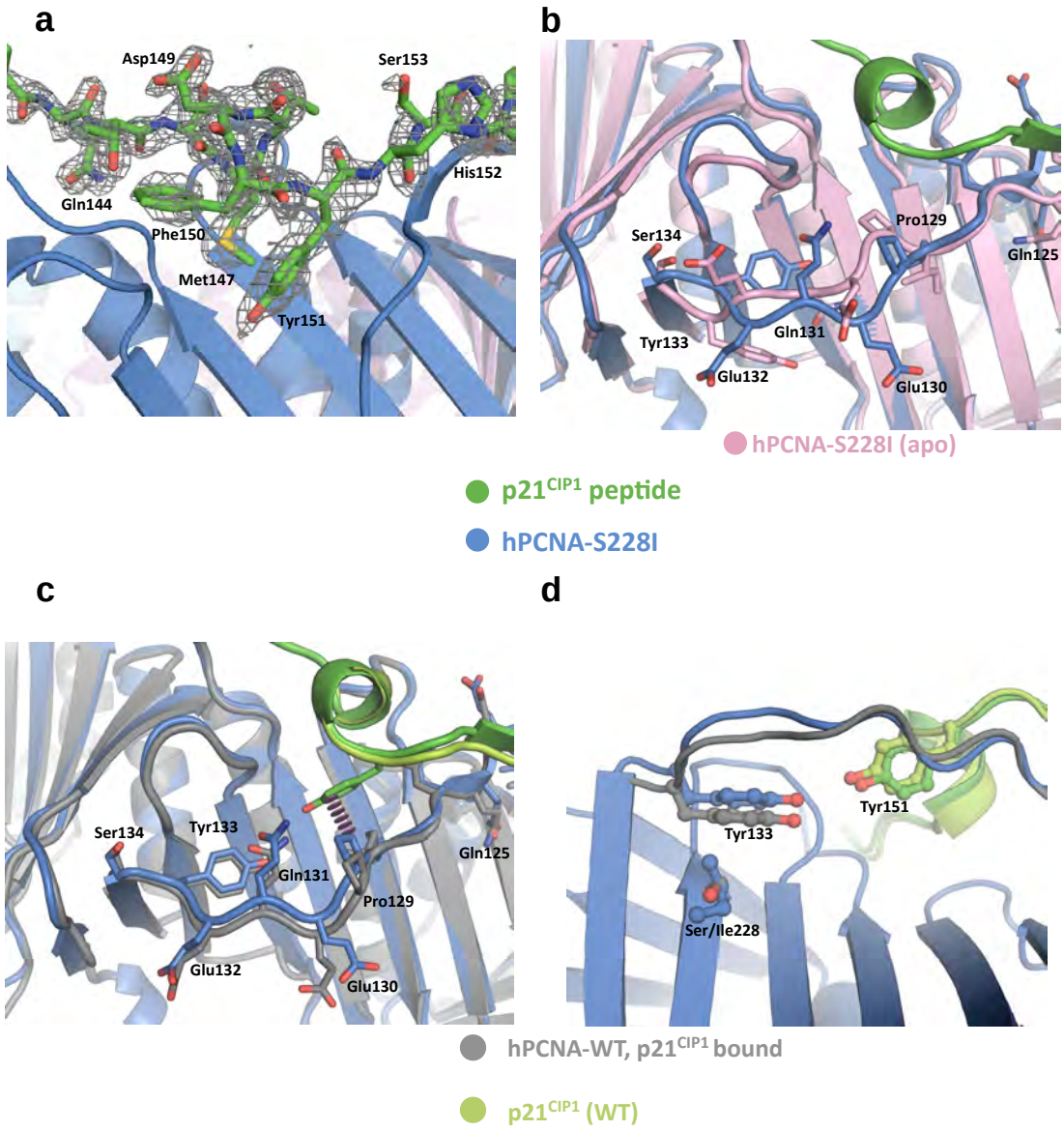


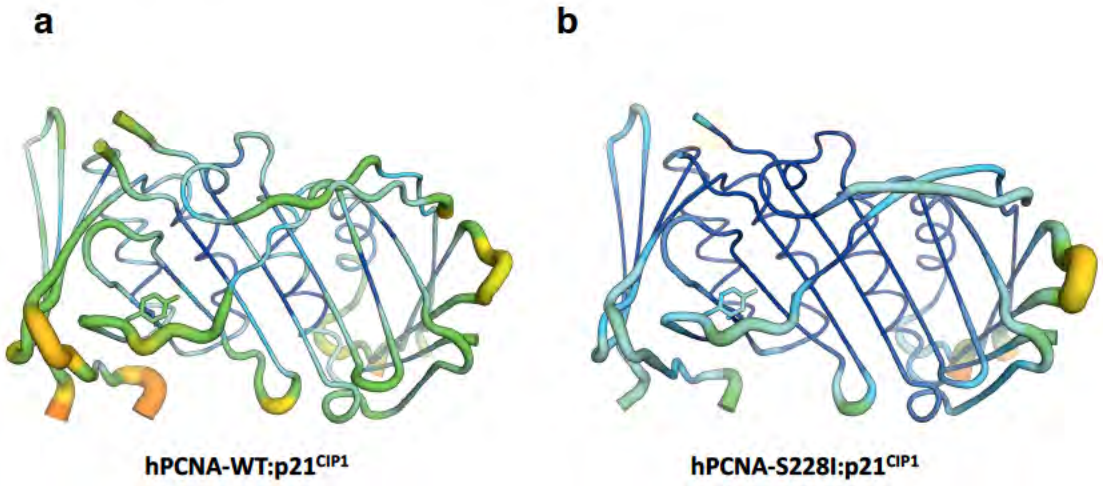
Figure 2.9: Structure of hPCNA-S228I bound to the p21^{CIP1} PIP-box motif.
A) p21^{CIP1} peptide is bound to hPCNA-S228I. Omit density (F_o-F_c) is contoured at 2.5 σ , which shows clear density for the peptide. p21^{CIP} PIP-box peptide is shown in green.
B) The hPCNA-S228I IDCL changes conformation upon p21^{CIP1} PIP-box binding. Superposition of apo and p21^{CIP1} PIP-box bound hPCNA-S228I (pink and blue, respectively) shows that the IDCL changes conformation substantially from the unbound state.
C) The hPCNA-S228I reverts to the canonical IDCL conformation upon p21^{CIP1} PIP-box binding. Superposition of p21^{CIP1} PIP-box bound WT- and hPCNA-S228I (gray and blue, respectively) indicates that all residues in the IDCL are in nearly the same conformation as in the canonical conformation. Hydrophobic interaction between Pro129 and Tyr151 shown with purple disks.
D) S228I mutation forces Tyr133 closer to the p21^{CIP1} peptide. The larger steric bulk of isoleucine 228 forces the Tyr133 sidechain upward toward the p21^{CIP1} peptide. Tyr133 packs tighter against Tyr151 of p21^{CIP1} than is observed in the hPCNA-WT structure (4.3 vs 4.6 Å).

Figure 2.9: Structure of hPCNA-S228I bound to the p21^{CIP1} PIP-box motif.



hydrogen bond between Tyr151 of $p21^{CIP}$ forms with Gln131 and Tyr133 of hPCNA, forcing the IDCL back into its canonical conformation. Tyr133 accounts for a significant amount of the binding energy of the $p21^{CIP}$ peptide and, for this reason, has been referred to as a molecular tether (165). We observe that the distance between Tyr133 of hPCNA and Tyr151 of $p21^{CIP}$ is shorter with the S228I mutant than with the wild-type (4.3 Å vs. 4.6 Å), suggesting tighter packing in this pocket. The close apposition of Tyr133 to the $p21^{CIP}$ peptide is due to the S228I mutation 'pushing' the sidechain of Tyr133 upward toward the ligand. In this way, S228I could 'assist' in regaining some of the binding energy that is lost due to rearrangement of the IDCL.

Figure 2.10: IDCL B-factors upon binding peptides. The thickness and color of the chain are scaled by the normalized B factor (low to high B factor = thin to thick, blue to red). A) & B) The IDCL relative B-factors are similar to the overall average for both WT- and hPCNA-S228I.



FEN1 PIP-box binding induces a novel conformation of the IDCL

To determine how the S228I mutation affects binding of other PIP-box motif sequences, we also crystallized a complex of the FEN1 PIP-box peptide (hereafter referred to as FEN1pep) with hPCNA-S228I. Although the binding thermodynamics of FEN1pep to hPCNA-S228I are greatly altered relative to hPCNA-WT (**Figure 2.7b**), at high peptide concentrations we were able to produce co-crystals whose growth was dependent on the presence of the peptide. The structure was solved to 2.1 Å resolution by molecular replacement and refined using twin law h,-h-k,-l (**Table 2.1**). In both copies of hPCNA in the asymmetric unit, FEN1pep is bound at high occupancy with clear density for nearly all residues in the peptide (**Figure 2.11a**).

The hPCNA-S228I:FEN1pep structure reveals a third and novel conformation of the IDCL. While the overall PCNA fold of the S228I variant is similar to the hPCNA-WT: FEN1pep structure (59) (C_α RMSD = 0.8 Å), the IDCL adopts a different conformation than has been observed in all other hPCNA structures, WT or mutant. A structural comparison of hPCNA-S228I:FEN1pep with either apo hPCNA-S228I or hPCNA-WT:FEN1pep reveals that the N-terminal portion of the IDCL (distal to the mutation; residues 120-125) adopts the canonical conformation as seen in structures of PCNA-WT (**Figure 2.11b&c**). However, the C-terminal portion (proximal to the mutation; residues 126-134) adopts a novel conformation that is distinct from the canonical conformation (**Figure 2.11b**). This conformation is somewhat similar to that seen in apo

hPCNA-S228I in that Tyr133 ‘flips-out’ into the non-canonical conformation; however, the backbone conformation is altered from that of the apo S228I form (**Figure 2.11c**). Thus, the IDCL architecture of hPCNA-S228I:FEN1pep represents a novel conformation.

Despite the large changes in the IDCL conformation in hPCNA-S228I, the interaction surface for FEN1pep is very similar to that in hPCNA-WT. The binding interface is rearranged in hPCNA-S228I such that the pocket for the central residue of the PIP-box motif (Leu340 of FEN1pep) is essentially identical to that seen in wild-type. The binding pocket for the second aromatic residue (Phe344 of FEN1pep) has a different conformation that closes more tightly to seal off the binding pocket (**Figure 2.11d&e**). However, this constriction does not alter the overall FEN1pep interaction surface area; FEN1pep buries a roughly equivalent surface area for both WT- and hPCNA-S228I (1380 $Å^2$ and 1390 $Å^2$, respectively). Moreover, the hydrogen bonding pattern with FEN1pep is similar in both S228I- and hPCNA-WT (**Table 2.3**). Thus, the interactions observed crystallographically are nearly identical between WT- and hPCNA-S228I with the exception that hPCNA-WT has a binding site that is largely prearranged whereas the hPCNA-S228I binding site must reorganize to accommodate FEN1pep. Despite the prearrangement of hPCNA-WT, the wild-type IDCL is significantly more dynamic than in hPCNA-S228I; the average IDCL B-factor is ~2.3-fold higher than the average for the entire hPCNA-WT protein, while the S228I IDCL is only ~1.3-fold higher than the protein average (**Figure 2.12**).

Figure 2.11: Structure of hPCNA-S228I bound to the FEN1 PIP-box motif.

A) FEN1 peptide is bound to hPCNA-S228I. Omit density (F_o-F_c) contoured at 2.5 σ shows clear density for the peptide. FEN1pep shown in tan color.

B) Comparison of WT- and hPCNA-S228I IDCL conformations upon FEN1pep binding. Superposition of FEN1pep bound WT- and hPCNA-S228I (gray and blue, respectively). The N-terminal half of the hPCNA-S228I IDCL reverts to the canonical IDCL conformation upon FEN1pep binding. However, the C-terminal region of the IDCL (in and around Tyr133) is distinct from the canonical conformation. Any residue that differs in C_α position by >2 Å is displayed with red arrows.

C) The entire hPCNA-S228I IDCL shifts upon FEN1pep binding. Superposition of apo and FEN1pep-bound hPCNA-S228I (pink and blue, respectively). Even though Tyr133 is in the 'flipped out' orientation, the conformation of the S228I IDCL is distinct from the apo state. Any residue that differs in C_α position by >2 Å is displayed with red arrows.

D) & E) Different shapes of the PIP binding groove for WT- and hPCNA-S228I upon FEN1pep binding. The molecular surface of WT- (panel D; gray) and hPCNA-S228I (panel E; blue) are displayed from above with FEN1pep shown as a translucent trace to allow visualization of the underlying pocket. The dotted oval indicates the major structural differences in the binding pockets of S228I and wild type. The hPCNA-S228I pocket is slightly more constricted but has no significant change in interaction surface area with FEN1pep.

Figure 2.11: Structure of hPCNA-S228I bound to the FEN1 PIP-box motif.

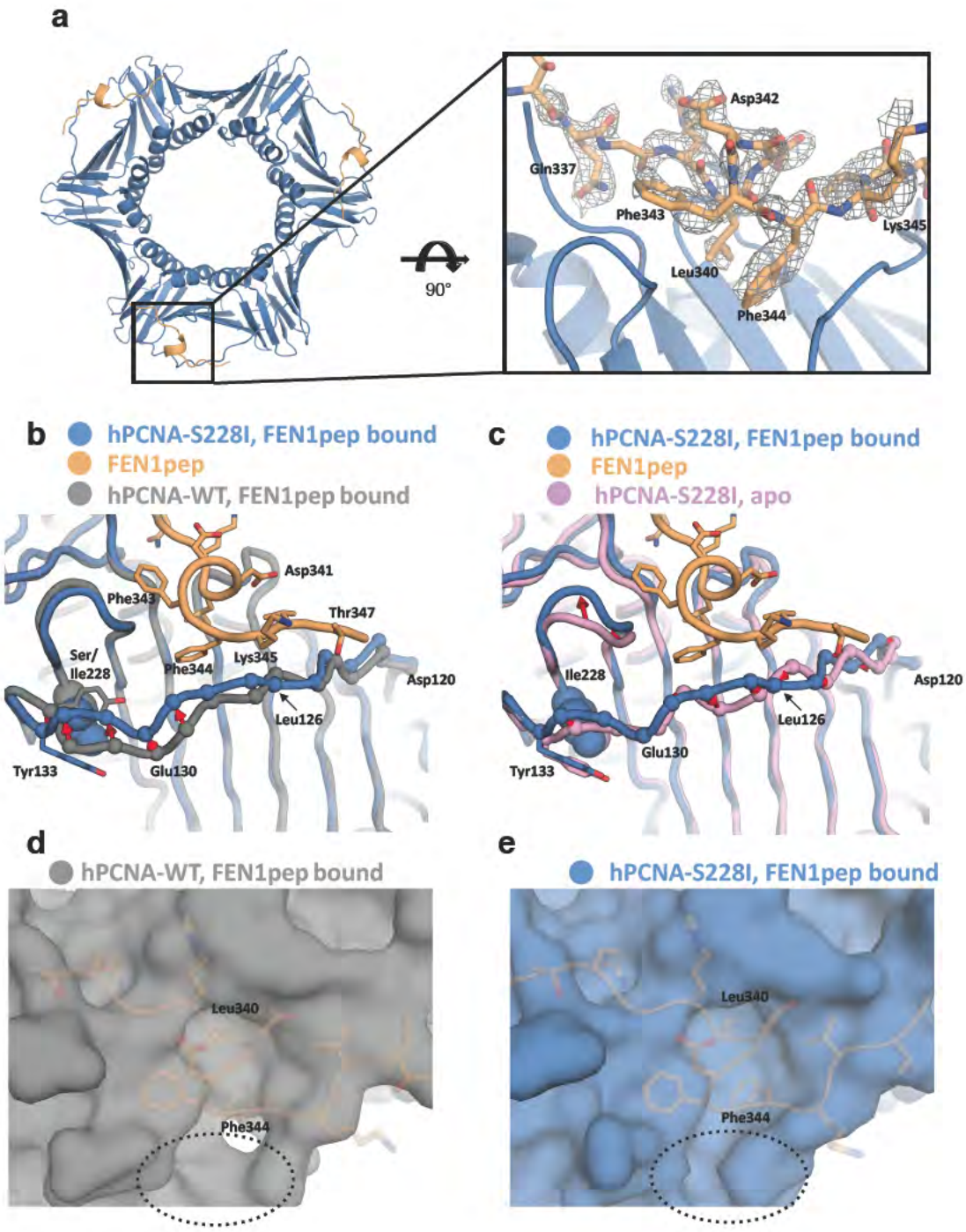


Table 2.3: Hydrogen bonding between hPCNA and FEN1pep

hPCNA-WT:FEN1pep (PDB: 1U7B)				
hPCNA residue	hPCNA Heavy Atom	Distance (Å)	FEN1pep Residue	FEN1pep Heavy Atom
HIS 44A	O	1.85	LEU 340B	N
PRO 253A	O	2.26	GLY 338B	N
ILE 255A	N	1.4	THR 336B	O
GLY 127A	N	2.3	LYS 345B	O
HIS 44A	ND1	2.64	ASP 341B	OD1
ALA 252A	O1	2.73	GLN 337B	NE2
SER 43A	O	2.83	ARG 339B	NH1
HIS 44A	ND1	2.64	ASP 341B	OD1

*Distances measured from donor hydrogen to acceptor atom

hPCNA-S228I:FEN1pep				
hPCNA residue	hPCNA Heavy Atom	Distance (Å)	FEN1pep Residue	FEN1pep Heavy Atom
HIS 44A	O	1.95	LEU 340C	N
GLN 125A	N	2.52	THR 347C	O
GLY 127A	O	1.86	LYS 345C	N
PRO 253A	O	2.04	GLY 338C	N
ILE 255A	N	1.79	THR 336C	O
GLN 125A	N	1.72	THR 347C	O
ASP 122A	OD2	3.16	GLY 348C	O
ALA 252A	O	2.9	GLN 337C	NE2
SER 43A	O	2.85	ARG 339C	NH1
HIS 44A	ND1	2.87	ASP 341C	OD1
HIS 44B	O	2.18	LEU 340D	N
GLY 127B	O	2.09	LYS 345D	N
PRO 253B	O	2.16	GLY 338D	N
ILE 255B	O	1.8	SER 335D	N
ILE 255B	N	1.67	THR 336D	O
HIS 44B	ND1	2.84	ASP 341D	OD
GLU 124B	OE1	2.72	THR 347D	O
SER 43B	O	2.68	ARG 339D	NH1
HIS 44B	ND1	2.84	ASP 341D	OD1

*Distances measured from donor hydrogen to acceptor atom

Discussion

Implications for PCNA conformation, dynamics and PIP binding

Our work illustrates that the disease-causing mutation S228I causes a major disruption in the PIP binding groove of hPCNA. This disruption is extensive, perturbing residues in the IDCL over 35 Å away from the site of the mutation. Because the pore of the hPCNA trimer is unaffected by the mutation, it is highly unlikely that DNA binding is altered by the mutation. (Although we did not test DNA-binding abilities of the mutant variant, we expect abolishing DNA-binding would be embryonic lethal.) Therefore, we propose that the mutation primarily exerts its effect by perturbing interactions with PCNA binding partners. The mutation causes a severe disruption of the PIP binding site (**Figure 2.6**) and alters PIP binding energetics (**Figure 2.7**), supporting this hypothesis. (Although known sites of post-translational modification in hPCNA are not disrupted by the mutation (87, 89, 168, 169), the mutation may exert its effect indirectly through

alterations of hPCNA post-translational modification, as enzymes that add or remove post-translational modifications often require a PIP-box motif for catalysis (89, 170).) The beta-branched side chain of Ile228 induces the rearrangement of Tyr133 and subsequent alteration of the IDCL, suggesting that beta-branched residues at this position are deleterious to hPCNA function. In support of this hypothesis, residue 228 is conserved to exclude beta-branched residues in known eukaryotic PCNA proteins (**Figure 2.13**).

Figure 2.12: IDCL B-factors upon binding FEN1pep. The thickness and color of the chain are scaled by the normalized B factor (low to high B factor = thin to thick, blue to red).
A) & B) Upon FEN1pep binding the relative B-factors for hPCNA-S228I are lower than those of hPCNA-WT.

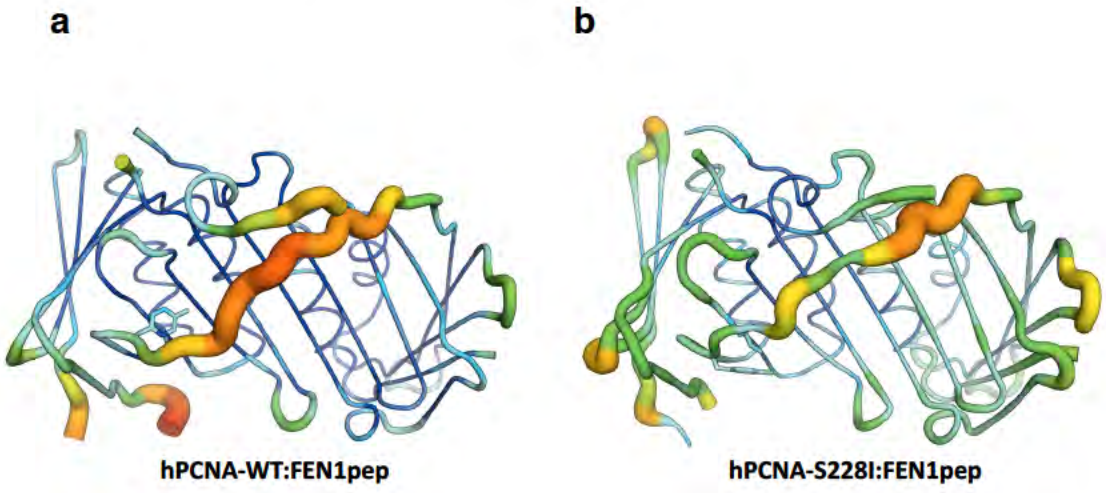


Figure 2.13: Conservation of key residues in PCNA.
Logo diagrams show conservation of residues within the IDCL (A) and near the site of mutation (B). Residue numbering is shown for human PCNA.
(A) Position 133 (outlined in cyan) is conserved to be an aromatic residue with a strong preference for tyrosine. Residues colored according to sidechain characteristics (black = hydrophobic, green = polar uncharged, red = acidic, blue = basic).
(B) Position 228 (outlined in cyan) is conserved to exclude beta branched residues. Beta branched residues are colored purple while all others are black.

How does the S228I mutation affect the energetics of PIP binding? The apo hPCNA-S228I structure reveals that the PIP binding groove is in a conformation incompatible with PIP binding (**Figure 2.6e**). Therefore, the IDCL must undergo an energetically unfavorable rearrangement to allow for PIP binding. Because we observe significant density in the apo S228I structure suggesting that Tyr133 can adopt the canonical conformation, we conclude that the IDCL oscillates between the two conformations with a relatively low energy gap (on the order of 1-2 K_bT). If the decrease in PIP-box binding affinity were simply the cost of rearranging the IDCL to assume a conformation compatible with PIP binding, we would predict that all PIPs would be equally disrupted. However, we observe idiosyncratic disruption of PIP binding affinity (**Figure 2.7**) and distinct conformations of hPCNA-S228I upon binding $p21^{CIP}$ and FEN1pep

(**Figures 2.9&2.11**). Therefore, the structural and sequence factors that control PIP binding in hPCNA-S228I are more complex than simple reversion of the IDCL conformation.

Our structural and thermodynamic investigation of p21, FEN1 and RNaseH2B binding highlights several factors that determine how PIPs are selectively perturbed by the S228I mutation. First, we observe a correlation between the overall binding affinity of the binding partner and the extent of its disruption; the high affinity PIP-box from $p21^{CIP}$ is the least disrupted while the low affinity interaction with the RNaseH2B PIP-box is the most disrupted. While high affinity binding by itself cannot prevent PIPs from being affected by the S228I mutation, work described here and elsewhere shows that high affinity PIP binding is associated with a larger interaction surface area between the PIP and PCNA (59, 61). A larger surface area provides more opportunity for new interactions between PCNA and the PIP to compensate for the energetic cost of IDCL reorganization. With this concept in mind, we will more carefully examine the binding features of $p21^{CIP}$ and FEN1pep below.

Why is $p21^{CIP}$ binding largely unaffected by the S228I mutation? We ascribe the minor reduction in $p21^{CIP}$ binding affinity to two effects: (1) the ability of the IDCL to revert to the canonical conformation, and (2) additional packing interactions between the IDCL and the $p21^{CIP}$ PIP-box peptide. Although reversion of the IDCL to the canonical conformation allows PCNA to maintain many interactions with $p21^{CIP}$ peptide (**Figure 2.14**), we observe slightly more

hydrogen bonds between $p21^{CIP}$ and hPCNA-WT than with hPCNA-S228I (**Table 2.4**). The small decrease in overall hydrogen bonding of S228I with $p21^{CIP}$ is predicted to reduce the enthalpy of binding, which must be compensated by the entropy of binding to keep the overall affinity near iso-energetic. The altered dynamics of S228I IDCL potentially contributes to this compensation. While our data are consistent with an entropy-enthalpy compensation mechanism, the measurement errors for the thermodynamic parameters are too large for a definitive conclusion. The second factor limiting the deleterious effects of the S228I mutation on $p21^{CIP}$ binding is a new interaction that is absent or weak in hPCNA-WT; Tyr133 shifts to accommodate the clash with Ile228 and in doing so packs more tightly against Tyr151 of $p21^{CIP}$. This augmented interaction may be especially important because Tyr151 of $p21^{CIP}$ is regarded as a 'molecular tether', critical for $p21^{CIP}$ affinity (165). We hypothesize that strengthening this interaction is a source of binding energy compensation for $p21^{CIP}$ with hPCNA-S228I. Tyr151 of $p21^{CIP}$ lies at the second aromatic position within the conserved PIP-box consensus sequence QxxØxxϑϑ, and is the only tyrosine at this position in the PIP-box in all known human PIP-box sequences (171). Therefore this proposed mechanism for binding energy compensation is likely to be unique to $p21^{CIP}$.

Table 2.4: Hydrogen bonding between hPCNA and p21^{CIP1}

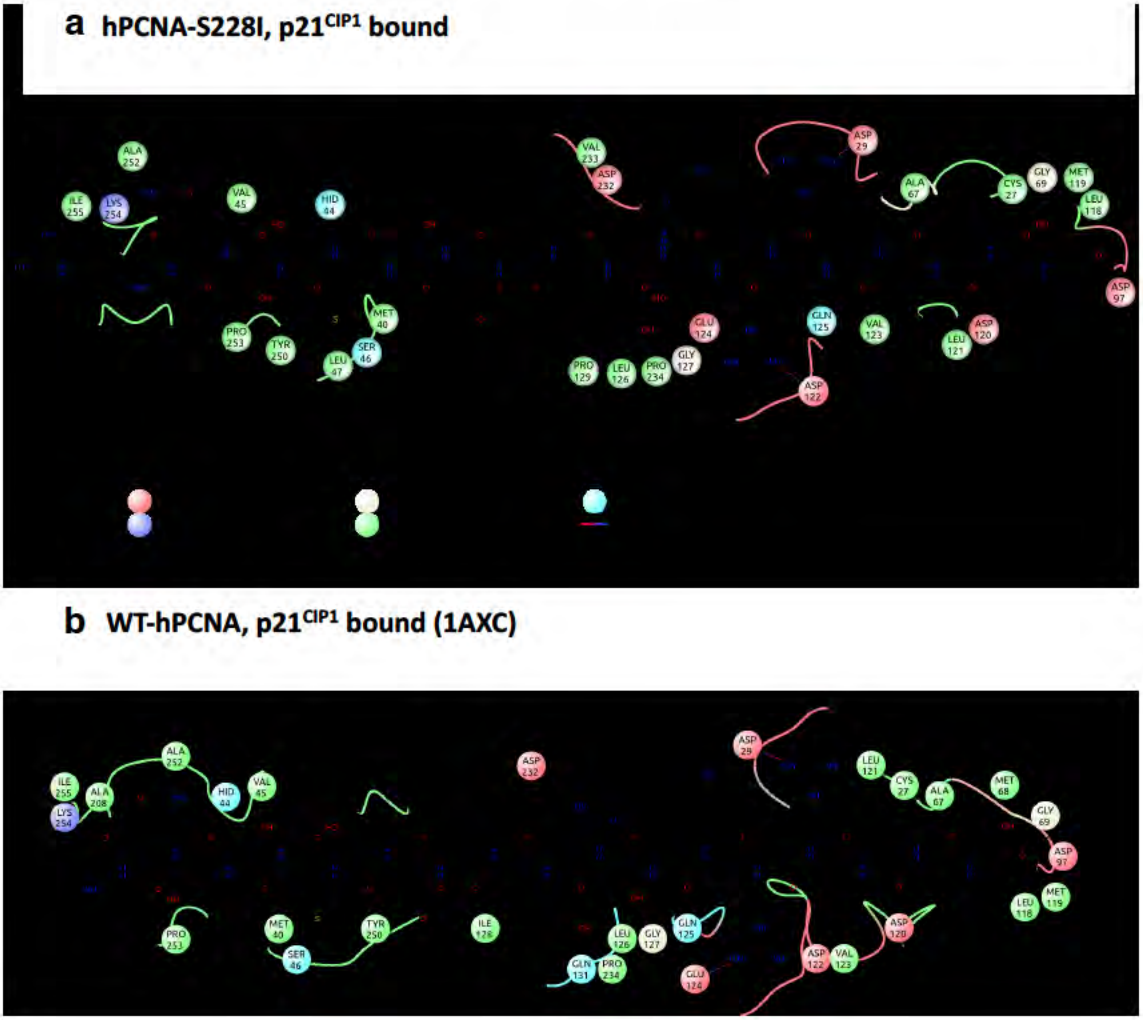
hPCNA-WT:p21^{CIP1} (PDB: 1AXC)				
hPCNA residue	hPCNA Heavy Atom	Distance (Å)*	p21^{CIP1} Residue	p21^{CIP1} Heavy Atom
HIS 44A	O	1.98	MET 147B	N
MET 119A	O	3.17	SER 160B	OG
LEU 121A	O	1.79	ILE 158B	N
VAL 123A	O	1.77	ARG 156B	N
GLN 125A	O	1.9	LYS 154B	N
PRO 253A	O	2.45	THR 145B	N
ILE 255A	N	1.68	ARG 143B	O
GLY 127A	N	1.74	HIS 152B	O
GLN 125A	N	2.08	LYS 154B	O
VAL 123A	N	2.08	ARG 156B	O
LEU 121A	N	2.0	ILE 158B	O
GLY 69A	N	2.77	SER 160B	OG
ASP 97A	OD1	2.72	SER 160B	O
ASP 120A	OD1	3.44	SER 160B	O
GLN 131A	NE2	2.59	TYR 151B	OH
ALA 252A	O	3.0	GLN 144B	NE2
PRO 253A	O	2.56	THR 145B	OG1
GLU 124A	OE1	2.97	ARG 155B	NH2
ASP 122A	O	3.16	ARG 155B	NH1
GLN 125A	OE1	3.13	ARG 156B	NH1
ASP 29A	OD1	2.88	ARG 156B	NH2
HIS 44C	O	1.94	MET 147D	N
LEU 121C	O	1.76	ILE 158D	N
VAL 123C	O	1.71	ARG 156D	N
GLN 125C	O	1.76	LYS 154D	N
PRO 253C	O	2.45	THR 145D	N
ILE 255C	N	1.78	ARG 143D	O
GLY 127C	N	1.87	HIS 152D	O
GLN 125C	N	1.83	LYS 154D	O
VAL 123C	N	2.14	ARG 156D	O
LEU 121C	N	1.94	ILE 158D	O
GLY 69C	N	2.86	SER 160D	OG
ASP 97C	OD1	2.5	SER 160D	O
ASP 122C	OD1	3.33	ARG 156D	O
GLN 131C	NE2	2.76	TYR 151D	OH
ALA 252C	O	2.92	GLN 144D	NE2
PRO 253C	O	2.69	THR 145D	OG1
GLU 124C	OE2	3.19	SER 153D	OG
ASP 122C	O	3.2	ARG 155D	NH1
GLU 124C	OE1	2.72	ARG 155D	NH2
ASP 29C	OD1	2.95	ARG 156D	NH2
ASP 29C	OD1	3.17	ARG 156D	NH1
ASP 120C	OD1	2.86	SER 160D	N
HIS 44E	O	1.93	MET 147F	N
LEU 121E	O	1.8	ILE 158F	N
VAL 123E	O	1.69	ARG 156F	N
GLN 125E	O	1.82	LYS 154F	N
PRO 253E	O	2.52	THR 145F	N
ILE 255E	N	1.98	ARG 143F	O
GLY 127E	N	1.76	HIS 152F	O
GLN 125E	N	1.9	LYS 154F	O
VAL 123E	N	2.05	ARG 156F	O
LEU 121E	N	2.21	ILE 158F	O
GLY 69E	N	3.11	SER 160F	OG
ASP 97E	OD1	3.12	SER 160F	O
ASP 120E	OD1	3.13	SER 160F	O
GLN 131E	NE2	2.72	TYR 151F	OH
ALA 252E	O	3.05	GLN 144F	NE2
PRO 253E	O	2.69	THR 145F	OG1
GLU 124E	OE2	3.49	SER 153F	OG
GLU 124E	OE1	2.87	ARG 155F	NH2
ASP 122E	O	2.99	ARG 155F	NH1
ASP 29E	OD1	2.8	ARG 156F	NH2
ASP 29E	OD2	3.11	ARG 156F	NH1
ASP 120E	OD1	3.06	SER 160F	N

*Distances measured from donor hydrogen to acceptor atom

hPCNA-S228I:p21^{CIP1}				
hPCNA residue	hPCNA Heavy Atom	Distance (Å)*	p21CIP1 Residue	p21CIP1 Heavy Atom
HIS 44A	O	2.03	MET 147D	N
LEU 121A	O	2.24	ILE 158D	N
VAL 123A	O	2.14	ARG 156D	N
GLN 125A	O	2.03	LYS 154D	N
PRO 253A	O	2.03	THR 145D	N
ILE 255A	N	1.55	ARG 143D	O
GLY 127A	N	1.98	HIS 152D	O
GLN 125A	N	2.15	LYS 154D	O
VAL 123A	N	2.15	ARG 156D	O
LEU 121A	N	2.2	ILE 158D	O
ASP 97A	OD1	2.83	SER 160D	O
ALA 252A	O	2.5	GLN 144D	NE2
PRO 253A	O	2.67	THR 145D	OG1
ASP 122A	OD2	3.3	ARG 155D	NH2
ASP 29A	OD1	2.48	ARG 156D	NH2
ASP 29A	OD2	2.98	ARG 156D	NH1
ASP 120A	OD1	2.19	SER 160D	N
HIS 44B	O	1.6	MET 147E	N
LEU 121B	O	2.15	ILE 158E	N
VAL 123B	O	1.9	ARG 156E	N
GLN 125B	O	1.76	LYS 154E	N
PRO 253B	O	2.38	THR 145E	N
ILE 255B	N	1.79	ARG 143E	O
GLY 127B	N	1.82	HIS 152E	O
GLN 125B	N	1.68	LYS 154E	O
VAL 123B	N	2.24	ARG 156E	O
LEU 121B	N	1.99	ILE 158E	O
ASP 97B	OD1	3.03	SER 160E	O
GLU 124B	OE2	2.44	SER 153E	OG
GLN 131B	NE2	3.26	TYR 151E	OH
LYS 254B	NZ	3.43	GLN 144E	OE1
ALA 252B	O	2.94	GLN 144E	NE2
PRO 253B	O	2.58	THR 145E	OG1
GLU 124B	OE2	2.44	SER 153E	OG
ASP 122B	OD2	3.12	ARG 155E	NH2
ASP 122B	OD1	3.45	ARG 155y	NE
ASP 29B	OD2	3.18	ARG 156y	NH1
ASP 29B	OD1	3.14	ARG 156y	NH2
HIS 44C	O	1.8	MET 147F	N
MET 119C	O	2.74	SER 160F	OG
LEU 121C	O	1.94	ILE 158F	N
VAL 123C	O	1.56	ARG 156F	N
GLN 125C	O	2.03	LYS 154F	N
PRO 253C	O	2.16	THR 145F	N
ILE 255C	N	1.28	ARG 143F	O
GLY 127C	N	2.43	HIS 152F	O
GLN 125C	N	2.15	LYS 154F	O
VAL 123C	N	2.02	ARG 156F	O
LEU 121C	N	1.8	ILE 158F	O
ASP120C	OD1	2.72	SER160F	O
ASP122C	OD1	3.22	ARG156F	O
GLU124C	OE2	2.2	SER153F	OG
PRO253C	N	3.3	THR145F	OG1
ALA252C	O	3.05	GLN144F	NE2
PRO253C	O	2.4	THR145F	OG1
GLU124C	OE2	2.2	SER153F	OG
ASP29C	OD2	3.24	ARG156F	NH1
ASP120C	OD1	2.22	SER160F	N

*Distances measured from donor hydrogen to acceptor atom

Figure 2.14: Schematic of the peptide interactions from p21^{CIP1} with PCNA. a) hPCNA-S228I and b) hPCNA-WT (PDB 1AXC) (58). Protein residues are shown as circles, interacting with the peptide, which has all bonds shown (created using Maestro ([Schrödinger Release 2015-3: Maestro, version 10.3, Schrödinger, LLC, New York, NY, 2015.]), with a cutoff of 4Å). Several residues within the peptide are labeled for clarity (black arrows).



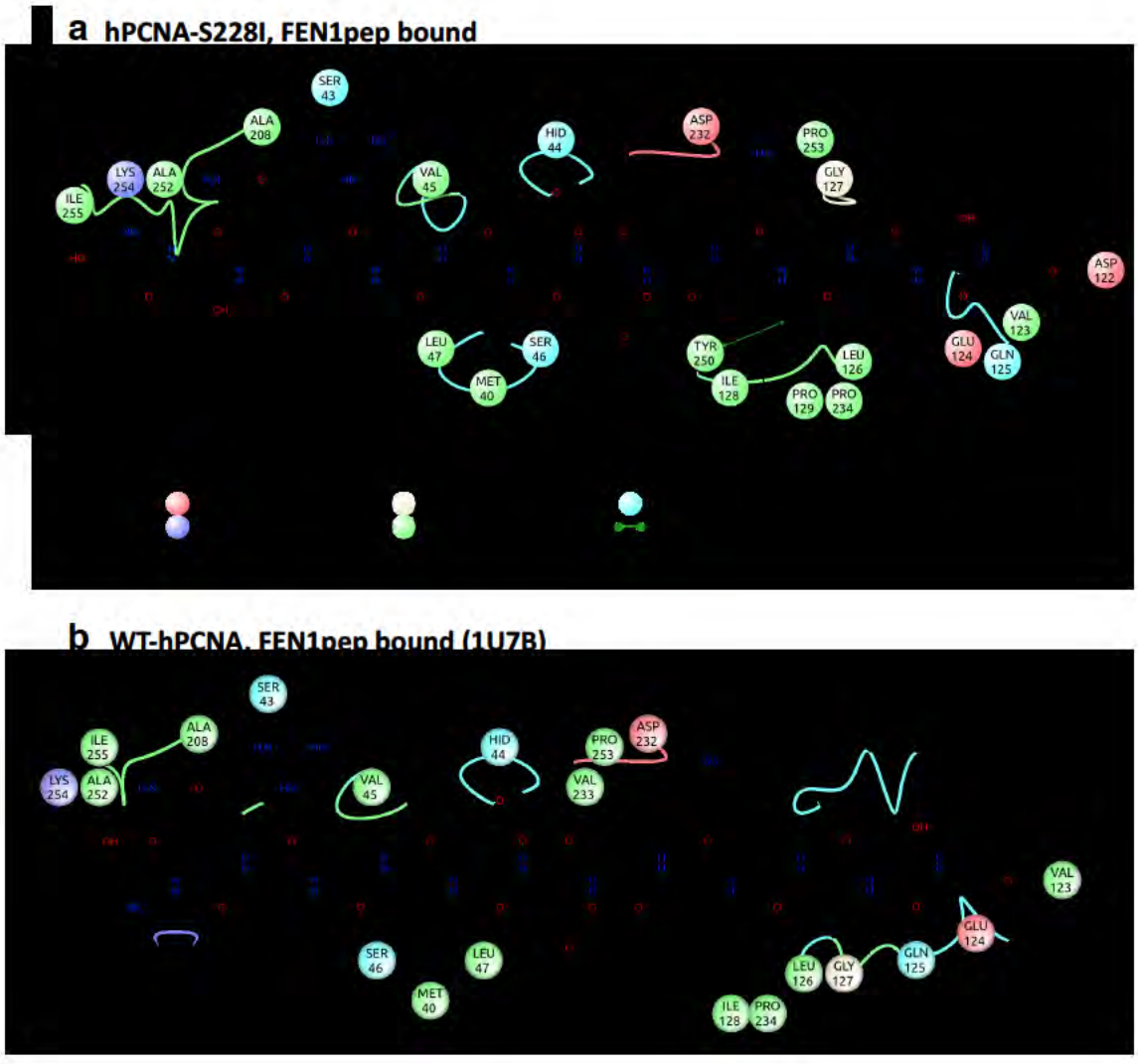
The forces driving interaction between FEN1pep and hPCNA-S228I are less clear. Whereas FEN1pep binding to hPCNA-WT evolves adequate heat for curve fitting in ITC, the enthalpy of binding is severely reduced for the S228I variant such that we cannot confidently determine the affinity. However, our structural data hints as to how the mutated binding site accommodates FEN1pep. The hPCNA-S228I:FEN1pep complex structure shows that the IDCL adopts a third conformation that is essentially a hybrid of the canonical and apo-S228I conformations, where Tyr133 adopts the “flipped out” conformation and the N-terminal region forming the canonical conformation. Although the entire IDCL shifts back to the canonical conformation upon $p21^{CIP}$ binding to hPCNA-S228I, only the regions with significant contact to FEN1pep adopt the canonical conformation; the rest forms a novel conformation. Therefore, we propose that the total surface area of the IDCL interaction with PIP-box peptide determines which conformation is adopted. Regardless of the large differences in IDCL conformation between WT- and hPCNA-S228I, the PIP binding sites appear to make comparable interactions with FEN1pep as measured by either buried surface area or number of hydrogen bonds formed (**Figure 2.15** and **Table 2.3**).

Our results indicate IDCL flexibility and dynamics control the binding of PIPs to hPCNA-S228I. We observe three different IDCL conformations indicating that the IDCL is dynamic and pliable in hPCNA-S228I. The apo form predominantly adopts a novel conformation with Tyr133 flipped away from the PIP binding pocket, while also populating the canonical Tyr133 conformation to a lesser

extent (**Figure 2.5a**). In contrast, $p21^{CIP}$ binding induces the IDCL to assume the canonical conformation at or near 100% occupancy (**Figure 2.5b**). FEN1pep binding induces a third conformation with Tyr133 fully in the non-canonical conformation (**Figure 2.5c**). Thus, the IDCL is energetically balanced between the canonical and non-canonical conformations of Tyr133; peptide binding can tip this balance in favor of either one or the other state. The inherent conformational flexibility of the IDCL allows it to wrap around different PIP-box motifs in different conformations in order to recover interaction surface area with the peptide. Thus, hPCNA-WT uses a classical 'lock-and-key' mechanism (172) in which the IDCL is predominantly prearranged to accept peptides with minimal conformational change (58, 59, 64), while hPCNA-S228I undergoes substantial backbone rearrangements to form the PIP-box binding pocket in an 'induced fit' mechanism (173). Additionally, our ITC results point to a larger entropic component to PIP binding in hPCNA-S228I, which again argues for a role of IDCL flexibility in PIP binding. Although the reduced enthalpy of binding for both FEN1pep and RNasH2pep makes binding affinity measurements unreliable or impossible, we propose that an entropy compensation mechanism for FEN1pep allows for productive binding at the high concentrations used in co-crystallization experiments. This potential entropic contribution to binding again points to flexibility and dynamics as playing an increased role in binding to hPCNA-S228I.

Figure 2.15 Schematic of the peptide interactions from FEN1pep with PCNA.

a) hPCNA-S228I and b) hPCNA-WT (PDB 1U7B) (59). Protein residues are shown as circles, interacting with the peptide, which has all bonds shown (created using Maestro ([Schrödinger Release 2015-3: Maestro, version 10.3, Schrödinger, LLC, New York, NY, 2015.]), with a cutoff of 4Å). Several residues within the peptide are labeled for clarity (black arrows).



Could the conformations of the variant reflect the flexibility of the IDCL in hPCNA-WT? Several lines of evidence indicate that the wild-type IDCL is quite flexible and undergoes regular dynamic transitions between different conformations. Crystallographic studies suggest that the wild-type IDCL is dynamic because B-factors in the IDCL are nearly two-fold higher than the average for the whole protein (64). Remarkably, we observe that the normalized crystallographic B factors for the hPCNA-S228I IDCL are lower than those for hPCNA-WT (**Figure 2.2d&e**). This result suggests that the IDCL is actually less mobile in hPCNA-S228I than in WT. Despite this fact, we observe multiple conformations in hPCNA-S228I, demonstrating significant flexibility for this region. Therefore, we hypothesize that the IDCL in hPCNA-WT is especially flexible, but this flexibility has been hidden in static x-ray crystal structures. In support of this idea, recent NMR studies of hPCNA-WT demonstrated that the IDCL exhibits lower order parameters for the IDCL relative to the average for the protein, indicating substantial dynamics on the ps-ns timescale (62, 174). Finally, molecular dynamics simulations of hPCNA, as well as those from budding yeast and the archaebacterium *Pyroccous furiosus*, directly show that the IDCL readily undergoes transitions to other conformational states (175) which may have important implications for controlling binding of PIPs (176). These results suggest the novel conformations that we observe crystallographically in hPCNA-S228I may be accessible for hPCNA-WT. We hypothesize that alternative conformations of the IDCL in hPCNA-WT contribute to the specificity of PIP

binding and could provide a mechanism for regulation. This hypothesis will be the focus of future studies.

Several unaddressed questions remain for how PIP binding affinity is affected by the S228I mutation. First, is the binding of APIM-containing PIPs perturbed by mutation? APIMs are found in only higher eukaryotic organisms (73) and have been suggested to play roles in DNA repair and damage response (83, 177) as well as a somewhat nebulous role in PCNA-directed kinase signaling and apoptosis in neutrophils (178, 179). Although the binding site for APIM sequences has not been identified, it has been suggested based on mutagenesis studies that the APIM and PIP-box motifs share the same binding site on PCNA (179). If this is indeed the case, APIM binding would be perturbed by the S228I mutation. Second, the structure of apo hPCNA-S228I illustrates that the IDCL would clash with the second aromatic residue in the PIP-box motif. All peptides tested here have an aromatic residue at this position of the motif, but many PIP-box sequences do not have an aromatic residue at this position (171). How does the S228I mutation affect binding of these motifs, if at all? Third, how does DNA affect binding of PIPs to hPCNA-S228I? In yeast PCNA, IDCL mutants that lose affinity for FEN1 in solution can partially restore binding when PCNA is bound to DNA (180). Thus, it remains unclear if DNA can partially rescue binding of FEN1 and/or RNaseH2B when it is loaded onto DNA. These questions will be the focus of future studies.

Implications for disease

Our results have important implications for the disease caused by the hypomorphic S228I mutation. The surprisingly strong alteration of the IDCL in hPCNA-S228I results in large deformation of the binding pocket relative to wild-type. Clearly, this deformation does not completely abrogate binding of all PIPs to the mutant PCNA because a complete loss of PIP binding would be expected to be fatal; after all, removal of just one PIP-box interaction (from FEN1) results in an embryonic lethal phenotype (181). Our results suggest that some PIPs have large changes in binding affinity while others are relatively unperturbed, with a possible correlation between overall binding affinity and the extent of IDCL perturbation by the S228I mutation. Our data point to the malleability of the IDCL as an important factor in mediating binding to certain PIPs. Despite the insights gleaned from our structural study, it will be difficult to elucidate which PIPs are important for determining the phenotypic spectrum of the hPCNA-S228I disorder because of the large diversity and number of PIPs. To tease apart the cellular and molecular basis for the hPCNA-S228I disorder will require careful *in vivo* studies coupled with *in vitro* biochemical, biophysical and structural characterization of binding.

The phenotype of the hPCNA-S228I disorder is reminiscent of defects in DNA repair or DNA damage response. In patient-derived cells, defects in nucleotide excision repair have been observed while DNA replication appears to be normal (35). Therefore the S228I mutation seems to affect the DNA repair and

damage response pathways more than DNA replication. Analogous to the hPCNA-S228I disorder, the Kolodner group identified two separate mutants of yeast PCNA that disrupt mismatch repair but do not seem to affect DNA replication (182). However, these two mutations do not affect the PIP binding pocket, but instead interfere with the PCNA inter-subunit interactions and the PCNA-DNA interaction surface, respectively. Thus the sites of the mutations suggest that these mutations cause PCNA to dissociate from DNA prematurely, possibly before mismatch repair proteins can act. The S228I mutation most likely mediates its affect in a different manner: by disrupting binding partners rather than disrupting interaction with DNA. Does the S228I mutation preferentially disturb interactions with PIPs involved in DNA repair? Or is DNA repair simply more sensitive to disruption than other PCNA-mediated pathways? These questions will be the focus of future studies.

Our studies with FEN1pep reinforce the proposal by Baple *et al.* (35) that the PCNA-S228I disorder is at least partially caused by defects in NER. Previous pull-down experiments indicate that the human xeroderma pigmentosum XPG repair endonuclease exhibits reduced binding with hPCNA-S228I (35). Mutations in XPG cause xeroderma pigmentosum, a DNA repair disease that shares many symptoms with the hPCNA-S228I disorder, among them heightened UV sensitivity. Based on the high level of sequence similarity between the PIP-box of XPG and FEN1 (84), we hypothesize a similar binding defect with XPG as we see with FEN1pep, which results in the decreased efficiency of NER observed in

patient-derived cells (35).

In addition to a disruption in NER, our observed lack of interaction with RNaseH2B raises the possibility that defects in ribonucleotide excision repair (RER) could be a major cause of disorder. RNaseH2B is a subunit of the RNaseH2 enzyme that carries out RER (86, 183). RER requires FEN1, RNaseH2 and PCNA loaded by RFC (86). Because RER is severely hindered in the absence of RNaseH2, RER is exquisitely sensitive to the presence and function of RNaseH2. In contrast FEN1 and Polδ can be replaced with Exo1 or Polε, respectively, illustrating the critical nature of RNaseH2B to the RER pathway (86). Thus, the loss of RNaseH2B interaction might be phenotypically fragile in that its loss has strong consequences for cellular homeostasis. Because mutations in RNaseH2 can give rise to a genetic disorder known as Aicardi–Goutières Syndrome, it is well-established that RNaseH2 activity is critical for human health (184). Interestingly, Aicardi–Goutières Syndrome shares some symptoms with the disorder caused by hPCNA-S228I. In particular, both disorders show progressive neurodegeneration and microcephaly (abnormally small head size). In the initial report of the hPCNA-S228I disorder (35), both of these symptoms are classified as being particularly strong and prevalent. Thus, our results would suggest that the possibility of an RER defect contributing to the hPCNA-S228I disorder.

CHAPTER III

Investigating the cellular effects of PCNA^{S228I} in a fission yeast model system

Introduction

The human sliding clamp PCNA (Proliferating cell nuclear antigen) is a ring-shaped protein that slides along DNA, orchestrating a variety of cellular processes through interactions with many protein partners (57). PCNA binding proteins typically use one or more of two conserved motifs, the PIP (PCNA-interacting protein) motif or the lesser-characterized APIM motif (Alk-B homolog 2 PCNA-interacting motif), to interact with PCNA (72, 73). Additionally, post-translational modifications of PCNA such as ubiquitination or SUMOylation also modulate binding of different partners to PCNA (56). Proper modulation of binding partners is essential to orchestrate processes such as the DNA damage response.

A recently identified variant of the human sliding clamp, $PCNA^{S228I}$, causes a novel disease called Ataxia telangiectasia-like disorder 2 (ATLD-2). Among the many functions modulated by PCNA, ATLD-2 is thought to arise from a defect in DNA repair, particularly nucleotide excision repair (35, 185). However, the molecular mechanism of this defect remains unknown. Furthermore, it is unclear how this mutation is tolerated given the many critical roles of PCNA.

Here, I describe my efforts to characterize $PCNA^{S228I}$ using a fission-yeast model system. Using standard methods for yeast genetics, I constructed a strain harboring *pcn1-S228I* as the sole copy of *pcn1* and confirmed the mutation was integrated into the genome. I tested the effect of this mutation on multiple DNA repair pathways including base excision repair, nucleotide excision repair, and

homologous recombination repair by treating cells with various DNA damaging drugs at multiple concentrations. Although I attempted to distinguish between potential defects in DNA repair mechanisms, this strain exhibits a subtle phenotype in preliminary studies using a small subset of DNA damaging agents. However, this work produced valuable reagents and laid the foundation for investigating the effects of this variant *in vivo*.

A subset of PCNA interactions in DNA repair

PCNA functions in almost all forms of DNA repair. PCNA helps localize DNA repair enzymes to sites of damage, and PCNA can help enhance enzymatic activities of these enzymes. I briefly highlight the role of PCNA in individual repair pathways below.

Base excision repair (BER) typically fixes bases damaged by alkylation, oxidation, deamination, or base mismatches. BER occurs via short- or long-patch pathways, both of which require PCNA (31). The first step in BER is formation of an apurinic/apyrmidinic (AP) site. Several glycosylases, each recognizing specific types of damaged bases, interact with PCNA for AP formation (186, 187). In short-patch BER, AP endonuclease (APE1 in humans) and DNA polymerase β (Pol β) excise the abasic sugar and fill 1-nucletotide gaps with the aid of the scaffolding protein XRCC1, which binds DNA ligase III. Pol β and APE1 use PIP-motifs to interact with PCNA (188-190), but surprisingly XRCC1 binding is not mediated by a PIP or APIM (191). Long-patch BER displaces 2-10 nucleotides

through the coordinated efforts of APE1 to cleave the abasic sugar, followed by PCNA-mediated DNA polymerases δ/ε and FEN1 endonuclease activity to synthesize a patch of the same size. All three long-patch BER enzymes use PIP box motifs to directly interact with PCNA, possibly simultaneously as in Okazaki fragment maturation (70). Thus PCNA orchestrates both the short- and long-patch BER pathways.

PCNA has critical roles in the nucleotide excision repair (NER) pathway that mediates removal and repair of bulky DNA lesions. In eukaryotes, six repair factors carry out NER: RPA, XPA, XPC, TFIIH, XPG, and XPF•ERCC1 complex (31). Repair occurs in three steps: 1) Damage is first recognized by XPA, RPA and XPC, 2) A short ~25 nucleotide oligomer is excised by the helicase activity of TFIIH and nuclease activities of XPG and XPF•ERCC1, and, 3) DNA polymerase ε/δ, with the aid of PCNA, synthesizes DNA to fill the gap (31). NER is dependent on PCNA for the DNA polymerase ε/δ-PCNA interactions, XPG-PCNA interaction (via PIP-motifs on polymerases ε/δ and XPG), and XPA-PCNA interaction (via an APIM motif on XPA) (83, 84).

PCNA mediates the ribonucleotide excision repair (RER) pathway that incises misincorporated rNTPs from genomic DNA. Misincorporation of ribonucleotides by the replicative DNA polymerases can occur on the order of 10^{-3}, making it the most abundant form of DNA damage (86, 192). The presence of the 2'-OH on RNA makes it more susceptible to hydrolysis, creating nicks that would stall or collapse the replication fork. Ribonuclease H2 (RNaseH2) incises

the backbone 5’ to a ribonucleotide, followed by strand displacement requiring the coordinated efforts of DNA polymerases ε/δ and PCNA, and finally FEN1 activity removes the flap produced by strand displacement and DNA Ligase1 performs ligation (85, 86). The catalytic subunit of the RNaseH2 complex, RNaseH2B, contains a PIP-motif (61) and depends on PCNA to be localized to sites of damage, and PCNA also enhances its enzymatic activity (86). As previously stated, PCNA coordinates with DNA polymerases ε/δ, FEN1 and DNA ligase I via PIP-motifs as well, demonstrating PCNA is critical to carry out RER.

PCNA is also involved in repairing, or preventing double strand breaks (DSBs). Stalling and subsequent collapse of a replication fork due to template lesions, reactive oxidative species or ionizing radiation can all lead to DSBs (30, 193, 194). DSBs are a particularly dangerous type of DNA damage because their repair can lead to gross chromosomal rearrangements or loss of genetic information (193). DSB are repaired primarily by two pathways, homologous recombination (HR) and nonhomologous end-joining (NHEJ) (195). In the HR pathway, the two broken ends of DNA are resected to generate 3’ single-stranded DNA onto which the recombinase, Rad51, can load. Rad51 mediates strand invasion of the sister chromatid as the homologous template for DNA polymerase to fill in the gap (196). PCNA is required for DNA polymerases ε/δ to synthesize DNA post-invasion (197, 198). Unlike HR, NEHJ does not use a template and instead rejoins the two broken ends together with minimal processing. The Ku complex (Ku70/Ku80 heterodimer) initiates NHEJ repair by

recognizing the two broken ends, and recruits DNA dependent protein kinases (DNA-PKcs) as well as a specialized ligase, DNA ligase IV, to join the two ends (31, 199). There are no reported roles for PCNA involved in the NHEJ pathway.

PCNA post-translational modifications allow for bypass replication or template switching mechanisms that help prevent DSBs. Cells use translesion synthesis (TLS) polymerases for bypass replication; these polymerases are error prone, but can synthesize through DNA lesions that stall the replicative polymerases ε/δ (200). In response to DNA damage such as UV radiation, the E3 ligase Rad6/Rad18 monoubiquitinates PCNA at K164, which helps recruit TLS polymerases η, κ or ι to PCNA-Ub for error-prone bypass replication (87). Polyubiquitinated-PCNA, formed by the combined actions of the Rad6/Rad18 and Rad5 E3 ligases, mediates the template-switching mechanism for error-free bypass replication (87). Finally, SUMOylated-PCNA at either K164 (as a primary site) or K127 (as a secondary site) by Ubc9/Siz1 is thought to prevent HR by recruiting the anti-recombinogenic Srs2 helicase (87, 90, 201, 202).

With a complicated role in DNA repair, it remains unclear which DNA repair pathway or pathways are affected by the S228I variant. Although preliminary experiments suggest an NER defect (35), other repair mechanisms may be perturbed by the mutation as well. In support of this hypothesis, ATLD-2 patients exhibit symptoms similar to diseases caused by disruption of DSB repair as well as NER disorders (35).

Our previous work (Chapter II) on human PCNA^{S228I} revealed the mutation reduces binding of three peptides from PCNA interacting partners: p21^{CIP1}, FEN1, and RNaseH2B (63). p21^{CIP1} is one of the tightest known binders to PCNA and is critical for cell cycle regulation. FEN1 endonuclease cleaves “flap” structures on DNA, and has essential roles for DNA damage repair and for Okazaki fragment maturation. Finally, RNaseH2B is part of the RNaseH2 complex that degrades transcription-related R-loops, and ribonucleotides improperly incorporated during DNA replication (61). While all three of these peptide-PCNA^{S228I} interactions were reduced compared to WT PCNA, FEN1 and RNaseH2B were the most affected (63). The loss-of-interaction with FEN1 and RNaseH2B supports the hypothesis that PCNA^{S228I} disrupts DNA repair while leaving other PCNA functions intact. However, neither study of ATLD2 (35, 63) have not revealed which specific DNA repair pathways are affected by this mutant.

Rationale for a yeast model system for ATLD2

In order to study the cellular effects of the PCNA^{S228I} mutant, we selected the fission yeast *Schizosaccharomyces pombe* as our model system. There is a relatively high overall sequence similarity between *S. pombe pcn1* and human PCNA (52% identical, **Figure 3.1a**), while *S. cerevisiae* PCNA is less well conserved (36% identical, **Figure 3.1b**) and therefore was not chosen for further study. Additionally, the site of the mutation, S228I, as well as key residues

affected by the mutation (63) are conserved in S. pombe (**Figure 3.1**). Finally, the relevant DNA repair pathways are conserved in *S. pombe* (203).

There is a wide assortment of small molecules that have been used in fission yeast to tease apart defects in DNA repair pathways. Because of the suggestion that S228I particularly effects excision repair, I selected small molecules that especially target these pathways. 4-nitroquinoline 1-oxide (4NQO) induces adducts on guanine and adenine, and can also cause purine to pyrimidine transversions (204). This type of damage induces the NER pathway. Methyl methanesulfonate (MMS) is an alkylating agent, and typically methylates guanine and adenine (205). These lesions can be repaired by excision repair, or if they lead to stalled replication forks, can require repair by HR (205-207). Hydroxyurea (HU) inhibits ribonucleotide reductase, leading to decreased dNTP pools. In checkpoint deficient backgrounds, decreased dNTP pools ultimately leads to fork collapse from stalled forks (208-210). HR and NHEJ repair this damage (210). Finally, ultraviolet (UV) light induces thymidine dimers, cyclobutane pyrimidine dimers (CPDs), and (6-4) pyrimidine pyrimidones (6-4 PPs) repaired by NER, or bypassed by using TLS polymerases (211). Using these DNA damage agents allows us to probe the affects of the $PCNA^{S228I}$ variant on specific repair pathways.

Figure 3.1: Amino acid alignment between human and yeast PCNA. Similarities between sequences are shaded in grey. Red boxes outline residues within the IDCL. A purple arrow points to the site of the mutation a) Alignment between *H. sapiens* and *S. pombe* is 51% identical. b) Alignment between *H. sapiens* and *S. cerevisiae* is 35% identical. UniProt identifiers are shown in blue text. Alignments were created using UniProt Align tool.

a

```
P12004 PCNA_HUMAN     1 MFEARLVQGSILKKVLEALKDLINEACWDISSSGVNLQSMDSSHVSLVQLTLRSEGFDTY   60
Q03392 PCNA_SCHPO     1 MLEARFQQAALLKKLLDAIKELVTDANFDCNDNGISLQAMDSSHVALVSMLIKSDGFEPY   60
                          *:***: *.::***:*:*:*:*:.:* :* ...*:.**:******:**.: ::*:**: *

P12004 PCNA_HUMAN    61 RCDRNLAMGVNLTSMSKILKCAGNEDIITLRAEDNADTLALVFEAPNQEKVSDYEMKLMD  120
Q03392 PCNA_SCHPO    61 RCDRNIALGINLNALSKVLRCAQNEDLVTLKAEDTPEVLNLVFESEKNDRISDYDVKLMD  120
                        *****:*:*:**.::**:*:** ***::**:***. :.* ****: ::::::***:****

P12004 PCNA_HUMAN   121 LDVEQLGIPEQEYSCVVKMPSGEFARICRDLSHIGDAVVISCAKDGVKFSASGELGNGNI  180
Q03392 PCNA_SCHPO   121 IDQEHLGIPDIEYDATITMPAAEFQRITRDLLTLSDSVTINASKEGVRFSCKGDIGNGST  180
                        :* *:****: **.:.:.**:.** ** ***  :.*:*.*..:*:**:**..*::***.

P12004 PCNA_HUMAN   181 KLSQTSNVDKEEEAVTIEMNEPVQLTFALRYLNFFTKATPLSSTVTLSMSADVPLVVEYK  240
Q03392 PCNA_SCHPO   181 TLKQHTDLSDQDQSIEISLIQAVTLTFSLKYLAQFTKATPLATRVTLSMSNDVPLLVEYK  240
                        .*.* :::..::::: *.:.: * ***:*:**  ********:: ****** ****:****

P12004 PCNA_HUMAN   241 IADMGHLKYYLAPKIEDEEGS                                    261
Q03392 PCNA_SCHPO   241 ME-SGFLRFYLAPKIGEEDEE                                    260
                        :   *.*::****** :*: .
```

b

```
P12004 PCNA_HUMAN     1 MFEARLVQGSILKKVLEALKDLINEACWDISSSGVNLQSMDSSHVSLVQLTLRSEGFDTY   60
P15873 PCNA_YEAST     1 MLEAKFEEASLFKRIIDGFKDCVQLVNFQCKEDGIIAQAVDDSRVLLVSLEIGVEAFQEY   60
                          *:**:: :.*:::*::::.:** :: . :: ...*:  *::*.*:* **.* :  *.*: *

P12004 PCNA_HUMAN    61 RCDRNLAMGVNLTSMSKILKCAGNEDIITLRAEDNADTLALVFEAPNQEKVSDYEMKLMD  120
P15873 PCNA_YEAST    61 RCDHPVTLGMDLTSLSKILRCGNNTDTLTLIADNTPDSIILLFEDTKKDRIAEYSIKLMD  120
                        ***: :::*::***:****:*. * * :** *::. *:: *:** ::::::::.:****

P12004 PCNA_HUMAN   121 LDVEQLGIPEQEYSCVVKMPSGEFARICRDLSHIGDAVVISCAKDGVKFSASGELGNGNI  180
P15873 PCNA_YEAST   121 IDADFLKIEELQYDSTLSLPSSEFSKIVRDLSQLSDSINIMITKETIKFVADGDIGSGSV  180
                        :*.: * * * :*.:.:.:**.**::* *****:.*:: *  :*: :** *.*::*.*.:

P12004 PCNA_HUMAN   181 KLSQTSNVDKEEEAVTIEMNEPVQLTFALRYLNFFTKATPLSSTVTLSMSADVPLVVEYK  240
P15873 PCNA_YEAST   181 IIKPFVDMEHPETSIKLEMDQPVDLTFGAKYLLDIIKGSSLSDRVGIRLSSEAPALFQFD  240
                          :.   :::: * ::.:**::**:***. :**  : *.: **. * : .:*::.* :.::.

P12004 PCNA_HUMAN   241 IADMGHLKYYLAPKIEDEEGS                                    261
P15873 PCNA_YEAST   241 L-KSGFLQFFLAPKFNDEE--                                    258
                        : . *.*:::*****::***
```

Materials and Methods

Plasmid construction

Genomic DNA was isolated from yFS105 (WT) cells using standard phenol-cholorform extraction (212). A fragment containing *pcn1*, with 1kb upstream and 0.5kb downstream flanking regions was PCR amplified from genomic *S. pombe* DNA using primers with the sequence 5'-CGGGTACCCATGGAATACTCGAAACCCGG-3' and 5'-GCGGATCCGCTCACAGAGTTGTGTCTAG-3'. This fragment was digested with *Kpn1* and *BamHI* and cloned into the *Kpn1* site of yFS117 target vector. Integration into the vector was confirmed through sequence analysis. Site-directed mutagenesis was performed using a protocol similar to QuikChange (147) using WT-*pcn1* in yFS117 and primers with the sequence 5'-GAGTCACACTTatcATGAGCAATGATGTTCCACTTCTTGTGG-3' and 5'-CATTGCTCATgatAAGTGTGACTCGGGTAGCTAAAGGAGTAGC-3'. The codon for isoleucine created a *BspHI* restriction endonuclease cleavage site within the modified *pcn1* gene, which was later used for diagnostic PCR (see below). Enzymes were purchased from New England BioLabs. Oligonucleotides were purchased from IDT.

Cell culture and preparation of DNA damaging agents

Cells were cultured using standard techniques (212). For rich media, cells were cultured in yeast extract plus supplements (YES; 0.5% w/v yeast extract,

3% w/v glucose, plus supplements of 225mg/L each of adenine, histidine, leucine, and uracil). For transformation selection, cells were cultured in Edinborough minimal media lacking uracil (EMM2-URA). To induce sporulation, diploids were patched onto malt extract plates (ME; 3% w/v malt extract plus supplements).

DNA-damaging agents were prepared by the following: 5mM stock solutions of 4NQO (Sigma-Aldrich N-8141) were solubilized in DMSO and stored at -20°C. 1M stock of hydroxyurea (Sigma-Aldrich H8627) was solubilized in sterile water and prepared fresh. 99% MMS (Sigma-Aldrich 129925) was diluted to a working stock concentration in water. DNA-damaging agents were added to cooled autoclaved YES media at specified concentrations in an Erlenmeyer flask and mixed prior to pouring plates. Fresh plates were prepared for each experiment and used within 24 hours. For UV-exposure, cells were spotted on YES plates and exposed to short-wavelength (254nm) UV using a Stratalinker (Stratagene).

Strain construction

The strains used are listed in Table 1. yNR45, yFS104, yFS105, and yFS189 were kindly provided by Dr. Nicholas Rhind. A strain containing *pcn1-S228I* as the sole copy of *pcn1* was constructed using a pop-in/pop-out strategy (213). The target vector, *pcn1-S228I* in yFS117 was linearized with *Pst1*, and transformed into the WT diploid yNR45 using standard lithium acetate protocols

(212, 214). Cells were plated on EMM2-URA select for transformants and incubated at 30°C. Single colonies were restreaked onto EMM2-URA+adenine to reduce false positives. Single colonies were patched onto ME plates to induce sporulation and incubated for 3-4 days at 25°C. After confirming crescent shaped tetrads containing 4 spores under a 40x light microscope, a single colony was resuspended in sterile water by toothpick, and the slurry was transferred onto a thin-layer YES plate in a horizontal line. Using a Zeiss Axioskop dissecting microscope, tetrads with uniform spores were moved 5mm away from the slurry streak. Plates were incubated at 37°C for 2-4 hours to allow the asci wall to break down before dissection. Plates were incubated at 25°C for 3-5 days until spores grew. Cells were replica-plated onto YES, EMM2-URA, and EMM2-URA supplemented with 1mg/mL 5'-fluroorotic acid (5'-FOA). 5'-FOA selects for cells in which the *ura4* gene has been lost (pop-out) (215). FOA^R colonies were re-streaked an additional time for single colonies to ensure a homologous population.

Table 3.1: List of *S. pombe* strains used in this study.

Strain	Phenotype	Genotype	Source
yNR45	WT (diploid)	*h+/h- leu1-32/leu1-32 ura4-D18/ura4-D18 ade6-210/ade6-216 his7-366/his7-366*	N. Rhind
yFS104	WT (haploid, *h+*)	*h+ leu1-32 ura4-D18*	N. Rhind
yFS105	WT (haploid, h-)	*h- leu1-32 ura4-D18*	N. Rhind
yFS189	*rad3Δ*	*h- leu1-32 ura4-? ade6-704 rad3::ura4*	N. Rhind
yBK06	*pcn1-S228I*	*h-? leu1-32 ura4-D18 ade6-? his7-366 pcn1-228*	This study
yBK07	*pcn1-S228I*	*h-? leu1-32 ura4-D18 ade6-? his7-366 pcn1-228*	This study

Strain confirmation using diagnostic PCR

To confirm *pcn1-S228I* was integrated into the genome, a 600-bp fragment containing the site of the mutation was PCR amplified using primers with the sequence 5'-GCCGAATTTCAACGCATTAC-3' and 5'-CAAGCTACAGCAGATGATTACGA-3'. Genomic DNA was isolated from single FOA^S colonies using a standard phenol-chloroform extraction (212). This PCR fragment was digested with *BspHI* and visualized on a 2% agarose gel. Cells which were successfully integrated with *pcn1-S228I* show two digested bands sized ~350bp and ~260bp. WT cells show a single band sized ~600bp.

Full-length *pcn1-S228I*, including 1kb upstream and 0.5kb downstream from the gene, was PCR amplified from cells which were positive for *BspHI* digest using primers with the sequence 5'-CGGGTACCCATGGAATACTCGAAACCCGG-3' and 5'-GCGGATCCGCTCACAGAGTTGTGTCTAG-3'. This ~2.7 kb fragment was sent for sequencing (Genewiz) to confirm the inserted mutation, and that no additional

mutations were introduced to the gene. Additional primers with the sequences 5'-CCAAATTCAATACCTCAGGTG-3', 5'-GCCGAATTTCAACGCATTAC-3', and 5'-CAAGCTACAGCAGATGATTACGA-3' were used for full sequencing coverage of the *pcn1* gene and flanking regions.

Serial dilution growth assay

Strains used in the serial dilution growth assay were grown in YES media overnight to saturation at 25°C (for initial temperature sensitivity studies) or 30°C (for all subsequent experiments). Cultures were diluted in YES to OD_{600} = 0.4 and allowed to grow 4 hours to an OD_{600} of ~ 1.0. 1.0 OD_{600} of each culture was spun down and resuspended in 250µL sterile water. 5-fold serial dilutions of culture were prepared using sterile water and transferred to a sterile 96-well plate. A 48-pin-replica plater for 96-well plate (Sigma-Aldrich R23823) was flame sterilized using 100% ethanol. Cells were transferred using the replica plater to YES agar plates, or YES+DNA-damaging agent plates. Plates were allowed to dry for 15 minutes before transferring to incubators at desired temperature (25°C, 30°C, 35°C, 37°C for initial temperature sensitivity studies; 30°C for all subsequent experiments). Plates were imaged on a Fujifilm LAS3000 (GE) after 72hrs.

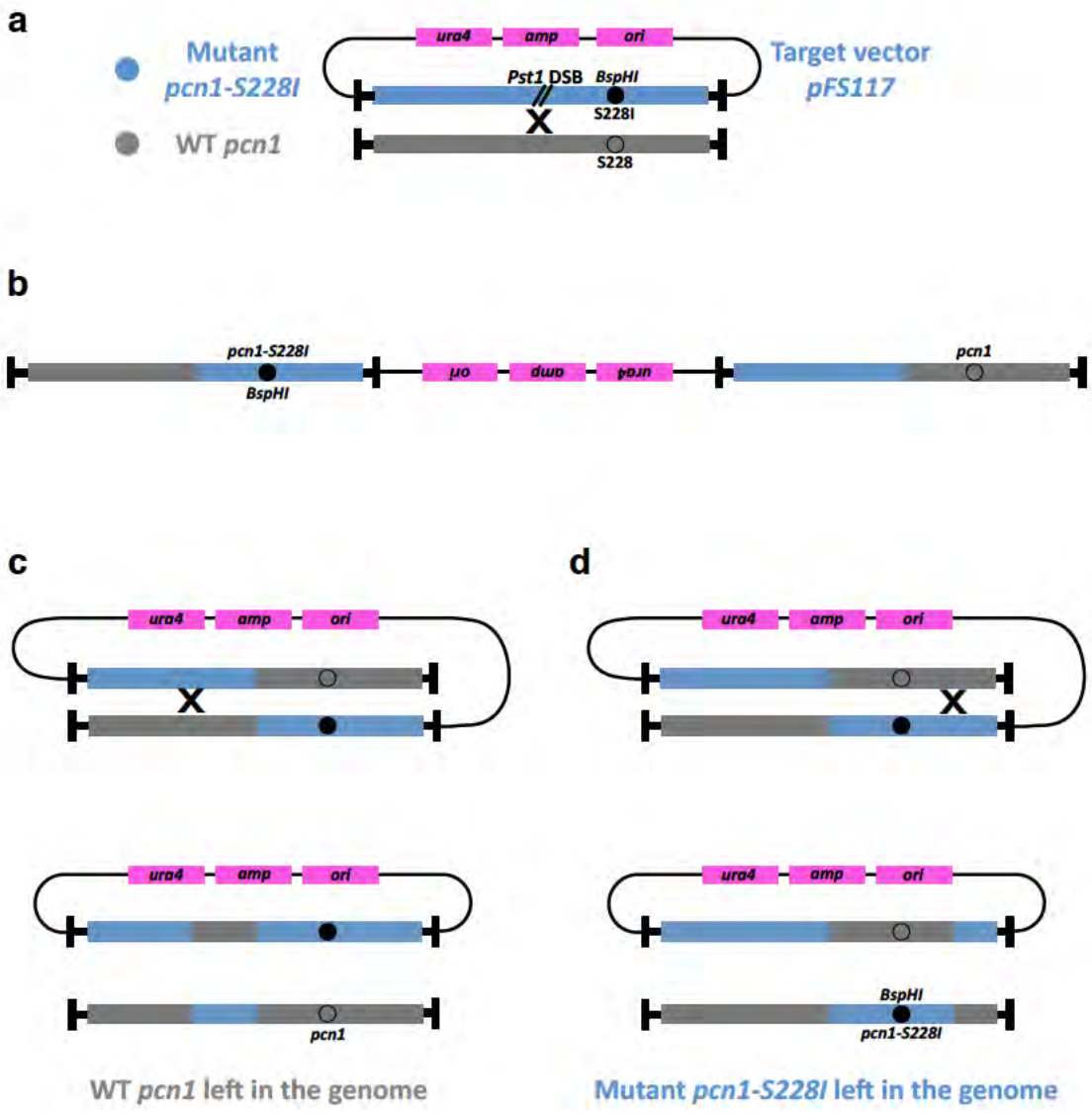
Results

pcn1-S228I is viable and not temperature sensitive

Using a pop-in/pop-out allele replacement strategy, I successfully constructed a haploid strain of *pcn1-S228I*. Briefly, a target vector lacking an autonomously replication sequence (ARS) containing *ura4+* as an auxotrophic marker was constructed containing *pcn1-S228I* and transformed into a diploid WT parent strain. The diploid parent strain was chosen because *pcn1* is an essential gene and we wanted to avoid viability issues in the case that *pcn1-S228I* was a lethal mutation.

The “pop-in” portion of the strategy takes advantage of the fact the target vector lacks an origin of replication, and *ura4*$^{+}$ colonies represent a successful integration of the target vector into the genome by homologous recombination (**Figure 3.2a**). The pop-in event results in an unstable intermediate in which a tandem duplication of *pcn1* exists in the genome: a WT copy and a copy harboring S228I (**Figure 3.2b**).

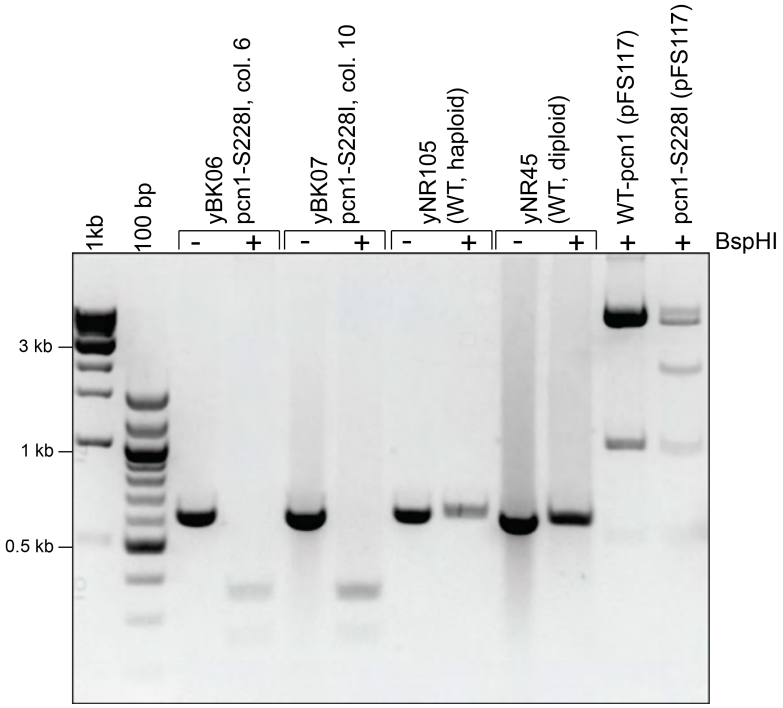
Figure 3.2: Fission yeast strain construction strategy. Steps required for *pcn1-S228I* integration into the genome. A) A target vector containing *pcn1-S228I* insert is transformed into the parent strain, also known as "pop-in". B) Integration results in unstable tandem gene duplication. C) Recombination downstream of the mutation results in a "pop-out" of the mutation. D) Recombination upstream of the mutation results in the mutation left in the genome.



The “pop-out” recombination event takes advantage of the negative selection against *ura4*$^{+}$ using the drug 5’-fluoroorotic acid (5’-FOA) (215). Pop-out events produce two outcomes, depending on which side of the modified allele recombination takes place. In the case of *pcn1-S228I*, recombination upstream of the mutation results in a WT copy of *pcn1* left in the genome (**Figure 3.2c**). Recombination downstream of the mutation results in *pcn1-S228I* left in the genome (**Figure 3.2d**). Cells were sporulated and dissected prior to pop-out. Successful tetrad dissection resulted in 4 viable spores on YES plates, which were plated on media selecting for *ura4*$^{+}$, followed by FOA selection. The viable FOA^{R} cells after pop-out indicate that *pcn1-S228I* is viable as the sole copy of *pcn1*.

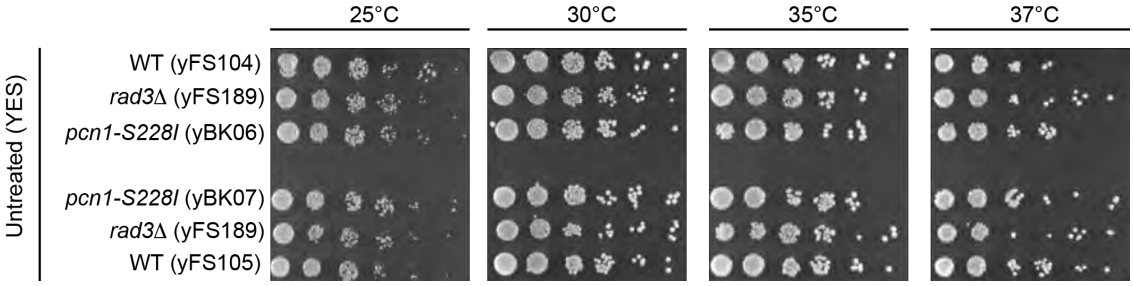
To confirm *pcn1-S228I* remained in the genome, genomic DNA was isolated from colonies that were FOA^{R}. A small fragment of *pcn1* was PCR amplified from genomic DNA and digested with *BspHI*. WT cells lack a *BspHI* site, whereas *pcn1-S228I* cells are susceptible to digest. I confirmed two strain isolates of *pcn1-S228I* by *BspHI* digest (**Figure 3.3**). Full-length *pcn1-S228I* was PCR amplified from genomic DNA from both strain isolates. Sequencing this fragment confirmed the S228I mutation, and also confirmed no additional mutations were introduced to the gene, or regulatory elements upstream and downstream of *pcn1*.

Figure 3.3: Restriction endonuclease genotyping confirms mutation integration into the genome. Genomic DNA was prepped from candidate strains (yBK06, yBK07) and WT controls (yFS105, yNR45). A 600bp fragment flanking the mutation site was amplified from gDNA, digested with *BspHI* and run on a 2% agarose gel. Candidate strains were positive for the digestion (~350bp and 260bp products), indicating the S228I-allele was integrated into the genome, whereas WT strains were negative for the digest. As an additional control, the targeting vector (confirmed via sequencing) used for the 'pop-in, pop-out' strategy (pFS117) was also *BspHI* treated: Mutant plasmid contains an additional cut site (approximately 2kb) compared to the WT plasmid.



Using a serial dilution growth assay, I confirmed that the two *pcn1-S228I* strains (yBK06 and yBK07, respectively) were viable on YES media and that growth is comparable to WT cells. After incubating *pcn1-S228I* at various temperatures, *pcn1-S228I* shows no temperature sensitivity and grows similarly to WT (**Figure 3.4**). Given the effect of the S228I variant on PCNA function (35, 63), I next wanted to investigate any responses *pcn1-S228I* has to DNA damaging agents.

Figure 3.4: *pcn1-S228I* is viable and not temperature sensitive. A 5-fold serial dilution of *S.pombe* strains were plated on YES and grown at indicated temperatures to assay for strain viability. Two WT strains (yFS104 and yFS105) serve as controls. *rad3Δ* serves as a negative control in the following DNA damage response experiments. The two isolated *pcn1-S228I* strains, yBK06 and yBK07, are viable, not temperature sensitive, and appear to grow similar to WT.



pcn1-S228I displays a subtle phenotype in response to select DNA damaging agents

Based on the results of Baple*, et al.* (35), we hypothesized *pcn1-S228I* would be sensitive to DNA damaging agents, particularly agents that induced the NER pathway. To test this hypothesis, yBK06 and yBK07 cells were treated with the DNA damaging agents MMS, 4NQO, HU, and UV light. WT strains yNR104 and yNR105 were included as a control. A strain in which the checkpoint kinase Rad3 was deleted served as a positive control, as *rad3Δ* would be sensitive to all damaging agents. Indeed, *rad3*Δ was inviable in the presence of all agents tested (**Figures 3.5-3.8**). yBK06 and yBK07 showed weak growth defects at high concentrations of MMS and HU (**Figures 3.5, 3.6**). These strains did not show any growth defects in response to 4NQO or UV light, or low concentrations of MMS or HU (**Figures 3.7, 3.8**).

Figure 3.5: *pcn1-S228I* is sensitive to high concentrations of hydroxyurea. A 5-fold serial dilution of *S. pombe* strains were plated on YES ± hydroxyurea. WT yFS105 and *rad3Δ* serves as controls. The two isolated *pcn1-S228I* strains, yBK06 and yBK07, show weak growth defects compared to WT in response to >5mM hydroxyurea.

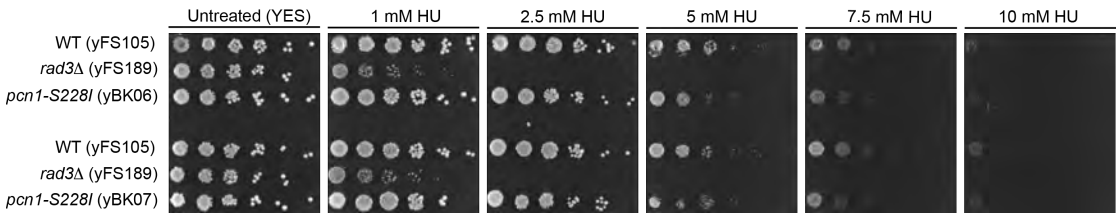


Figure 3.6: *pcn1-S228I* is sensitive to high concentrations of MMS.
5-fold serial dilution of *S. pombe* strains were plated on YES ± MMS. WT yFS105 and *rad3Δ* serves as controls. The two isolated *pcn1-S228I* strains, yBK06 and yBK07, show weak growth defects compared to WT in response to >0.0075% MMS.

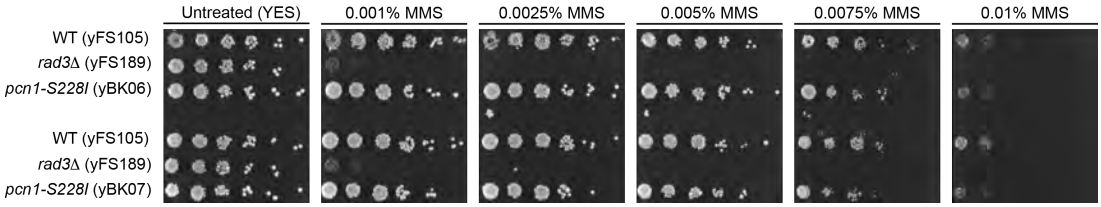


Figure 3.7: *pcn1-S228I* is unaffected by 4NQO.

5-fold serial dilutions of S. pombe strains were plated on YES ± 4NQO. WT yFS105 and *rad3Δ* serves as controls. The two isolated *pcn1-S228I* strains, yBK06 and yBK07, show no growth defects compared to WT in response to any dose of 4NQO.

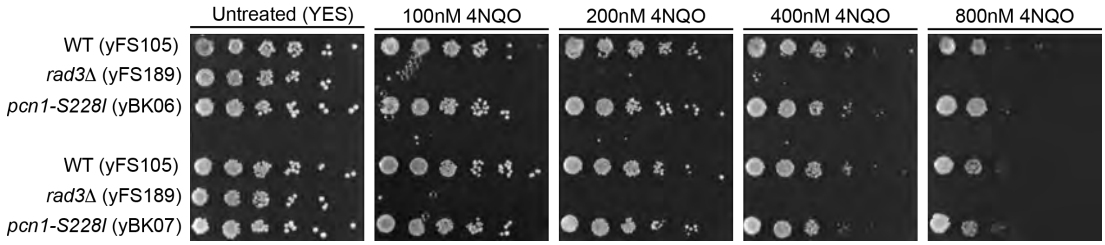
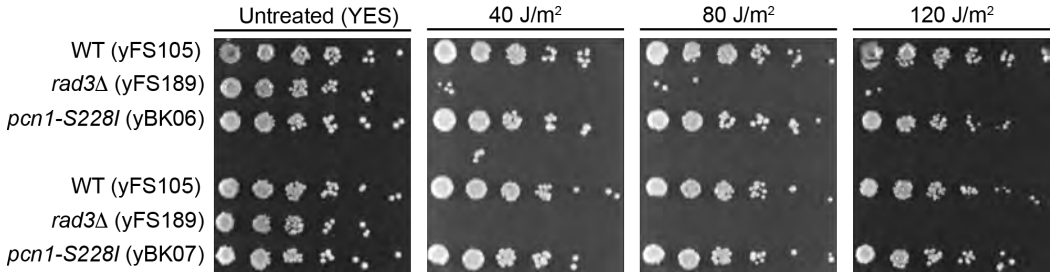


Figure 3.8: *pcn1-S228I* is unaffected by UV light.

5-fold serial dilution of *S. pombe* strains were plated on YES and exposed to 254-nm UV light at indicated doses. WT yFS105 and *rad3Δ* serves as controls. The two isolated *pcn1-S228I* strains, yBK06 and yBK07, show no growth defects compared to WT in response to UV light.



Discussion

I generated a fission yeast model system to further investigate the cellular effects of *pcn1-S228I*. This mutation is viable as the sole copy of *pcn1* (**Figure 3.4**, and there were subtle growth defects in response to DNA damaging agents (**Figures 3.5, 3.6**). *pcn1-S228I* cells grew similarly to WT where NER inducing damaging agents were introduced (UV light and 4NQO) suggesting NER is not affected by the mutation (**Figures 3.7, 3.8**).

S. pombe has an alternate excision pathway to repair UV damage (216, 217) which could potentially compensate for the loss of function associated with the PCNA S228I mutation. The UV damage excision repair pathway (UVER) uses an endonuclease, UVDE, encoded by the *uvde* gene, to excise both the major types of UV induced lesions, CPDs and 6-4 photoproducts, using the help of *rad2*, the FEN1 homolog in *S. pombe* (218, 219). UVDE can also nick AP sites, acting as an AP endonuclease (220). If *pcn1-S228I* does induce NER, perhaps the lack of a phenotype seen in response to UV light or 4NQO is indicative of a compensatory mechanism where the UVER pathway takes over. It is possible PCNA may mediate UVER through *rad2*/FEN1, but to date, there are no reported roles of *pcn1* acting in the alternate UVER pathway. Repeating these experiments in an *uvde*$^{-}$ deficient strain would test UVER compensates for the S228I variant.

The weak phenotype in response to high concentrations of MMS and HU may be attributed to a defect in HR. Both MMS and HU stall replication forks,

which can lead to fork collapse, requiring repair by HR. This requirement suggests *pcn1-S228I* cannot be modified by ubiquitin or SUMO because these PCNA modifications are necessary to prevent folk collapse by initiating bypass replication or template-switching. This could be tested by immunoprecipitating *pcn1* and measuring modifications using Ub- or SUMO-specific antibodies.

The HU phenotype may also reflect PCNA's role in the RER pathway. HU inhibits ribonucleotide reductase, depleting the dNTP pool (206, 210). Consequentially, the ratios of rNTP/dNTP increase, raising the possibility of ribonucleotide incorporation into DNA and triggering the RER response. A similar mechanism was observed in studies with human DNA polymerase μ; changing ratios of rNTP/dNTP lead to increased rNTP incorporation (221). Our previous work shows $PCNA^{S228I}$ abolishes interaction with RNaseH2B, a critical factor in RER. Mutations in RNaseH cause Aicardi Goutieres syndrome which shares clinical similarities to patients with ATLD-2 (184). Therefore we proposed a loss-of-interaction induced by the S228I mutation may cause an RER defect that contributes to ATLD-2. The weak phenotype *pcn1-S228I* shows in response to HU supports this putative defect in RER. A defect in RER leading to increased rNTP incorporation can be assessed from alkali-treated genomic DNA from *pcn1-S228I*; if *pcn1-S228I* perturbs RER, it should be more highly sensitive to alkali treatment than WT and display alkaline gel electrophoresis mobility similar to a strain with a deletion in RNase H (*S. pombe* gene *rnh1*) (222, 223).

Disruption of DNA damage checkpoints could also explain the weak phenotype of *pcn1-S228I*. These experiments were not performed under conditions of cell cycle control, so we cannot reasonably speculate if the weak phenotype is due to a defect in DNA damage checkpoints unless repeated under synchronized conditions. However, ataxia telangiectasia (AT), which shares clinical phenotypes with ATLD-2, is caused by mutations in ATM, a master checkpoint kinase (224). ATM interacts with PCNA, despite lacking a PIP-motif (225). If *pcn1-S228I* disrupts DNA damage checkpoints, it could be due to a loss of interaction with ATM.

A compensatory mechanism may alleviate deleterious effects caused by a loss of interaction between the S228I mutant and ATM. An alternative clamp loader, Rad17-RFC, and the alternative sliding clamp, Rad9-Rad1-Hus1 complex (9-1-1), also function during checkpoint responses (226). 9-1-1 also interacts with the damage sensor kinases ATM and ATR (Tel1 and Rad3 in *S. pombe*, respectively) in response to DNA damage (227). The alternative clamp, 9-1-1, which is functionally and structurally similar to PCNA (226, 228, 229), could compensate for *pcn1-S228I*, which would explain the weak phenotype in response to DNA damaging agents. If this compensatory mechanism is responsible for the weak phenotype then knocking out the 9-1-1 clamp in both WT and *pcn1-S228I* backgrounds and synchronizing cells prior to DNA damage should result in a more severe phenotype.

The subtle nuances of exactly how DNA damage response pathways are affected by the S228I mutation require further investigation. This fission yeast strain will be a valuable reagent to use in synthetic lethality screens to identify which pathways are affected by this mutation. Other model systems could also be considered for investigating the cellular effects of this mutant, particularly *in vitro* cell culture systems using cell lines derived from affected patients, or *Drosophila melanogaster* S2-cell lines due to the high homology shared between D. melanogaster mus209 and human PCNA (71% identical). However, the tractability of an *in vivo* yeast system, as well as the reagents generated here will be useful to further understand the mechanisms that are the underlying cause of ATLD-2, as well as further elucidate the vast network of PCNA-interactions.

CHAPTER IV

Mechanisms for primer/template binding and ATP hydrolysis in the DNA sliding clamp loader

Introduction

The sliding clamp loader is required for DNA replication

Most DNA polymerases require a sliding clamp (PCNA), a ring shaped protein that prevents polymerase dissociation from DNA, in order to achieve high-speed, accurate DNA replication. The clamp, an essential component of the replication fork, increases the processivity of polymerases by several orders of magnitude (4, 13). Because the clamp is a closed ring, it must be actively loaded onto primed DNA by the sliding clamp loader (51, 99, 100). The sliding clamp loader is a member of the AAA+ (ATPase associated with various cellular activities) family of ATPases (106), and couples primer/template binding and subsequent ATP hydrolysis to clamp closure and clamp loader ejection (**Figure 4.1a**).

Sliding clamp loaders are evolutionarily conserved, and share an overall similar architecture across all branches of life (99, 100). Unlike related hexameric AAA+ ATPases, the clamp loader is a spiral-shaped pentamer that possesses a gap between the first and fifth subunit where the missing sixth subunit would be (101, 103, 116). Viral clamp loaders contain 2 proteins in a 4:1 stoichiometry, bacterial clamp loaders contain 3 proteins in a 3:1:1 stoichiometry, and eukaryotic clamp loaders are heteropentamers. A universal naming scheme A-E describes the position of each subunit, starting with the subunit at the open interface and proceeding counter-clockwise around the ring (102) (**Figure 4.1b**).

Elegant structural and biochemical studies have provided a detailed understanding of the clamp loader reaction (103, 116) (**Figure 4.2**). In the absence of ATP, the clamp loader cannot organize into a spiral shape (101). Upon ATP-binding, the clamp loader undergoes a conformational change to adopt spiral conformation that allows for clamp binding and opening (101, 102). The clamp loader and open clamp match the helical symmetry of DNA, allowing primer/template binding (103). Because the clamp opening is not wide enough to accommodate dsDNA (103, 121), it is hypothesized that the single stranded 5'-template overhang binds first, allowing the rest of the primer/template to corkscrew up into the clamp:clamp loader complex. DNA binding promotes ATP hydrolysis, causing clamp closure (103, 116). ATP hydrolysis is thought to initiate in the B-subunit (103). Subsequent ATP hydrolysis in the C- and D-subunits promotes clamp loader ejection. Recent crystal structures helped reveal an unprecedented view of the clamp loader reaction, it is still not understood how different nucleic acid duplexes modulate the energetics of binding, nor how ATP hydrolysis is coordinated within the context of the AAA+ spiral.

Figure 4.1: Clamp loaders and sliding clamps.
(A) Clamp-loading reaction. The clamp loader has low affinity for both clamp and primer/template DNA in the absence of ATP. Upon binding ATP, the clamp loader can bind the clamp and open it. The binding of primer/template DNA activates ATP hydrolysis, leading to ejection of the clamp loader. (B) Three classes of clamp loaders. Bacterial clamp loaders are pentamers consisting of three proteins: δ (A position), γ (B, C, and D positions), and δ' (E position). Eukaryotic clamp loaders (RFCs) consist of five different proteins, with the A subunit containing an A´ domain that bridges the gap between the A and E AAA+ modules. The T4 bacteriophage clamp loader consists of two proteins: gp44 (the B, C, D, and E subunits) and gp62 (the A subunit). (103)

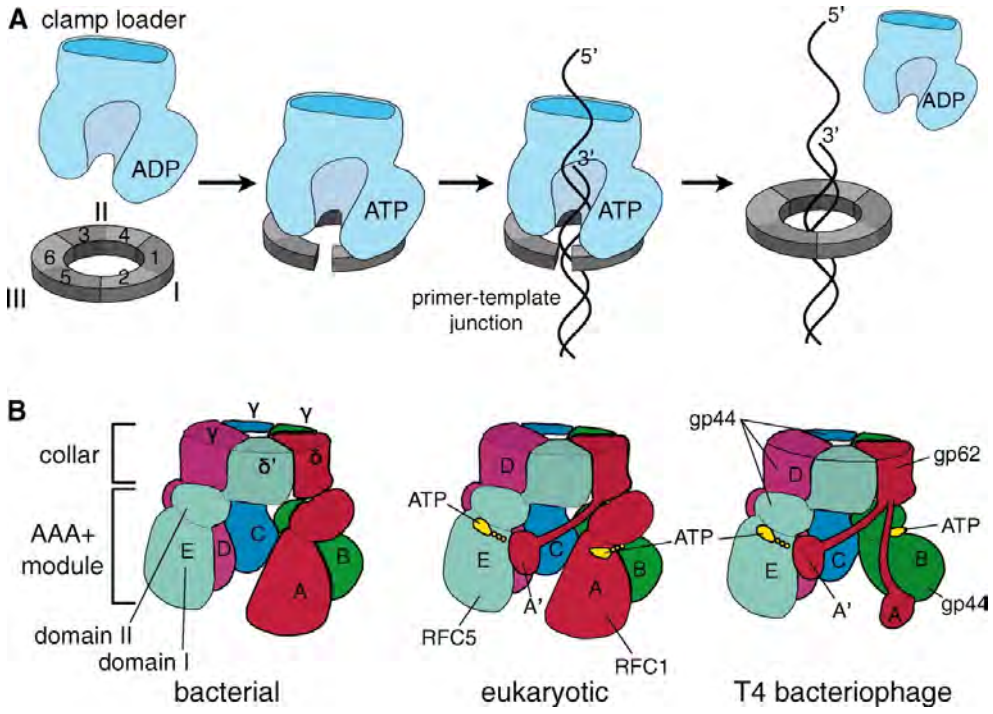
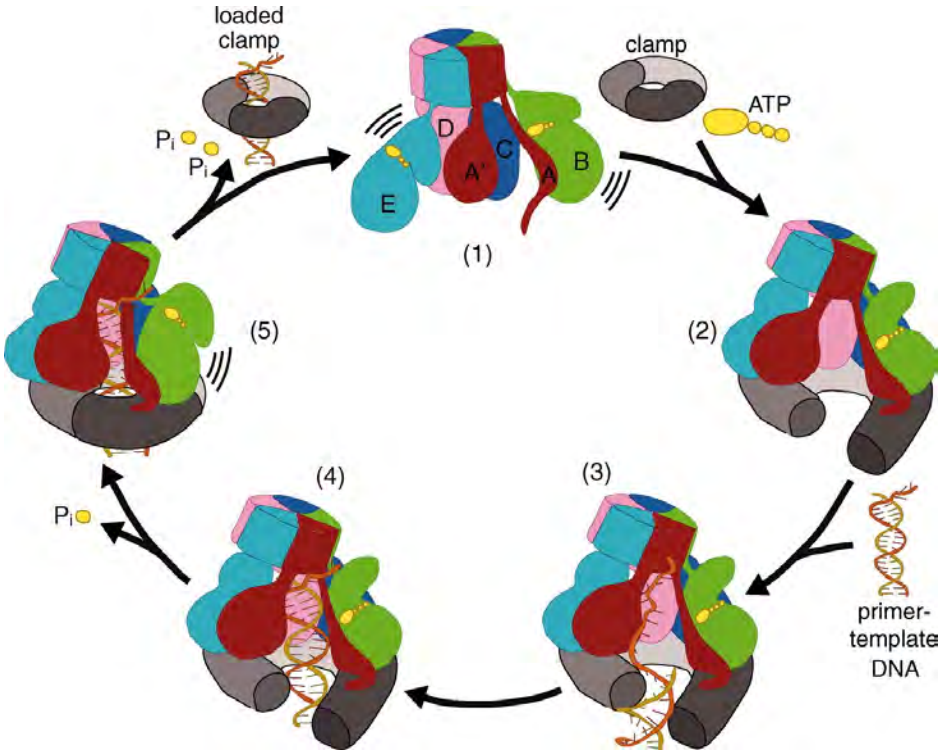


Figure 4.2: A detailed mechanism for the clamp loading reaction.
The reaction cycle for the T4 clamp loader is shown as a schematic diagram. (1) In the absence of ATP, the clamp loader AAA+ modules cannot organize into a spiral shape. (2) Upon ATP binding, the AAA+ modules form a spiral that can bind and open the clamp. (3) Primer/template DNA must thread through the gaps between the clamp subunits I and III and the clamp loader A and A´ domains. (4) Upon DNA binding in the interior chamber of the clamp loader, ATP hydrolysis is activated, most likely through flipping of the switch residue and release of the Walker B glutamate. (5) ATP hydrolysis at the B subunit breaks the interface at the AAA+ modules of the B and C subunits and allows closure of the clamp around primer/template DNA. Further ATP hydrolyses at the C and D subunits dissolve the symmetric spiral of AAA+ modules, thus ejecting the clamp loader because the recognition of DNA and the clamp is broken. The clamp is now loaded onto primer/template DNA, and the clamp loader is free to recycle for another round of clamp loading.
Intermediates (1), (4), and (5) are based on crystal structures from Jeruzalmi *et al.* 2001, Kelch *et al.* 2011, and Kelch *et al.* 2011, respectively. (101, 103)

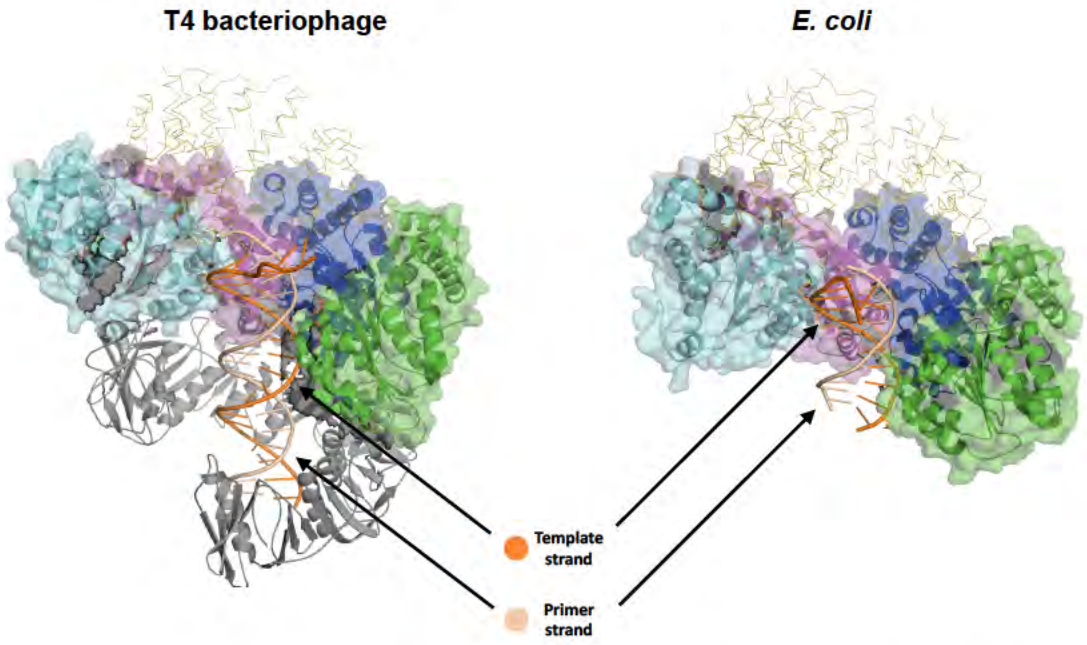


How does the clamp loader engage at primer/template junctions?

Because DNA polymerases cannot synthesize DNA *de novo*, they require a short primer to be synthesized at the replication fork by primases (22). Primases synthesize RNA-containing primers at the replication fork and their actions are required at the beginning of each Okazaki fragment on the lagging strand. In bacteria, archaeal and viruses, this RNA/DNA primer/template is the substrate for clamp loader action (22, 230, 231). Despite the primer/template consisting of a RNA/DNA duplex, all previous structural studies and most biochemical experiments investigating the clamp loader:primer/template interaction have used a primer/template consisting solely of DNA. The lack of data utilizing a RNA/DNA hybrid duplex demonstrates a gap in understanding the clamp loader reaction.

The significance of using RNA/DNA hybrids during replication and Okazaki fragment maturation requires that the clamp and clamp loader must accommodate both A- and B-form DNA. Double-stranded DNA predominantly exists as B-form in solution, while dsRNA and RNA/DNA duplexes are A-form (29, 125, 232, 233). Ribonucleotides cannot adopt B-form because the 2'-OH would clash with the 3'-OH in the B-form conformation (29). Indeed, the introduction of even a single ribonucleotide in a DNA strand converts B-DNA to A-DNA (234). Based on the several crystal structures of the sliding clamp from various organisms, the width of its inner pore is approximately 34Å, wide enough to accommodate both A- and B-forms (43, 48, 49).

Figure 4.3: The template strand is highly contacted by the clamp loader. Primer/template DNA bound to the T4 bacteriophage and *Escherichia coli* clamp loader show the primer and template strands make different contacts to the loader. The clamp loader subunits match the helical symmetry of the template strand (orange) and the template strand is highly contacted by the clamp loader, but the primer strand (beige) does not appear to make any specific interaction with the clamp loader. For clarity, the A-subunit is hidden and the collar domains of the B-E subunits are shown as ribbons (PDB: 3u60, 3GLF). (103, 116)



Whereas the clamp can accommodate both A- and B-form DNA, the question remained as to whether clamp loader prefers one conformation. Several features of primer/template strand binding to the clamp loader are showcased in two crystal structures: 1) the co-crystal structure of the T4 clamp loader bound to the clamp and primer/template DNA in the presence of an ATP analog, and 2)

the *E. coli* clamp loader bound to primer/template DNA (103, 116). Both structures show the primer strand lacks significant contacts to the clamp loader, suggesting that the loader can accommodate either DNA or RNA primer (**Figure 4.3**). The strongest interaction between the primer strand and the clamp loader is at residue S112 in the T4 gp44 B-subunit, but this interaction is not found in all subunits within the assembly (**Figure 4.4**). The clamp loader also induces a conformational change in the bound primer/template DNA from B-form to nearly A-form (**Figure 4.5**), again suggesting an RNA primer could be accommodated (103, 116). A model of ideal B-form primer/template DNA into the T4 clamp loader structure revealed there would be steric clashes between the primer strand, and S112 within B-E subunits of the clamp loader (103). Additionally, modeling an A-form RNA/DNA hybrid into the *E. coli* clamp loader structure showed no clashes (116). Because RNA/DNA hybrids are already A-form in solution, these primer/template junctions would not need to reorganize upon clamp loader binding. In contrast, B-form DNA-DNA duplexes would be expected to use some of the binding energy to reorganize the double-helix into A-form. Therefore, we hypothesize a RNA/DNA primer/template interacts more favorably with the clamp loader than a DNA/DNA primer/template.

Figure 4.4: The primer strand hardly contacts the clamp loader subunits. The strongest contact the primer strand makes with the clamp loader is a weak hydrogen bond (3.4Å) to S112 in the B-subunit. However, this distance increases in the remaining subunits.

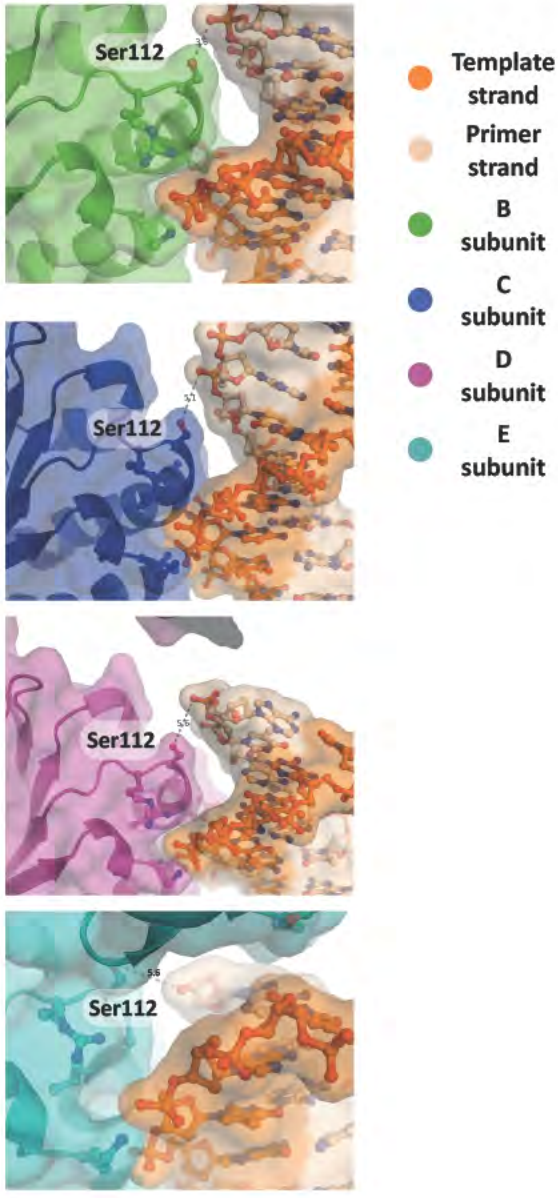
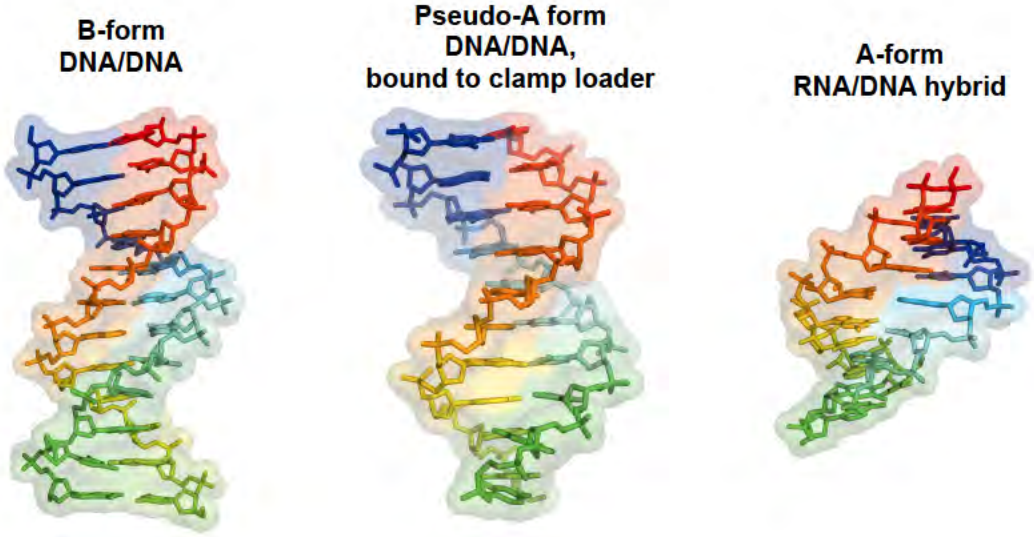


Figure 4.5: DNA primer/template bound to the clamp loader is pseudo-A form.

Left: A crystal structure of pure B-form dsDNA (PDB: 1bna; 12 bp)

Middle: A duplex (10bp) region of primer/template DNA bound to the T4 clamp loader and open clamp is converted from B-form, and is closer to A-form.

Right: A crystal structure of a pure A-form RNA:DNA hybrid (PDB: 124d; 8 bp). (103, 125, 232)



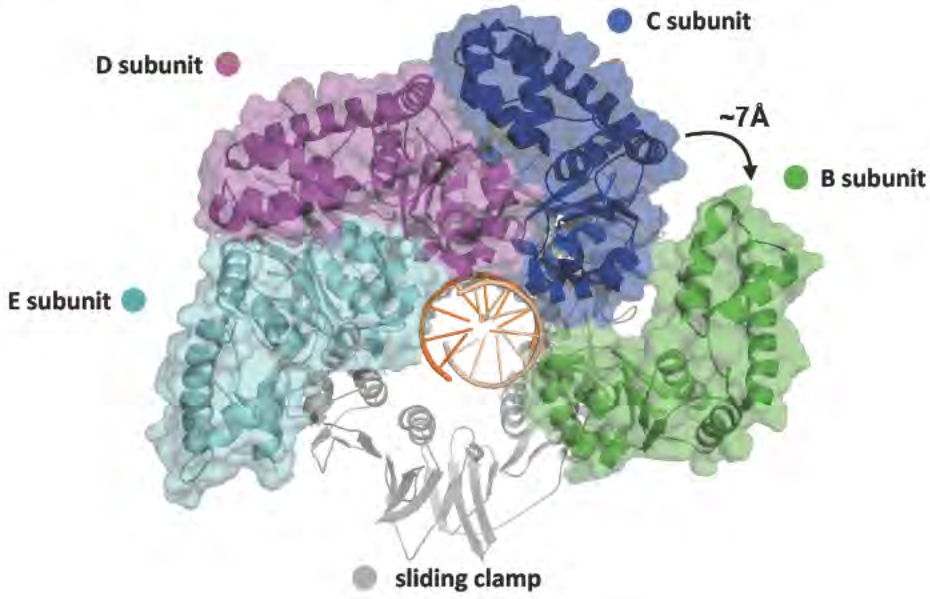
How does ATP hydrolysis occur around the clamp loader spiral?

Primer/template binding is directly coupled to ATP hydrolysis (118, 126, 127). An important residue responsible for initiating ATP hydrolysis is a conserved glutamate in the Walker B motif, which is common to all AAA-ATPases (106). This catalytic glutamate orients a water molecule for nucleophilic attack on the γ-phosphate of ATP. However, prior to primer/template binding, the Walker B catalytic glutamate is held in an inactive conformation by a conserved basic “switch” residue (102, 103, 126). This switch residue, Lys80 in T4, interacts with the backbone of the catalytic glutamate and orients it away from the catalytic water, preventing ATP hydrolysis. Upon primer/template DNA binding, the switch residue changes conformation to directly contact the phosphate backbone of the template strand, releasing the glutamate to assume a conformation competent to catalyze ATP hydrolysis (103).

A crystal structure of the T4 clamp loader in a partially hydrolyzed state suggested a mechanism by which ATP hydrolysis occurs around the ring (103). This structure shows only the B-subunit in a post-hydrolysis state, and as a result, the clamp is closed (**Figure 4.6**). The ATP hydrolyzed B-subunit moves away from its neighboring C-subunit, creating space for a similar conformational change upon ATP hydrolysis in the C-subunit. Therefore, we hypothesize that ATP hydrolysis occurs sequentially around the spiral, starting from the B subunit. Evidence for a sequential ATP hydrolysis mechanism has been seen in other

ASCE ATPases, including the DNA packaging motor from phi29 (132, 235), E1 helicase (133, 236), and Rho helicase (237).

Figure 4.6: A crystal structure shows one clamp loader subunit in a post-hydrolysis state. T4 clamp loader shows post-hydrolysis state in the B-subunit only. Upon ATP-hydrolysis, the B-subunit moves ~7Å away from the C-subunit. This action breaks contacts with the sliding clamp, which is closed. (View is from the top looking down; A-subunit and collar domains are hidden for clarity) (PDB:3u61) (103).



Approach

In this chapter, I attempt to understand two important aspects of the clamp loader reaction: 1) how nucleic acid composition of the primer template junction modulates clamp loader binding and catalysis, and 2) whether the ATP hydrolysis mechanism is sequential or random for subunits within the clamp loader spiral. While my crystallographic attempts were unsuccessful, I generated biochemical data to partially address both of these questions using a combination of *in vitro* assays. This data will be useful for further understanding the clamp loader reaction, as well as serve as a model for additional members of the AAA+ family.

Materials and Methods

Mutant T4 clamp loader constructs:

Site-directed mutagenesis for T57A was performed using a protocol similar to QuikChange (147) using primers with the sequence 5'—GAACAGGTAAAgcaACTGTAGCAAAAGCATTATGTCATGATGTAAATGC—3' and 5'—GCTTTTGCTACAGTtgcTTTACCTGTTCCTGGAGAAGGAGAATGAAG—3'.

Mutant T4 clamp loader expression:

Constructs harboring S112A, S112K or T57A mutations in gp44 were transformed into BLR(DE3) cells, and maintained in the presence of kanamycin (Kan). Single colonies were inoculated in 100mL LB+Kan and grew overnight at 37℃. 10mL of saturated culture was inoculated into TB+Kan and incubated at 37℃ until an OD_{600} of ~0.8 was reached. Cultures were moved to 4℃ for 30 minutes before being induced with 1x Isopropyl β-D-1-thiogalactopyranoside (IPTG), and transferred to 18℃ overnight. Cultures were pelleted at 5000 rpm for 15 minutes at 4℃. Pellets were resuspended in 25mL of Buffer A (50mM Tris pH 7.5, 20mM imidazole, 250mM NaCl, 10% glycerol, 5mM β-mercaptoethanol) before being flash frozen in liquid nitrogen and stored at -80℃.

T4 clamp loader purification:

Pellets were thawed in water and resuspended in Buffer A before being passed through a cell disruptor at 80 psi (Microfluidics Corporation). Lysates were spun at 15,000 rpm, for 40 minutes at 4℃. Cleared lysate was pre-filtered by syringe using a 5µm filter, followed by vacuum filtration using a 0.45µm filter. Filtered lysates were applied to three tandem HisTrap 5mL affinity columns (GE) pre-equilibrated in Buffer A. Loaded columns were washed with Buffer A, and protein was eluted in 100% Buffer B (50mM Tris pH 7.5, 500mM imidazole, 250mM NaCl, 10% glycerol, 5mM β-mercaptoethanol). Fractions containing gp44/gp22 were pooled and dialyzed overnight in Buffer HepA (20mM Tris pH 7.5, 150mM NaCl, 10% glycerol, 2mM 1,4-Dithiothreitol (DTT)) containing

PreScission protease at 4℃. Dialysates were filtered using a 0.45µm filter before being loaded onto a 5mL HiTrap-SP column (GE), followed by three tandem Heparin 5mL affinity columns (GE) pre-equilibrated in Buffer HepA. After loading protein, the SP column was removed. HepA columns were washed with 50mL of Buffer HepA. Protein was eluted in a gradient of Buffer HepB (20mM Tris, pH 7.5, 1 M NaCl, 10% glycerol, 2mM DTT); gp44/gp62 eluted between 10-25% Buffer HepB. Fractions containing gp44/62 were pooled and concentrated using an Amicon 10K MWCO spin concentrator to ~5mL before being injected onto a Superdex 200 (GE) gel filtration column pre-equilibrated in gel filtration buffer (20mM HEPEs, pH 7.5, 150mM NaCl, 10mM DTT). Fractions containing gp44/gp62 were pooled and concentrated using Amicon 10K MWCO spin concentrator to 60-70mg/mL. Aliquots were flash frozen in liquid nitrogen before being stored at -80℃.

DNA anisotropy assays:

Anisotropy assays to assess primer/template binding were performed at 20℃. Reagents were added to the specified concentration to a single cuvette in the following order: Buffer A (50mM Tris, pH7.5, 250mM potassium glutamate, 2mM DTT, 10% v/v glycerol), 1 mM ADP, 10 mM NaF, 2 mM $BeCl_2$, 1 mM $MgCl_2$, 100 nM 20d5p10-TAMRA primer/template (20d5p10 indicates a 20 bp duplex, followed by 10nt 5’ overhang), and 2 µM gp45. The clamp loader gp44/gp62 was added last and titrated across 0-4 µM such that the total protein concentration

only varied by less than 5%. Primer/templates having a 20 base pair double-stranded duplex and a 10 nucleotide overhang were constructed using oliognucleotides with the sequences 5'—TTTTTTTTTTTATGTACTCGTAGTGTCTGC-3' (template) and either 5'—GCAGACACTACGAGTACATA–3' (DNA primer) or 5' GCAGACACUACGAGUACAUA –3' (RNA primer). Primer/templates were annealed by heating to 95°C, followed by a slow cooling (>3 hours) to room temperature. A Fluoromax-4 fluorimeter (Horibal Scientific) equipped with a Peltier-temperature controlled sample chamber was set up for a single point anisotropy method, excitation at 550nm and emission at 580nm, using slit-widths of 3.5nm and 9nm, respectively, and an integration time of 5 seconds. Anistropy was calculated from using the equation $(I_{II} - I\perp)/(I_{II} + 2I\perp)$. Data were fit to a tight-binding equation using the tight-binding equation $\theta = \frac{C_T+D_T+K_D \pm \sqrt{(C_T+D_T+K_D)^2-4(C_T D_T)}}{2D_T}(B_{max})$ where θ= fraction bound, C_T= total clamp loader concentration, D_T= total DNA/DNA concentration (100nM), K_D is clamp loader concentration that yields half-maximum bound, and B_{max} is the change in amplitude.

Enzyme-coupled ATPase assay:

Coupled enzyme assays (238) were performed at room temperature using the following concentrations (unless otherwise noted): 6 U/mL pyruvate kinase, 6 U/mL lactate dehydrogenase, 1 mM phosphoenolpyruvate, 340 μM NADH, 50

mM Tris (pH 7.5), 500 µM Tris(2-carboxyethyl) phosphine, 100 mM potassium glutamate, 5 mM magnesium chloride, 0.5 µM gp45, and 7.5nM gp44/gp62.

For primer/template titrations, 1mM ATP was also included. DNA/DNA or RNA/DNA primer/template were titrated across 0-1µM. For ATP or ATPγS titrations, DNA primer/template was included at 600nM, and nucleotide was titrated across 0-1mM. For ATPγS doping experiments the nucleotide concentration was held constant at 1mM, and the ratio of ATP/ATPγS ranged from 0-100%. For mutant clamp loader doping experiments (T57A), the clamp loader concentration was held constant at 7.5nM. T57A and WT were mixed and allowed to equilibrate for 24 hours before initiating the reaction with ATP. Absorbance was measured in a 96-well format with a Perkin–Elmer Victor3 1420 multichannel counter using an excitation filter centered at 355 nm, with a bandpass of 40 nm. In this assay, every NADH oxidized to NAD+ corresponds to one ATP hydrolyzed. Initial slopes were linearly fit to obtain rates, which were plotted as a function of primer/template concentration. Data comparing the rate of ATP hydrolysis to substrate concentration (either primer/template for DNA binding assays or ATP or ATPγS for ATP binding assay) were fit using the equation: $v = V_{max} * [S] / (K_{half\text{-}maximal} + [S])$, where V_{max} is the maximum rate and $K_{half\text{-}maximal}$ is the substrate concentration that yields half-maximal activity. For examining cooperatively, data were fit to a Hill equation: $v = V_{max} * [S]^n / (K_{half\text{-}maximal}^n + [S]^n)$, where n is the Hill coefficient. Data from ATPγS doping assay were fit to an equation describing a random, sequential, or concerted model,

below. The models assumes all binding sites are occupied and three ATP hydrolysis events. In these models, [A] and [G] refer to concentrations of ATP and ATPγS, respectively. The models also take into consideration the respective affinities for ATP and ATPγS binding to the clamp loader as $K_{d,ATP}$ and $K_{d,ATP\gamma S}$, respectively. For a random model, the order of ATP hydrolysis events is not taken into account. For a sequential model, activity will be reduced depending on the subunit where ATPγS is bound; when ATPγS is bound in an upstream subunit, all subsequent ATP hydrolysis will be blocked. For a concerted model, ATP hydrolysis will occur in an "all or none" mechanism; the only ATPase activity occurs when ATP occupies all sites.

Random:

$$E_T * k_{cat}\left(\frac{(\frac{3[A]^3}{K_{D,ATP}^3}) + (\frac{2[A]^2[G]}{K_{D,ATP}^2 K_{D,ATP\gamma S}} * 3) + (\frac{[A][G]^2}{K_{D,ATP} K_{D,ATP\gamma S}^2} * 3)}{1 + \left(\frac{[A]^3}{K_{D,ATP}^3}\right) + \left(\frac{[G]^3}{K_{D,ATP\gamma S}^3}\right) + \left(\frac{[A]^2[G]}{K_{D,ATP}^2 K_{D,ATP\gamma S}} * 3\right) + \left(\frac{[A][G]^2}{K_{D,ATP} K_{D,ATP\gamma S}^2} * 3\right) + \left(\frac{[A][G]}{K_{D,ATP} K_{D,ATP\gamma S}} * 6\right) + \left(\frac{[G]^2}{K_{D,ATP\gamma S}^2} * 3\right) + \left(\frac{[A]^2}{K_{D,ATP}^2} * 3\right) + \left(\frac{[G]}{K_{D,ATP\gamma S}} * 3\right) + \left(\frac{[A]}{K_{D,ATP}} * 3\right)}\right)$$

Sequential:

$$E_T * k_{cat}\left(\frac{(\frac{3[A]^3}{K_{D,ATP}^3}) + (\frac{2[A]^2[G]}{K_{D,ATP}^2 K_{D,ATP\gamma S}}) + (\frac{[A]^2[G]}{K_{D,ATP}^2 K_{D,ATP\gamma S}}) + (\frac{[A][G]^2}{K_{D,ATP} K_{D,ATP\gamma S}^2})}{1 + \left(\frac{[A]^3}{K_{D,ATP}^3}\right) + \left(\frac{[G]^3}{K_{D,ATP\gamma S}^3}\right) + \left(\frac{[A]^2[G]}{K_{D,ATP}^2 K_{D,ATP\gamma S}} * 3\right) + \left(\frac{[A][G]^2}{K_{D,ATP} K_{D,ATP\gamma S}^2} * 3\right) + \left(\frac{[A][G]}{K_{D,ATP} K_{D,ATP\gamma S}} * 6\right) + \left(\frac{[G]^2}{K_{D,ATP\gamma S}^2} * 3\right) + \left(\frac{[A]^2}{K_{D,ATP}^2} * 3\right) + \left(\frac{[G]}{K_{D,ATP\gamma S}} * 3\right) + \left(\frac{[A]}{K_{D,ATP}} * 3\right)}\right)$$

Concerted:

$$E_T * k_{cat}\left(\frac{(\frac{3[A]^3}{K_{D,ATP}^3})}{1 + \left(\frac{[A]^3}{K_{D,ATP}^3}\right) + \left(\frac{[G]^3}{K_{D,ATP\gamma S}^3}\right) + \left(\frac{[A]^2[G]}{K_{D,ATP}^2 K_{D,ATP\gamma S}} * 3\right) + \left(\frac{[A][G]^2}{K_{D,ATP} K_{D,ATP\gamma S}^2} * 3\right) + \left(\frac{[A][G]}{K_{D,ATP} K_{D,ATP\gamma S}} * 6\right) + \left(\frac{[G]^2}{K_{D,ATP\gamma S}^2} * 3\right) + \left(\frac{[A]^2}{K_{D,ATP}^2} * 3\right) + \left(\frac{[G]}{K_{D,ATP\gamma S}} * 3\right) + \left(\frac{[A]}{K_{D,ATP}} * 3\right)}\right)$$

Results

DNA vs. RNA primer/template binding

A critical step in the clamp loader reaction is binding of the clamp/clamp loader to primer/template DNA. To date, all crystal structures of primer/templates bound to the clamp loader were composed only of DNA duplexes, despite the fact that most systems use a RNA primer (103, 116). Therefore, we hypothesized RNA/DNA primer/template-junctions would act as better substrates for the clamp loader than DNA/DNA primer/template-junctions. We sought to investigate biochemical and structural differences between DNA vs. RNA primer strands.

Binding affinities of DNA versus RNA substrates would quantify any preference of the clamp loader possesses for either substrate. We used an anisotropy-binding assay in order to determine the dissociation constant for DNA primer template to WT T4 clamp loader, as has been used previously (47, 103, 116, 124). T4 clamp loader was titrated into TAMRA-labeled primer/template DNA incubated with the T4 sliding clamp and non-hydrolyzable ATP analog, ADP•BeF_3. Anistropy increased as clamp loader was titrated in (**Figure 4.7**). Surprisingly, the data show a steep dependence of anistropy on clamp loader concentration and the data showed a poor fit to an equation describing a bimolecular reaction. We therefore fit the data to a tight-binding equation and the measured K_D was approximately 56 nM ± 23nM. Because the concentration of DNA in this assay was higher than the K_D (100nM), we were indeed under tight binding conditions. We therefore switched to an enzyme-coupled ATPase assay

where we could decrease concentrations of components in the assay and monitor binding in a reasonable amount of time (**Figure 4.8**) (238).

Using the enzyme-coupled ATPase assay, we determined apparent dissociation constants for both DNA/DNA and RNA/DNA primer/templates bound to WT T4 clamp loader in the presence of the clamp and ATP (**Figure 4.9**). DNA/DNA bound with a $K_{D,app}$ = 25 ± 2 nM and RNA/DNA bound with a $K_{D,app}$ = 10 ± 2 nM. In support of our hypothesis, the RNA/DNA primer/template hybrid bound with ~2-fold higher affinity than a DNA primer/template. The RNA primer also appears to reduce the maximal fraction bound compared to DNA. One possible explanation for this is the quantification for nucleic acid concentration was inaccurate and the RNA/DNA primer/template was less concentrated than DNA/DNA primer/template.

Figure 4.7: DNA/DNA primer/template binding to the T4 clamp loader:clamp complex is under tight binding conditions. T4 clamp loader was titrated into TAMRA-labeled DNA/DNA primer/template and clamp complex in the presence of ADP•BeF_3. Data could not be fit to an equation describe a bimolecular reaction and was fit to a tight-binding equation and the measured K_D = 56.1 ± 23 nM. Error bars represent standard error of the mean, SEM, for n=3.

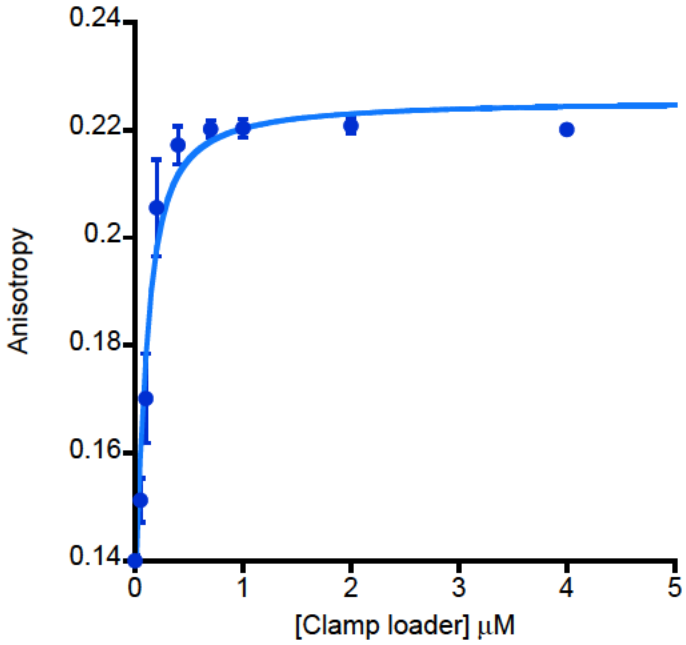
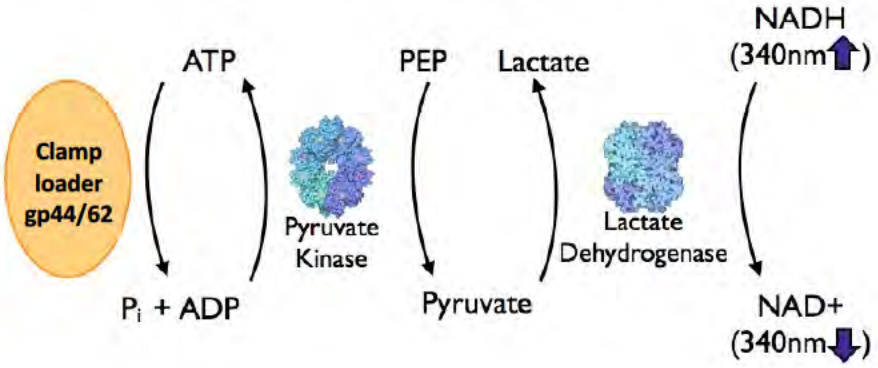
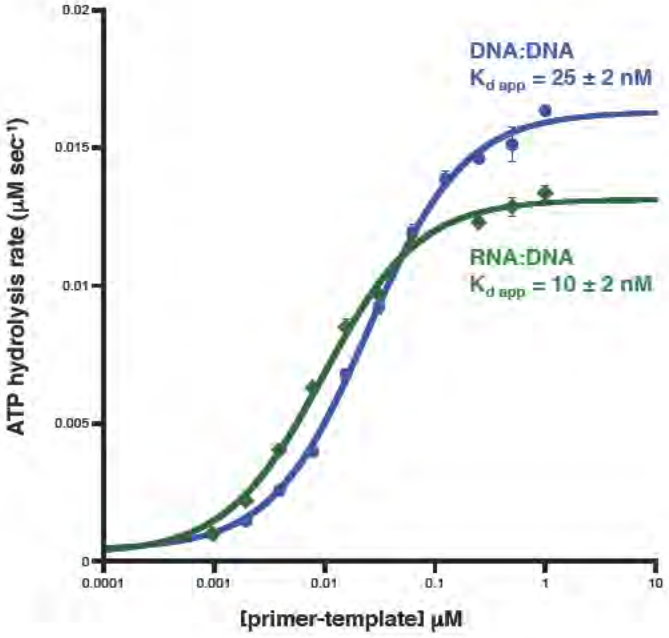


Figure 4.8: Coupled-enzyme ATPase assay approach. An *in vitro* assay acts as a readout for several events in the clamp loader reaction. ATP hydrolysis produces ADP and inorganic phosphate. ADP acts as a substrate in the next steps of this reaction, eventually providing a spectroscopic readout of a decreased absorbance at 340nm following the conversation of NADH→NAD^+.



To characterize the differences in binding between RNA/DNA and DNA/DNA hybrids, we attempted to co-crystallize the RNA/DNA primer with the clamp loader:clamp complex in the presence of ADP•BeF_3. However, these attempts were unsuccessful. The only crystals resulting from the screen yielded a structure of the isolated sliding clamp in a closed conformation (data not shown). Attaining the RNA/DNA:Clamp loader complex structure will require further screening for optimal crystallization conditions or an alternate method such as cryo-electron microscopy.

Figure 4.9: RNA/DNA primer template hybrid binds with 2-fold tighter affinity to the T4 clamp loader:clamp complex. Primer/template was titrated into T4 clamp loader:clamp complex bound with ATP. Data was fit to an equation describing a bimolecular reaction. $K_{d,app,\ DNA/RNA}$ = 10 ± 2 nM and $K_{d,app,\ DNA/DNA}$ = 25 ± 2 nM. Error bars represent standard error of the mean, SEM, for n=3.



Does ATP hydrolysis occur sequentially?

Primer/template binding triggers ATP hydrolysis in the pentameric clamp loader spiral (118, 126, 127). Because a recent crystal structure of the T4 clamp loader:clamp bound to primer/template DNA and the non-hydrolyzable analog ADP•BeF_3 showed the B-subunit in a post-hydrolysis state (103), we hypothesized that ATP hydrolysis initiates in the B subunit, and proceeds sequentially around the ring from B-subunit to C-subunit to D-subunit. This mechanism is in contrast to a random order, where subunits hydrolyze ATP independently of one another, or a concerted mechanism in which all active sites hydrolyze ATP simultaneously.

We used the weakly hydrolyzable ATPγS analog in a doping assay to distinguish between different mechanisms for ATP hydrolysis activity. As increasing amounts of ATPγS bind to the available sites in the clamp loader, there is a higher probability of decreased ATPase activity. If ATP hydrolysis is random, ATPase activity should decrease linearly with respect to ATPγS concentration. However, if ATPase activity is sequential, ATPase activity will have a steeper dependence on the ratio of ATP/ATPγS. Finally, if ATPase activity is concerted, ATPase activity will have a sharp dependence on the ratio of ATP/ATPγS, as the only ATP hydrolysis events will occur simultaneously if ATP is bound to all active sites. To quantify these differences, we generated mathematical models to distinguish between these mechanisms (**Figure 4.10a**). These models assumed three ATP hydrolysis events, and also assume all

binding sites are occupied by ATP or ATPγS. These models require the K_m for nucleotide binding to the clamp loader, which I separately measured to be 79 ± 7 μM and 48 ± 12 μM for ATP and ATPγS, respectively (**Figure 2.11**). I again used the enzyme-coupled ATPase assay, modified as a doping experiment wherein the total nucleotide concentration is constant, but the ratio of ATP to ATPγS is varied. The results of this experiment show a decidedly non-linear decrease in ATPase activity (**Figure 2.10b**), but the data I collected results in an even more pronounce curvature than all three of my models. The curvature of my data rules out a random model, leaving a sequential or concerted ATP hydrolysis mechanism as possibilities. The structure of the T4 clamp loader showing the B-subunit in a post-hydrolysis state argues against a concerted model, suggesting a sequential ATP hydrolysis mechanism. I will further discuss the discrepancies between my data and the derived models below.

Figure 4.10: ATP hydrolysis likely occurs sequentially.

a) Expressions describing a random (*top*) sequential (*middle*), or concerted (*bottom*) mechanism for ATP hydrolysis. Each expression assumes 3 occupied binding sites for T4 clamp loader. For a random model, the order of hydrolysis events is not taken into account. For a sequential model, activity is reduced depending on the subunit where ATPγS is bound. For a concerted model, the only ATPase activity will occur when ATP is bound in all three sites, and will simultaneously hydrolyze. [A] and [G] refer to the concentrations of ATP and ATPγS, respectively.

b) Clamp loader ATPase activity in the presence of increasing amounts of ATPγS/ATP (*blue circles*) shows a non-linear rate dependence on ATPγS. The models for ATP hydrolysis are shown in (*red*), sequential (*green*), or concerted (*orange*). A random model shows activity decreases linearly; sequential or concerted models show activity decreases non-linearly. Error bars show standard error of the mean, n=3.

Figure 4.10: ATP hydrolysis occurs sequentially.

a

Random:

$$E_T * k_{cat}\left(\frac{(\frac{3[A]^3}{K_{D,ATP}^3}) + (\frac{2[A]^2[G]}{K_{D,ATP}^2 K_{D,ATP\gamma S}} * 3) + (\frac{[A][G]^2}{K_{D,ATP} K_{D,ATP\gamma S}^2} * 3)}{1 + \left(\frac{[A]^3}{K_{D,ATP}^3}\right) + \left(\frac{[G]^3}{K_{D,ATP\gamma S}^3}\right) + \left(\frac{[A]^2[G]}{K_{D,ATP}^2 K_{D,ATP\gamma S}} * 3\right) + \left(\frac{[A][G]^2}{K_{D,ATP} K_{D,ATP\gamma S}^2} * 3\right) + \left(\frac{[A][G]}{K_{D,ATP} K_{D,ATP\gamma S}} * 6\right) + \left(\frac{[G]^2}{K_{D,ATP\gamma S}^2} * 3\right) + \left(\frac{[A]^2}{K_{D,ATP}^2} * 3\right) + \left(\frac{[G]}{K_{D,ATP\gamma S}} * 3\right) + \left(\frac{[A]}{K_{D,ATP}} * 3\right)}\right)$$

Sequential:

$$E_T * k_{cat}\left(\frac{(\frac{3[A]^3}{K_{D,ATP}^3}) + (\frac{2[A]^2[G]}{K_{D,ATP}^2 K_{D,ATP\gamma S}}) + (\frac{[A]^2[G]}{K_{D,ATP}^2 K_{D,ATP\gamma S}}) + (\frac{[A][G]^2}{K_{D,ATP} K_{D,ATP\gamma S}^2})}{1 + \left(\frac{[A]^3}{K_{D,ATP}^3}\right) + \left(\frac{[G]^3}{K_{D,ATP\gamma S}^3}\right) + \left(\frac{[A]^2[G]}{K_{D,ATP}^2 K_{D,ATP\gamma S}} * 3\right) + \left(\frac{[A][G]^2}{K_{D,ATP} K_{D,ATP\gamma S}^2} * 3\right) + \left(\frac{[A][G]}{K_{D,ATP} K_{D,ATP\gamma S}} * 6\right) + \left(\frac{[G]^2}{K_{D,ATP\gamma S}^2} * 3\right) + \left(\frac{[A]^2}{K_{D,ATP}^2} * 3\right) + \left(\frac{[G]}{K_{D,ATP\gamma S}} * 3\right) + \left(\frac{[A]}{K_{D,ATP}} * 3\right)}\right)$$

Concerted:

$$E_T * k_{cat}\left(\frac{(\frac{3[A]^3}{K_{D,ATP}^3})}{1 + \left(\frac{[A]^3}{K_{D,ATP}^3}\right) + \left(\frac{[G]^3}{K_{D,ATP\gamma S}^3}\right) + \left(\frac{[A]^2[G]}{K_{D,ATP}^2 K_{D,ATP\gamma S}} * 3\right) + \left(\frac{[A][G]^2}{K_{D,ATP} K_{D,ATP\gamma S}^2} * 3\right) + \left(\frac{[A][G]}{K_{D,ATP} K_{D,ATP\gamma S}} * 6\right) + \left(\frac{[G]^2}{K_{D,ATP\gamma S}^2} * 3\right) + \left(\frac{[A]^2}{K_{D,ATP}^2} * 3\right) + \left(\frac{[G]}{K_{D,ATP\gamma S}} * 3\right) + \left(\frac{[A]}{K_{D,ATP}} * 3\right)}\right)$$

b

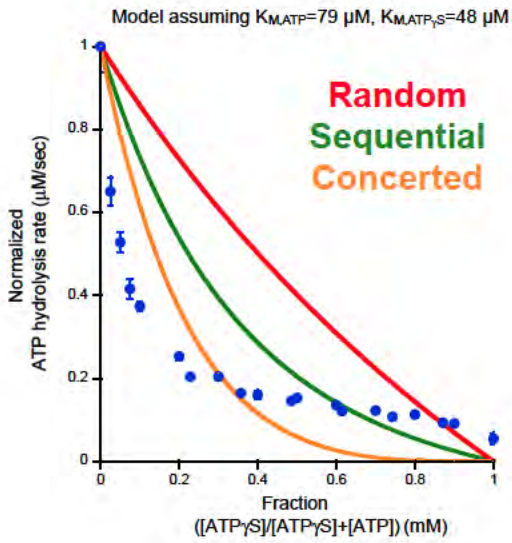
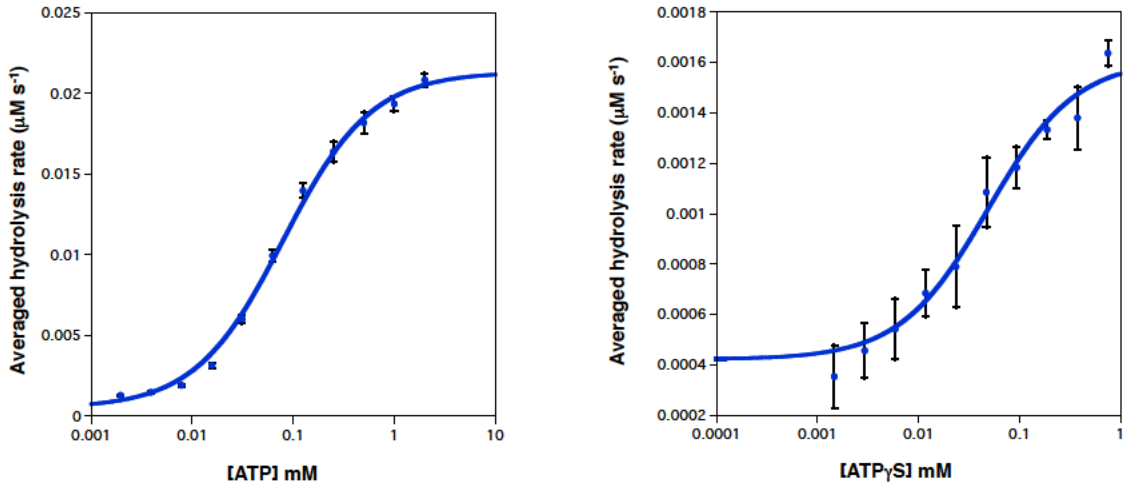


Figure 4.11: Respective affinities for ATP and ATPγS bound to T4 clamp loader.

ATP (*left*) or the non-hydrolyzable analog ATPγS (*right*) was titrated into T4 clamp loader:clamp complex bound with primer/temper DNA. $K_{m,ATP}$ = 79 ± 7uM; $K_{m,ATP\gamma S}$ = 48 ± 12 uM. Error bars show standard error of the mean, n=3



To further test if ATP hydrolysis is sequential, we performed a similar doping assay in which we use a clamp loader mutant that is deficient in ATP hydrolysis. In this assay, the total clamp loader concentration is held constant at 7.5nM, but the ratio of mutant/WT is varied. Based on the T4 clamp loader structure, Thr57 in the B C D and E subunits of T4 clamp loader helps coordinate Mg^{2+} in the active site (**Figure 4.12a**). The analogous threonine (T157) in the E coli clamp loader is required for ATP hydrolysis but is dispensable for DNA binding (116). Mutation of the analogous threonine residue in the sensor-1 motif of the related ATPase Hsp104 disrupts ATP hydrolysis (239). Therefore, we mutated the T4 clamp loader Thr57 to alanine, as we hypothesized this mutation

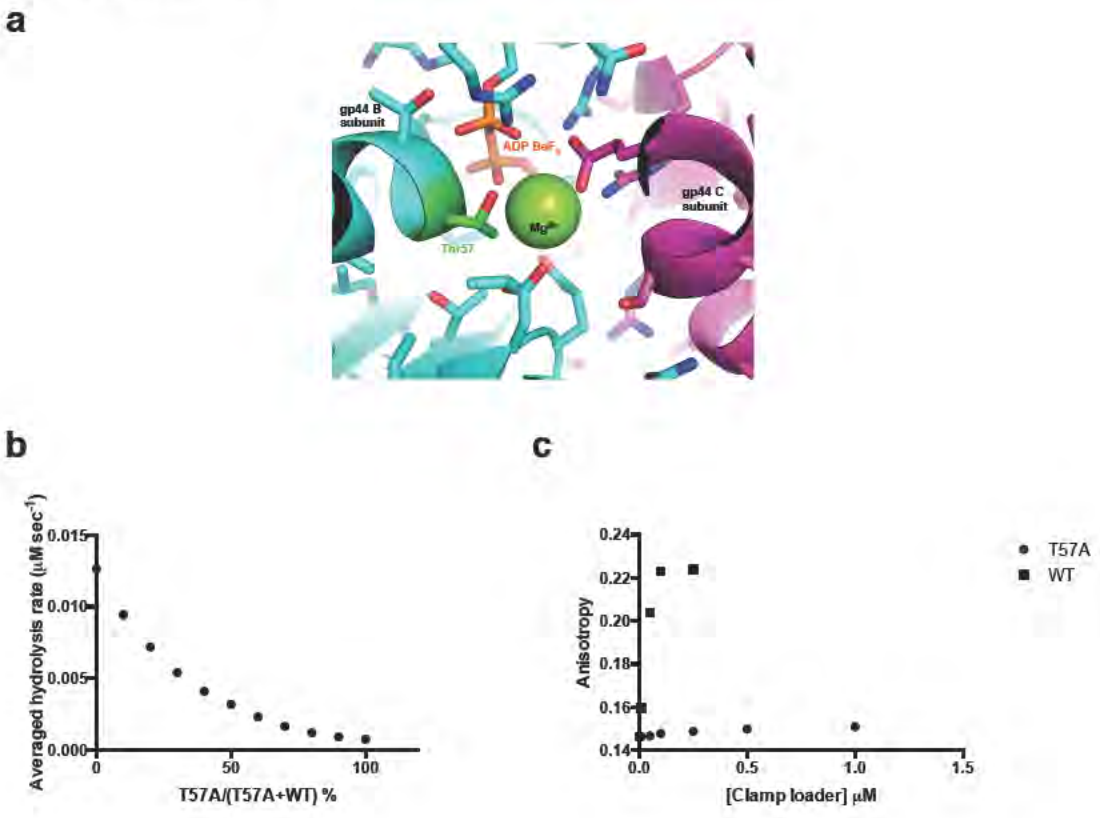
would disrupt ATP hydrolysis. In a modified doping assay, I held the overall clamp loader concentration constant, but increased the ratio of T57A/WT titrated into the assay. To allow for sufficient subunit exchange, especially given the low concentration, aliquots were mixed 24 hours prior to running the experiment. We expected to see a decrease in ATPase activity as we increased the ratio of T57A/WT because there is a higher probability of subunits deficient in ATP hydrolysis in the complex as we increase the mutant. Indeed, increasing the ratio of T57A/WT disrupts ATP hydrolysis. The steep dependence on the T57A/WT ratio also supports a sequential mechanism (**Figure 4.12b**). However, the T57A variant does not bind DNA with a measurable affinity (**Figure 4.12c**). Because primer/template binding is required for efficient ATP hydrolysis, the steep dependence on the T57A/WT ratio could be due to the hybrid clamp loaders inability to bind DNA effectively. Therefore, we cannot reasonably conclude a sequential hydrolysis mechanism using this mutant.

Figure 4.12: A clamp loader mutant disrupts hydrolysis but does not bind DNA

a) A conserved threonine residue in the clamp loader, T57A, coordinates Mg^{2+} in the active site

b) Doping T57A into WT clamp loader shows a non-linear decrease in ATPase activity. Error bars show standard error of the mean, n=3.

c) T57A does not bind TAMRA-labeled DNA primer/template compared to WT clamp loader (n=1).



Discussion

The clamp loader is essential for loading the sliding clamp onto primer/template junctions in an ATP-dependent manner. The field has satisfied our understanding for how the clamp loader engages with a primer composed of DNA (47, 103, 116), we lack the same insight for how a clamp loader engages with its preferred substrate, a RNA/DNA primer/template. Previous studies have studied how ATP hydrolysis occurs in the clamp loader (103), but there is a gap in knowledge for how ATP hydrolysis relates to the mechanism of clamp loading and clamp loader ejection.

An RNA primer binds the clamp loader tighter than a DNA primer

We aimed to structurally and biochemically characterize how an RNA primer interacts with the clamp loader. We first examined primer/template binding using an anisotropy assay that had been previously used (103, 124). The anisotropy data showed a steep dependence on the clamp loader concentration, indicating we were under tight binding conditions (**Figure 4.7**). Indeed, when fit to a tight-binding equation, the K_D was ~56 nM, lower than the concentration of DNA/DNA (100nM) in the assay. To avoid tight binding conditions, I could have used a lowered concentration of DNA; however, the binding reaction was very slow (each independent replicate took one day), so decreasing the DNA concentration would only aggravate this problem. Switching to a an enzyme-coupled ATPase assay, where we could more easily assay both primer/template

binding, as well as investigate ATPase activity further facilitated determining if an RNA primer was in fact a better substrate.

The data using ATPase activity as readout indicate that a RNA primer binds to the clamp loader with a 2-fold tighter affinity than a DNA primer (**Figure 4.9**). The higher affinity supports our hypothesis that a RNA primer would interact more favorably than a DNA primer. We believe the increase in affinity occurs because a RNA/DNA primer/template is already A-form in solution, and would not have to undergo conformational change like a B-form DNA primer. However, it has been estimated that there is a Gibbs free energy difference of ~2.8 kcal/mol for the conversion of B-form DNA to A-form for a DNA hexamer (240). This difference in free energy supports our hypothesis that an RNA/DNA would bind with tighter affinity, this energy difference actually corresponds to a ~113-fold difference in K_D at room temperature. This difference is much higher than my data only showing a 2-fold difference in K_D, making a crystal structure showing the form of RNA/DNA bound to the clamp loader of great interest. We also learned that the RNA/DNA primer/template also reduces the maximal fraction bound, which raised several questions. An important technical aspect to consider is perhaps quantification of our RNA/DNA concentration was actually less than observed, which would explain the reduction in bound primer/template. If this is not the case, several questions still remain—is the RNA primer reducing clamp loader enzymatic activity? Does the RNA primer bind more slowly than DNA? Is primer/template release slower with an RNA primer? A crystal structure of clamp

loader bound to an RNA/DNA hybrid would reveal how the RNA primer activates ATP hydrolysis.

An RNA/DNA, A-form crystal structure would potentially reveal a network of contacts that better represents the primary clamp loader interactions. The existing T4 clamp loader structure is in a pseudo-A form (103) and therefore may not possess interactions that occur with a true A-form duplex. Therefore a structure with an RNA primer may possess slightly different template strand interactions with the clamp loader, consequently altering contact with the switch residue, which would explain the possible decrease in enzymatic activity seen by the ATPase assay. Additionally, a structure would help explain energetic differences from potential conversions of B-form to A-form nucleic acid structures. If there is a 2.8 kcal/mol difference in free energy between B- and A-form, we would expect RNA/DNA binding to be much tighter than we observed. Perhaps a RNA/DNA primer/template undergoes an energetically unfavorable conformational change to pseudo-A form, which may account for the differences in binding. However, my crystallization attempts were unsuccessful using both the T4 and *E.coli* clamp loaders, so this remains an open question.

From examining the T4 co-crystal structure, we have identified a residue (gp44-S112) that may directly interact with an RNA primer more favorably than DNA primer. To test this hypothesis, we have prepared two variant clamp loaders S112A and S112K that we predict to perturb primer/template junction binding. S112A is predicted to remove the RNA-specific interaction; therefore, the affinity

for an RNA/DNA hybrid would be predicted to be reduced, but the affinity for a DNA/DNA hybrid would be unchanged. S112K is predicted to disrupt the interaction with an RNA/DNA hybrid, but provide a new interaction with a DNA/DNA hybrid. Thus, these two mutations make testable predictions with regards to the mechanism of primer/template junction binding. We have expressed S112A and S112K, but not yet tested for any differences for binding to a DNA or RNA/DNA primer/template. These experiments will be a focus of future studies.

ATP hydrolysis likely occurs sequentially

I sought to understand how ATP hydrolysis occurs around the clamp loader spiral. Based on a co-crystal structure of the T4 clamp loader showing the B-subunit in a post-hydrolysis state, we hypothesized ATP hydrolysis begins at the B-subunit and propagates around the spiral. Our preliminary data suggests that ATP hydrolysis in clamp loader occurs sequentially and not randomly. Using an ATPγS doping experiment, we observe a steep dependence of ATPase activity on ATP/ATPgS ratio. This data rules out a random ATP hydrolysis model. The measurements are in accord with the sequential model or concerted models, albeit with a more pronounced curvature. However, an “all or none” concerted model is not supported by previous studies in the clamp loader field; a study from T4 clamp loader using kinetic measurements showed two molecules of ATP hydrolyze sequentially, followed by additional ATP hydrolysis events (241). This

data argues against ATP hydrolysis occurring at all sites simultaneously. Additionally, a crystal structure of the T4 clamp loader that shows only one subunit has hydrolyzed ATP does not support a concerted mechanism (103). In a concerted mechanism, all molecules of ATP will hydrolyze simultaneously. We therefore rule out concerted as a probable ATP hydrolysis mechanism.

We derived the sequential model assuming three hydrolysis events. Previous studies have also reported that three hydrolysis events are sufficient for clamp loading activity (113). Surprisingly, a model describing four hydrolysis events would fit to this data better. Even though all four of the gp44 subunits in the T4 clamp loader can bind ATP, it is unlikely that all four can hydrolyze ATP due to the lack of an “arginine finger” contributing to the active site in the E-subunit. The arginine finger is a conserved basic residue that functions *in trans* to activate ATP hydrolysis in a neighboring subunit. The crystal structure of the T4 clamp loader does not reveal any arginines or lysines that are in position to coordinate catalysis in the E subunit (103). However, there are two basic residues in the A’ domain, K129 and R130, that are somewhat proximal to the bound ADP and Mg^{2+} in the E-subunit active site. These residues are too far (>18Å) from the gamma phosphate of ATP for any direct interaction, but it is possible that the A’ domain can undergo a large ~35° rotation to position either of these residues for catalysis. If so, either K129 or R130 may act as an arginine finger, which would support a four-hydrolysis event model. This could easily be investigated further via mutation and testing using the ATPase assay; if one of

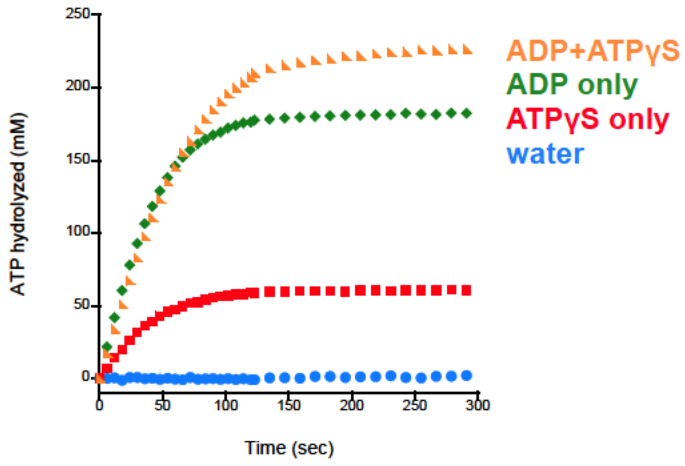
these residues was acting as an arginine finger, ATP hydrolysis rate would decrease by ¼ relative to the WT.

There are several alternative explanations for the steepness of the data due to technical aspects of the ATPγS doping experiment. We show that the clamp loader is still active in the presence of ATPγS, indicating that it hydrolyzes ATPγS, albeit slowly. Once converted to ADP, pyruvate kinase will convert ADP to ATP, which will perturb the ATP/ATPγS ratio. Moreover, we determined our stock of ATPγS contains contaminating ADP, which will be converted to ATP by pyruvate kinase in the coupled enzyme mix. We monitored the effect of nucleotide on pyruvate kinase in the assay in the absence of clamp loader (**Figure 4.13**). The signal from ATPγS alone suggests contaminating ADP in our stock. When ATPγS is mixed with ADP, the results are additive compared to ADP or ATPγS alone. We estimate that there is a ~25% contamination of ADP in the ATPγS stock, which make the interpretation of our ATPγS doping assay nearly impossible. Because of the success from data I discussed in Chapter II, I abandoned these studies to focus on the sliding clamp variant.

A previous study on the T4 clamp loader suggested ATP hydrolysis is cooperative in the presence of primer/template DNA, which supports a sequential hydrolysis model; if hydrolysis occurs at one subunit, it could promote hydrolysis at an adjacent site (242). Conversely, our doping experiments do not appear to be cooperative; when ATPγS binding data was fit to a Hill equation, the Hill coefficient was 0.98 ± 0.28 μM. Likewise, when ATP binding data was fit to a Hill

equation, the Hill coefficient was 1.2 ± 0.1 μM. Therefore we cannot reasonably conclude ATP binding is cooperative given the errors of the Hill coefficients. However, the previous study did not fit their data to a Hill equation, and their data shows only a slight sigmoidal shape, so the extent of cooperativity is not quantified.

Figure 4.13: ATPγS contains contaminating ADP. Water (control), ATPγS, ADP, or ATPγS+ADP were added to ATPase assay enzyme mix and activity was monitored over time. As expected, a water-only negative control yielded no activity. Both ADP and ATPγS showed significant activity over water-only. When ADP+ATPγS were mixed, there was an additive effect



Because of the technical limitations of the ATPγS doping assay, we designed an experiment with potentially more discriminating power. Based on previously published mutants (116, 239), we mutated a conserved threonine residue within the active site, gp44-T57A, to disrupt hydrolysis. We modified the doping experiment to increase the ratio of T57A/WT and observed a non-linear decrease in ATPase activity. This data would support a sequential ATP hydrolysis mechanism, but unfortunately, this mutant does not bind DNA. The abolished binding was surprising, given the equivalent mutation in the *E. coli* clamp loader is fully competent to bind DNA (116). Regardless, we cannot conclude the results from this doping assay are due to a sequential ATP hydrolysis mechanism.

Several other studies have also proposed a sequential ATP hydrolysis mechanism in clamp loaders (243, 244). However these studies suffer from similar difficulties, particularly disruption of DNA binding, similar to my results with the T57A variant. In one study using the E. coli clamp loader, the arginine finger in the γ (R169) and δ’ (R158) subunits, equivalent to the B/C/D and E subunits respectively, were mutated to alanine. While γ-R169A is inactive, δ’-R158A retains some ATPase activity, due to its location in the complex (equivalent to the E-position). The authors argued that because the mutated arginine finger, δ’-R158A, would be involved in the final hydrolysis event of the sequence, some activity remained from ATP hydrolysis occurring in the subunits at the bottom of the spiral. However, as previously stated, these mutants do not

bind DNA, bringing into question the validity of the interpretation. A second study uses the eukaryotic clamp loader, the RFC complex, and mutated the Walker-A lysine residue to alanine. While their results are also consistent with a sequential hydrolysis mechanism, these mutants also fail to bind DNA (244). Additionally, this study did not account for ATPγS being weakly hydrolysable.

The eukaryotic clamp loader, RFC, is a heteropentamer, so designing mutations to target specific subunits is more easily facilitated in this system. Thus, if we could identify a mutation that disrupts ATP hydrolysis but not DNA binding, then we could determine an order for ATP hydrolysis by reconstituting RFC clamp loader variants with individual subunits inactivated for ATP hydrolysis. We would predict that inactivation of the B subunit would have the greatest reduction on ATPase activity. Conversely, inactivation of downstream subunits would have less of an effect on ATPase activity. A strong candidate for such a variant is the conserved sensor-1 motif threonine mutated to alanine in RFC1-T360A, RFC2-T72A, RFC3-T60A, and RFC4-T56A. These mutations in this system may therefore be able to definitively address the order of ATP hydrolysis.

To conclude, the clamp loader reaction contains several critical steps, and our investigation supports potential mechanisms for two of those steps; first, how primer/templates bind to the clamp:clamp loader complex, and second, if ATP hydrolysis occurs sequentially. My data here show RNA/DNA primer/template binds with 2-fold tighter affinity, but raises several questions about how the

primer strand interacts with the clamp loader. I demonstrate that ATP hydrolysis is likely sequential, but more work is required to tease apart the order of ATP hydrolysis around the clamp loader spiral. Future work using the structural and biochemical approaches described herein will help reveal these mechanisms, which would further elucidate general mechanisms of AAA+ machines.

CHAPTER V

Discussion

Discussion

My overall goals for this thesis were to structurally and biochemically characterize regulatory roles for two protein complexes essential for DNA replication: the sliding clamp and the sliding clamp loader. The clamp surrounds DNA and acts as a scaffold for interacting partners to carry out their function on DNA. As such, the clamp interacts with a vast network of proteins, orchestrating many cellular processes including DNA replication and repair, cell cycle regulation, chromatin remodeling, and epigenetic inheritance (57). The ring-shaped clamp does not interact with DNA on its own accord; it must be actively placed onto primed DNA by the sliding clamp loader in an ATP-dependent reaction (100). In the presence of ATP, the sliding clamp loader binds and opens a closed clamp. The clamp loader couples primer/template binding, and subsequent ATP hydrolysis to clamp closure around DNA (100). Although much is understood about the clamp and clamp loading process, I sought to address gaps in the current understanding. I summarize my findings for the sliding clamp and sliding clamp loader, and discuss open questions and future directions, below.

Summary: The sliding clamp

In Chapter II, I investigated the structural and biochemical consequences of a recently identified and novel disease-causing mutant in the human sliding clamp, PCNA (proliferating cell nuclear antigen). This mutant, $PCNA^{S228I}$, had

been previously hypothesized to cause defects in DNA repair (35, 185), with clinical symptoms that mimic other related DNA repair disorders. However, the mechanism underlying the disease was not understood. Of particular interest was that this mutation, in a protein as critical as PCNA, was not lethal.

My structural studies reveal that the disease mutation disrupts a promiscuous binding site called the IDCL (interdomain connecting linker) in PCNA (**Figures 2.2, 2.6**). Because PCNA interacts with >100 partner proteins at the IDCL, I investigated whether the disrupted binding site was still compatible for known PCNA interacting partner (PIP) binding. Structural and biochemical data revealed the IDCL of the S228I variant is pliable and accommodates PIP ligands idiosyncratically (**Figures 2.7, 2.9, 2.11**) (63).

Of the three PIPs we tested for binding, FEN1 and RNaseH2B showed substantial loss of binding affinity to the mutant variant. This data supports the claim from a previous study that the underlying cause of the disease is a defect in DNA repair (35, 185), but the previous authors specifically described a defect in nucleotide excision repair (NER), and neither FEN1 nor RNaseH2B function in NER. The RNaseH2B data also suggests a defect in ribonucleotide excision repair (RER) could contribute to the disease. While our major findings in Chapter II were specific to the mutant PCNA, our comparison with WT PCNA revealed important insights into normal PCNA function that could have implications for PCNA dynamics, regulation, and function, which I discuss further below.

In order to identify the source of the disease phenotype, I developed an *in vivo* model system to further understand how the PCNA variant affects cellular pathways. In Chapter III, I describe a fission-yeast model system to study the cellular effects of the S228I mutant variant. Using a strain harboring *pcn1-S288I* as the sole copy of *pcn1*, I conducted growth assays after treating cells with DNA damaging agents. I am limited in making strong conclusions about potential defects in DNA repair pathways because the DNA damaging agents only revealed a weak phenotype (**Figures 3.5-3.8**). We were initially surprised that there was no response to UV light, given patients are highly sensitive to sunlight; however, there is an additional DNA UV repair pathway in fission yeast that likely protects *pcn1-S228I* from significant UV damage (217, 218).

Open Questions: The sliding clamp

A major unanswered question in the PCNA field is how binding of PIPs are regulated in a spatiotemporal manner. Because $PCNA^{S228I}$ disrupts a major binding site, we sought to use this mutant as a tool to probe what drives binding specificity. We obtained new insights into normal PCNA function from studying this mutant, including that the IDCL binding site may be more flexible than previously indicated from static crystal structures. Additionally, during our analysis of clamp structures, particularly when analyzing the crystal structure of WT human PCNA bound to the PIP-box peptide from polymerase κ, we observed that PCNA contains an uncharacterized disulfide bond (60). We believe this

disulfide bond may regulate binding specificity. I will discuss implications of these findings, and how we can further investigate them below.

What are the implications of a flexible binding site?

The data from Chapter II reveals that the IDCL is more flexible than current structural models suggest, likely because alternate conformations are often masked in static crystal structures. Are the novel conformations we observed with $PCNA^{S228I}$ structures accessible for the WT protein? Additionally, what are the physical constraints to which conformations the IDCL can adopt? In the disease-causing mutant variant, the IDCL mobility is actually dampened, as evidenced by a reduction in normalized B-factors relative to other PCNA structures (63, 64). These results suggest flexibility is important for cellular fitness. Knowing that PCNA is a promiscuous binding partner, we hypothesize that dynamics within the IDCL regulate partner binding, and that a flexible IDCL region is important for PCNA to maintain interactions with all of its binding partners. We can use several strategies to test these hypotheses.

Recently developed x-ray crystallography techniques can help probe differences in IDCL conformations. Low-temperature (~100K) data acquisition strategies typically identify only the most populated state (245-247). As such, most conformational heterogeneity is lost. Alternatively, data acquisition at room temperature can identify multiple higher-energy states (246). For PCNA, these higher-energy states might reveal dynamics within the IDCL that regulate binding

interactions. Room temperature data collection does cause higher radiation damage, but recent advances in data acquisition (namely, larger x-ray dosage combined with faster data collection) have been shown to reduce radiation damage (247). Additionally, advances in ensemble x-ray crystallographic refinement can help identify alternative conformations for data collected at low temperature or room temperature: qFit software automates building and refinement of alternative conformations within models (248, 249), while Ringer analysis reveals additional conformational diversity in electron density maps (below 1σ) that would otherwise be overlooked as noise with traditional model building and refinement (246, 250). However, in order to take advantage of these techniques, we must improve our crystals of WT and mutant PCNA to yield higher resolution diffraction data (<2Å). We have some preliminary data (see Appendix A1) on our efforts to improve crystal quality. I have identified multiple crystal forms grown in various conditions, which we can readily screen at our home source. This maximizes the potential to improve resolution to the range required to study the IDCL dynamics. This will be the focus of future studies.

What are the cellular effects of changes in the IDCL region? Based on data in Chapter II, we know that the mutant variant disrupts a promiscuous binding site in PCNA, and that the binding site is especially flexible. Surprisingly, the IDCL in the disease-causing variant is actually less mobile than the WT, suggesting that the IDCL flexibility is important for fitness. If mutations in the IDCL disrupt fitness *in vivo*, I would expect these mutations would reduce

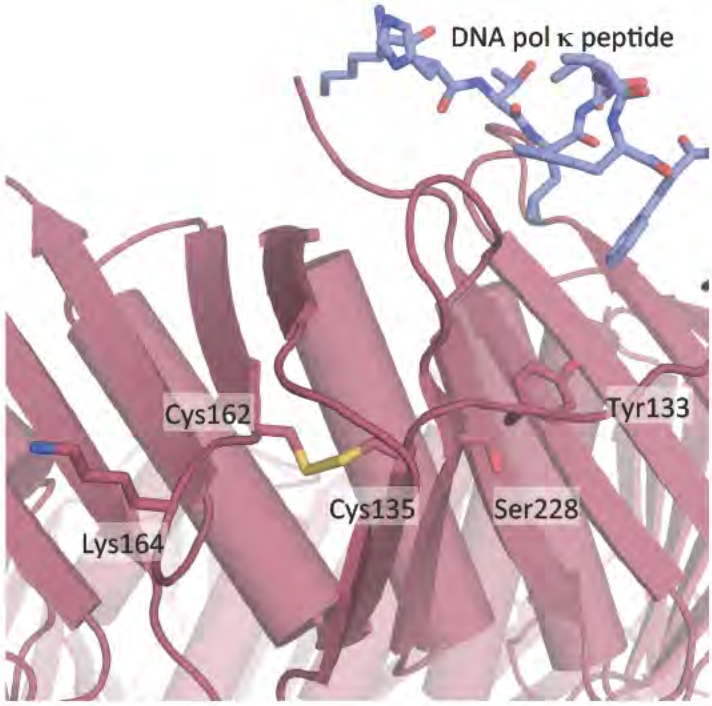
flexibility within the IDCL. Additionally, decreased fitness might also identify particularly important residues within the IDCL that are critical for PIP binding. We can measure fitness resulting from changes within the IDCL using a systematic approach called EMPIRIC (Exceedingly Meticulous and Parallel Investigation of Randomized Individual Codons) (251-253). EMPIRIC would allow us to investigate all possible amino acid substitutions in the IDCL region and their effects on growth rate in yeast. Ideally, we could identify regions within the binding site that are critical for dynamics, as well as partner interactions. This would comprehensively establish how PCNA interacts with so many proteins in a regulated manner.

In Chapter III, I describe the generation of a valuable tool for further investigation of any PCNA-mediated pathways disrupted by *pcn1-S228I*. Future studies could selectively perturb particular DNA repair pathways to determine which are affected by the S228I variant. As discussed in Chapter III, we could target RER to investigate if the loss of RNaseH2B binding to the PCNA variant causes an RER defect. Future studies using this strain could take advantage of high-throughput techniques like synthetic genetic arrays (SGA). SGA identifies synthetic lethality pairs from a query allele (*pcn1-S228I*) when crossed to an ordered array of gene deletion strains, which are then assayed for growth defects (254). Because this mutant variant appears to have a specific defect in DNA repair, it would be informative to learn how cellular networks, particularly DNA repair networks, are connected in the mutant background.

Can the redox state of PCNA regulate binding specify?

During our initial analysis of the PCNA^{S228I} structure, we identified a pair of cysteines, Cys135 and Cys162, which are proximal to the site of the mutation, Ser228. The thiol groups of the cysteines are close (~3.5 Å) to one another, with geometry favorable for adopting a disulfide bond. These cysteines are also near Lys164, which serves as a site for post-translational modification by mono- and poly-ubiquitnaylation and SUMOylation (56). Upon reviewing the available crystal structures of WT PCNA, we discovered that the co-crystal structure of PCNA bound to the PIP box peptide from Pol κ was modeled with these cysteines in a disulfide bond (**Figure 5.1**) (60). Several proteins involved in DNA repair pathways have been previously shown to act as redox sensors, including AP endonuclease 1 and XRCC1 (255, 256). These proteins have been shown to directly interact with PCNA (189, 191). Additionally, the oxidation state of these proteins has also been shown to regulate binding partners. Intriguingly, Ubc9, which is the enzyme that SUMOylates PCNA at Lys164 (87), was recently shown to be controlled by redox conditions in the cell (257). These studies suggest that this disulfide motif in PCNA may control the binding of PIPs. Additionally, because XRCC1 lacks a PIP-box motif (191), this interaction may be specifically regulated by the oxidation state of PCNA. To our knowledge, there are no reports characterizing PCNA as a redox sensor. Therefore, I hypothesize the oxidation state of PCNA modulates binding specificity to various PIPs.

Figure 5.1: A previously uncharacterized disulfide bond may act as a redox sensor for regulating PCNA interactions.
A crystal structure of human PCNA bound to the PIP-box peptide from polymerase κ. Cys135 and 162, shown in sticks, are modeled as a disulfide bond. The site of the mutation, PCNAS228I, is proximal to the disulfide. Also shown is the site for PCNA post-translational modifications, Lys164.



This hypothesis could be investigated in multiple ways. To identify if this disulfide bond forms under physiological conditions, we can pulldown PCNA from established cell lines (HeLa cells, for example), and compare protein mobility on reducing or non-reducing SDS-PAGE gels. As another approach we can purify WT, $PCNA^{C135A}$, $PCNA^{C162A}$, or the double mutant $PCNA^{C135A/C162A}$ in either the absence or presence of reducing agent. I have already established a protocol for PCNA purification in the absence of reducing agent. With purified protein in the oxidized or reduced state, we can test for differences in binding to PIP-box peptides using my established protocols for PCNA isothermal titration calorimetry. Given oxidative species generate responses from DNA repair pathways (258), I expect that PIPs involved in DNA repair would bind to oxidized PCNA with tighter affinity.

If we have success with identifying alternate conformations of the IDCL using multi-conformation ensemble refinement techniques, it would also be of interest to investigate if a disulfide bond in PCNA alters IDCL conformations in conjunction with altering binding affinity for PIPs. Advances in data collection would also help circumvent issues for x-ray radiation damage that can destroy disulfide bonds (259).

Given the vast number of clamp binding partners, I would expect there would be multiple ways to modulate these interactions. If we can further understand how binding is regulated in the WT protein, we can easily apply similar approaches to determine why binding is abrogated for the disease-

causing variant PCNA^{S228I}. These findings would provide insight for not only the underlying mechanism for the disease and identifying possible therapeutic approaches, but also as a means of characterizing the hierarchy of how PCNA interactions occur. Finally, reagents generated here will be useful for teasing apart how these actions are coordinated spatially and temporally in the cell.

Summary: The sliding clamp loader

In Chapter IV, I investigated two critical steps in the clamp loader reaction: primer/template binding and ATP hydrolysis. Most previous biochemical and all structural studies investigating primer/template binding to the clamp loader used a DNA primer, when, in fact, most clamp loaders primarily use an RNA primer (22). I sought to understand how the T4 bacteriophage clamp loader engages with an RNA/DNA primer/template. A co-crystal structure of T4 clamp loader:clamp:primer/template duplex DNA showed that DNA was A-form, instead of the more common B-form (29, 103). We hypothesized a RNA primer would interact more favorably than a DNA primer because RNA/DNA duplexes are already A-form. Binding studies show that an RNA primer does bind with 2-fold tighter affinity than a DNA primer, but may reduce the maximal fraction bound (**Figure 4.9**). Crystallization of the full complex proved difficult and I was unsuccessful in obtaining a structure that would help explain these differences.

In addition to primer/template binding, I sought to understand the sequence of ATP hydrolysis around clamp loader spiral. Based on a co-crystal

structure showing the B-subunit in the clamp loader in a post-hydrolysis state (103), we hypothesized that ATP hydrolyzes sequentially around the spiral, beginning from the B-subunit. I generated mathematical models to distinguish between random, sequential, and concerted hydrolysis mechanisms (**Figure 4.10a**). Data from an ATPγS doping experiment ruled out a random ATP hydrolysis mechanism (**Figure 4.10b**), and previous studies have ruled out a concerted mechanism (103, 241). Although the data suggests a sequential mechanism, the data does not perfectly fit our model, which is likely due to technical aspects of the assay, but could also be due to faulty assumptions in the model, which I will discuss below. I tried to corroborate these results through an alternate mutant doping strategy, but the T57A mutant does not bind DNA (**Figure 4.12c**). Therefore we could not conclude a sequential ATP hydrolysis mechanism using this mutant due to the coupling between primer/template binding and ATP hydrolysis (118, 126, 127).

Open questions: The sliding clamp loader

The limited data I collected in Chapter IV does not fully address the two goals I hoped to accomplish. For RNA/DNA primer/template binding, we are limited in making conclusions for how RNA interacts with the clamp loader without a crystal structure. By no means have we exhausted options for crystallography; my attempts to crystallize the clamp loader bound to RNA/DNA primer template were not comprehensive. My approach, including

primer/template design, were modeled after the T4 clamp loader structure (103). Future experiments should vary primer/template lengths, which may aid crystallization. My efforts to crystallize clamp loader components from *E. coli* were brief and focused exclusively on conditions established for a DNA/DNA primer/template (116). Future studies could expand upon the sparse screening conditions or utilize an orthologous clamp loader. Structural studies would help address how an RNA primer interacts with the clamp loader, but they also could reveal a yet-unknown intermediate in the clamp loader reaction: what a clamp loader bound to an open clamp looks like in the absence of DNA.

Optimizing the biochemical approach may provide insight into clamp loader RNA versus DNA primer recognition. I abandoned the anisotropy assay to measure binding primarily due to the tight binding of duplexes relative to the concentrations used in the assay, but the binding reaction was also very slow. Future work could explore utilizing lower DNA concentrations, despite the risk of a reduced signal and slower reaction times. Exploring conjugation of different fluorophores to the substrate DNA may provide a better signal to measure binding at lower DNA concentration. Optimization of the anisotropy assay and further utilization of the enzyme-coupled assay provide a dual approach to investigating clamp loader mutants. Future experiments could thus identify residues critical the RNA/DNA binding as well as discriminate between the different duplexes.

My studies into sequential versus random ATP hydrolysis can also be expanded. My data demonstrate that hydrolysis is likely sequential. However, a four-hydrolysis event model is in better agreement with the data relative to a three-hydrolysis model, which was puzzling given three subunits, B, C and D, hydrolyze ATP and are sufficient for clamp loading; the A-subunit lacks the AAA+ fold and does not bind ATP (103, 110, 113). This is due to a conserved feature of AAA+ ATPases called the arginine finger (106). The arginine finger reaches across subunit interfaces into the active site of a neighboring subunit to stabilize the negative charge on the γ-phosphate of ATP and aid catalysis. For example, the arginine finger in subunit C reaches into the active site of subunit B. Because the clamp loader is a pentamer, with a gap between subunits A and E, there is no arginine finger to present to the active site of subunit E. However, from examining the T4 clamp loader crystal structure, there are two putative arginine fingers located in the A' domain. As discussed in Chapter IV, these two basic residues, K129 or R130, may support a fourth ATP hydrolysis event, adding to the proposed three events. Ideally, to discriminate between a three- vs. four-event hydrolysis mechanism, future studies should probe these possibilities using a mutational approach; if either of these basic residues were acting at the arginine finger, ATP hydrolysis rate would decrease ¼ relative to the WT. The ideal mutant to identify for a doping experiment would disrupt ATP hydrolysis, but maintain DNA binding because of the coupling between these two events in the clamp loader reaction. A potential candidate for this mutant is the Walker-B

catalytic glutamate, E108 in T4 clamp loader gp44 subunits. Several other AAA+ ATPases, include the membrane trafficking proteins AAA+ NSF, p97, VPS4, or the unfoldase ClpX have shown that mutating the catalytic glutamate to glutamine maintains structure and function, but disrupts ATP hydrolysis (260-262).

Overall, the work described herein describes regulatory roles for the sliding clamp loader. RNA binds to the clamp loader more tightly than DNA, which is expected given it is the most common substrate. ATP hydrolysis does not occur randomly and is likely sequential, in line with other related ATPases. Future structural studies will reveal the mechanisms for these findings in atomic detail.

Appendix 1

Crystallization optimization of WT PCNA and

PCNA^{S228I}

Crystallization optimization of WT PCNA and PCNA^{S228I}

In Chapter II, we investigated the structural consequences of a disease-causing variant of PCNA, PCNA^{S228I} (63). A crystal structure of the mutant variant shows the mutation induces a large-scale reorganization of the major binding site in PCNA, called the IDCL (interdomain connecting loop). Two additional PCNA^{S228I} crystal structures bound to peptides from PCNA interacting partners (PIPs) each reveal a unique conformation in the IDCL, which suggests this region is unusually flexible. To support this observation, we utilized normalized B-factors of the IDCL from WT and PCNA^{S228I} structures to investigate flexibility. The PCNA^{S228I} analysis revealed that the highest degree of motion occurs in the IDCL region (**Figure 2.2d&e**). However, when compared with the relative B-factors of the WT structure, the WT IDCL is nearly 2-fold higher than the mutant (63, 64). Despite having higher relative B-factors, and presumably higher flexibility than PCNA^{S228I}, all existing WT PCNA structures exhibit the same IDCL architecture (58-62) (63, 64) (**Figure 2.4**). The mutant IDCL therefore appears more rigid and less able to bind some protein partners.

We hypothesize that the IDCL is exceptionally flexible, but these dynamics have been hidden in static crystal structures. Studies of PCNA^{S228I} highlight structural diversity within the IDCL binding site, yet the mutant exhibits reduced binding to some PIPs (63). Taken together, this suggests that the IDCL conformation is especially important for PCNA binding partners. We hypothesize that alternative conformations of the IDCL in hPCNA-WT contribute to the

specificity of PIP binding. We aim to probe differences in IDCL conformations using multi-conformation ensemble refinement techniques (See Chapter V for discussion) (249, 263). However, these techniques require high-resolution diffraction (<2Å), and must withstand elevated radiation damage that occurs during room-temperature data acquisition.

Here I present crystallography conditions to improve the quality of WT and mutant PCNA crystals for multi-conformation ensemble refinement. Previously published structures of *apo* WT PCNA are at 2.3Å resolution (64). The structure I solved of *apo* $PCNA^{S228I}$ is at 2.7Å resolution, but I did not pursue any optimization of conditions (i.e., crystal growth conditions, freezing conditions, cryogenic buffers, etc.), so it is reasonable to expect improved quality crystals after optimization.

Materials and Methods

All proteins were crystallized using the hanging-drop vapor diffusion method. Drops contained 0.9µL:0.9µL of protein:buffer. To improve crystals for $PCNA^{S228I}$, hand trays were set up to screen around 100mM sodium acetate (pH 4.0, pH 4.5, pH 4.6) or 100mM Tris (pH 8.5) and between 1.7-2.2M ammonium sulfate. Suitable cryogenic buffers contain 25% (v/v) glycerol and elevated concentrations of ammonium sulfate. Drops contained 5-20mg/mL protein.

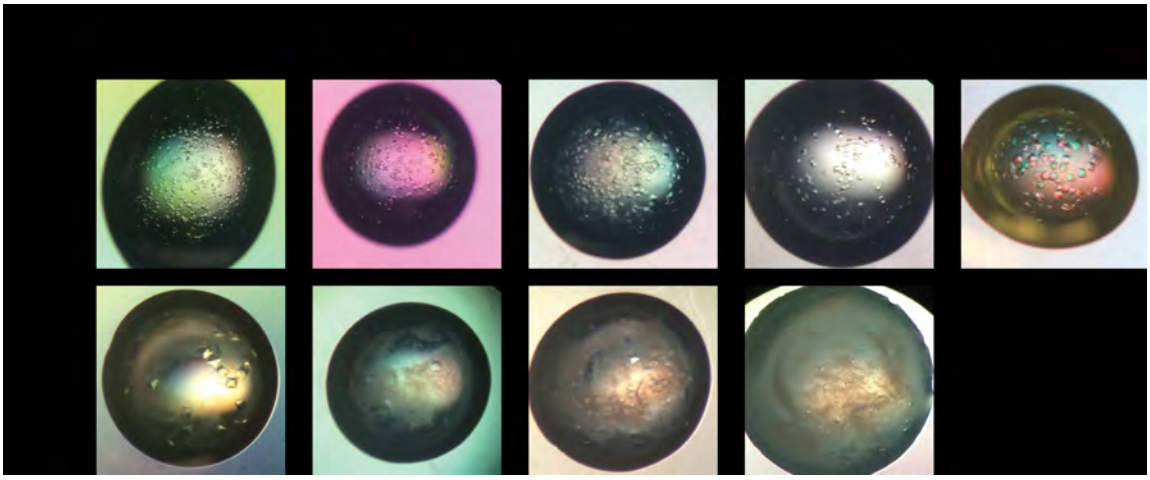
To improve crystals for WT PCNA, hand trays were set up to screen around 100mM sodium acetate (pH 4-5.5), 15-30% PEG 400 (which also serves

as a cryoprotectant), and 100mM calcium acetate. Drops contained 5-15mg/mL protein.

Results and Discussion

PCNA^{S228I} crystals form in a range of acidic pH conditions. Diffraction quality crystals formed in sodium acetate buffer between pH 4.5 and 4.6 (**Figure A1.1**). Small, jagged plate crystals form at low pH (4.0) (**Figure A1.3**), and crystals do not form at pH 5.0-5.5, or Tris pH 8.5 (not pictured). Within the pH 4-5.5 range, PCNA^{S228I} forms diamond shaped crystals.

Figure A1.1: PCNA^{S228I} crystal optimization.
PCNA^{S228I} crystallizes in ammonium sulfate between sodium acetate buffers, pH4.5-4.6



PCNA^{S228I} crystal nucleation and growth rate can be controlled through modulating precipitant concentration. In high precipitant conditions crystals form immediately after setting up drops, and most conditions yield crystals within 48 hours. Decreasing the amount of ammonium sulfate slowed the crystal formation rate (crystals formed within one week), and significantly increased crystal size likely because of fewer nucleation events. Increasing crystal size could potentially provide protection against radiation damage.

PCNA^{S228I} crystals grown in hanging drops with 5-20mg/mL protein did not exhibit a strong correlation between protein concentration and crystal size. In order to optimize crystallization and improve diffraction, I recommend optimizing drop volume to check for volume dependency, experimenting with various protein:buffer ratios, including additive screens, and testing crystal growth at varying temperatures. By optimizing these variables, one can grow large single crystals that diffract well. Additionally, I also recommend screening different cryogenic protectants for freezing PCNA^{S228I} crystals. Cryo-protectants beyond 25% v/v glycerol were not extensively explored. Ideal freezing is critical for achieving high-resolution diffraction. While I have proposed room-temperature data collection, optimizing freezing conditions could still prove beneficial; the ensemble refinement strategies I discussed in Chapter V can potentially identify alternate IDCL conformations from data collected at either low temperature or room temperature.

WT crystals form between pH 4.5-5.0 at protein concentrations between 5-15mg/mL. Interestingly, all WT crystals require at least 100mM calcium acetate for crystal formation. Calcium acetate concentration also requires optimization. We observe multiple crystal forms with WT conditions; we see starbursts at higher protein concentrations or increased precipitant concentration, and cubes of various sizes at lower precipitant concentration (**Figures A1.2, A1.3a**). The current starburst-form crystals diffract out to ~2.6Å (data was collected at low-temperature), and are a distinct, novel crystal form from the previously published WT structure (**Figure A1.3b**; courtesy of B. Hilbert) (64). Being novel suggests optimization may allow us to achieve diffraction higher than the published WT crystal form.

Figure A1.2: WT PCNA crystal optimization.
WT PCNA crystallizes in a range of PEG400 between sodium acetate buffers, pH4.0-5.5. 100mM calcium acetate is required for crystallization.

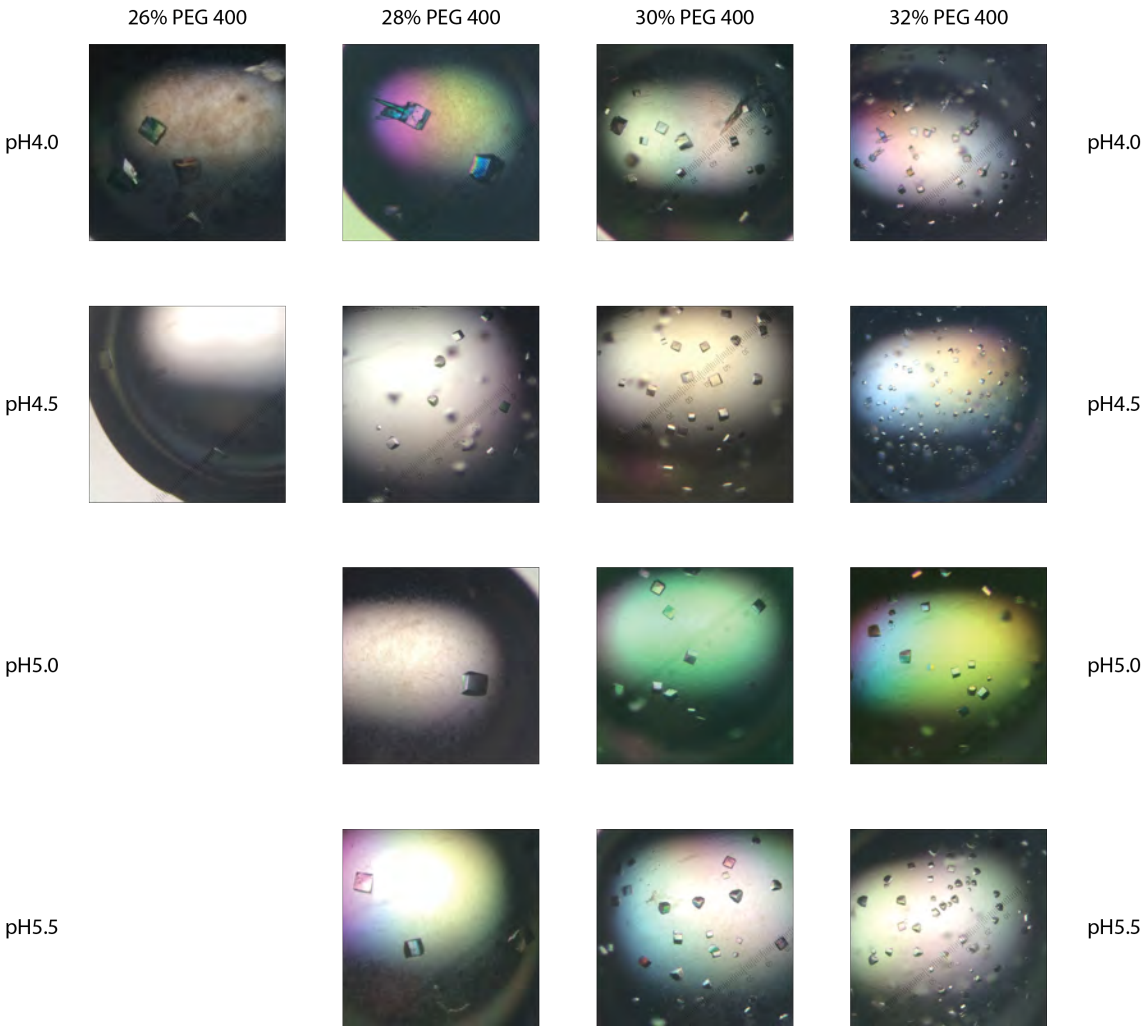
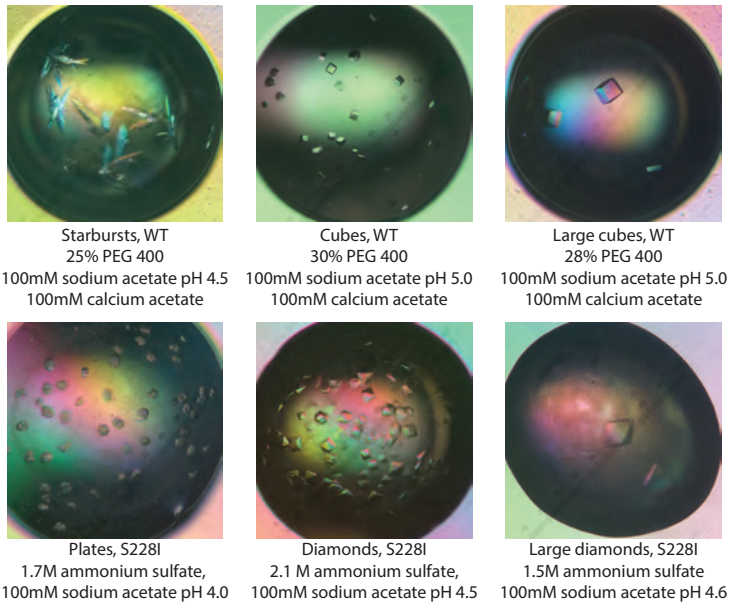
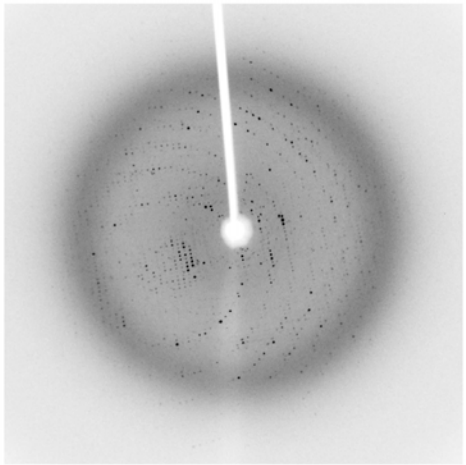


Figure A1.3: Summary of PCNA crystal forms and preliminary diffraction data.
a) Two crystal forms are seen with each WT PCNA and PCNA^{S228I}
b) A starburst crystal (from 25% PEG 400, 1mmM sodium acetate pH4.5, 100mM calcium acetate) of WT PCNA diffracts to ~2.6Å (courtesy of B. Hilbert).

a



b



Appendix 2

Purification of the intact exocyst complex

Purification of the intact exocyst complex

The following work was done in Dr. Mary Munson's laboratory, or collaborators' laboratories (noted in text). CMD contributed the following to the attached publication (264):

I constructed a set of C-terminally Protein A (PrA) tagged exocyst subunits, integrated into the genomic locus of *Saccharomyces cerevisiae*: SEC3, SEC5, SEC6, SEC8, SEC10, SEC15, EXO70, and EXO84. I validated these strains by sequencing, western blot, and MALDI-MS. I also constructed a partial set of C-terminally GFP-tagged exocyst strains crossed to PrA-tagged exocyst subunits, which were validated by western blot. I confirmed the tagged subunits in these strains were functional using a serial dilution growth assay. Finally, I constructed a set of strains in which one exocyst subunit was PrA-tagged, in the background of a temperature-sensitive exocyst allele.

Using these strains, I optimized an exocyst purification protocol, and determined conditions to purify intact complexes from PrA-tagged yeast strains, as well as optimize protease cleavage to yield native, untagged complexes. The purification strategy was done in collaboration with members of Dr. Michael Rout's laboratory (Rockefeller University). In particular, I optimized lysis conditions, as well as purification buffers to ensure complex integrity. In order to examine partially destabilized exocyst complexes, I investigated the effects of

detergents, chaotropic agents, or high salt on the complex. In order to selectively destabilize only one exocyst subunit at a time, I purified complexes from strains in which a single exocyst subunit was encoded by a temperature-sensitive allele at both permissive and non-permissive temperatures.

Finally, I initiated negative stain electron microscopy studies on the native, purified complex (post-protease cleavage). These studies were done in collaboration with members of Dr. Adam Frost's laboratory (formally at University of Utah; currently at University of California, San Francisco). Although I did not collect a full dataset of negative stain micrographs, this work produced the first look at the intact native complex from *S. cerevisiae*.

Subunit connectivity, assembly determinants and architecture of the yeast exocyst complex

Margaret R Heider[1], Mingyu Gu[2], Caroline M Duffy[1], Anne M Mirza[1], Laura L Marcotte[1,6], Alexandra C Walls[1], Nicholas Farrall[2], Zhanna Hakhverdyan[3], Mark C Field[4], Michael P Rout[3], Adam Frost[2,5] & Mary Munson[1]

The exocyst is a hetero-octameric complex that has been proposed to serve as the tethering complex for exocytosis, although it remains poorly understood at the molecular level. Here, we purified endogenous exocyst complexes from *Saccharomyces cerevisiae* and showed that they are stable and consist of all eight subunits with equal stoichiometry. Using a combination of biochemical and auxin induced–degradation experiments in yeast, we mapped the subunit connectivity, identified two stable four-subunit modules within the octamer and demonstrated that several known exocyst-binding partners are not necessary for exocyst assembly and stability. Furthermore, we visualized the structure of the yeast complex by using negative-stain electron microscopy; our results indicate that the exocyst exists predominantly as a stable, octameric complex with an elongated architecture that suggests that the subunits are contiguous helical bundles packed together into a bundle of long rods.

Exocytosis is the evolutionarily conserved pathway by which protein and lipid cargos are trafficked from intracellular compartments to the plasma membrane in membrane-bound vesicles. This pathway is essential for cellular growth and division as well as for specialized processes such as cell migration, ciliogenesis and autophagy[1]. To maintain the fidelity of the secretory pathway, numerous conserved protein families regulate every step of the process[2]. Tethering factors, including the multisubunit tethering complexes (MTCs), serve as the first long-range, reversible connection between a vesicle and its target membrane[3,4]. However, in many cases, experimental evidence demonstrating tethering by these factors is lacking[5]. Tethers have been proposed to provide specificity for vesicle targeting, but they may also assume a more active role in regulating SNARE-mediated membrane fusion[3,6–8].

The exocyst complex is the MTC for secretory vesicles at the plasma membrane, and it contains eight subunits—Sec3, Sec5, Sec6, Sec8, Sec10, Sec15, Exo70 and Exo84—all of which have orthologs in eukaryotes including yeasts and humans[9–14]. Yeast exocyst mutants display severe growth and secretion defects and accumulate post-Golgi secretory vesicles in the cytoplasm[15,16]. Similarly, null mutants in mice and flies exhibit embryonic and larval lethality, respectively[17,18]. Although previous studies have revealed requirements for the exocyst in many critical cellular processes involving polarized vesicle trafficking, the structure and mechanisms of tethering by the exocyst remain unresolved[1].

Similarly to other tethering factors, the exocyst is a peripheral membrane protein complex that interacts with numerous GTPases, SNAREs, phospholipids and the vesicle-transport motor myosin V[1,3,19,20]. The exocyst has been proposed to interact with vesicles through Sec15 binding to the Rab GTPase Sec4 and myosin V, as well as Sec6 binding to the v-SNARE Snc[16,19,21]. On the target-membrane side, both Sec3 and Exo70 interact with Rho GTPases and phosphatidylinositol 4,5-bisphosphate[22–26], and Sec6 may interact with an as-yet-unidentified 'anchor' factor at the plasma membrane[27]. It is through this myriad of connections that the exocyst has been predicted to selectively capture secretory vesicles and tether them to the plasma membrane. A current model for exocyst function proposes that a subcomplex of exocyst subunits in *S. cerevisiae* is carried on vesicles to another subcomplex at the plasma membrane, and assembly of these subcomplexes drives vesicle tethering[28]. However, this model has not yet been validated biochemically, nor have the putative subcomplexes been identified. Whether regulated assembly of the exocyst is required for tethering and SNARE-complex regulation in yeast or other organisms, and whether these mechanisms differ between different species, are important unanswered questions.

Mechanistic models for exocyst function must be informed by the structural arrangement of exocyst subunits. Crystal structures of several exocyst subunits reveal a strikingly similar motif of contiguous helical bundles that pack together into long rods, thus classifying the exocyst within the evolutionarily conserved complexes associated with tethering containing helical rods (CATCHR) family[3,20]. Numerous pairwise subunit interactions have been identified via yeast two-hybrid assays, immunoprecipitations and *in vitro* binding experiments using recombinant and *in vitro*–translated proteins[20,29]. To examine the architecture and regulation of exocyst assembly, we developed a new robust exocyst purification method to reproducibly

[1]Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, Massachusetts, USA. [2]Department of Biochemistry, University of Utah, Salt Lake City, Utah, USA. [3]Laboratory of Cellular and Structural Biology, The Rockefeller University, New York, New York, USA. [4]Division of Biological Chemistry and Drug Discovery, University of Dundee, Dundee, UK. [5]Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, California, USA. [6]Present address: Department of Natural Sciences, Assumption College, Worcester, Massachusetts, USA. Correspondence should be addressed to M.M. (mary.munson@umassmed.edu).

Received 15 May; accepted 19 November; published online 14 December 2015; doi:10.1038/nsmb.3146

isolate stable exocyst complexes from *S. cerevisiae*. Using an auxin-inducible degradation (AID) system to deplete single subunits, we mapped the connectivity of the eight subunits and determined that most of the subunits are required for the association of two assembly modules within the exocyst. In contrast, depletion of known binding partners had no effect on the assembly status of the exocyst. Here we present the first structure, to our knowledge, of a fully assembled CATCHR MTC—we determined the structure of the fully assembled exocyst through negative-stain EM and two-dimensional (2D) averaging. Furthermore, we demonstrate that exocyst complexes are stoichiometric with no detectable subcomplexes; therefore, we propose that the yeast exocyst functions predominantly as a fully assembled complex.

RESULTS

Purification of intact yeast exocyst complexes

Biochemical and structural studies of the intact exocyst complex have previously been limited by preparations with poor yield, stability and purity (refs. 9,10,30–32 and M.M., unpublished data). In order to answer critical questions regarding the architecture of the yeast exocyst complex and its putative assembly dynamics, we have developed an improved protocol for isolating the entire native complex from yeast extract[33,34]. To maintain endogenous expression levels and function, we fused C-terminal Protein A (PrA) affinity tags to each exocyst subunit individually by integrating DNA encoding PrA at each genomic locus, creating eight different tagged haploid *S. cerevisiae* strains (**Supplementary Table 1**). The C-terminal PrA tags did not confer growth defects (**Supplementary Fig. 1a**), thus demonstrating that each of the tagged subunits was functional. We grew yeast strains, harvested them in log phase as frozen noodles and lysed them with a planetary ball mill grinder (Online Methods). We then resuspended the lysate powder in a physiological buffer, incubated the lysate with rabbit IgG–conjugated magnetic beads and eluted bound proteins from the beads either by proteolytic digestion or by denaturation with SDS loading buffer (**Fig. 1**). We confirmed exocyst subunit identities on the basis of the molecular-weight shift of the PrA tag in SDS-PAGE (**Fig. 1**), MALDI-MS and western blot analyses (data not shown).

We isolated intact exocyst complexes from yeast extracts by using each of the eight subunits as the PrA-tagged purification handle. The eight exocyst subunits copurified with equal stoichiometry, as detected by both Coomassie-stained SDS-PAGE and densitometry with Krypton fluorescent protein stain (**Fig. 1**), consistently with results from earlier reports[9,31]. We next asked whether the complexes purified by this method undergo disassembly and reassembly during the purification. When we mixed Sec10-GFP lysate with either Sec3-PrA or Exo70-PrA lysates and subsequently purified the exocyst complexes, we detected no Sec10-GFP in either pulldown, thus indicating that no exchange or assembly of subunits occurred during the incubation (1 h at 4 °C) (**Supplementary Fig. 1b**); these results are consistent with those from our previous studies[27]. Therefore, the purified complexes represent the state of the endogenous complex at the time of cell lysis.

The improved yield and purity of our exocyst preparations are the result of decreased proteolysis from cryogenic lysis (**Supplementary Fig. 1c**) and the use of rabbit IgG–conjugated magnetic beads, which have a high affinity for PrA[35,36]. Additionally, protease cleavage allowed for increased purity and native elution of untagged complexes for structural studies (**Supplementary Fig. 1d**). We detected substoichiometric levels of copurifying proteins by MS and krypton fluorescent protein staining, but they appeared to primarily be highly

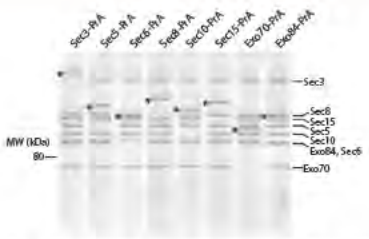


Figure 1 Purification of intact yeast exocyst complexes. Purified complexes were separated by SDS-PAGE and visualized by Krypton staining. The asterisks correspond to the PrA-tagged exocyst subunit used as a purification handle (which shifts the protein molecular weight (MW) by 25 kDa). Owing to phosphorylation, both the Sec3 and Exo84 protein bands often migrate as multiple species, which appear as slightly smeared bands on SDS-PAGE. The resuspension buffer used was 50 mM HEPES, pH 7.4, and 300 mM NaCl, plus protease inhibitors. Full-size images for this and subsequent cropped gel images are shown in Supplementary Data Set 1.

expressed, nonspecific contaminants or previously detected binding partners, including Rtn1 (ref. 31).

We next tested the functionality of our exocyst preparations by western blotting for known exocyst-interacting partners (**Supplementary Fig. 2**). The improved yield and rapid, gentle purification procedure allowed detection of binding of Sec1, Myo2 and Snc1/2 (redundant paralogs) to the exocyst. Previous studies have revealed an interaction of the exocyst subunit Sec6 with both Sec1 and Snc2 (refs. 7,21), and Sec15 with Myo2 (ref. 19). Here, we show that these proteins can be pulled down with tagged exocyst subunits that are not their direct binding partners, thus suggesting that these interactions occur within the context of the assembled complex.

Using Sec15-PrA as the purification handle, we monitored exocyst integrity under a variety of pH and salt conditions (**Fig. 2a**). The presence of reducing agents had no effect on complex recovery, and the complex was stable across a range of pH solutions, in contrast to results from previous studies[30]. Increasing the pH above 8.5 rendered purified exocyst complexes sensitive to salt concentrations ≥300 mM. In Tris, pH 8.5, and ≥500 mM salt, only Sec15 and Sec10 remained bound together, thus indicating a strong physical interaction between these two subunits that is consistent with results of earlier studies[16].

The exocyst complex peripherally associates with vesicles and the plasma membrane[37]. We therefore tested the effect of detergents, particularly whether the stoichiometry changes because of the solubilization of membrane-bound subcomplexes or the disruption of intersubunit interactions. We tested several nonionic detergents including NP-40 (IGEPAL), Tween-20 and Triton X-100, and none of the detergents affected the overall yield of assembled exocysts or the relative stoichiometry of the subunits (**Fig. 2b**). In contrast, the exocyst was severely disrupted by sodium cholate, a strong anionic detergent. Together, these results indicate that varying the ionic strength of the resuspension buffer has a pronounced effect on exocyst integrity, suggesting that ionic interactions may be a major stabilizing force for intersubunit connections.

We used our pulldown assay to identify stable intracomplex interactions within the endogenous exocyst complex, using partially destabilizing buffer conditions, with each of the eight PrA-tagged exocyst subunits (**Fig. 2c**). Several stable subunit pairs emerged: Sec3–Sec5, Sec6–Sec8 and Sec10–Sec15. Neither Exo70 nor Exo84 bound tightly to any of the other subunits under these

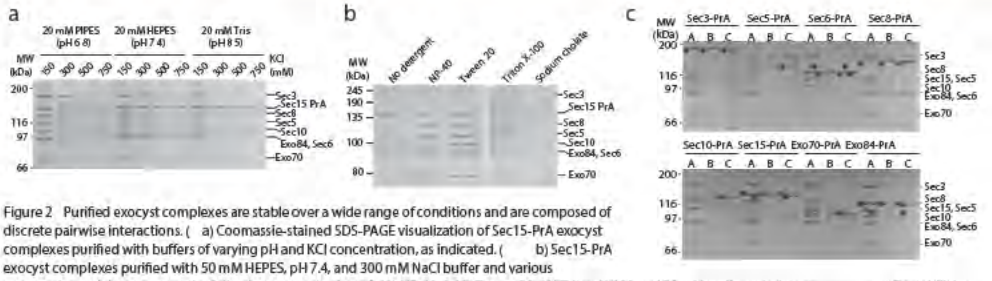


Figure 2 Purified exocyst complexes are stable over a wide range of conditions and are composed of discrete pairwise interactions. (a) Coomassie-stained SDS-PAGE visualization of Sec15-PrA exocyst complexes purified with buffers of varying pH and KCl concentration, as indicated. (b) Sec15-PrA exocyst complexes purified with 50 mM HEPES, pH 7.4, and 300 mM NaCl buffer and various commonly used detergents at the following concentrations: 0.1% NP-40, 0.1% Tween-20, 1% Triton X-100 and 20 mM sodium cholate. (c) Destabilizing buffer conditions were used with each exocyst subunit as a PrA purification handle to isolate subcomplexes and stable subunit pairs. A, 20 mM PIPES, pH 6.8, and 300 mM KCl; B, 20 mM Tris, pH 8.5, and 500 mM KCl; C, 20 mM Tris, pH 8.5, 300 mM KCl, and 500 mM urea. Asterisks correspond to the PrA-tagged subunit.

destabilizing conditions. Although several of these pairwise interactions had previously been identified[16,29,38,39], the relative stabilities of the subunit pairs compared to other intersubunit interactions were unknown.

Subunit connections and intracomplex assembly determinants

We applied a more targeted approach to answer additional architectural questions: How are these pairs of subunits assembled into the overall connectivity map of the assembled exocyst? Which of these intersubunit interactions are functionally important for maintaining exocyst integrity? Are some subunits more important for interactions with binding partners on the plasma membrane and vesicle? We decided to selectively eliminate individual exocyst subunits to define their role in maintaining overall complex assembly. All exocyst subunits except Sec3 are encoded by essential genes and therefore cannot be deleted from the yeast genome[39,40]. We tested the temperature-sensitive (ts) mutants *sec3-2*, *sec5-24*, *sec6-4*, *sec8-6* and *sec10-2* and used Sec15-PrA as the purification handle. Only *sec8-6* had a major effect on exocyst integrity at the restrictive temperature (data not shown). These results were difficult to interpret, however, because the ts alleles vary in their severity and in the amount of destabilization or degradation of the mutant protein. Previous studies using a similar panel of exocyst ts mutants have shown greater disassembly for several of the mutants than we observed, even at the permissive temperature[10]. These differences are probably due to proteolysis of exocyst subunits during spheroplasting lysis, which destabilizes the complex (**Supplementary Fig. 1c**). To overcome these challenges, we used an AID system to specifically remove each individual exocyst subunit.

This degron system uses the IAA17 AID sequence from *Arabidopsis thaliana*, which is fused to each exocyst subunit. When this tagged protein is coexpressed with the auxin receptor OsTIR1, exposure to the plant hormone auxin leads to rapid proteosomal degradation of the tagged subunit[41,42] (**Fig. 3a**). Addition of these tags to the C-terminal ends of exocyst subunits conferred no growth defects on their own, but when cells were grown on plates containing auxin (indole-3-acetic acid, IAA), all exocyst-AID strains were inviable except for Sec3-AID (**Fig. 3b**). We confirmed rapid and specific IAA-induced degradation of individual exocyst subunits in liquid culture by western blot analyses of yeast lysates. Each exocyst subunit was degraded to <12% of the starting level within 60 min of IAA treatment (**Fig. 3c**), whereas the protein levels of the remaining subunits were mostly unchanged (**Supplementary Fig. 3**).

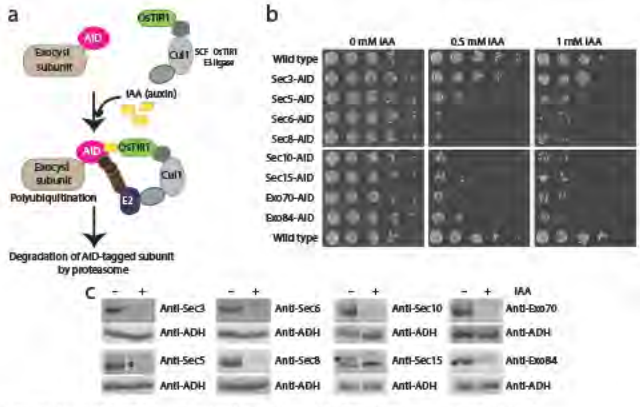


Figure 3 Use of the auxin-inducible degron (AID) system to selectively degrade essential exocyst proteins from yeast. (a) Schematic of the AID system. The AID tag from A. thaliana was fused to the C terminus of exocyst subunits at their genomic locus in yeast strains constitutively expressing OsTIR1 (F-box transport inhibitor response 1) protein. Upon treatment with the natural plant hormone auxin (indole-3-acetic acid (IAA)), the SCF–OsTIR1 E3 ubiquitin ligase complex is activated and subsequently recruits E2 ubiquitin conjugating enzymes for polyubiquitination of the AID-tagged protein. The AID-tagged protein is then rapidly degraded by the proteasome[41,42]. (b) Serial-dilution growth assay of AID-tagged exocyst strains on yeast peptone dextrose (YPD) plates containing the indicated amount of IAA. Suppressor colonies can be seen in some dilutions. (c) Western blotting of lysates, confirming degradation of exocyst subunits in these strains. Minus denotes untreated strains, and plus denotes strains treated with IAA. All subunits were degraded to <10–12% of the starting protein level. Asterisks indicate the AID-tagged exocyst subunit in blots in which antibodies also bind nonexocyst proteins.

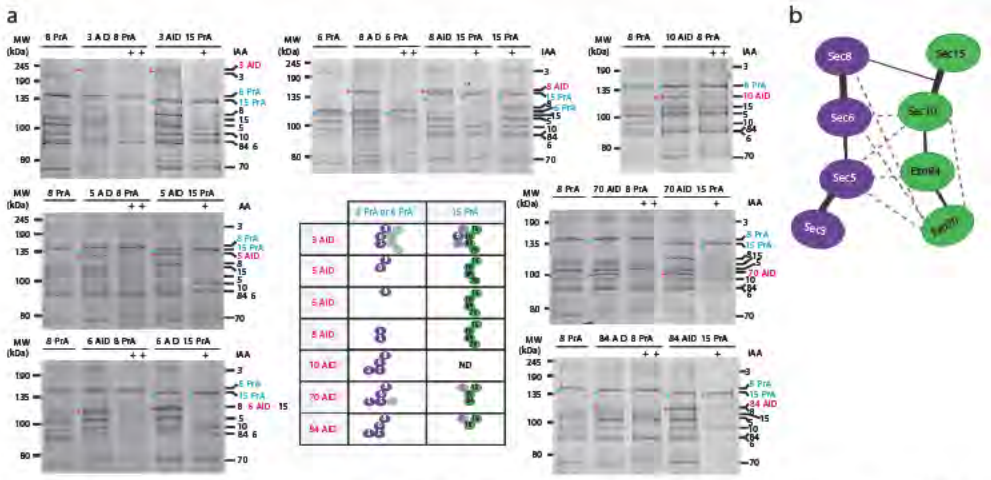


Figure 4 Most exocyst subunits are critical for maintaining the assembly of two four-subunit modules within the full octameric complex. (a) SDS-PAGE and Coomassie staining of exocyst complexes purified with the indicated PrA purification handles (blue) from yeast strains in which one AID-tagged subunit (magenta) was degraded. Minus denotes untreated strains, and plus denotes strains treated with IAA. For simplicity, exocyst subunits are denoted by their associated number. Degradation of six of the subunits tested led to the complete separation of exocyst complexes into two four-subunit modules: 3–5–6–8 and 10–15–70–84, with the connections depicted in the central table. The combination of Sec10-AID and Sec15-PrA was not determined (ND). Faded ovals represent subunits that showed partial loss from the complex. (b) Model depicting the subunit connectivity within and between each exocyst module (green and purple). Thick lines indicate the strong pairwise connections required for stability of the assembled exocyst. The thin line depicts a putative connection between Sec8 and Sec10–Sec15 identified in the AID studies, although Sec8's direct binding partner within this pair is not known. Dashed lines represent weaker interactions identified in previous in vitro studies (summarized in ref. 20).

To assess the role of each individual subunit in maintaining the assembly of the endogenous exocyst complex, we combined this AID system with our PrA-tag purification approach. We added genomic C-terminal PrA tags to Sec8, Sec15 or Sec6 in strains already expressing an AID-tagged exocyst subunit and OsTIR1. We tested two different PrA-tag handles for each AID-tagged subunit to determine the fate of each of the exocyst subunits. Most of the dual-tagged exocyst strains grew normally but were inviable on IAA plates, as expected (**Supplementary Fig. 4a**). Unexpectedly, the strain with Sec10-AID and Sec15-PrA showed no growth defect on IAA plates and no loss of Sec10-AID in IAA-containing liquid culture; similarly, Sec15-AID was not degraded in combination with PrA-tagged exocyst subunits (data not shown). We speculate that the lack of degradation in these strains may be due to masking of the AID tag by the 25-kDa PrA tag on a neighboring exocyst subunit.

We purified the exocyst complex from both untreated and IAA-treated cultures for each exocyst-AID-PrA combination strain and visualized the complexes by Coomassie staining and western blotting (**Fig. 4a** and **Supplementary Fig. 4b**). Surprisingly, the loss of Sec5, Sec6, Sec8, Sec10, Exo70 or Exo84 resulted in the exocyst complex splitting into two distinct, stable modules: Sec3–Sec5–Sec6–Sec8 (3–5–6–8) and Sec10–Sec15–Exo70–Exo84 (10–15–70–84). The results from the different combinations of AID and PrA tags indicate the division of the exocyst structure into two modules (summarized in **Fig. 4a**). Loss of Sec3 had the least destabilizing effect on exocyst-complex assembly. Degradation of each of the other subunits had distinct effects on its own module, depending on the strength and connectivity of its interactions with its partners, but had no effect on the integrity of the opposing module.

We found that the individual assembly of each module is predominantly based on the association of three stable subunit pairs (3–5, 6–8 and 10–15), instead of requiring the cooperative assembly of all four subunits together. If exocyst assembly were cooperative, we would expect to observe complete disassembly of all four subunits from each module upon loss of one subunit; instead, we generally found subcomplexes containing two or three subunits (for example, Sec6 and Sec5 remain bound after Sec8 is degraded). This finding is consistent with our earlier biochemistry results demonstrating that these subunit pairs are stable enough to be copurified (**Fig. 2c**). Therefore, the most robust interactions within the complex exist between pairs of subunits, and the overall assembly appears to be mediated by a network of weaker interactions. Several additional rules for exocyst assembly can be drawn from these results (**Fig. 4b**). Sec8 requires Sec6 for assembly into the complex. Sec5 is required for Sec3's assembly and for the stable interaction of Sec3 with Sec6 and Sec8. In the absence of Sec8, there was also loss of Sec3 from Sec5–Sec6, thus suggesting either a potential interaction between Sec3 and Sec8 or a potential conformational change that weakens Sec3's association with Sec5–Sec6. In the case of the other module, Sec10 and Sec15 are a stable pair that require Exo84 for their association with Exo70. Although we were unable to test this hypothesis, we predict that degradation of Sec15 would not disrupt Sec10's connection with Exo84 and Exo70, because Sec15's only known stable exocyst partner is Sec10 (ref. 16) (**Fig. 2**).

These studies provide only a few clues as to the interconnections between the modules. All subunits are required for the assembly of the two modules, including Exo70 and Exo84, which is perhaps surprising

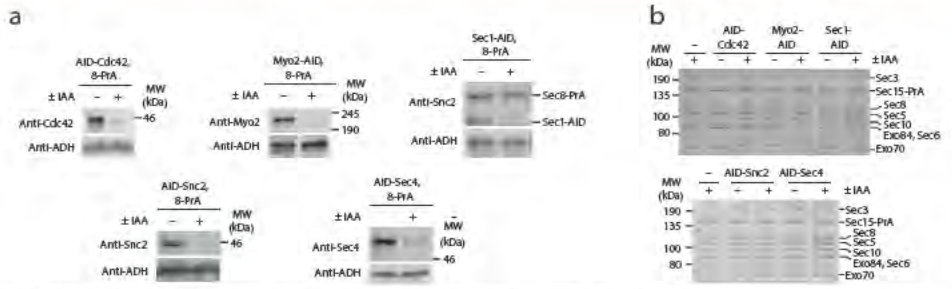


Figure 5 Depletion of known exocyst-binding partners does not affect the assembly of the exocyst complex. (a) Western blotting of exocyst-binding partners Cdc42, Myo2, Sec1, Snc2 (in the snc1 Δ strain background) and Sec4, which were AID-tagged in strains expressing Sec8-PrA and constitutively expressing OsTIR1. Minus denotes untreated strains, and plus denotes strains treated with IAA for 60 min. Western blots with antibodies specific to the AID-tagged protein of interest demonstrate degradation of these proteins from yeast lysate. In the Sec1 blot, the Sec1 antibody also reacts with the PrA tag on Sec8-PrA. (b) Exocyst complexes purified with Sec8-PrA as the purification handle from untreated (–) and IAA-treated (+) yeast lysates.

in light of our biochemical studies, which demonstrated that Exo70 and Exo84 are not tightly associated with any other subunits of the complex (Fig. 2). We propose that the interconnections between the modules are made up of a network of weaker subunit-subunit interactions, although we cannot rule out that the degradation of a subunit from one module may alter the structure of its respective subcomplex, thus making it incompatible for binding the opposing module. Other previously identified subunit interactions may contribute to this intermodule network[20,29], but their relative contributions remain to be tested (Fig. 4b).

Exocyst-binding partners have no effect on exocyst assembly

We wondered whether any additional binding partners would be necessary to maintain this stable assembly. However, we detected only substoichiometric amounts of known binding partners in our exocyst preparations, thus suggesting that these partners do not need to remain bound to maintain exocyst integrity (Fig. 1 and Supplementary Fig. 2).

A major unresolved question is how the exocyst assembles *in vivo* and whether additional factors are required for regulating this assembly. Selective elimination of individual exocyst-interacting partners along the late secretory pathway might identify subcomplexes, thereby indicating a failure of the complex to fully assemble. To test this idea, we again used our AID-tag approach to deplete the master polarity regulator Cdc42 (ref. 43), the type V myosin motor Myo2 (ref. 19), the SNARE regulator Sec1 (ref. 44), the v-SNARE Snc2 (ref. 21) and the Rab GTPase Sec4 (ref. 16) (Supplementary Fig. 5a).

The functional consequences of each of these interactions are not known, and it is unclear at which stage of exocytosis these interactions occur[7,23,39,45].

We treated the AID-tagged partner strains with IAA for 1 h, a time sufficient for numerous rounds of vesicle delivery and fusion in *S. cerevisiae*[46]. Degradation of Sec1 induced a severe vesicle accumulation phenotype, as expected[47], whereas degradation of Myo2 and Cdc42 caused a milder secretion defect consistent with results from previous reports[43,48] (Supplementary Fig. 5b,c). N-terminal AID-tagging of Snc and Sec4 resulted in severe vesicle accumulation even before IAA treatment, thus suggesting that these N-terminal tags partially impair protein function (Supplementary Fig. 5b,c). Using a PrA tag on Sec8, we pulled down exocyst complexes after degradation of these partners (Fig. 5). For each of the proteins tested, we observed that the exocyst complexes were fully assembled and stoichiometric, and they could be recovered with the same yield.

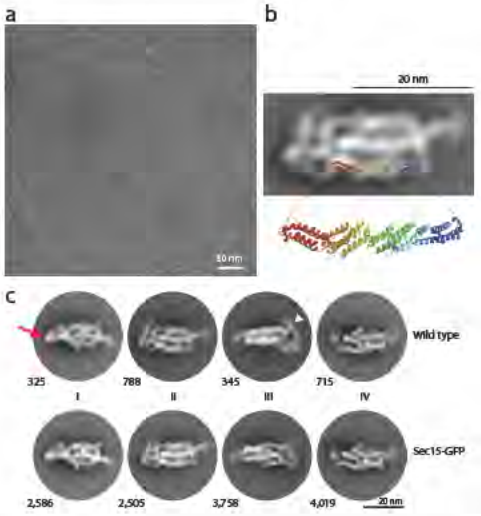


Figure 6 Negative-stain EM of purified exocyst complexes. (a) A representative transmission electron micrograph of Sec15-GFP exocyst complexes after negative staining in uranyl acetate. Scale bar, 50 nm. (b) 2D class average (Sec15-GFP), overlaid with a ribbon diagram of the structure of yeast Exo70 (residues 67–623), PDB 2B1 E[51]. The orientation and position of Exo70 were arbitrarily chosen to illustrate the similarities in the length and width of the legs of the complex and Exo70. (c) Highly populated 2D class averages generated by unsupervised classification for both wild-type and Sec15-GFP image data sets. The number of particles per class is indicated next to each 2D average. Four apparent faces of the complex are labeled I–IV. The red arrow points to the more 'compact' end of the complex in class I, and the white arrowhead points to the more open, flexible end in class III. Scale bar, 20 nm.

These results indicate that none of these components are required for driving or stabilizing the assembly of exocyst complexes. Together with the preceding observations that the exocyst subunits copurify in stoichiometric complexes, these data support a model wherein the exocyst functions predominantly in a fully assembled state in actively growing cells, even under conditions in which vesicles are not being transported, and the exocyst is not interacting with its partners.

Visualization of exocyst structure by electron microscopy

Our new purification method for the yeast exocyst complex allowed us to obtain pure complexes for structural studies. We purified both Sec15-GFP and wild-type complexes and analyzed them through negative-stain EM. Raw micrographs revealed distinct particles with an ellipsoid structure approximately 25 nm in length (**Fig. 6a**). Iterative rounds of unsupervised 2D classification and class averaging revealed multiple coherent views of the exocyst complex resolved between 17-Å and 25-Å resolution (**Fig. 6b,c** and **Supplementary Fig. 6**). However, this averaging failed to reveal a unique density attributable to GFP, thus precluding identification of Sec15's location within the structure. At this resolution, the orientations and overall architecture of the exocyst were indistinguishable between these biologically and technically independent data sets (**Fig. 6c** and **Supplementary Fig. 6**). We observed no apparent density or class averages for smaller particles, such as subcomplexes.

The 2D class averages resolved into roughly four distinct views of the complex (**Fig. 6c** and **Supplementary Fig. 6**), which may represent four 'faces' of the complex as it interacts with the EM grid. One end of the structure (left side of each of the 2D images) appears to be more tightly packed and ordered than the other end, which appears to be more flexible, often containing a long looping 'leg' wrapping around the end (right side of each of the 2D images). Two of the faces of the complex (I and II) appear wider and contain three to four legs or columns of density packed together, whereas the two slightly narrower faces (III and IV) appear to have only two to three legs each. We speculate that the more tightly packed end of the long axis of the complex may be composed of many of the N-terminal ends of exocyst subunits, because they generally have not been amenable to biochemical studies in isolation[49]. The C-terminal ends, therefore, would be present in the more flexible, 'open' end of the structure; these regions contain many of the regions involved in binding GTPases and the plasma membrane[23,24,50,51]. The exception is Sec3, whose membrane-interaction domain is located at its N-terminal end[26,52] and may therefore lie at the flexible open end of the exocyst, in an opposite orientation to the others.

Each of the individual legs observed in the 2D class averages of the exocyst complex are ~3 nm wide. Although the N- and C-terminal ends of the subunits cannot be unambiguously identified at this resolution, we can estimate the length of the legs in the range of ~15–35 nm, with the additional long leg at the flexible end being ~25 nm longer than the others. When the crystal structure of nearly full-length yeast Exo70 (residues 67–623) is superimposed onto an arbitrarily chosen leg, the width and length of the leg are consistent with the structure, which is ~16 nm long and ~3.0–3.5 nm wide[51] (**Fig. 6b**). Exo70 is the smallest exocyst subunit (71 kDa), and the others range from 84 kDa to 155 kDa. The large size of Sec3 (155 kDa; estimated extended helical-bundle length of ~38 nm) suggests that it may be the subunit that wraps around the end of the complex (**Fig. 6c**). The other available crystal structures (Exo84CT, Sec6CT and Sec15CT) also revealed similar CATCHR family helical bundles that are ~3 nm wide; the other subunits have been predicted to have similar folds[49–51,53,54]. The subunits of the complex appear to lie in a roughly parallel arrangement to each other, as suggested by previous interaction studies[20,51,54]. Our interpretation of the 2D averages suggests that this structure represents a fully assembled complex with an estimated volume of ~1,800–2,200 nm^3. Using the volume and molecular weight of the structure of Exo70, and the assumption that all the subunits have roughly similar helical-bundle structures, we calculated a comparable volume of ~1,900 nm^3 for the octameric complex. Therefore, we suggest that our structure contains all eight subunits, consistently with the biochemical and AID experiments. Furthermore, we speculate that the wider faces containing three or four legs represent the two distinct modules identified in our AID studies, with one module as the top face and the other as the bottom face. However, we cannot rule out that the 2D averages could actually be showing the same face in alternative conformations; higher-resolution data will therefore be necessary to resolve these models.

DISCUSSION

In this study, we used biochemical, genetic and structural methods to dissect the architecture of the yeast exocyst complex and examined mechanisms for its assembly and function. We purified endogenous, intact exocyst complexes from *S. cerevisiae* (**Fig. 1**), and our biochemical and structural characterization demonstrated an intrinsically stable, intact octameric complex (**Figs. 2** and **6**). Our results with the AID system indicated that the presence of most of the exocyst subunits is critical to complex integrity and stability (**Fig. 4**). Degradation of six out of the seven AID-tagged subunits tested, except Sec3, triggered complete separation of the exocyst into two modules (**Fig. 4**). Each of these modules (3–5–6–8 and 10–15–70–84) is assembled by several critical pairwise interactions (3–5, 6–8 and 10–15) with weaker contributions from 5–6, 70–84, 84–10 and 8–10 or 8–15 (**Figs. 2** and **4**); furthermore, the disassembly of one module does not affect the integrity of the other. Consistently with this, our negative-stain EM 2D class averages demonstrated a stable, homogenous, octameric complex (**Fig. 6**). The assembly and stability of the exocyst structure is independent of the known binding partners Sec4, Snc1/2, Myo2, Sec1 and Cdc42 (**Fig. 5**). These components are not stable, stoichiometric partners of the exocyst complex, nor is their binding necessary to assemble or stabilize the exocyst complex during vesicle transport, tethering or fusion. We propose that the role of these interactions is to modulate the function, rather than the assembly, of the exocyst complex.

Our results do not support previous hypotheses that have suggested a requirement for polarized vesicle transport in driving the assembly of a subcomplex of exocyst subunits (for example, Sec15–Sec10–Sec6–Sec8–Exo84) on vesicles with a subgroup (Sec3 and Exo70) serving as a 'landmark' on the plasma membrane; assembly of these two subgroups would subsequently drive vesicle tethering[28]. Under physiological conditions, we did not detect any stable subcomplexes in our pulldowns. It is possible that we detected only stoichiometric complexes because uncomplexed subunits or unstable subcomplexes are degraded during the purification; however, our biochemical and AID experiments do not support this possibility, because we could easily purify individual subunits and subcomplexes from yeast lysate with a yield equal to that of assembled complexes (**Figs. 2** and **4**). Furthermore, under conditions in which we disrupted vesicle transport, cell polarity and exocyst binding to vesicles, no subcomplexes were detectable (**Fig. 5**). We cannot rule out the presence of either low levels of subcomplexes or free pools of exocyst subunits below our limit of detection (<5–10%), however, the majority of the exocyst exists in the fully assembled state. In contrast, subcomplexes appear to be present in mammalian cells: the components identified thus

far (Exo84–Sec10 and Sec5–Sec6 in opposing groups) are consistent with the modules identified here[55,56]. Similarly, differences in subunit localization patterns in the growing hyphae of *Neurospora crassa*, in *A. thaliana* and in different *Drosophila melanogaster* tissues suggest putative subgroups of exocyst subunits[14,57,58]. Regulated assembly and disassembly of the exocyst in different organisms may be an important mechanism by which the exocyst complex participates in a diverse array of processes in a variety of cell types.

Negative-stain EM revealed the first evidence, to our knowledge, of the ellipsoid-shaped structure of the yeast exocyst complex, with its distinct helical bundle–shaped legs packed together (Fig. 6). Overall, the yeast exocyst structure is roughly similar to those of the mammalian exocyst complexes previously imaged with rotary shadowing EM[32]. However, unlike the individual Y-shaped structures observed with glutaraldehyde-fixed mammalian exocyst particles, our yeast 2D averages do not appear to have the same short 'arms'. The arms may be too flexible or heterogeneous to be observed in our 2D averages or may represent mammalian-specific domains (for example, Ral-binding domains in Sec5 and Exo84); alternatively, perhaps the mammalian exocyst was partially disassembled during processing. Future efforts will require the use of higher-resolution data and other strategies to uniquely identify each exocyst subunit within the structure.

Members of the CATCHR family of MTCs, including the exocyst, COG, GARP and Dsl1, share functional similarity as well as structural similarity at the individual subunit level. Thus, they might be expected to assemble into comparable quaternary structures, although they contain different numbers of subunits[3]. Similarly to the exocyst modules identified here, COG consists of two structurally and functionally distinct subassemblies with four subunits each[59]. However, in terms of their overall shapes, as determined by negative-stain EM, the exocyst differs markedly from that of both COG and Dsl1. The COG and Dsl1 structures consist of ~3-nm-wide legs emanating from a central flexible 'joint'[59,60], whereas the exocyst's legs fold alongside each other to form a compact ellipsoid structure. It is possible that the COG and Dsl1 complexes might adopt more compact structures with all their subunits present or that they might represent a different, biologically relevant conformation that is not captured in the exocyst EM particles. It will be interesting to determine whether there is a conserved distance for vesicle capture by MTCs at the target membrane and whether all MTCs undergo conformational changes to bring vesicles into closer proximity for SNARE assembly and vesicle fusion, as has previously been suggested for the Dsl1 complex[60].

We propose that, in contrast with models proposing that assembly of subcomplexes is required for exocyst function, the yeast exocyst complex functions as a stably assembled octamer in the cell. The subunits pack together into an elongated structure. This structure could be a single conformation that functions through changing interactions with various partner proteins. Alternatively, the exocyst may undergo conformational changes in response to binding its protein or membrane partners. Defining the subunit positions and binding of partners at higher resolution is necessary for elucidating the mechanisms of vesicle tethering and SNARE-complex regulation at the plasma membrane. This knowledge is also critical in determining whether the MTCs function by similar mechanisms and how they are uniquely suited to specific trafficking pathways and cell types. Importantly, the ability to purify stable yeast exocyst complexes will now enable functional studies to obtain a detailed molecular understanding of the exocyst's role in vesicle tethering and SNARE-complex regulation.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank P. Brennwald (University of North Carolina, Chapel Hill), C. Carr (Texas A&M University) and L. Weisman (University of Michigan) for antibodies, P. Novick (University of California, San Diego) for gifts of yeast strains, the Yeast Genome Resource Center in Japan for the AID-system reagents and the Wendland laboratory (Johns Hopkins University) for technical advice. Thanks to W. Holmes, M. Jacques, R. Kalia, L. Hassinger and members of the University of Massachusetts Medical School Core EM Facility for technical assistance. Thanks to R. Gilmore, S. Ryder, P. Pryciak, C. Carr and members of M.M.'s laboratory for critical reading of this manuscript and advice. Work in our laboratories is supported by US National Institutes of Health grants GM068803 (M.M. and A.F.), 1DP2GM110772 (A.F.), U54 GM103511 and P41 GM109824 (M.P.R.), and a Searle Scholars Award (A.F.).

AUTHOR CONTRIBUTIONS

M.R.H. and M.M. conceived the study, designed the biochemical and cell biology experiments and wrote the manuscript; M.R.H. made yeast strains and performed most of the biochemistry and cell biology experiments with assistance from A.M.M., C.M.D., L.L.M. and A.C.W.; M.G. and A.F. designed, performed and analyzed the EM experiments; early EM optimization work was done by N.F.; development of the purification method was done by Z.H., C.M.D., M.P.R. and M.C.F.; all authors contributed to discussion and approved the final manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Heider, M.R. & Munson, M. Exorcising the exocyst complex. Traffic 13, 898–907 (2012).
2. Wickner, W. & Schekman, R. Membrane fusion. Nat. Struct. Mol. Biol. 15, 658–664 (2008).
3. Yu, I.M. & Hughson, F.M. Tethering factors as organizers of intracellular vesicular traffic. Annu. Rev. Cell Dev. Biol. 26, 137–156 (2010).
4. Chia, P.Z. & Gleeson, P.A. Membrane tethering. F1000Prime Rep. 6, 74 (2014).
5. Brunet, S. & Sacher, M. Are MTCs tethering complexes or trafficking complexes that may act as tethers? Traffic 15, 1282–1287 (2014).
6. Sivaram, M.V., Saporita, J.A., Furgason, M.L.M., Boettcher, A.J. & Munson, M. Dimerization of the exocyst protein Sec6p and its interaction with the t-SNARE Sec9p. Biochemistry 44, 6302–6311 (2005).
7. Morgera, F. et al. Regulation of exocytosis by the exocyst subunit Sec6 and the SM protein Sec1. Mol. Biol. Cell 23, 337–346 (2012).
8. Laufman, O., Hong, W. & Lev, S. The COG complex interacts with multiple Golgi SNAREs and enhances fusogenic assembly of SNARE complexes. J. Cell Sci. 126, 1506–1516 (2013).
9. TerBush, D.R., Maurice, T., Roth, D. & Novick, P. The Exocyst is a multiprotein complex required for exocytosis in Saccharomyces cerevisiae. EMBO J. 15, 6483–6494 (1996).
10. TerBush, D.R. & Novick, P. Sec6, Sec8, and Sec15 are components of a multisubunit complex which localizes to small bud tips in Saccharomyces cerevisiae. J. Cell Biol. 130, 299–312 (1995).
11. Guo, W., Grant, A. & Novick, P. Exo84p is an exocyst protein essential for secretion. J. Biol. Chem. 274, 23558–23564 (1999).
12. Hsu, S.C. et al. The mammalian brain rsec6/8 complex. Neuron 17, 1209–1219 (1996).
13. Koumandou, V.L., Dacks, J.B., Coulson, R.M. & Field, M.C. Control systems for membrane fusion in the ancestral eukaryote; evolution of tethering complexes and SM proteins. BMC Evol. Biol. 7, 29 (2007).
14. Riquelme, M. et al. The Neurospora crassa exocyst complex tethers Spitzenkörper vesicles to the apical plasma membrane during polarized growth. Mol. Biol. Cell 25, 1312–1326 (2014).
15. Novick, P., Field, C. & Schekman, R. Identification of 23 complementation groups required for post-translational events in the yeast secretory pathway. Cell 21, 205–215 (1980).
16. Guo, W., Roth, D., Walch-Solimena, C. & Novick, P. The exocyst is an effector for Sec4p, targeting secretory vesicles to sites of exocytosis. EMBO J. 18, 1071–1080 (1999).
17. Friedrich, G.A., Hildebrand, J.D. & Soriano, P. The secretory protein Sec8 is required for paraxial mesoderm formation in the mouse. Dev. Biol. 192, 364–374 (1997).

18. Murthy, M., Garza, D., Scheller, R.H. & Schwarz, T.L. Mutations in the exocyst component Sec5 disrupt neuronal membrane traffic, but neurotransmitter release persists. Neuron 37, 433–447 (2003).
19. Jin, Y. et al. Myosin V transports secretory vesicles via a Rab GTPase cascade and interaction with the exocyst complex. Dev. Cell 21, 1156–1170 (2011).
20. Munson, M. & Novick, P. The exocyst defrocked, a framework of rods revealed. Nat. Struct. Mol. Biol. 13, 577–581 (2006).
21. Shen, D. et al. The synaptobrevin homologue Snc2p recruits the exocyst to secretory vesicles by binding to Sec6p. J. Cell Biol. 202, 509–526 (2013).
22. Wu, H., Rossi, G. & Brennwald, P. The ghost in the machine: small GTPases as spatial regulators of exocytosis. Trends Cell Biol. 18, 397–404 (2008).
23. Wu, H., Turner, C., Gardner, J., Temple, B. & Brennwald, P. The Exo70 subunit of the exocyst is an effector for both Cdc42 and Rho3 function in polarized exocytosis. Mol. Biol. Cell 21, 430–442 (2010).
24. He, B., Xi, F., Zhang, X., Zhang, J. & Guo, W. Exo70 interacts with phospholipids and mediates the targeting of the exocyst to the plasma membrane. EMBO J. 26, 4053–4065 (2007).
25. Zhang, X. et al. Membrane association and functional regulation of Sec3 by phospholipids and Cdc42. J. Cell Biol. 180, 145–158 (2008).
26. Baek, K. et al. Structure-function study of the N-terminal domain of exocyst subunit Sec3. J. Biol. Chem. 285, 10424–10433 (2010).
27. Songer, J.A. & Munson, M. Sec6p anchors the assembled exocyst complex at sites of secretion. Mol. Biol. Cell 20, 973–982 (2009).
28. Boyd, C., Hughes, T., Pypaert, M. & Novick, P. Vesicles carry most exocyst subunits to exocytic sites marked by the remaining two subunits, Sec3p and Exo70p. J. Cell Biol. 167, 889–901 (2004).
29. Katoh, Y., Nozaki, S., Hartanto, D., Miyano, R. & Nakayama, K. Architectures of multisubunit complexes revealed by a visible immunoprecipitation assay using fluorescent fusion proteins. J. Cell Sci. 128, 2351–2362 (2015).
30. Terbush, D.R., Guo, W., Dunkelbarger, S. & Novick, P. Purification and characterization of yeast exocyst complex. Methods Enzymol. 329, 100–110 (2001).
31. De Craene, J.O. et al. Rtn1p is involved in structuring the cortical endoplasmic reticulum. Mol. Biol. Cell 17, 3009–3020 (2006).
32. Hsu, S.C. et al. Subunit composition, protein interactions, and structures of the mammalian brain sec6/8 complex and septin filaments. Neuron 20, 1111–1122 (1998).
33. Oeffinger, M. et al. Rrp17p is a eukaryotic exonuclease required for 5′ end processing of Pre-60S ribosomal RNA. Mol. Cell 36, 768–781 (2009).
34. Hakhverdyan, Z. et al. Rapid, optimized interactomic screening. Nat. Methods 12, 553–560 (2015).
35. Richman, D.D., Cleveland, P.H., Oxman, M.N. & Johnson, K.M. The binding of staphylococcal protein A by the sera of different animal species. J. Immunol. 128, 2300–2305 (1982).
36. Oeffinger, M. et al. Comprehensive analysis of diverse ribonucleoprotein complexes. Nat. Methods 4, 951–956 (2007).
37. Bowser, R., Müller, H., Govindan, B. & Novick, P. Sec8p and Sec15p are components of a plasma membrane-associated 19.5S particle that may function downstream of Sec4p to control exocytosis. J. Cell Biol. 118, 1041–1056 (1992).
38. Roth, D., Guo, W. & Novick, P. Dominant negative alleles of SEC10 reveal distinct domains involved in secretion and morphogenesis in yeast. Mol. Biol. Cell 9, 1725–1739 (1998).
39. Wiederkehr, A., De Craene, J.O., Ferro-Novick, S. & Novick, P. Functional specialization within a vesicle tethering complex: bypass of a subset of exocyst deletion mutants by Sec1p or Sec4p. J. Cell Biol. 167, 875–887 (2004).
40. Haarer, B.K. et al. SEC3 mutations are synthetically lethal with profilin mutations and cause defects in diploid-specific bud-site selection. Genetics 144, 495–510 (1996).
41. Nishimura, K., Fukagawa, T., Takisawa, H., Kakimoto, T. & Kanemaki, M. An auxin-based degron system for the rapid depletion of proteins in nonplant cells. Nat. Methods 6, 917–922 (2009).
42. Nishimura, K. & Kanemaki, M.T. Rapid depletion of budding yeast proteins via the fusion of an auxin-inducible degron (AID). Curr. Protoc. Cell Biol. 64, 20.9 (2014).
43. Adamo, J.E. et al. Yeast Cdc42 functions at a late step in exocytosis, specifically during polarized growth of the emerging bud. J. Cell Biol. 155, 581–592 (2001).
44. Hashizume, K., Cheng, Y.S., Hutton, J.L., Chiu, C.H. & Carr, C.M. Yeast Sec1p functions before and after vesicle docking. Mol. Biol. Cell 20, 4673–4685 (2009).
45. Zhang, X. et al. Cdc42 interacts with the exocyst and regulates polarized secretion. J. Biol. Chem. 276, 46745–46750 (2001).
46. Donovan, K.W. & Bretscher, A. Myosin-V is activated by binding secretory cargo and released in coordination with Rab/exocyst function. Dev. Cell 23, 769–781 (2012).
47. Novick, P. & Schekman, R. Secretion and cell-surface growth are blocked in a temperature-sensitive mutant of Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. USA 76, 1858–1862 (1979).
48. Govindan, B., Bowser, R. & Novick, P. The role of Myo2, a yeast class V myosin, in vesicular transport. J. Cell Biol. 128, 1055–1068 (1995).
49. Croteau, N.J., Furgason, M.L., Devos, D. & Munson, M. Conservation of helical bundle structure between the exocyst subunits. PLoS One 4, e4443 (2009).
50. Wu, S., Mehta, S.Q., Pichaud, F., Bellen, H.J. & Quiocho, F.A. Sec15 interacts with Rab11 via a novel domain and affects Rab11 localization in vivo. Nat. Struct. Mol. Biol. 12, 879–885 (2005).
51. Dong, G., Hutagalung, A.H., Fu, C., Novick, P. & Reinisch, K.M. The structures of exocyst subunit Exo70p and the Exo84p C-terminal domains reveal a common motif. Nat. Struct. Mol. Biol. 12, 1094–1100 (2005).
52. Yamashita, M. et al. Structural basis for the Rho- and phosphoinositide-dependent localization of the exocyst subunit Sec3. Nat. Struct. Mol. Biol. 17, 180–186 (2010).
53. Hamburger, Z.A., Hamburger, A.E., West, A.P. Jr. & Weis, W.I. Crystal structure of the S.cerevisiae exocyst component Exo70p. J. Mol. Biol. 356, 9–21 (2006).
54. Sivaram, M.V., Furgason, M.L.M., Brewer, D.N. & Munson, M. The structure of the exocyst subunit Sec6p defines a conserved architecture with diverse roles. Nat. Struct. Mol. Biol. 13, 555–556 (2006).
55. Moskalenko, S. et al. Ral GTPases regulate exocyst assembly through dual subunit interactions. J. Biol. Chem. 278, 51743–51748 (2003).
56. Bodemann, B.O. et al. RalB and the exocyst mediate the cellular starvation response by direct activation of autophagosome assembly. Cell 144, 253–267 (2011).
57. Fendrych, M. et al. Visualization of the exocyst complex dynamics at the plasma membrane of Arabidopsis thaliana. Mol. Biol. Cell 24, 510–520 (2013).
58. Murthy, M. et al. Sec6 mutations and the Drosophila exocyst complex. J. Cell Sci. 118, 1139–1150 (2005).
59. Lees, J.A., Yip, C.K., Walz, T. & Hughson, F.M. Molecular organization of the COG vesicle tethering complex. Nat. Struct. Mol. Biol. 17, 1292–1297 (2010).
60. Ren, Y. et al. A structure-based mechanism for vesicle capture by the multisubunit tethering complex Dsl1. Cell 139, 1119–1129 (2009).

ONLINE METHODS

Yeast methods. The strains used in this study are listed in **Supplementary Table 1**. Standard methods were used for yeast media and genetic manipulations. Cells were grown in YPD medium containing 1% Bacto-yeast extract (Fisher Scientific), 2% Bacto-peptone (Fisher Scientific), and 2% glucose (Sigma-Aldrich). All protein A (PrA) tags were integrated at the genomic loci in haploid yeast strains (BY4741 or BY4742) by integration of linear PCR products. PrA products were amplified from a plasmid (pProtAHIS5, M.P.R.'s laboratory) encoding a PreScission Protease (PPX) site upstream of the PrA tag and a *Schizosaccharomyces pombe HIS5* selection marker[36]. Approximately 60 bp of sequence homologous to the 5′ end of the coding sequence and 60 bp of sequence homologous to the 3′ flanking sequence were used for homologous recombination. All exocyst PrA tags were added at the C-terminal ends. AID tags (IAA17) and linker were amplified from BYP6740 (pMK43, Yeast Genome Resource Center (YGRC), Japan). For C-terminal AID-tag strains, tags were added at the genomic locus of the strain BY25598 (YGRC), which expresses OsTIR1 under the ADH1 promoter (parent w303-1a), with linear PCR products and *kanMX* selection. N-terminal AID tags (*SNC2*, *SEC4*, and *CDC42* only) were integrated at the genomic locus of BY4742 with the pRS306 integrating plasmid[61]. Inserts were amplified by overlap extension of PCR products to generate a product consisting of ~300 bp of 5′ regulatory element, AID tag, linker, and sequence matching the 5′ end of the gene of interest, and this was then inserted into pRS306 through NotI and XhoI restriction sites. The plasmids were linearized with restriction enzymes specific to the 5′ regulatory elements of each gene (*SNC2*, MluI; *SEC4*, BsrGI; *CDC42*, HpaI) before yeast transformation. For the AID-Snc2 strain, *SNC1* was deleted by replacing the genomic locus with the *kanMX* cassette. Finally, for all N-terminal AID-tag strains, the OsTIR1 gene was integrated at the *MET15* locus with a *URA3* marker and an *ADH1* promoter. The plasmid BYP6744 (pNHK53, YGRC) was used as a template for generating the OsTIR1 PCR product, and sequence homologous to the *MET15* regulatory elements was added to the ends. For serial-dilution growth assays, yeast cells were grown in YPD to an OD of 1.5 and serially diluted ten-fold across YPD plates or YPD plates containing the indicated concentrations of IAA (VWR). Yeast plates were incubated at 30 °C for 2 d before imaging on Fujifilm LAS3000 (GE).

Exocyst protein A purification. Yeast cells (2 L) were grown in YPD at 30 °C to an OD of 1.3–1.5. Cells were washed with water, extruded through a syringe as frozen noodles into liquid nitrogen, and stored at −80 °C until lysis[36]. Noodles were lysed in a 50-ml stainless steel Komfort jar with stainless-steel ball bearings, prechilled in liquid nitrogen with a PM100 machine (Retsch). The resulting yeast powder was stored at −80 °C. 150 mg of yeast powder was added to 1.5-ml microcentrifuge tubes prechilled in liquid nitrogen. 600 µl of resuspension buffer (50 mM HEPES, pH 7.4, and 150 mM NaCl unless noted otherwise in the text, with 1× complete Mini EDTA-free protease-inhibitor solution (Roche Life Science)) was added to the tube (with buffer composition varying by experiment, as noted in the relevant figures) and was then vortexed and pipetted briefly to achieve complete resuspension. Spheroplasting and bead-beating lysis were performed as previously described[7] with 50 mM HEPES, pH 7.4, and 300 mM KCl lysis buffer. The use of NaCl versus KCl had no effect on exocyst preparations. Tubes were spun at 14,000*g* for 10 min at 4 °C, and the supernatant was added to 5 µl homemade rabbit IgG magnetic bead slurry[34,36]. Binding was carried out for 45 min at 4 °C on a nutating platform. The beads were washed in resuspension buffer and eluted either in 1× SDS loading buffer or by 1 h treatment with PreScission Protease (GE Healthcare) at 4 °C for a native elution. Samples were run on SDS-PAGE and stained with Coomassie Blue or Krypton fluorescent protein stain (Thermo Fisher Scientific). Western blot analyses were performed with rabbit polyclonal antibodies to Sec6, Sec8, Sec10, Exo70, and Exo84 (refs. 7,27). Rabbit polyclonal antibodies to Sec3, Sec15, and Sec5, and mouse monoclonal antibodies to Cdc42 and Sec4 were gifts from P. Brennwald (University of North Carolina, Chapel Hill). Rabbit polyclonal antibodies to Sec1 and Snc were gifts from C. Carr (Texas A&M University). Goat polyclonal antibody to Myo2 was a gift from L. Weisman (University of Michigan). Rabbit polyclonal antibody to ADH was purchased from Abcam (ab20994). Mouse monoclonal antibody to GFP was purchased from Clontech (632380). Western blot analyses of exocyst protein levels in input versus unbound samples showed that ~60% of exocyst complexes were bound to the beads (varying slightly by bead preparation). The IgG beads were saturated in these experiments, however, because the exocyst complexes remaining in the lysates could be pulled down by sequential bead incubations. Krypton staining of the resulting gels showed no differences in stoichiometry in sequential pulldowns of either Sec5-PrA or Sec15-PrA. Coomassie-stained gels were imaged on an LAS 4000 (GE Healthcare Life Sciences), and Krypton gels were imaged on a Typhoon FLA9000 (GE Healthcare Life Sciences). Western blots were treated with ECL and imaged on an LAS 4000. Full-size gels and western blots are available in **Supplementary Data Set 1**.

Auxin-induced degradation of exocyst subunits and exocyst regulators. Yeast cells (2 L) were grown in YPD at 30 °C to an OD of 1.0. IAA (VWR), dissolved in 100% ethanol at 500 mM, was added to yeast cultures for a final concentration of 0.7 mM. The cells were allowed to grow in IAA for 45 min (with 15 min for post-processing) at 30 °C until an OD of about 1.5 was reached. The cells were then washed with water, harvested as frozen noodles, and lysed as described above in the purification section. NaOH/SDS lysis was used for visualizing IAA-induced degradation in yeast lysates for **Figure 3c** and **Supplementary Figure 3**. Briefly, 2.5 OD units of yeast was incubated in 100 mM NaOH for 5 min, centrifuged to remove the NaOH, resuspended in SDS loading buffer with DTT, and heated at 95 °C before loading onto gels for SDS-PAGE and western blotting.

Bgl2 secretion assay. AID strains were grown at 30 °C in YPD and treated for 1 h with 0.7 mM IAA before harvesting. Bgl2 secretion assays were performed as described in Adamo *et al.*[62]. Internal Bgl2 levels were quantified by western blotting and normalized to internal ADH levels. All strains were normalized relative to internal Bgl2 levels of the appropriate untreated wild-type-strain control.

Thin-section electron microscopy. EM on wild-type and AID-tagged yeast strains was performed as previously described[63]. Briefly, yeast cells were grown in YPD at 30 °C and treated with 0.7 mM IAA for 1 h. 10 OD units were harvested, fixed for 1 h at room temperature with 3% glutaraldehyde, 2.5% sucrose, 5 mM $CaCl_2$, and 5 mM MgCl in 0.1 M sodium cacodylate, pH 7.4. Cells were spheroplasted with buffer containing 10% β-glucuronidase and 0.5 mg/ml zymolyase for 30 min at 30 °C, washed in 0.1 M cacodylate/1 M sorbitol, resuspended in 0.1 M sodium cacodylate, pH 6.8/1 M sorbitol, and embedded in 2% agarose. Agarose pieces were stained with 1% OsO_4 and 1% potassium ferrocyanide in 0.1 M sodium cacodylate, pH 6.8, for 30 min, then washed completely and stained in 1% thiocarbohydrazide for 5 min at room temperature. After being washed completely, samples were treated for 5 min with 1% OsO_4/1% potassium ferrocyanide and washed again. After ethanol dehydration and embedding in Epon resin (Electron Microscopy Science), thin sections were cut at 70 nm and added to uncoated copper grids. Grids were post-stained with uranyl acetate and lead citrate. Samples were viewed on a Philips CM10 at 80 kV and recorded with a Gatan Erlangshen 785 CCD digital camera.

Negative-stain electron microscopy and image analysis. Sec15-PrA and Sec15-GFP–Sec6-PrA complexes were purified in 20 mM PIPES, pH 6.8, and 300 mM KCl. The complexes were released from IgG beads after PPX cleavage to produce purified wild-type and Sec15-GFP complexes. Those complexes were absorbed to glow-discharged carbon-coated copper grids and stained with 1% uranyl acetate. Micrographs of wild-type complex were collected on an FEI Tecnai F20 electron microscope operated at 200 kV and 20,400× nominal magnification. The defocus value ranged from 0.5 to 2.0 µm. Images were collected with a Gatan K2 summit direct detector with final pixel size of 2.45 Å. We semiautomatically picked 67,509 Sec15-GFP particles and 24,891 wild-type particles, and performed grayscale normalization with Relion-1.3 (ref. 64). Micrographs of the Sec15-GFP complex were collected on an FEI Titan Krios electron microscope operated at 300 kV and 29,000× nominal magnification. The defocus value ranged from 0.5 to 3.0 µm. Images were collected automatically with EPU (FEI) with final pixel size of 2.87 Å. Particles were selected manually and grayscale normalized with BOXER as implemented in EMAN2 (ref. 65). For the Sec15-GFP data set, there were 2,568 unique micrographs and 67,509 particles picked; 60,751 particles survived. For the untagged wild-type data set, there were 298 unique micrographs and 24,891 particles picked; 17,420 particles survived. Contrast transfer function (CTF) estimation was performed with CTFFIND3 (ref. 66). CTF correction, 2D classification and averaging were performed via Maximum a posteriori refinement as implemented in RELION[64]. The negative-stain EM data are available upon request.

61. Sikorski, R.S. & Hieter, P. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae . Genetics 122 , 19–27 (1989).
62. Adamo, J.E., Rossi, G. & Brennwald, P. The Rho GTPase Rho3 has a direct role in exocytosis that is distinct from its role in actin polarity. Mol. Biol. Cell 10 , 4121–4133 (1999).
63. Perkins, E.M. & McCaffery, J.M. in Mitochondria 467–483 (Springer, 2007).
64. Scheres, S.H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. J. Struct. Biol. 180 , 519–530 (2012).
65. Ludtke, S.J., Baldwin, P.R. & Chiu, W. EMAN: semiautomated software for high-resolution single-particle reconstructions. J. Struct. Biol. 128 , 82–97 (1999).
66. Mindell, J.A. & Grigorieff, N. Accurate determination of local defocus and specimen tilt in electron microscopy. J. Struct. Biol. 142 , 334–347 (2003).

doi:10.1038/nsmb.3146

Bibliography

1. Yao NY, Georgescu RE, Finkelstein J, O'Donnell ME (2009) Single-molecule analysis reveals that the lagging strand increases replisome processivity but slows replication fork progression. *Proc Natl Acad Sci USA* 106(32):13236–13241.

2. Georgescu RE, et al. (2015) Reconstitution of a eukaryotic replisome reveals suppression mechanisms that define leading/lagging strand operation. *Elife* 4:e04988.

3. Baker TA, Bell SP (1998) Polymerases and the replisome: machines within machines. *Cell*.

4. Benkovic SJ, Valentine AM, Salinas F (2001) Replisome-mediated DNA replication. *Annu Rev Biochem* 70(1):181–208.

5. Georgescu R, Langston L, O'Donnell M (2015) A proposal: Evolution of PCNA's role as a marker of newly replicated DNA. *DNA Repair (Amst)* 29:4–15.

6. Neuwald AF, Aravind L, Spouge JL, Koonin EV (1999) AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. *Genome Res* 9(1):27–43.

7. Kaguni JM (2011) Replication initiation at the Escherichia coli chromosomal origin. *Current Opinion in Chemical Biology* 15(5):606–613.

8. Erzberger JP, Mott ML, Berger JM (2006) Structural basis for ATP-dependent DnaA assembly and replication-origin remodeling. *Nat Struct Mol Biol* 13(8):676–683.

9. Messer W, Blaesing F, Jakimowicz D, Krause M (2001) Bacterial replication initiator DnaA. Rules for DnaA binding and rolesof DnaA in origin unwinding and helicase loading. *Biochimie*.

10. Méchali M (2010) Eukaryotic DNA replication origins: many choices for appropriate answers. *Nat Rev Mol Cell Biol* 11(10):728–738.

11. Palzkill TG, Newlon CS (1988) A yeast replication origin consists of multiple copies of a small conserved sequence. *Cell*. doi:10.1016/0092-

8674(88)90164-X.

12. Masai H, Matsumoto S, You Z, Yoshizawa-Sugata N, Oda M (2010) Eukaryotic Chromosome DNA Replication: Where, When, and How? *http://dxdoiorg/101146/annurevbiochem052308103205* 79(1):89–130.

13. O'Donnell M, Langston L, Stillman B (2013) Principles and Concepts of DNA Replication in Bacteria, Archaea, and Eukarya. *Cold Spring Harb Perspect Biol* 5(7):a010108–a010108.

14. Bell SP, Stillman B (1992) ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. *Nature* 357(6374):128–134.

15. Speck C, Stillman B (2007) Cdc6 ATPase activity regulates ORC x Cdc6 stability and the selection of specific DNA sequences as origins of DNA replication. *Journal of Biological Chemistry* 282(16):11705–11714.

16. Bell SP, Kaguni JM (2013) Helicase Loading at Chromosomal Origins of Replication. *Cold Spring Harb Perspect Biol* 5(6):a010124–a010124.

17. Arias EE, Walter JC (2007) Strength in numbers: preventing rereplication via multiple mechanisms in eukaryotic cells. *Genes Dev* 21(5):497–518.

18. Marszalek J, Kaguni JM (1994) DnaA protein directs the binding of DnaB protein in initiation of DNA replication in Escherichia coli. *J Biol Chem* 269(7):4883–4890.

19. Ilves I, Petojevic T, Pesavento JJ, Botchan MR (2010) Activation of the MCM2-7 helicase by association with Cdc45 and GINS proteins. *Mol Cell* 37(2):247–258.

20. Moyer SE, Lewis PW, Botchan MR (2006) Isolation of the Cdc45/Mcm2-7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase. *PNAS* 103(27):10236–10241.

21. Kanemaki M, Labib K (2006) Distinct roles for Sld3 and GINS during establishment and progression of eukaryotic DNA replication forks. *EMBO J* 25(8):1753–1763.

22. Frick DN, Richardson CC (2001) DNA primases. *Annu Rev Biochem*. doi:10.1146/annurev.biochem.70.1.39.

23. Smith DJ, Whitehouse I (2012) Intrinsic coupling of lagging-strand synthesis to chromatin assembly. *Nature* 483(7390):434–438.

24. O'Donnell M, Kuriyan J (2006) Clamp loaders and replication initiation. *Curr Opin Struct Biol* 16(1):35–41.

25. Sinha NK, Morris CF, Alberts BM (1980) Efficient in vitro replication of double-stranded DNA templates by a purified T4 bacteriophage replication system. *J Biol Chem* 255(9):4290–4293.

26. Yao NY, O'Donnell M (2008) Replisome dynamics and use of DNA trombone loops to bypass replication blocks. *Mol Biosyst* 4(11):1075–1084.

27. Gao D, McHenry CS (2001) Tau binds and organizes Escherichia coli replication proteins through distinct domains. Domain III, shared by gamma and tau, binds delta delta ' and chi psi. *Journal of Biological Chemistry* 276(6):4447–4453.

28. O'Donnell M, Li H (2016) The Eukaryotic Replisome Goes Under the Microscope. *Current Biology* 26(6):R247–R256.

29. Kornberg A, Baker TA, Kornberg A (2005) *DNA Replication* (University Science Books).

30. Ciccia A, Elledge SJ (2010) The DNA damage response: making it safe to play with knives. *Mol Cell* 40(2):179–204.

31. Aziz Sancar, Laura A Lindsey-Boltz, Keziban Ünsal-Kaçmaz, Linn S (2004) Molecular Mechanisms of Mammalian DNA Repair and the DNA Damage Checkpoints. *http://dxdoiorg/101146/annurevbiochem73011303073723* 73(1):39–85.

32. McKinnon PJ (2004) ATM and ataxia telangiectasia. *EMBO reports* 5(8):772–776.

33. Cleaver JE (1968) Defective repair replication of DNA in xeroderma pigmentosum. doi:10.1038/218652a0.

34. Moriwaki S (2016) Human DNA repair disorders in dermatology: A historical perspective, current concepts and new insight. *J Dermatol Sci* 81(2):77–84.

35. Baple EL, et al. (2014) Hypomorphic PCNA mutation underlies a human DNA repair disorder. *J Clin Invest* 124(7):3137–3146.

36. Fay PJ, Johanson KO, McHenry CS, Bambara RA (1981) Size classes of products synthesized processively by DNA polymerase III and DNA

polymerase III holoenzyme of Escherichia coli. *J Biol Chem* 256(2):976–983.

37. O'Donnell ME, Kornberg A (1985) Dynamics of DNA polymerase III holoenzyme of Escherichia coli in replication of a multiprimed template. *J Biol Chem* 260(23):12875–12883.

38. Mok M, Marians KJ (1987) The Escherichia coli preprimosome and DNA B helicase can form replication forks that move at the same rate. *J Biol Chem* 262(34):16644–16654.

39. McInerney P, Johnson A, Katz F, O'Donnell M (2007) Characterization of a triple DNA polymerase replisome. *Mol Cell* 27(4):527–538.

40. Huang CC, Hearst JE, Alberts BM (1981) Two types of replication proteins increase the rate at which T4 DNA polymerase traverses the helical regions in a single-stranded DNA template. *J Biol Chem* 256(8):4087–4094.

41. Tsurimoto T, Stillman B (1989) Multiple replication factors augment DNA synthesis by the two eukaryotic DNA polymerases, alpha and delta. *EMBO J* 8(12):3883–3889.

42. Stukenberg PT, Studwell-Vaughan PS, O'Donnell M (1991) Mechanism of the sliding beta-clamp of DNA polymerase III holoenzyme. *J Biol Chem* 266(17):11328–11334.

43. Kong XP, Onrust R, O'Donnell M, Kuriyan J (1992) Three-dimensional structure of the beta subunit of E. coli DNA polymerase III holoenzyme: a sliding DNA clamp. *Cell* 69(3):425–437.

44. Kuriyan J, O'Donnell M (1993) Sliding clamps of DNA polymerases. *J Mol Biol* 234(4):915–925.

45. Ivanov I, Chapados BR, McCammon JA, Tainer JA (2006) Proliferating cell nuclear antigen loaded onto double-stranded DNA: dynamics, minor groove interactions and functional implications. *Nucleic Acids Res* 34(20):6023–6033.

46. Georgescu RE, et al. (2008) Structure of a sliding clamp on DNA. *Cell* 132(1):43–54.

47. McNally R, Bowman GD, Goedken ER, O'Donnell M, Kuriyan J (2010) Analysis of the role of PCNA-DNA contacts during clamp loading. *BMC Struct Biol* 10(1):3.

48. Shamoo Y, Steitz TA (1999) Building a replisome from interacting pieces: sliding clamp complexed to a peptide from DNA polymerase and a polymerase editing complex. *Cell* 99(2):155–166.

49. Krishna TS, Kong XP, Gary S, Burgers PM, Kuriyan J (1994) Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA. *Cell* 79(7):1233–1243.

50. Leipe DD, Aravind L, Koonin EV (1999) Did DNA replication evolve twice independently? *Nucleic Acids Res* 27(17):3389–3401.

51. Hedglin M, Kumar R, Benkovic SJ (2013) Replication clamps and clamp loaders. *Cold Spring Harb Perspect Biol* 5(4):a010165–a010165.

52. Williams GJ, et al. (2006) Structure of the heterotrimeric PCNA from Sulfolobus solfataricus. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 62(Pt 10):944–948.

53. Matsumiya S, Ishino Y, Morikawa K (2001) Crystal structure of an archaeal DNA sliding clamp: proliferating cell nuclear antigen from Pyrococcus furiosus. *Protein Sci* 10(1):17–23.

54. Moarefi I, Jeruzalmi D, Turner J, O'Donnell M, Kuriyan J (2000) Crystal structure of the DNA polymerase processivity factor of T4 bacteriophage. *J Mol Biol* 296(5):1215–1223.

55. Kong X-P, Onrust R, O'Donnell M, Kuriyan J (1992) Three-dimensional structure of the β subunit of E. coli DNA polymerase III holoenzyme: A sliding DNA clamp. *Cell* 69(3):425–437.

56. Dieckman LM, Freudenthal BD, Washington MT (2012) PCNA structure and function: insights from structures of PCNA complexes and post-translationally modified PCNA. *Subcell Biochem* 62(Chapter 15):281–299.

57. Moldovan G-L, Pfander B, Jentsch S (2007) PCNA, the maestro of the replication fork. *Cell* 129(4):665–679.

58. Gulbis JM, Kelman Z, Hurwitz J, O'Donnell M, Kuriyan J (1996) Structure of the C-terminal region of p21(WAF1/CIP1) complexed with human PCNA. *Cell* 87(2):297–306.

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

60. Hishiki A, et al. (2009) Structural basis for novel interactions between human translesion synthesis polymerases and proliferating cell nuclear antigen. *J Biol Chem* 284(16):10552–10560.

61. Bubeck D, et al. (2011) PCNA directs type 2 RNase H activity on DNA replication and repair substrates. *Nucleic Acids Res* 39(9):3652–3666.

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

63. Duffy CM, Hilbert BJ, Kelch BA (2015) A Disease-Causing Variant in PCNA Disrupts a Promiscuous Protein Binding Site. *J Mol Biol* 428(6):1023–1040.

64. Kontopidis G, et al. (2005) Structural and biochemical studies of human proliferating cell nuclear antigen complexes provide a rationale for cyclin association and inhibitor design. *PNAS* 102(6):1871–1876.

65. Altieri AS, et al. (2016) A small protein inhibits proliferating cell nuclear antigen by breaking the DNA clamp. *Nucleic Acids Res*:gkw351.

66. Georgescu RE, et al. (2008) Structure of a small-molecule inhibitor of a DNA polymerase sliding clamp. *Proc Natl Acad Sci USA* 105(32):11116–11121.

67. Fuchs R (2002) How DNA lesions are turned into mutations within cells? *Oncogene*.

68. Indiani C, McInerney P, Georgescu R, Goodman MF, O'Donnell M (2005) A sliding-clamp toolbelt binds high- and low-fidelity DNA polymerases simultaneously. *Mol Cell* 19(6):805–815.

69. Dionne I, Nookala RK, Jackson SP, Doherty AJ, Bell SD (2003) A heterotrimeric PCNA in the hyperthermophilic archaeon Sulfolobus solfataricus. *Mol Cell* 11(1):275–282.

70. Beattie TR, Bell SD (2012) Coordination of multiple enzyme activities by a single PCNA in archaeal Okazaki fragment maturation. *EMBO J* 31(6):1556–1567.

71. Dovrat D, Stodola JL, Burgers PMJ, Aharoni A (2014) Sequential switching of binding partners on PCNA during in vitro Okazaki fragment maturation. *Proc Natl Acad Sci USA* 111(39):14118–14123.

72. Warbrick E (1998) PCNA binding through a conserved motif. *Bioessays* 20(3):195–199.

73. Gilljam KM, et al. (2009) Identification of a novel, widespread, and functionally important PCNA-binding motif. *J Cell Biol* 186(5):645–654.

74. Fu D, Samson LD, Hübscher U, van Loon B (2015) The interaction between ALKBH2 DNA repair enzyme and PCNA is direct, mediated by the hydrophobic pocket of PCNA and perturbed in naturally-occurring ALKBH2 variants. *DNA Repair (Amst)* 35:13–18.

75. Blanco L, Salas M (1985) Replication of phage phi 29 DNA with purified terminal protein and DNA polymerase: synthesis of full-length phi 29 DNA. *PNAS* 82(19):6404–6408.

76. Andraos N, Tabor S, Richardson CC (2004) The Highly Processive DNA Polymerase of Bacteriophage T5 ROLE OF THE UNIQUE N AND C TERMINI. *Journal of Biological Chemistry* 279(48):50609–50618.

77. Langston LD, Indiani C, O'Donnell M (2009) Whither the replisome: Emerging perspectives on the dynamic nature of the DNA replication machinery. *Cell Cycle* 8(17):2686–2691.

78. Tabor S, Huber HE, Richardson CC (1987) Escherichia coli thioredoxin confers processivity on the DNA polymerase activity of the gene 5 protein of bacteriophage T7. *Journal of Biological Chemistry* 262(33):16212–16223.

79. Freudenthal BD, Gakhar L, Ramaswamy S, Washington MT (2010) Structure of monoubiquitinated PCNA and implications for translesion synthesis and DNA polymerase exchange. *Nat Struct Mol Biol* 17(4):479–484.

80. Freudenthal BD, Brogie JE, Gakhar L, Kondratick CM, Washington MT (2011) Crystal structure of SUMO-modified proliferating cell nuclear antigen. *J Mol Biol* 406(1):9–17.

81. Umar A, et al. (1996) Requirement for PCNA in DNA Mismatch Repair at a Step Preceding DNA Resynthesis. *Cell* 87(1):65–73.

82. Ko R, Bennett SE (2005) Physical and functional interaction of human nuclear uracil-DNA glycosylase with proliferating cell nuclear antigen. *DNA Repair (Amst)* 4(12):1421–1431.

83. Gilljam KM, Müller R, Liabakk NB, Otterlei M (2012) Nucleotide Excision

Repair Is Associated with the Replisome and Its Efficiency Depends on a Direct Interaction between XPA and PCNA. *PLoS ONE* 7(11):e49199–11.

84. Gary R, Ludwig DL, Cornelius HL, MacInnes MA, Park MS (1997) The DNA repair endonuclease XPG binds to proliferating cell nuclear antigen (PCNA) and shares sequence elements with the PCNA-binding regions of FEN-1 and cyclin-dependent kinase inhibitor p21. *J Biol Chem* 272(39):24522–24529.

85. Williams JS, Kunkel TA (2014) Ribonucleotides in DNA: Origins, repair and consequences. *DNA Repair (Amst)* 19:27–37.

86. Sparks JL, et al. (2012) RNase H2-Initiated Ribonucleotide Excision Repair. *Mol Cell* 47(6):980–986.

87. Hoege C, Pfander B, Moldovan G-L, Pyrowolakis G, Jentsch S (2002) RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. *Nature* 419(6903):135–141.

88. Kannouche PL, Wing J, Lehmann AR (2004) Interaction of human DNA polymerase eta with monoubiquitinated PCNA: a possible mechanism for the polymerase switch in response to DNA damage. *Mol Cell* 14(4):491–500.

89. Park JM, et al. (2014) Modification of PCNA by ISG15 plays a crucial role in termination of error-prone translesion DNA synthesis. *Mol Cell* 54(4):626–638.

90. Pfander B, Moldovan G-L, Sacher M, Hoege C, Jentsch S (2005) SUMO-modified PCNA recruits Srs2 to prevent recombination during S phase. *Nature* 436(7049):428–433.

91. Gali H, et al. (2012) Role of SUMO modification of human PCNA at stalled replication fork. *Nucleic Acids Res* 40(13):6049–6059.

92. Arias EE, Walter JC (2006) PCNA functions as a molecular platform to trigger Cdt1 destruction and prevent re-replication. *Nature Cell Biology* 8(1):84–90.

93. Waga S, Hannon GJ, Beach D, Stillman B (1994) The p21 inhibitor of cyclin-dependent kinases controls DNA replication by interaction with PCNA. *Nature* 369(6481):574–578.

94. Vairapandi M, Azam N, Balliet AG, Hoffman B, Liebermann DA (2000) Characterization of MyD118, Gadd45, and proliferating cell nuclear

antigen (PCNA) interacting domains. PCNA impedes MyD118 AND Gadd45-mediated negative growth control. *Journal of Biological Chemistry* 275(22):16810–16819.

95. Scott M, et al. (2001) UV-induced binding of ING1 to PCNA regulates the induction of apoptosis. *J Cell Sci* 114(Pt 19):3455–3462.

96. Iida T, et al. (2002) PCNA clamp facilitates action of DNA cytosine methyltransferase 1 on hemimethylated DNA. *Genes Cells* 7(10):997–1007.

97. Mortusewicz O, Schermelleh L, Walter J, Cardoso MC, Leonhardt H (2005) Recruitment of DNA methyltransferase I to DNA repair sites. *PNAS* 102(25):8905–8909.

98. Zhang K, et al. (2016) A DNA binding winged helix domain in CAF-1 functions with PCNA to stabilize CAF-1 at replication forks. *Nucleic Acids Res*:gkw106.

99. Kelch BA, Makino DL, O'Donnell M, Kuriyan J (2012) Clamp loader ATPases and the evolution of DNA replication machinery. *BMC Biol* 10(1):34.

100. Kelch B (2016) The Lord of the Rings: Structure and mechanism of the sliding clamp loader. *Biopolymers*:n/a–n/a.

101. Jeruzalmi D, O'Donnell M, Kuriyan J (2001) Crystal structure of the processivity clamp loader gamma (gamma) complex of E. coli DNA polymerase III. *Cell* 106(4):429–441.

102. Bowman GD, O'Donnell M, Kuriyan J (2004) Structural analysis of a eukaryotic sliding DNA clamp-clamp loader complex. *Nature* 429(6993):724–730.

103. Kelch BA, Makino DL, O'Donnell M, Kuriyan J (2011) How a DNA polymerase clamp loader opens a sliding clamp. *Science* 334(6063):1675–1680.

104. Miyata T, et al. (2005) Open clamp structure in the clamp-loading complex visualized by electron microscopic image analysis. *Proc Natl Acad Sci USA* 102(39):13795–13800.

105. Hanson PI, Whiteheart SW (2005) AAA+ proteins: have engine, will work. *Nat Rev Mol Cell Biol* 6(7):519–529.

106. Erzberger JP, Berger JM (2006) Evolutionary relationships and structural mechanisms of AAA+ proteins. *Annu Rev Biophys Biomol Struct* 35:93–114.

107. Guenther B, Onrust R, Sali A, O'Donnell M, Kuriyan J (1997) Crystal structure of the delta' subunit of the clamp-loader complex of E. coli DNA polymerase III. *Cell* 91(3):335–345.

108. Marceau AH, et al. (2011) Structure of the SSB-DNA polymerase III interface and its role in DNA replication. *EMBO J* 30(20):4236–4247.

109. Seybert A, Scott DJ, Scaife S, Singleton MR, Wigley DB (2002) Biochemical characterisation of the clamp/clamp loader proteins from the euryarchaeon Archaeoglobus fulgidus. *Nucleic Acids Res* 30(20):4329–4338.

110. Jarvis TC, Paul LS, Hippel von PH (1989) Structural and enzymatic studies of the T4 DNA replication system. I. Physical characterization of the polymerase accessory protein complex. *J Biol Chem* 264(21):12709–12716.

111. Maki S, Kornberg A (1988) DNA polymerase III holoenzyme of Escherichia coli. II. A novel complex including the gamma subunit essential for processive synthesis. *J Biol Chem* 263(14):6555–6560.

112. Onrust R, et al. (1995) Assembly of a chromosomal replication machine: two DNA polymerases, a clamp loader, and sliding clamps in one holoenzyme particle. I. Organization of the clamp loader. *J Biol Chem* 270(22):13348–13357.

113. Schmidt SL, Gomes XV, Burgers PM (2001) ATP utilization by yeast replication factor C. III. The ATP-binding domains of Rfc2, Rfc3, and Rfc4 are essential for DNA recognition and clamp loading. *J Biol Chem* 276(37):34784–34791.

114. Hingorani MM, O'Donnell M (1998) ATP binding to the Escherichia coli clamp loader powers opening of the ring-shaped clamp of DNA polymerase III holoenzyme. *J Biol Chem* 273(38):24550–24563.

115. Turner J, Hingorani MM, Kelman Z, O'Donnell M (1999) The internal workings of a DNA polymerase clamp-loading machine. *EMBO J* 18(3):771–783.

116. Simonetta KR, et al. (2009) The mechanism of ATP-dependent primer-template recognition by a clamp loader complex. *Cell* 137(4):659–671.

117. Kazmirski SL, Podobnik M, Weitze TF, O'Donnell M, Kuriyan J (2004) Structural analysis of the inactive state of the Escherichia coli DNA polymerase clamp-loader complex. *Proc Natl Acad Sci USA* 101(48):16750–16755.

118. Gomes XV, Schmidt SL, Burgers PM (2001) ATP utilization by yeast replication factor C. II. Multiple stepwise ATP binding events are required to load proliferating cell nuclear antigen onto primed DNA. *Journal of Biological Chemistry* 276(37):34776–34783.

119. Yao N, et al. (1996) Clamp loading, unloading and intrinsic stability of the PCNA, β and gp45 sliding clamps of human, E. coli and T4 replicases. *Genes to Cells* 1(1):101–113.

120. Binder JK, et al. (2014) Intrinsic stability and oligomerization dynamics of DNA processivity clamps. *Nucleic Acids Res* 42(10):6476–6486.

121. Jeruzalmi D, et al. (2001) Mechanism of processivity clamp opening by the delta subunit wrench of the clamp loader complex of E. coli DNA polymerase III. *Cell* 106(4):417–428.

122. Alley SC, Abel-Santos E, Benkovic SJ (2000) Tracking sliding clamp opening and closing during bacteriophage T4 DNA polymerase holoenzyme assembly. *Biochemistry* 39(11):3076–3090.

123. Bowman GD, Goedken ER, Kazmirski SL, O'Donnell M, Kuriyan J (2005) DNA polymerase clamp loaders and DNA recognition. *FEBS Lett* 579(4):863–867.

124. Goedken ER, Kazmirski SL, Bowman GD, O'Donnell M, Kuriyan J (2005) Mapping the interaction of DNA with the Escherichia coli DNA polymerase clamp loader complex. *Nat Struct Mol Biol* 12(2):183–190.

125. Fedoroff OYu, Salazar M, Reid BR (1993) Structure of a DNA:RNA hybrid duplex. Why RNase H does not cleave pure RNA. *J Mol Biol* 233(3):509–523.

126. Ason B, et al. (2003) Mechanism of loading the Escherichia coli DNA polymerase III beta sliding clamp on DNA. Bona fide primer/templates preferentially trigger the gamma complex to hydrolyze ATP and load the clamp. *J Biol Chem* 278(12):10033–10040.

127. Jarvis TC, Paul LS, Hockensmith JW, Hippel von PH (1989) Structural and enzymatic studies of the T4 DNA replication system. II. ATPase properties of the polymerase accessory protein complex. *J Biol Chem*

264(21):12717–12729.

128. Badaracco G, Bianchi M, Valsasnini P, Magni G, Plevani P (1985) Initiation, elongation and pausing of in vitro DNA synthesis catalyzed by immunopurified yeast DNA primase: DNA polymerase complex. *EMBO J* 4(5):1313–1317.

129. Pellegrini L (2012) The Pol α-primase complex. *Subcell Biochem* 62(Chapter 9):157–169.

130. Hendrix RW, Smith M, Burns RN (1999) Evolutionary relationships among diverse bacteriophages and prophages: all the world'sa phage doi:10.1073/pnas.96.5.2192.

131. Wommack KE, Colwell RR (2000) Virioplankton: viruses in aquatic ecosystems. *Microbiol Mol Biol Rev* 64(1):69–114.

132. Chistol G, et al. (2012) High Degree of Coordination and Division of Labor among Subunits in a Homomeric Ring ATPase. *Cell* 151(5):1017–1028.

133. Enemark EJ, Joshua-Tor L (2008) On helicases and other motor proteins. *Curr Opin Struct Biol* 18(2):243–257.

134. Adelman JL, et al. (2006) Mechanochemistry of transcription termination factor Rho. *Mol Cell* 22(5):611–621.

135. Johnson A, O'Donnell M (2005) Cellular DNA replicases: components and dynamics at the replication fork. *Annu Rev Biochem* 74(1):283–315.

136. Fay PJ, Johanson KO, McHenry CS, Bambara RA (1982) Size classes of products synthesized processively by two subassemblies of Escherichia coli DNA polymerase III holoenzyme. *Journal of Biological Chemistry* 257(10):5692–5699.

137. Moggs JG, et al. (2000) A CAF-1-PCNA-mediated chromatin assembly pathway triggered by sensing DNA damage. *Mol Cell Biol* 20(4):1206–1218.

138. Schermelleh L, et al. (2007) Dynamics of Dnmt1 interaction with the replication machinery and its role in postreplicative maintenance of DNA methylation. *Nucleic Acids Res* 35(13):4301–4312.

139. Liang Z, Diamond M, Smith JA, Schnell M, Daniel R (2011) Proliferating cell nuclear antigen is required for loading of the SMCX/KMD5C histone demethylase onto chromatin. *Epigenetics Chromatin* 4(1):18.

140. Pacaud R, et al. (2014) The DNMT1/PCNA/UHRF1 disruption induces tumorigenesis characterized by similar genetic and epigenetic signatures. *Sci Rep* 4:1–9.

141. Witko-Sarsat V, et al. (2010) Proliferating cell nuclear antigen acts as a cytoplasmic platform controlling human neutrophil survival. *J Exp Med* 207(12):2631–2645.

142. De Chiara A, et al. (2013) Characterization of cytosolic proliferating cell nuclear antigen (PCNA) in neutrophils: antiapoptotic role of the monomer. *Journal of Leukocyte Biology* 94(4):723–731.

143. Yin L, et al. (2015) The S-Nitrosylation Status of PCNA Localized in Cytosol Impacts the Apoptotic Pathway in a Parkinson's Disease Paradigm. *PLoS ONE* 10(2):e0117546.

144. Yang J, Zhuang Z, Roccasecca RM, Trakselis MA, Benkovic SJ (2004) The dynamic processivity of the T4 DNA polymerase during replication. *PNAS* 101(22):8289–8294.

145. Mailand N, Gibbs-Seymour I, Bekker-Jensen S (2013) Regulation of PCNA–protein interactions for genome stability. *Nature Publishing Group* 14(5):269–282.

146. Fridman Y, et al. (2010) Subtle Alterations in PCNA-Partner Interactions Severely Impair DNA Replication and Repair. *PLoS Biol* 8(10):e1000507–15.

147. Liu H, Naismith JH (2008) An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotechnol* 8(1):91.

148. Otwinowski Z, Minor W (1997) [20] Processing of X-ray diffraction data collected in oscillation mode. *Macromolecular Crystallography Part A*, Methods in Enzymology. (Elsevier), pp 307–326.

149. McCoy AJ, et al. (2007) Phaser crystallographic software. *J Appl Crystallogr* 40(Pt 4):658–674.

150. Emsley P, Cowtan K, IUCr (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(12):2126–2132.

151. Adams PD, et al. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol*

Crystallogr 66(Pt 2):213–221.

152. Maiti R, Van Domselaar GH, Zhang H, Wishart DS (2004) SuperPose: a simple server for sophisticated structural superposition. *Nucleic Acids Res* 32(Web Server):W590–W594.

153. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372(3):774–797.

154. Willard L, et al. (2003) VADAR: a web server for quantitative evaluation of protein structure quality. *Nucleic Acids Res* 31(13):3316–3319.

155. Fukuda K, et al. (1995) Structure-function relationship of the eukaryotic DNA replication factor, proliferating cell nuclear antigen. *J Biol Chem* 270(38):22527–22534.

156. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *PNAS* 98(18):10037–10041.

157. Sakurai S, et al. (2005) Structural basis for recruitment of human flap endonuclease 1 to PCNA. *EMBO J* 24(4):683–693.

158. Punchihewa C, et al. (2012) Identification of small molecule proliferating cell nuclear antigen (PCNA) inhibitor that disrupts interactions with PIP-box proteins and inhibits DNA replication. *J Biol Chem* 287(17):14289–14300.

159. Inoue A, et al. (2014) A small molecule inhibitor of monoubiquitinated Proliferating Cell Nuclear Antigen (PCNA) inhibits repair of interstrand DNA cross-link, enhances DNA double strand break, and sensitizes cancer cells to cisplatin. *J Biol Chem* 289(10):7109–7120.

160. Chovancova E, et al. (2012) CAVER 3.0: A Tool for the Analysis of Transport Pathways in Dynamic Protein Structures. *PLoS Comput Biol* 8(10):e1002708–12.

161. Gartel AL, Radhakrishnan SK (2005) Lost in transcription: p21 repression, mechanisms, and consequences. *Cancer Res* 65(10):3980–3985.

162. Reijns MAM, Jackson AP (2014) Ribonuclease H2 in health and disease. *Biochem Soc Trans* 42(4):717–725.

163. Balakrishnan L, Bambara RA (2013) Flap Endonuclease 1. *Annu Rev*

Biochem 82(1):119–138.

164. Park S-S, et al. (2012) Effective correction of experimental errors in quantitative proteomics using stable isotope labeling by amino acids in cell culture (SILAC). *Journal of Proteomics* 75(12):3720–3732.

165. Kroker AJ, Bruning JB (2015) p21 Exploits Residue Tyr151 as a Tether for High-Affinity PCNA Binding. *Biochemistry* 54(22):3483–3493.

166. Flores-Rozas H, et al. (1994) Cdk-interacting protein 1 directly binds with proliferating cell nuclear antigen and inhibits DNA replication catalyzed by the DNA polymerase delta holoenzyme. *PNAS* 91(18):8655–8659.

167. Wang Y, et al. (2011) Intrinsic disorder mediates the diverse regulatory functions of the Cdk inhibitor p21. *Nature Chemical Biology*:1–8.

168. Wang S-C, et al. (2006) Tyrosine phosphorylation controls PCNA function through protein stability. *Nature Cell Biology* 8(12):1359–1368.

169. Cazzalini O, et al. (2014) CBP and p300 acetylate PCNA to link its degradation with nucleotide excision repair synthesis. *Nucleic Acids Res* 42(13):8433–8448.

170. Han J, et al. (2014) SIVA1 directs the E3 ubiquitin ligase RAD18 for PCNA monoubiquitination. *J Cell Biol* 205(6):811–827.

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

172. Fischer E (1894) Einfluss der Configuration auf die Wirkung der Enzyme. II. *Berichte der deutschen chemischen Gesellschaft* 27(3):3479–3483.

173. Koshland DE (1958) Application of a Theory of Enzyme Specificity to Protein Synthesis. *PNAS* 44(2):98–104.

174. De Biasio A, et al. (2011) Reduced stability and increased dynamics in the human proliferating cell nuclear antigen (PCNA) relative to the yeast homolog. *PLoS ONE* 6(2):e16600.

175. Kazmirski SL, Zhao Y, Bowman GD, O'Donnell M, Kuriyan J (2005) Out-of-plane motions in open sliding clamps: molecular dynamics simulations of eukaryotic and archaeal proliferating cell nuclear antigen. *Proc Natl Acad Sci USA* 102(39):13801–13806.

176. Pedley AM, Lill MA, Davisson VJ (2014) Flexibility of PCNA-Protein

Interface Accommodates Differential Binding Partners. *PLoS ONE* 9(7):e102481–12.

177. Elledge ACANYHIHYALISPBAJYSKDLLHS, et al. (2012) Polyubiquitinated PCNA Recruits the ZRANB3 Translocase to Maintain Genomic Integrity after Replication Stress. *Mol Cell* 47(3):396–409.

178. Olaisen C, Müller R, Nedal A, Otterlei M (2015) PCNA-interacting peptides reduce Akt phosphorylation and TLR-mediated cytokine secretion suggesting a role of PCNA in cellular signaling. *Cellular Signalling* 27(7):1478–1487.

179. Müller R, et al. (2013) Targeting Proliferating Cell Nuclear Antigen and Its Protein Interactions Induces Apoptosis in Multiple Myeloma Cells. *PLoS ONE* 8(7):e70430–12.

180. Gomes XV, Burgers PM (2000) Two modes of FEN1 binding to PCNA regulated by DNA. *EMBO J* 19(14):3811–3821.

181. Larsen E, et al. (2008) Early-onset lymphoma and extensive embryonic apoptosis in two domain-specific Fen1 mice mutants. *Cancer Res* 68(12):4571–4579.

182. Lau PJ, Flores-Rozas H, Kolodner RD (2002) Isolation and characterization of new proliferating cell nuclear antigen (POL30) mutator mutants that are defective in DNA mismatch repair. *Mol Cell Biol* 22(19):6669–6680.

183. Reijns MAM, et al. (2012) Enzymatic Removal of Ribonucleotides from DNA Is Essential for Mammalian Genome Integrity and Development. *Cell* 149(5):1008–1022.

184. Crow YJ, et al. (2006) Mutations in genes encoding ribonuclease H2 subunits cause Aicardi-Goutières syndrome and mimic congenital viral brain infection. *Nat Genet* 38(8):910–916.

185. Green CM, Baple EL, Crosby AH (2014) PCNA mutation affects DNA repair not replication. *Cell Cycle* 13(20):3157–3158.

186. Xia L, et al. (2005) Human 3-Methyladenine-DNA Glycosylase: Effect of Sequence Context on Excision, Association with PCNA, and Stimulation by AP Endonuclease. *J Mol Biol* 346(5):1259–1274.

187. Otterlei M, et al. (1999) Post-replicative base excision repair in replication foci. *EMBO J* 18(13):3834–3844.

188. Kedar PS, et al. (2002) Direct interaction between mammalian DNA polymerase beta and proliferating cell nuclear antigen. *J Biol Chem* 277(34):31115–31123.

189. Irina I Dianova, Vilhelm A Bohr A, Grigory L Dianov (2001) *Interaction of Human AP Endonuclease 1 with Flap Endonuclease 1 and Proliferating Cell Nuclear Antigen Involved in Long-Patch Base Excision Repair* (American Chemical Society).

190. Montecucco A, et al. (1998) DNA ligase I is recruited to sites of DNA replication by an interaction with proliferating cell nuclear antigen: identification of a common targeting mechanism for the assembly of replication factories. *EMBO J* 17(13):3786–3795.

191. Fan J, Otterlei M, Wong HK, Tomkinson AE, Wilson DM (2004) XRCC1 co-localizes and physically interacts with PCNA. *Nucleic Acids Res* 32(7):2193–2201.

192. Nick McElhinny SA, et al. (2010) Abundant ribonucleotide incorporation into DNA by yeast replicative polymerases. *Proc Natl Acad Sci USA* 107(11):4949–4954.

193. Hiom K (2010) Coping with DNA double strand breaks. *DNA Repair (Amst)* 9(12):1256–1263.

194. Chance B, Sies H, Boveris A (1979) Hydroperoxide metabolism in mammalian organs. *Physiological reviews*.

195. Kasparek TR, Humphrey TC (2011) DNA double-strand break repair pathways, chromosomal rearrangements and cancer. *Semin Cell Dev Biol* 22(8):886–897.

196. Sung P, Robberson DL (1995) DNA strand exchange mediated by a RAD51-ssDNA nucleoprotein filament with polarity opposite to that of RecA. *Cell* 82(3):453–461.

197. Li X, Stith CM, Burgers PM, Heyer W-D (2009) PCNA Is Required for Initiation of Recombination-Associated DNA Synthesis by DNA Polymerase δ. *Mol Cell* 36(4):704–713.

198. Holmes AM, Haber JE (1999) Double-Strand Break Repair in Yeast Requires Both Leading and Lagging Strand DNA Polymerases. *Cell* 96(3):415–424.

199. Lieber MR (2010) The Mechanism of Double-Strand DNA Break Repair

by the Nonhomologous DNA End Joining Pathway. *Annu Rev Biochem* 79(1):181–211.

200. Prakash S, Johnson RE, Prakash L (2005) Eukaryotic translesion synthesis DNA polymerases: specificity of structure and function. *Annu Rev Biochem* 74:317–353.

201. Parker JL, Ulrich HD (2012) A SUMO-interacting motif activates budding yeast ubiquitin ligase Rad18 towards SUMO-modified PCNA. *Nucleic Acids Res* 40(22):11380–11388.

202. Veaute X, et al. (2003) The Srs2 helicase prevents recombination by disrupting Rad51 nucleoprotein filaments. *Nature* 423(6937):309–312.

203. Ulrich HD (2007) Conservation of DNA damage tolerance pathways from yeast to humans. *Biochem Soc Trans* 35(Pt 5):1334–1337.

204. Fronza G, Campomenosi P, Iannone R, Abbondandolo A (1992) The 4-nitroquinoline 1-oxide mutational spectrum in single stranded DNA is characterized by guanine to pyrimidine transversions. *Nucleic Acids Res* 20(6):1283–1287.

205. Memisoglu A, Samson L (2000) Contribution of base excision repair, nucleotide excision repair, and DNA recombination to alkylation resistance of the fission yeast Schizosaccharomyces pombe. *J Bacteriol* 182(8):2104–2112.

206. Noguchi E, Ansbach AB, Noguchi C, Russell P (2009) Assays used to study the DNA replication checkpoint in fission yeast. *Methods Mol Biol* 521(Chapter 28):493–507.

207. Morishita T, Tsutsui Y, Iwasaki H, Shinagawa H (2002) The Schizosaccharomyces pombe rad60 gene is essential for repairing double-strand DNA breaks spontaneously occurring during replication and induced by DNA-damaging agents. *Mol Cell Biol* 22(10):3537–3548.

208. Enoch T, Carr AM, Nurse P (1992) Fission yeast genes involved in coupling mitosis to completion of DNA replication. *Genes Dev* 6(11):2035–2046.

209. Koc A, Wheeler LJ, Mathews CK, Merrill GF (2004) Hydroxyurea Arrests DNA Replication by a Mechanism That Preserves Basal dNTP Pools. *Journal of Biological Chemistry* 279(1):223–230.

210. Lundin C, et al. (2002) Different roles for nonhomologous end joining and

homologous recombination following replication arrest in mammalian cells. *Mol Cell Biol* 22(16):5869–5878.

211. Sinha RP, Häder D-P (2002) UV-induced DNA damage and repair: a review. *Photochem Photobiol Sci* 1(4):225–236.

212. Forsburg SL, Rhind N (2006) Basic methods for fission yeast. *Yeast* 23(3):173–183.

213. Gao J, et al. (2014) Rapid, efficient and precise allele replacement in the fission yeast Schizosaccharomyces pombe. *Curr Genet* 60(2):109–119.

214. Sabatinos SA, Forsburg SL (2010) Molecular genetics of Schizosaccharomyces pombe. *Meth Enzymol* 470:759–795.

215. Boeke JD, Trueheart J, Natsoulis G, Fink GR (1987) 5-Fluoroorotic acid as a selective agent in yeast molecular genetics. *Meth Enzymol* 154:164–175.

216. Freyer GA, et al. (1995) An alternative eukaryotic DNA excision repair pathway. *Mol Cell Biol* 15(8):4572–4577.

217. Yonemasu R, et al. (1997) Characterization of the alternative excision repair pathway of UV-damaged DNA in Schizosaccharomyces pombe. *Nucleic Acids Res* 25(8):1553–1558.

218. McCready SJ, Osman F, Yasui A (2000) Repair of UV damage in the fission yeast Schizosaccharomyces pombe. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* 451(1-2):197–210.

219. Murray JM, et al. (1994) Structural and functional conservation of the human homolog of the Schizosaccharomyces pombe rad2 gene, which is required for chromosome segregation and recovery from DNA damage. *Mol Cell Biol* 14(7):4878–4888.

220. Kanno S, Iwai S, Takao M, Yasui A (1999) Repair of apurinic/apyrimidinic sites by UV damage endonuclease; a repair protein for UV and oxidative damage. *Nucleic Acids Res* 27(15):3096–3103.

221. Nick McElhinny SA, Ramsden DA (2003) Polymerase mu is a DNA-directed DNA/RNA polymerase. *Mol Cell Biol* 23(7):2309–2315.

222. Miyabe I, Kunkel TA, Carr AM (2011) The Major Roles of DNA Polymerases Epsilon and Delta at the Eukaryotic Replication Fork Are

Evolutionarily Conserved. *PLoS Genet* 7(12):e1002407.

223. McElhinny SAN, et al. (2010) Genome instability due to ribonucleotide incorporation into DNA. *Nature Chemical Biology* 6(10):774–781.

224. Abraham RT (2001) Cell cycle checkpoint signaling through the ATM and ATR kinases. *Genes Dev* 15(17):2177–2196.

225. Gamper AM, et al. (2012) ATM protein physically and functionally interacts with proliferating cell nuclear antigen to regulate DNA synthesis. *J Biol Chem* 287(15):12445–12454.

226. Melo J, Toczyski D (2002) A unified view of the DNA-damage checkpoint. *Curr Opin Cell Biol* 14(2):237–245.

227. Chen MJ, Lin YT, Lieberman HB, Chen G, Lee EY (2001) ATM-dependent phosphorylation of human Rad9 is required for ionizing radiation-induced checkpoint activation. *Journal of Biological Chemistry* 276(19):16580–16586.

228. Shiomi Y, et al. (2002) Clamp and clamp loader structures of the human checkpoint protein complexes, Rad9-1-1 and Rad17-RFC. *Genes to Cells* 7(8):861–868.

229. Venclovas C, Thelen MP (2000) Structure-based predictions of Rad1, Rad9, Hus1 and Rad17 participation in sliding clamp and clamp-loading complexes. *Nucleic Acids Res* 28(13):2481–2493.

230. Arai K, Kornberg A (1979) A general priming system employing only dnaB protein and primase for DNA replication. *PNAS* 76(9):4308–4312.

231. Kurosawa Y, Okazaki T (1979) Structure of the RNA portion of the RNA-linked DNA pieces in bacteriophage T4-infected Escherichia coli cells. *J Mol Biol* 135(4):841–861.

232. Drew HR, et al. (1981) Structure of a B-DNA dodecamer: conformation and dynamics. *PNAS* 78(4):2179–2183.

233. Xiong Y, Sundaralingam M (2000) Crystal structure of a DNA.RNA hybrid duplex with a polypurine RNA r(gaagaagag) and a complementary polypyrimidine DNA d(CTCTTCTTC). *Nucleic Acids Res* 28(10):2171–2176.

234. Ban C, Ramakrishnan B, Sundaralingam M (1994) A Single 2′-Hydroxyl Group Converts B-DNA to A-DNA: Crystal Structure of the DNA-RNA

Chimeric Decamer Duplex d (CCGGC) r (G) d (CCGG) with a Novel …. *J Mol Biol*. doi:10.1006/jmbi.1994.1134.

235. Chemla YR, et al. (2005) Mechanism of Force Generation of a Viral DNA Packaging Motor. *Cell* 122(5):683–692.

236. Enemark EJ, Joshua-Tor L (2006) Mechanism of DNA translocation in a replicative hexameric helicase. *Nature* 442(7100):270–275.

237. Thomsen ND, Berger JM (2009) Running in reverse: the structural basis for translocation polarity in hexameric helicases. *Cell* 139(3):523–534.

238. Kornberg A, Pricer WE (1951) Enzymatic phosphorylation of adenosine and 2,6-diaminopurine riboside. *Journal of Biological Chemistry* 193(2):481–495.

239. Hattendorf DA, Lindquist SL (2002) Cooperative kinetics of both Hsp104 ATPase domains and interdomain communication revealed by AAA sensor-1 mutants. *EMBO J* 21(1-2):12–21.

240. Banavali NK, Roux B (2005) Free energy landscape of A-DNA to B-DNA conversion in aqueous solution. *J Am Chem Soc* 127(18):6866–6876.

241. Trakselis MA, Berdis AJ, Benkovic SJ (2003) Examination of the role of the clamp-loader and ATP hydrolysis in the formation of the bacteriophage T4 polymerase holoenzyme. *J Mol Biol* 326(2):435–451.

242. Pietroni P, Hippel von PH (2008) Multiple ATP binding is required to stabilize the “activated” (clamp open) clamp loader of the T4 DNA replication complex. *J Biol Chem* 283(42):28338–28353.

243. Johnson A, O'Donnell M (2003) Ordered ATP hydrolysis in the gamma complex clamp loader AAA+ machine. *J Biol Chem* 278(16):14406–14413.

244. Marzahn MR, Hayner JN, Finkelstein J, O'Donnell M, Bloom LB (2014) The ATP sites of AAA+ clamp loaders work together as a switch to assemble clamps on DNA. *J Biol Chem* 289(9):5537–5548.

245. Ramanathan A, Savol A, Burger V, Chennubhotla CS, Agarwal PK (2014) Protein conformational populations and functionally relevant substates. *Acc Chem Res* 47(1):149–156.

246. Fraser JS, et al. (2011) Accessing protein conformational ensembles using room-temperature X-ray crystallography. *Proc Natl Acad Sci USA*

108(39):16247–16252.

247. Warkentin M, et al. (2013) Global radiation damage: temperature dependence, time dependence and how to outrun it. *J Synchrotron Radiat* 20(Pt 1):7–13.

248. Van Den Bedem H, Dhanik A, Latombe JC, Deacon AM, IUCr (2009) Modeling discrete heterogeneity in X-ray diffraction data by fitting multi-conformers. *Acta Crystallogr D Biol Crystallogr* 65(10):1107–1117.

249. Keedy DA, Fraser JS, van den Bedem H (2015) Exposing Hidden Alternative Backbone Conformations in X-ray Crystallography Using qFit. *PLoS Comput Biol* 11(10):e1004507.

250. Lang PT, et al. (2010) Automated electron-density sampling reveals widespread conformational polymorphism in proteins. *Protein Science* 19(7):1420–1431.

251. Jiang L, Mishra P, Hietpas RT, Zeldovich KB, Bolon DNA (2013) Latent effects of Hsp90 mutants revealed at reduced expression levels. *PLoS Genet* 9(6):e1003600.

252. Hietpas RT, Jensen JD, Bolon DNA (2011) Experimental illumination of a fitness landscape. *Proc Natl Acad Sci USA* 108(19):7896–7901.

253. Roscoe BP, Thayer KM, Zeldovich KB, Fushman D, Bolon DNA (2013) Analyses of the Effects of All Ubiquitin Point Mutants on Yeast Growth Rate. *J Mol Biol* 425(8):1363–1377.

254. Baryshnikova A, et al. (2010) Synthetic Genetic Array (SGA) Analysis in Saccharomyces cerevisiae and Schizosaccharomyces pombe. *Guide to Yeast Genetics: Functional Genomics, Proteomics, and Other Systems Analysis*, Methods in Enzymology. (Elsevier), pp 145–179.

255. Luo M, et al. (2012) Characterization of the redox activity and disulfide bond formation in apurinic/apyrimidinic endonuclease. *Biochemistry* 51(2):695–705.

256. Gabel SA, et al. (2014) Characterization of the redox transition of the XRCC1 N-terminal domain. *Structure* 22(12):1754–1763.

257. Stankovic Valentin N, Drzewicka K, König C, Schiebel E, Melchior F (2016) Redox regulation of SUMO enzymes is required for ATM activity and survival in oxidative stress. *EMBO J* 35(12):1312–1329.

258. Storr SJ, Woolston CM, Zhang Y, Martin SG (2013) Redox Environment, Free Radical, and Oxidative DNA Damage. *http://dxdoiorgezproxyumassmededu/101089/ars20124920* 18(18):2399–2408.

259. Ravelli R, McSweeney SM (2000) The “fingerprint”that X-rays can leave on structures. *Structure*.

260. Dalal S, Rosser MFN, Cyr DM, Hanson PI (2004) Distinct roles for the AAA ATPases NSF and p97 in the secretory pathway. *Mol Biol Cell* 15(2):637–648.

261. Babst M, Wendland B, Estepa EJ, Emr SD (1998) The Vps4p AAA ATPase regulates membrane association of a Vps protein complex required for normal endosome function. *EMBO J* 17(11):2982–2993.

262. Hersch GL, Burton RE, Bolon DN, Baker TA, Sauer RT (2005) Asymmetric interactions of ATP with the AAA+ ClpX6 unfoldase: allosteric control of a protein machine. *Cell* 121(7):1017–1027.

263. Burnley BT, Afonine PV, Adams PD, Gros P, Brunger AT (2012) Modelling dynamics in protein crystal structures by ensemble refinement. *Elife* 1:e00311.

264. Heider MR, et al. (2016) Subunit connectivity, assembly determinants and architecture of the yeast exocyst complex. *Nat Struct Mol Biol* 23(1):59–66.