A novel 5′ DNA binding site in RFC facilitates PCNA loading for gap DNA repair

Fengwei Zheng[1], Roxana E. Georgescu[2], Nina Y. Yao[2], Huilin Li[1,*], and Michael E. O'Donnell[2,3,*]

[1] Department of Structural Biology, Van Andel Institute, Grand Rapids, Michigan, USA

[2] DNA Replication Laboratory, The Rockefeller University, New York, New York, USA

[3] Howard Hughes Medical Institute, The Rockefeller University, New York, New York, USA

[*] Correspondence should be addressed to H.L (huilin.li@vai.org) or M.E.O (odonnel@rockefeller.edu)

Keywords: PCNA, RFC, sliding clamp, DNA repair, DNA replication

Abstract

RFC uses ATP to assemble PCNA onto primed sites for replicative DNA polymerases δ and ε. The RFC pentamer forms a central chamber that binds 3′ ss/ds DNA junctions to load PCNA onto DNA during replication. We show here five structures revealing a novel 2nd DNA binding site in RFC that binds a 5′ duplex. This 5′ DNA site is located in the N-terminal BRCT domain and AAA+ module of the large Rfc1 subunit. It appears 5′ DNA binds after 3′ DNA binding to RFC. Our structures reveal ideal binding to a 7-nt gap, which includes 2 bp unwound by the clamp loader. Biochemical studies show enhanced binding to 5 and 10 nt gaps, consistent with the structural results. We propose that the 5′ site facilitates RFC's PCNA loading activity at a DNA damage-induced gap to recruit gap-filling polymerases. These findings are consistent with genetic studies showing that base excision repair of gaps greater than 1 base requires PCNA and involves the 5′ DNA binding domain of Rfc1. We further observe that a 5′ end facilitates PCNA loading at an RPA coated 50-nt gap, suggesting a potential role of the RFC 5′ DNA site in longer gaps including during lagging strand DNA synthesis.

INTRODUCTION

All three domains of life possess a DNA clamp and clamp loader that are essential for cell growth and proliferation (Bell and Labib, 2016; Yao and O'Donnell, 2016b). The clamp and clamp loader function in genomic DNA replication, DNA damage repair, and other DNA metabolic processes (Burgers and Kunkel, 2017; Harrison and Haber, 2006; Ohashi and Tsurimoto, 2017; Sancar et al., 2004; Su, 2006; Waga and Stillman, 1998). The eukaryotic PCNA clamp and its clamp loader complex are key components of the replisome (Ellison and Stillman, 2001; Hedglin et al., 2013; Hingorani and O'Donnell, 2000; Indiani and O'Donnell, 2006; Yao and O'Donnell, 2012; Yao and O'Donnell, 2016a; Zhang and O'Donnell, 2016). PCNA needs to be cracked open by a clamp loader to encircle the genomic DNA because it is a topologically closed ring. The eukaryotic clamp loader is Replication Factor C (RFC), composed of Rfc1-5, belonging to a large family of proteins involved in diverse cellular activities, referred to as the AAA+ (ATPases Associated with various cellular Activities) family (**Fig. 1a)** (Cullmann et al., 1995; Davey et al., 2002; Miller and Enemark, 2016; Snider et al., 2008).

Structural studies on clamp loaders of *E. coli*, T4 phage, yeast, and human have revealed a conserved architecture of a five-subunit spiral with two tiers (Bowman et al., 2004; Gaubitz et al., 2020; Kelch et al., 2011; Simonetta et al., 2009). The upper tier is a sealed ring consisting of five collar domains - one from each subunit, whereas the lower tier is composed of five AAA+ modules and is an open spiral with a gate between the first and the fifth subunits. However, the first subunit of the eukaryotic and T4 phage clamp loaders contain an additional A' domain at their respective C termini that contacts the fifth subunit and closes the spiral gap (Kelch et al., 2012b). The five ATP binding sites are located at the interfaces between adjacent subunits, with one subunit contributing the Walker A (P-loop) and Walker B (DExx box, where the x can be any residue) motifs, and the other contributing the SRC motif that harbors an essential "arginine finger" residue (Johnson et al., 2006). All

studies thus far have revealed that primer/template (P/T) DNA with a 3′-recessed end binds inside the chamber of the ATP-bound clamp loader (**Fig. 1b**) (Bowman et al., 2005; Kelch et al., 2011). Moreover, clamp loaders of bacterial and eukaryotic systems only need ATP binding to crack open the clamp ring; neither ATP hydrolysis nor DNA is required (Kelch et al., 2012a,b; Turner et al., 1999). DNA binding appears to properly engage the ATPase sites for hydrolysis, leading to RFC dissociation from DNA and PCNA closure around the duplex (Gaubitz et al., 2021; Kelch et al., 2012a).

The replication clamps are ring-shaped complexes that encircle dsDNA and slide freely on the duplex, conferring processivity to their respective replicative DNA polymerase (Yao and O'Donnell, 2021). In eukaryotes, PCNA (Proliferating Cell Nuclear Antigen) functions with replicative polymerases Pol δ on the lagging strand and Pol ε on the leading stand (Kunkel and Burgers, 2014; Kunkel and Burgers, 2017; Nick McElhinny et al., 2008; Pursell et al., 2007; Yao and O'Donnell, 2021). Structural studies showed that, the *E. coli* β-clamp is a homodimer, while the T4 phage clamp and archaeal and eukaryotic PCNA clamps are homotrimers, all of which display 6 domains having similar chain folds and remarkable pseudo six-fold symmetric architecture with an obvious ancestor in evolution (Gulbis et al., 1996; Kong et al., 1992; Krishna et al., 1994; Matsumiya et al., 2001; Yao and O'Donnell, 2016b). The monomers are arranged in head-to-tail manner, with 12 α-helices that line the inner cavity and a continuous layer of antiparallel β-sheet forming the outside perimeter, even at the protomer interfaces. In each case, the domains within each monomer are linked by an interdomain connecting loop (IDCL). The diameter of the PCNA inner lumen is about 38 Å, and the inner surface is also basic, thus suitable to encircle a 20-Å wide dsDNA and is shown to slide on DNA freely (Bloom, 2009; Li et al., 2021). All replication clamps characterized so far mainly interact with their partners by their respective carboxy-terminal “front” faces. PCNA interacts with numerous proteins in replication, repair and cell cycle control (Kelman, 1997; Maga and Hubscher, 2003). The “front” face of each monomer possesses a hydrophobic pocket that interacts with the PIP motif (PCNA interacting peptide), or modified PIP motifs of clamp interactive proteins (Altieri and Kelman, 2018; Fernandez-Leiro et al., 2015; Gulbis et al., 1996; Lancey et al., 2020; Madru et al., 2020; Prestel et al., 2019; Zheng et al., 2020).

A recent cryo-EM study captured yeast RFC in a series of intermediate states during loading of the PCNA clamp onto DNA (Gaubitz et al., 2021), including an intermediate in which the PCNA ring is open by 20 Å, sufficiently wide for dsDNA to pass through. This conformation was not observed in the crystallographic study of the T4 phage clamp−loader−DNA complex (Kelch et al., 2011). The new structure confirms that clamp opening does not need ATP hydrolysis. However, PCNA closure around dsDNA induced conformation changes that tighten up the RFC ATP binding sites at subunit interfaces and likely stimulate the ATPase. ATP hydrolysis presumably releases RFC from PCNA, making PCNA available to bind its other partners, such as Pols δ and ε for DNA synthesis, and FEN1 and ligase for Okazaki fragment maturation (Bell and Labib, 2016).

Rad24-RFC is an alternative clamp loader in which Rad24 replaces Rfc1 of the RFC clamp loader and loads the DNA damage signaling 9-1-1 clamp onto a 5′-recessed DNA end (Ellison and Stillman, 2003; Majka et al., 2006; Majka and Burgers, 2003). We and others recently determined the cryo-EM structures of the Rad24-RFC−9-1-1 clamp complex bound to a 5′-ss/ds DNA (Castaneda et al., 2021; Zheng et al., 2021). We observe two structures in which the 9-1-1 clamp is either closed or had a 27 Å opening. These structures revealed that the 5′-junction DNA is held specifically by the Rad24 subunit in a long basic groove between the AAA+ module and collar domain, and the 3′ ssDNA overhang enters the inner chamber of the alternative loader. This DNA interaction mode is unexpected, as it is entirely different from the DNA binding pattern of RFC which binds the 3′-recessed dsDNA inside the central chamber, while the 5′ ssDNA overhang extends to the outside of the clamp loader as reported in the T4 phage and yeast structures (Kelch et al., 2011).

The surprising DNA binding site in the Rad24 subunit of Rad24-RFC led us to ask if RFC also contains a second DNA binding site above the clamp ring. If true, the second binding site may explain the previously reported 5′-junction DNA binding by the human RFC1 BRCT domain (Kobayashi et al., 2010; Kobayashi et al., 2006). In addition, dual binding of 3′ and 5′ DNA ends by RFC suggests that RFC may bind DNA gaps, congruent with genetic studies demonstrating that mutants in the N-terminal region of Rfc1 are defective in excision gap repair (Aboussekhra et al., 1995; Gomes et al., 2000; Li et al., 1994; McAlear et al., 1996; Shivji et al., 1992).

These clues to dual DNA binding sites in RFC have inspired us to examine the structure of *S. cerevisiae* RFC in the presence of a double-tailed DNA in which both the 3′-recessed and 5′-recessed ends are present on the

same molecule, as we used in our previous Rad24-RFC study (Zheng et al., 2021). Indeed, we found that RFC can bind up to two DNA molecules at the same time, one being the well-established site wherein a 3′ ss/ds primed DNA binds inside the pentamer central chamber, and a second site in which 5′ DNA binds on the Rfc1 "shoulder" outside the central chamber. This surprising dual DNA site binding mode of RFC resembles the newly discovered DNA binding by Rad24-RFC (Castaneda et al., 2021; Zheng et al., 2021). We also examine the affinity and PCNA clamp loading efficiency of RFC at different sized DNA gaps. The results reveal enhanced RFC binding at 5-50 nt gaps, and 5' end stimulation in loading of PCNA at a 10-nt gap, consistent with a role of RFC 5' end binding in gap repair. The results also show a measurable enhancement in RFC loading of PCNA at an RPA-coated 50-nt gap, suggesting that RFC may act in longer excision tracks, and possibly binds both 3′ and 5′ ends of an Okazaki fragment fill-in site, perhaps to protect the 5′ end at Okazaki fragments.

RESULTS

1. RFC can engage a 3′ DNA site in the central chamber simultaneously with a 5′ DNA site located outside on the Rfc1 "shoulder"

The *S. cerevisiae* PCNA clamp and RFC clamp loader were separately expressed and purified from *E. coli* (see **Materials and Methods**). We assembled the loading complexes in vitro by directly mixing the purified proteins with a double-tailed primer/template DNA with both 3′-recessed and 5′-recessed ends (**Fig. 1a**) in the presence of 0.5 mM slowly hydrolysable ATP analog ATPγS (see details in **Materials and Methods**). Use of DNA having both 3′ and 5′ ssDNA ends strategically enables the unbiased binding of either end to sites on RFC. 2D classification of cryo-EM images demonstrated the assembly of a variety of loading intermediates, including RFC bound to either one or two DNA molecules, and with the loader either engaged or not with PCNA (**Supplementary Fig. 1a-b**). Subsequent 3D classification and 3D variability analysis (3DVA) led to the capture of five 3D maps: 1) RFC bound to a DNA in the central chamber (termed 3′-DNA1 herein) at 3.3-Å resolution, 2) RFC bound to two DNA molecules at 3.3-Å resolution, with 3′-DNA1 in the central chamber plus a 5′ ss/ds DNA bound to the Rfc1 shoulder (referred to here as 5′-DNA2). The 5′-DNA2 is nearly parallel with and above the position that is occupied by the PCNA ring, 3) RFC-closed planar PCNA with 3′-DNA1 bound in the central chamber and threaded through the PCNA ring at 3.3-Å resolution, 4) RFC-open PCNA with 3′-DNA1 bound in the central chamber and threaded through an open PCNA ring that assumes a right-hand spiral with a 14-Å opening in the clamp at 3.4-Å resolution, and 5) RFC-closed PCNA with 3′-DNA1 in the central chamber plus a molecule of 5′-DNA2 on the Rfc1 shoulder at 3.1-Å resolution (**Fig. 1c, Supplementary Figs. 2-3**). The fifth map is a composite map produced by combining the 3.1-Å map of RFC−3′-DNA1−5′-DNA2−PCNA with a locally refined map focusing around Rfc1 NTD (including the BRCT domain and AAA+ module) and 5′-DNA2.

Interestingly, the RFC structure is highly similar among the five structures captured here, with main chain Cα RMSDs (root-mean-square deviation) ranging from 0.6 Å to 1.0 Å (**Supplementary Fig. 4a-d, Supplementary Video 1**). In all five structures, the Rfc1 AAA+ module sits above the PCNA protomer at the right side of the DNA entry gate (viewing down from the C-terminal collar), while the Rfc5 AAA+ module is above the PCNA protomer at the left side of the DNA entry gate (**Supplementary Fig. 4a-d**). These structural features are consistent with previous studies (Bowman et al., 2004; Gaubitz et al., 2020; Kelch et al., 2011). The closed-ring PCNA structure of RFC−3′-DNA1−PCNA is comparable with the RFC−PCNA structures determined in the absence of DNA (Bowman et al., 2004; Gaubitz et al., 2020) (**Supplementary Fig. 4a-b**). However, DNA makes the structure significantly more compact, as PCNA tilts up by 10° toward RFC, forming a tighter interface (**Supplementary Fig. 4b**). In the open PCNA structure, Rfc1 and Rfc5 each bind one different PCNA protomer, holding the two protomers 14-Å apart in-plane and 13-Å apart out-of-plane (**Fig. 1c**, middle panel). Because the 14-Å gap in PCNA is narrower than the 20-Å width of a dsDNA, and 3′-DNA1 is already inside the RFC central chamber, this is likely a post-entry intermediate in which the DNA gate has partially closed. These open and closed PCNA states resemble the open and closed clamp states observed in the T4 phage clamp-clamp loader complex (Kelch et al., 2011) and the recent RFC−PCNA−DNA structure (Gaubitz et al., 2021) (**Supplementary Fig. 4c**).

This is the first time that RFC is found to possess two DNA binding sites. Interestingly, we found that RFC interacts with DNA either in the absence of PCNA (RFC−3′-DNA1 and RFC−3′-DNA1+5′-DNA2), or in the presence of PCNA (RFC−3′-DNA1−open spiral PCNA, and RFC−3′ DNA1−closed ring PCNA, and RFC−3′-DNA1+5′-DNA2−PCNA) (**Figs. 1c, 2a-d)**. However, we did not observe RFC−PCNA in our reaction mixture that lacked DNA, and only observed RFC-PCNA when DNA was present. This observation suggests that

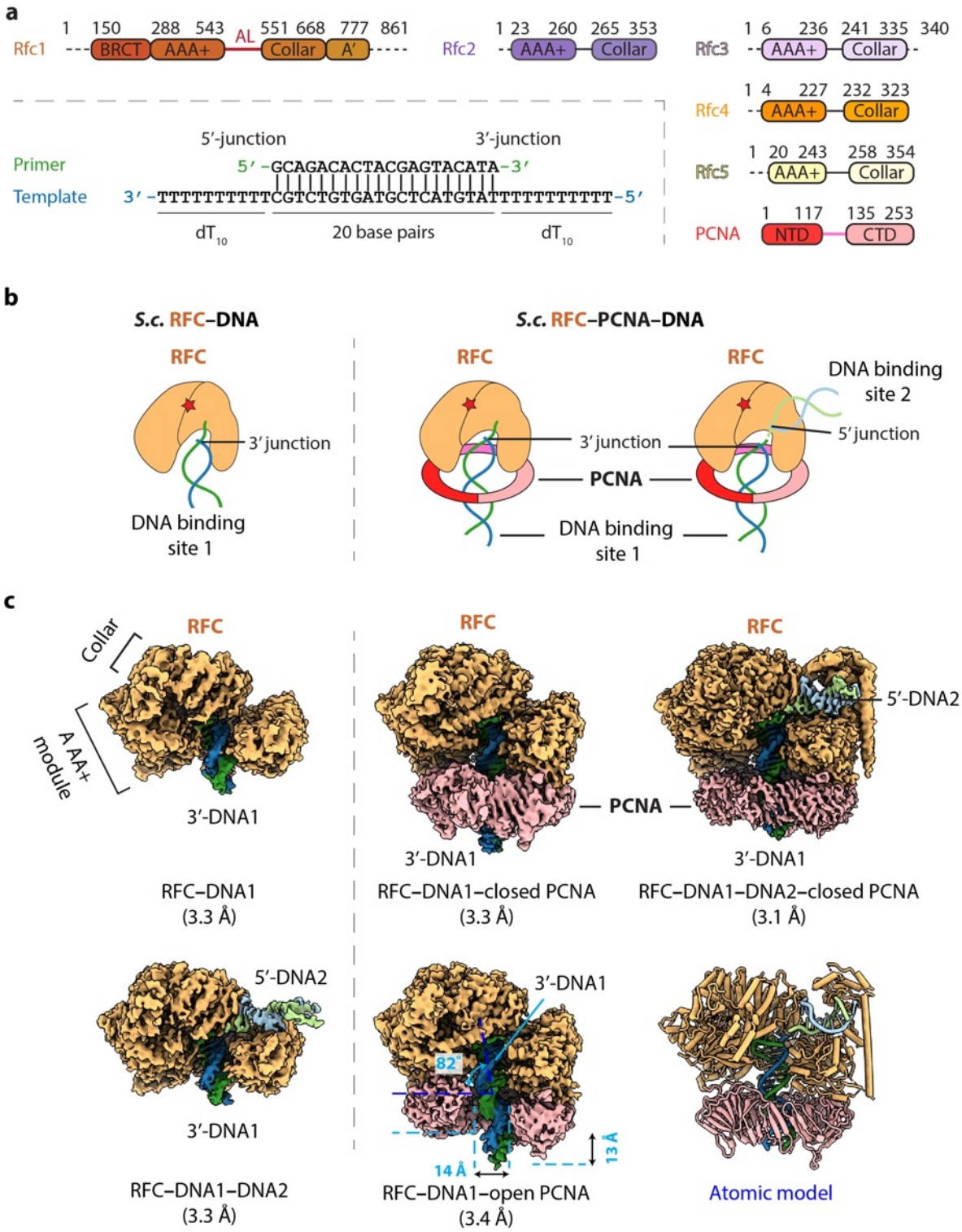


Fig. 1. Cryo-EM captured multiple S.c. RFC–(PCNA)–DNA complexes. a) Domain architecture of RFC, PCNA, and the template/primer DNA. Dashed lines indicate unsolved regions. AL is the alternative linker between Rfc1 AAA+ and collar domain. Pink line between PCNA NTD and CTD represents the inter-domain connecting loop (IDCL). DNA used in this study harbors both 5′-DNA and 3′-DNA junctions. **b**) Sketch illustrating the previously established DNA1 binding site recognizing the 3′-DNA junction in the RFC central chamber (left) and the novel 5′-DNA2 site recognizing the 5′-DNA junction on Rfc1 shoulder. **c**) 3D maps of the five loading intermediates: RFC–3′-DNA1 and RFC–3′-DNA1–5′-DNA2 (left), RFC–3′-DNA1–PCNA in closed and open spiral form (middle), and RFC–3′-DNA1–5′-DNA2–PCNA (right). These maps are displayed at the same threshold (0.2). RFC is in orange, PCNA pink. DNA strands are individually colored.

RFC can interact with DNA independent of PCNA, which we presume reflects the stability of this structure using ATPγS, compared to freely hydrolysable ATP as determined by earlier studies (Gomes and Burgers, 2001).

2. 3′-DNA1 binding induced large conformational changes in RFC thus forming the 5′-DNA2 binding site. RFC bound to 3′-DNA1 is in the shape of a two-tiered right-handed spiral with a closed collar ring on the top and an open spiral AAA+ tier on the bottom (**Fig. 2a-b**). The five RFC subunits are arranged counterclockwise from Rfc1, Rfc4, Rfc3, Rfc2, to Rfc5 when viewed from the top (i.e., from the C-terminal collar domain). Among the five nucleotide binding sites, four were occupied by ATPγS and one by ADP (between Rfc5 and Rfc1), illustrated in **Fig. 2a,** viewed from the C-terminal domains and upon removal of the C-terminal collar domains. ADP is possibly an ATPγS hydrolysis product by Rfc5. The Rfc5 subunit contains a "β-hairpin plug" element that protrudes into the central chamber. In the absence of 3′-DNA1, the Rfc5 plug is largely disordered, and there is a gap at the ATP binding interface between Rfc4 and Rfc3 in the RFC−PCNA structures in the presence of ATPγS and Mg^{2+} (Bowman et al., 2004; Gaubitz et al., 2020) (**Fig. 2b**). In the 3′-DNA1- or 3′-DNA1−5′-DNA2-bound structures, the Rfc5 β-hairpin plug becomes ordered and inserts into the major groove of 3′-DNA1 located in the central chamber (**Fig. 2a, Supplementary Fig. 5a-d**). Furthermore, the Rossman fold domains of Rfc3, 2 and 5 undergo a large rotation of 63° to open a gate, referred to as the A-gate (Gaubitz et al., 2021), between the AAA+ module and the A' domain of Rfc1 and to tighten up the gapped ATP-binding interface between Rfc4 and Rfc3, perhaps stimulating the ATPase activity at this site (**Fig. 2b**). However, the Rfc1 A' domain becomes largely flexible upon 3′-DNA1 binding, and only the bridge helix – the α-helix connecting the collar and A' domains (aa 669-689) – remained stable (**Fig. 2a-b**).

The 3′-DNA1 binding site in the chamber is largely pre-formed in RFC, but the 5′-DNA2 site, located between the collar domain and the AAA+ module of Rfc1, is not formed in the absence of 3′-DNA1 binding. Specifically, the Rfc1 collar region is very close to its AAA+ module with only a single residue (Leu-549) linking these two domains (**Fig. 2c**). Our structures reveal that, concomitant to the large-scale rotation of the Rfc3-2-5 half ring upon binding 3′-DNA1, as described above, 3′-DNA1 binding also causes conformation changes in the last α-helix (Asn-535 to Ala-548) of the Rfc1 AAA+ module and the first helix of the collar domain. Specifically, in Rfc1, the second half of the last AAA+ helix (Trp-543 to Ala-548) and the first residue of the collar helix (Lys-550) unwind and they, together with Leu-549, become an extended loop. This loop allowed the counter-clockwise movement of the Rfc1 collar and A' domains, leading to the formation of the 5′-DNA2-binding cleft (**Fig. 2c, Supplementary Fig. 6a-b, Supplementary Video 2**). We have termed the loop that forms upon binding 3′-DNA1 as the "alternative linker" (AL), because it is reminiscent of the Rad24 long linker between the AAA+ module and collar domain which allows the DNA binding at the Rad24 shoulder (**Supplementary Fig. 4d)**.

Among the eight 3D subclasses of RFC−DNA derived from 3D variability analysis (3DVA), two subclasses had either 3′-DNA1 alone or had both 3′-DNA1 and 5′-DNA2 sites occupied (**Supplementary Fig. 2**); each type accounted for about 31% of the ~ 1 million particle population. The remaining four subclasses had 3′-DNA1 and partially occupied 5′-DNA2 sites, and they account for 38% of the population. However, we did not observe an RFC−5′-DNA2 structure in which the 5′-DNA2 site is occupied, and the 3′-DNA1 site is empty. Based on our structural analysis, we suggest that RFC binds 3′-DNA1 inside the central chamber first, and that 3′-DNA1 binding induces the above-described conformation changes that configure the 5′-DNA2 binding site.

3. Rfc1 BRCT domain assists 5′-DNA2 binding

In the RFC−3′-DNA1−5′-DNA2−PCNA structure, 5′-DNA2 binding stabilized the otherwise flexible Rfc1 BRCA1 C-terminal homology (BRCT) domain (**Fig. 3a-b**). The yeast Rfc1 BRCT domain resembles a clenched hand on an arm and is composed of a core subdomain (hand) followed by a vertical long α-helix (arm) (**Fig 3a**). The core adopts the canonical BRCT domain folding and is similar to the NMR structure of the human RFC1 BRCT domain (Kobayashi et al., 2010) (**Fig. 3b**). The BRCT core contains a 4-stranded (β1-4) β-sheet flanked by three α-helices, with α1 and α3 at one side and α2 at the other side, and an additional α1'-helix at the N-terminus. The loop connecting α1' and β1 in our yeast Rfc1 BRCT structure (Rfc1 aa 113-161; 49 aa) is twice the length of the corresponding loop in the human RFC1 BRCT domain.

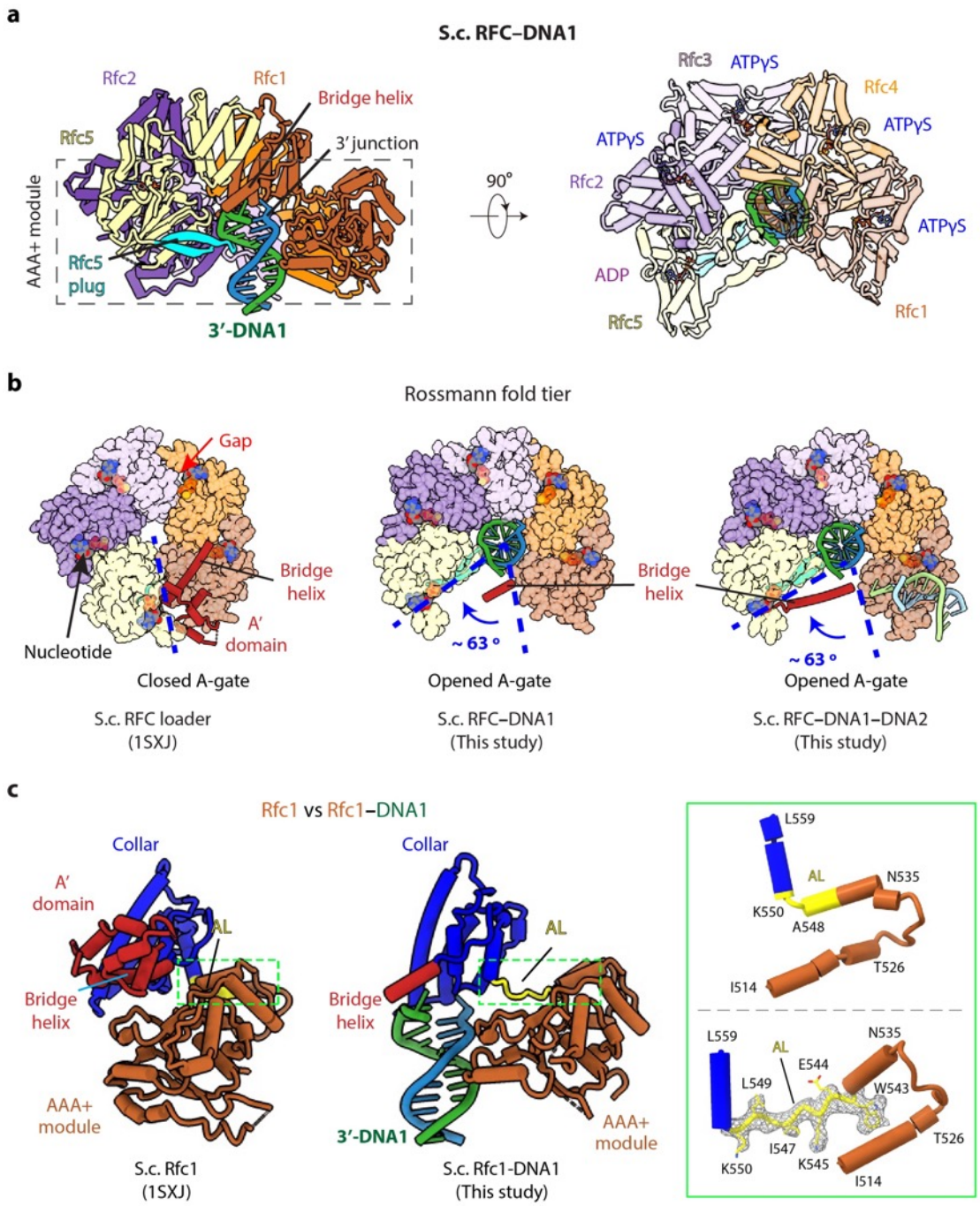


Fig. 2. 3′-DNA1 binding induces conformation changes and formation of the 5′-DNA2 site. a) Structure of RFC–3′-DNA1 complex colored by individual subunits and DNA strands (template in green and primer in steel blue) in a side (left) and a top view (right). The Rfc5 plug inserted into the DNA major groove is cyan. The four ATPγS and one ADP (only at the Rfc5:1 interface) are in sticks. **b**) Comparison of the five AAA+ modules of RFC in the RFC–PCNA structure in the absence of DNA (left) or bound to 3′-DNA1 (middle) or bound to 3′-DNA1 and 5′-DNA2 (right). **c**) Rfc1 in the absence of DNA (PDB entry 1SXJ) is superimposed with the AAA+ module with Rfc1 in the presence of 3′-DNA1 (this study), revealing large scale leftward movements of the collar and A' domains (along with the 3′-DNA1). Region in the dashed green box is enlarged in the right panels showing the conformational changes of the alternative linker (AL), the EM density of the changed AL is shown in meshes.

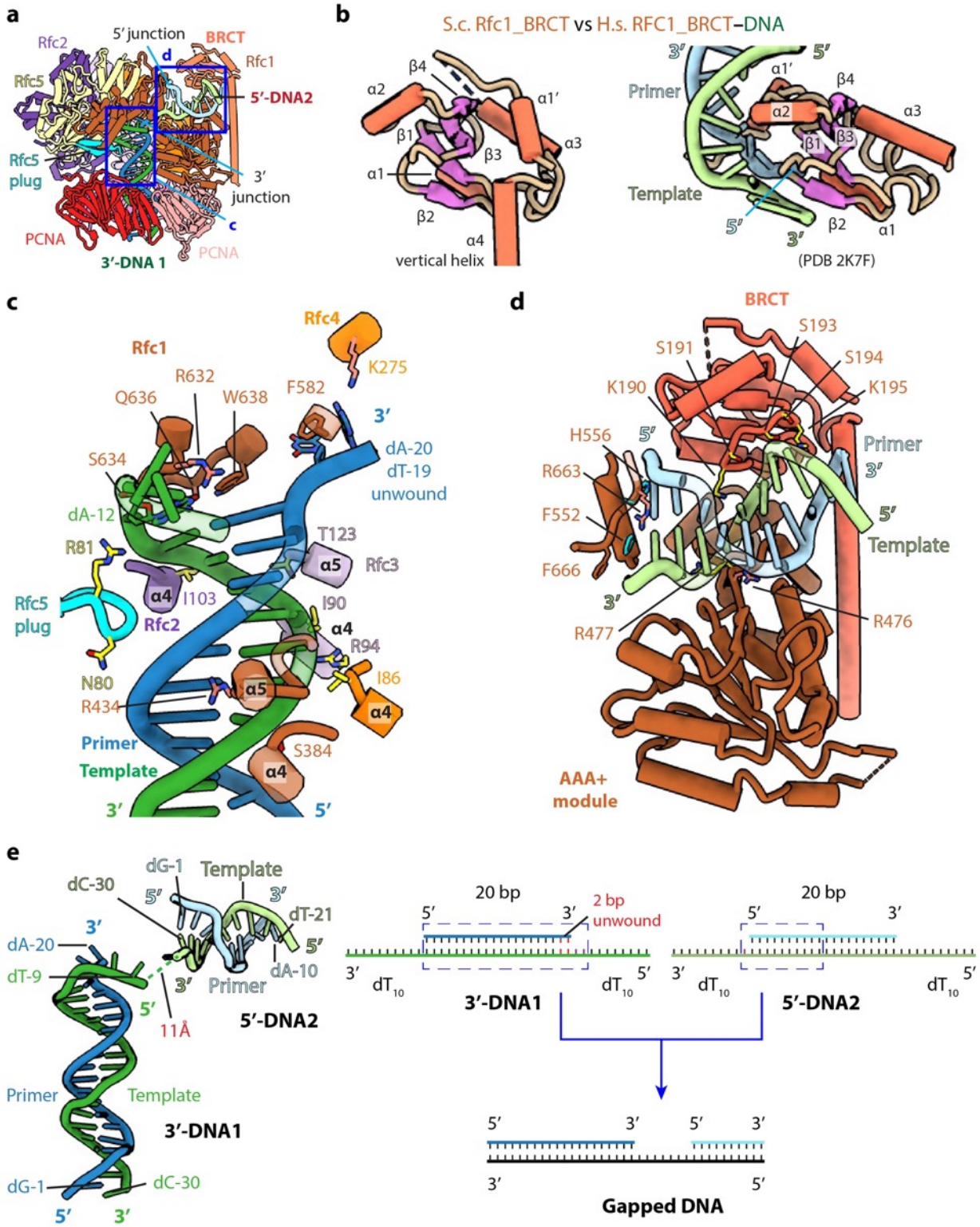


Fig. 3. The BRCT domain of Rfc1 stabilizes 5′-DNA2 in the quaternary RFC–3′-DNA1–5′-DNA2–PCNA complex. a) The atomic model of the quaternary RFC–3′-DNA1–5′-DNA2–PCNA complex. The Rfc1 BRCT domain caps 5′-DNA2. Regions in the two blue boxes are enlarged in (c) and (d). **b**) Structure of the yeast Rfc1 BRCT domain in the context of the full RFC complex bound to DNA (DNA not shown) compared with the human RFC1 BRCT domain (aa 375-480; PDB entry 2K7F). The human RFC1 BRCT was computationally modeled to recognize the 5′-junction of dsDNA by the end region of β1-2. **c)** Enlarged view showing interactions between RFC and 3′-DNA1. Helices α4 and α5 of Rossmann fold of Rfc1, 4, 3 and 2 wrap around the template strand. DNA is shown as stubs with the separated base pair (dA-12:dT-19) and base dA-20 in primer strand shown in sticks. Residues Arg-632 and Gln-636 (from Rfc1), Lys-275 (from Rfc4) contacting the bases of DNA are in

salmon sticks. Residues H-bonding with DNAs are in yellow. The main chain nitrogen atoms of Ile-86, Ile-90, and Ile-103 in the α4 helices of Rfc4, 3 and 2 form H-bonds with template DNA. **d)** Enlarged view of the 5′-DNA2 binding region. The electrostatic surface shows that the positively charged Rfc1 BRCT domain on top and positively charged AAA+ module from the bottom stabilize the 5′-DNA2. Key residues surrounding 5′-DNA2 are labelled. The α-helix of Rfc1 collar domain harboring Phe-552 and His-556 (cyan sticks) blocks the 5′-DNA2 5′-junction. The "separation pin" residue Phe-666 located at the DNA2 5′-junction, Arg-476 and Arg-663 (yellow stick) insert into the DNA minor groove, while Arg-477 point to the DNA major groove. For clarity, only partial protein secondary structures are shown. Unless noted otherwise, the same color and rendering scheme is used in all figures. **e**) The arrangement of 3′-DNA1 and 5′-DNA2 in RFC (left) resembles a gapped DNA (right).

4. RFC interactions with 3′-DNA1 and 5′-DNA2

The RFC interaction with 3′-DNA1 is similar in all five captured structures. 3′-DNA1 is recognized by a series of α-helices in the RFC−3′-DNA1−5′-DNA2 structure (**Fig. 3c**). Helices α4 and α5 of Rossmann fold domains of Rfc1-4 bind and spiral around the template 3′-DNA1 strand, similar to the T4 clamp-clamp loader system (Kelch et al., 2011). In particular, the α4 Ser-384 of Rfc1; α4 Ile-86 of Rfc4; α4 Ile-90, Arg-94 and α5 Thr-123 of Rfc3 H-bond with the phosphate backbone of the template strand; the α4 Ile-103 of Rfc2 and Arg-81 in the β-hairpin of Rfc5 (Rfc5 plug) also H-bond with the phosphate backbone of the template strand. Rfc4 Lys-275 and the Rfc5 plug Asn-80 H-bond with the primer strand. The α5 Arg-434 of Rfc1 inserted into the minor groove of DNA1, the separation pin, Rfc1 Trp-638, is equivalent to the separation pin observed in the bacterial β-clamp loader (Simonetta et al., 2009), disrupting dT-19 and dA-20 at the 3′-end of the primer from the template, and the two separated bases are stabilized by the Rfc1 Phe-582 and Rfc4 Lys-275, respectively. In addition, Rfc1 Arg-632 and Gln-636 also H-bonds with the unwound template dA-12 (**Fig. 3c**). Single base unwinding by RFC at the 3′-junction DNA has been reported previously (Gaubitz et al., 2021).

In overview, RFC mainly contacts the template strand of 3′-DNA1, but there are a few contacts to the primer strand: specifically, the Rfc5 plug touches both strands, and Rfc1 Phe-582 and Rfc4 Lys-275 also contact the primer strand. Regardless, the main contact of RFC to the primer/template DNA is with the template strand, as it is in the *E. coli* and T4 phage clamp loader-DNA complexes (Kelch et al., 2011; Simonetta et al., 2009).

Only the first 10 bp of the duplex portion of the 5′-DNA2 containing the 5′-recessed end is stably bound to RFC and modelled. 5′-DNA2 is sandwiched between the BRCT domain on the top and the Rfc1 AAA+ module on the bottom (**Fig. 3a, d**). Therefore, 5′-DNA2 is primarily bound by the first clamp loader subunit, Rfc1, in a manner similar to Rad24 DNA binding in Rad24-RFC (**Supplementary Fig. 6c-d**) (Castaneda et al., 2021; Zheng et al., 2021). However, the specific contacts are distinct. Four basic residues on the top of the Rfc1 AAA+ module interact with the double strand region of 5′-DNA2: Lys-314 and Asn-459 H-bond with the phosphate backbone of the primer and template respectively, and Arg-476 and Arg-477 sandwich the template strand, with Arg-476 inserted in the minor groove. The 5′-recessed end of the 5′-DNA2 is blocked by an α-helix of the collar domain and is stabilized by Phe-552 and His-556. Phe-666 is precisely located at the 5′-junction, perhaps functioning as the "separation pin" for 5′-DNA2. Arg-663 in the minor groove appears to stabilize the 5′-junction. The inner surface of this 5′-DNA2-binding groove is basic, lined by a series of alkaline residues including Lys-190 and Lys-195 of the BRCT domain and Arg-476 and Arg-477 of the AAA+ module. The 5′-DNA2 interaction with Rfc1 BRCT in the full RFC complex as observed in our structure is very different from the previous prediction based on the NMR structure of the truncated human RFC1 BRCT domain (**Fig. 3b, d**). It is currently unclear if the predicted binding mode is an in vitro artifact of using an isolated domain or might be an intermediate preceding the stable binding observed in our structure.

5. 3′-DNA1 and 5′-DNA2 arrangement in RFC mimics a single strand DNA gap

Even though both 3′-DNA1 and 5′-DNA2 are identical, and both have 3′ and 5′ junctions, our structural analysis reveals that one DNA (3′-DNA1) engages the RFC central chamber with the 3′-recessed end, and the other DNA (5′-DNA2) is bound by Rfc1 such that the 5′ end is pointed towards the central chamber of the RFC (**Figs. 1b-c, 3c-d**). Surprisingly, we found that the two identical DNA molecules, when bound to RFC, resemble a single gapped DNA (**Fig. 3e**). As we described above, RFC stabilizes 2 unpaired and unwound 2 additional bases in the 3′-DNA1, and Rfc1 stabilized 1 unpaired base in the 5′-DNA2. Furthermore, the 11 Å space between the 5′

end of the 3′-DNA1 template strand and the 3′-end of the 5′-DNA2 template strand can accommodate 2 additional bases. Therefore, DNA1 and DNA2 are held by RFC like a gapped DNA with a gap size of about 7 bases. We suggest that the use of two DNAs has provided an unbiased view that RFC will most easily bind a gapped DNA with a minimal gap size of 5 bases, which will need an additional two 3′ base pairs unwound to form an ideal 7-base gapped DNA substrate. Longer gaps could likely be easily accommodated as the ssDNA region could simply loop outside of RFC.

6. Biochemical analysis of RFC−PCNA binding different sized gaps

To examine the gapped DNA binding of RFC−PCNA, we first measured binding affinities of RFC alone to gel-purified DNAs having specific gap sizes using ATPγS. One may anticipate that the smallest gap that gives the tightest affinity may reflect the optimal gap size. Gaps that are smaller than the optimal gap size would not bind as tight because they would expend some binding energy for their observed affinity. For binding experiments,

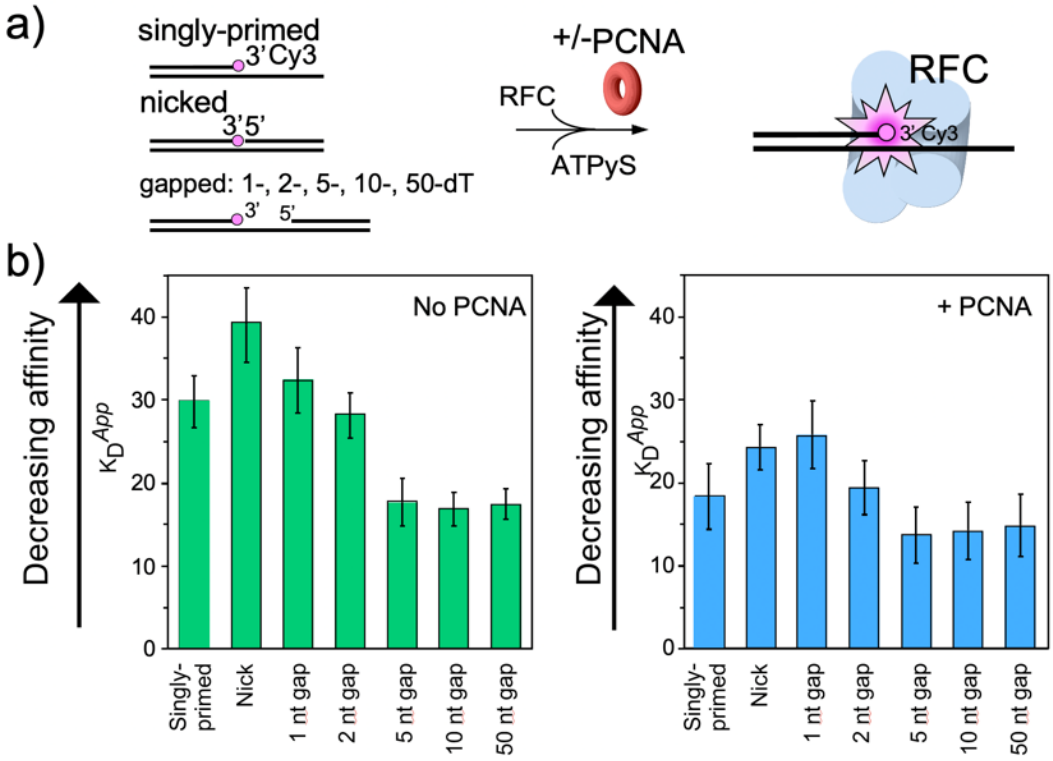


Fig. 4. RFC/PCNA affinity for DNAs with different gap sizes. a) Scheme for the affinity assay of RFC +/- PCNA for gapped DNAs. The top strand is composed of two constant length DNA oligonucleotides, that are hybridized to a variable length bottom strand to generate different sized gaps. The left most oligonucleotide had a 3′ fluorophore. **b)** Histograms for K_d measurements of RFC binding assays to each DNA template +/- PCNA. In the presence of ATPγS. The histograms in the absence of PCNA are shown to the left, and in the presence of PCNA are shown to the right. The raw data of the titrations are shown in **Supplementary Figures 7 and 8**. The data are in triplicate, and the error bars show the standard deviation. The arrows indicates that the larger the K_d value, the weaker the binding.

we used a 3′ oligo having a fluorophore, as in our earlier studies (Johnson et al., 2006). The fluorescence intensity change during titration of RFC into the DNA was measured to determine the K_d of RFC binding to primed DNA, as well as DNAs containing either a nick or gaps of 1, 2, 5, 10 or 50 nt. We found that RFC binds all of the DNAs but exhibits an approximately 2-fold tighter affinity (i.e., lower K_d) for DNAs having a gap of 5 bases or greater (**Fig. 4a; Supplementary Fig S7).** The cryo-EM structure predicted an optimal gap of 7 bases, but when one takes into account that RFC melts 2 bp at the 3' end, the optimal binding to a 5-nucleotide gap is consistent with the cryo-EM data. RFC binds gaps larger than 5 nucleotides with comparable affinity, likely due to simply looping out the extra single-strand DNA. The same binding experiments in the presence of PCNA showed a similar outcome, except the affinities were all tighter (lower K_d), likely due to the assistance of PCNA that binds RFC and holds it to DNA (**Fig. 4b, Supplementary Fig. S8**).

The gapped DNAs used here contained a 5′-OH. The RFC−PCNA−DNA structure reveals a region (formed by Rfc1 His-556, His-659 and Arg-663, **Fig.3d**) in the 5′-DNA2 site having positively charged side chains that might bind the 5′-P of a 5′ ss/ds duplex DNA. Therefore, we compared the binding affinity of RFC to a 10-nt gap in which the 5′ oligo used to form that gap has either a 5′-OH or 5′-P (**Supplementary Fig. S9**). The K_d values were not measurably different for a 5′-P versus a 5′-OH. On hindsight this result seems reasonable considering that the 5′-DNA2 site on Rfc1 binds 10 bp of dsDNA with numerous contacts to Rfc1 (described above), that would overshadow any affinity derived from a single 5′-P.

To gain additional evidence that the difference in K_d values for gaps versus nicked or singly primed DNA was due to RFC binding a 5′ ss/dsDNA duplex, we compared the affinity of full-length RFC with $RFC^{\Delta N282}$, a truncated RFC (tRFC) that contains an Rfc1 lacking the amino terminal 282 amino acids including the BRCT domain. As expected, the $RFC^{\Delta 282}$ bound a 10-nt gapped DNA with similar affinity to a singly primed DNA compared to the approximately 2-fold enhanced affinity of wt RFC containing the Rfc1 5′ DNA site (**Supplementary Figure 10**).

7. PCNA loading analysis at a 10-nt gap and at an RPA-coated 50-nt gap

We next compared PCNA loading at a singly primed site lacking a downstream 5′ duplex compared to the presence of a downstream 5′ duplex that provides a 10-nt gap. To monitor PCNA loading, we labeled PCNA with ^{32}P via an N-terminal 6 residue kinase motif as we have described in detail earlier (Kelman et al., 1995). Since PCNA slides off linear DNA after it is loaded, we blocked PCNA sliding off DNA by attaching the biotinylated end to magnetic streptavidin beads, and the other end of the template strand contained a DIG moiety, to which we added an antibody to DIG (see scheme in **Fig. 5a**). We then added RFC, ^{32}P-PCNA and ATP and enabled loading for 20 s before quenching with EDTA. Magnetic beads were then washed and the retained ^{32}P-PCNA was removed with SDS and analyzed by liquid scintillation. The results, in **Fig. 5b**, show that the downstream 5′ duplex facilitates loading of PCNA by about 5 folds. This result is consistent with the 5′-DNA2 site within RFC facilitating PCNA loading at a short ssDNA gap.

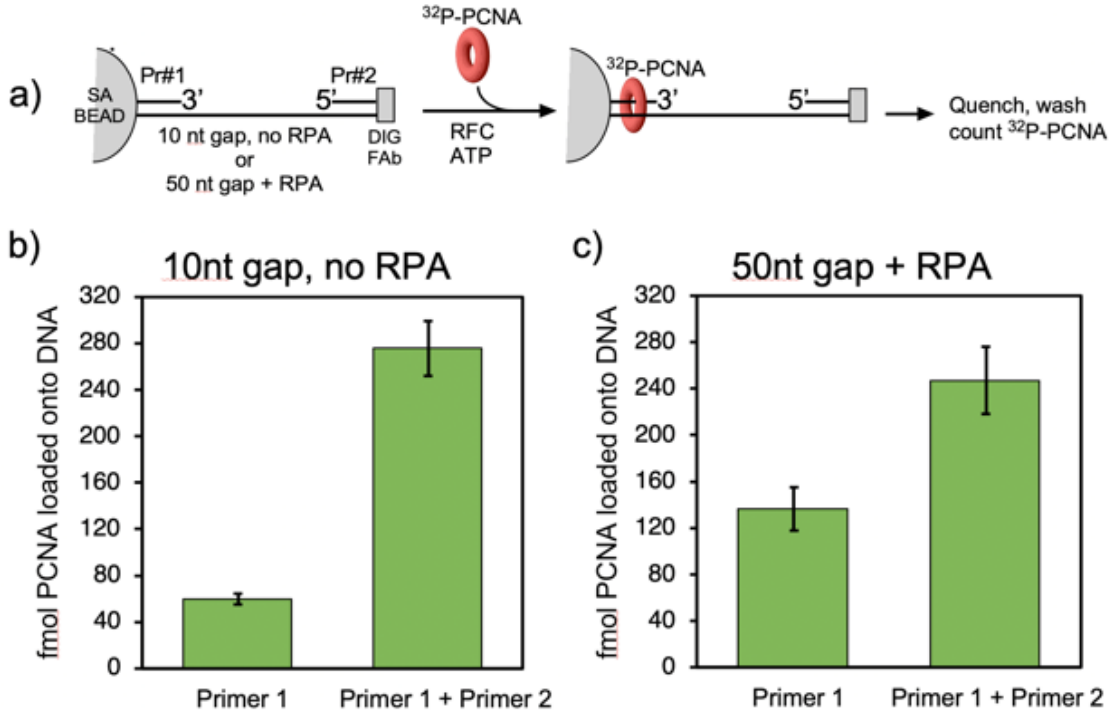


Fig. 5. Efficiency of RFC loading PCNA onto a gapped DNA versus a singly primed DNA. a) Scheme of the PCNA loading assay. The DNA template strand contains a 3′ biotin and a 5′ Dig moiety. The DNA's were bound to streptavidin magnetic beads, then incubated with antibody to DIG to prevent PCNA sliding off DNA. Then RFC and ^{32}P-PCNA were added with ATP for 20s, quenched with EDTA, and beads were washed two times and ^{32}P-PCNA that had been loaded onto DNA was counted by liquid scintillation counting. **b**) Results of ^{32}P-PCNA loading by RFC onto the 10 nt gap DNA substrate compared to singly primed DNA lacking primer #2. **c)** As in panel b except using a 50 nt gapped DNA and in the presence of RPA. All experiments were performed in triplicate and the error bars represent the SEM.

We also examined a 50 nt gap DNA (**Fig. 5c**). For these experiments we added RPA considering that RPA can bind 30 nt of ssDNA and thus would bind gaps of this size in the cell (Chen and Wold et al., 2014). The results show a measurable stimulation of PCNA loading achieved by the 5' DNA site at an RPA-coated 50-nt gap, but it is less than at a 10-nt gap. This may suggest a greater role of 5' binding by RFC in short gap repair (see Discussion).

DISCUSSION

1. Function of the 5′-DNA2 site of RFC. The fact that 3′-DNA1 and 5′-DNA2 are arranged in RFC like a single gapped DNA, as revealed by our structures, indicates a DNA gap repair function of the 5′-DNA2 binding site. We suggest that the 5′-DNA2 binding site functions in concert with the 3′-DNA1 site in nucleotide excision repair. DNA damage caused by base methylation or bulky UV-induced pyrimidine dimers and chemical carcinogen-induced adducts (Cadet et al., 2005; Nohmi et al., 2005) trigger the base and nucleotide excision repair pathways to remove lesion-containing bases, leaving a gap in DNA (Boiteux and Jinks-Robertson, 2013). Pol δ, Pol ε, Pol β, or possibly translesion (TLS) Pols are recruited by PCNA to fill-in the gap (Overmeer et al., 2010) (illustrated in **Fig. 6a)**. Furthermore, early genetic studies demonstrated that mutants in the N-terminal region of Rfc1 are defective in excision gap repair (Aboussekhra et al., 1995; Gomes et al., 2000; Li et al., 1994; McAlear et al., 1996; Shivji et al., 1992). These genetic results are supported by our demonstration that RFC binds and loads PCNA with higher affinity and rate at gapped DNA than a single 3′-recessed primed template DNA. Our proposal is consistent with several previous reports indicating that the N-terminal region of Rfc1 is important for DNA damage gap repair, as further explained below.

Possible coordinated 3′ and 5′ sites in RFC repair activity. Previous genetic and structural studies have provided evidence that the N-terminal region of Rfc1 (i.e. the N-terminal DNA binding BRCT domain/ligase homology region) binds 5′ DNA and is demonstrated by genetic experiments to be important to DNA repair (Aboussekhra et al., 1995; Allen et al., 1998; Gomes et al., 2000; Li et al., 1994; Shivji et al., 1992; Uchiumi et al., 1999), for which PCNA is required (Aboussekhra et al., 1995; Li et al., 1994; Shivji et al., 1992). The predominant evidence for involvement of the N-terminal Rfc1 in repair are studies of methylated nucleotide bases using methylmethane sufonate (MMS), that mostly methylates G, forming 7-methyl (7MeG), but also methylates A, forming 3-methyladenine (3MeA) (McAlear et al., 1996). These repair defects are well documented in a double mutant of Rfc1 in yeast at amino acids G185E and P234L (McAlear et al., 1996). In support of these findings, a Rfc1 mutant containing an insertion in the 5′ DNA binding region of Rfc1 also showed increased sensitivity to DNA damaging reagents that require excision repair (Xie et al., 1999). It is demonstrated that alkylated bases are repaired by base excision repair (Xiao and Samson, 1993). PCNA (and therefore also RFC) is required for base excision repair extending over 1 bp (Aboussekhra et al., 1995; Li et al., 1994; Shivji et al., 1992). It is proposed that the involvement of PCNA is not for processivity, given the short gap size of this type of repair, but is instead utilized to attract DNA polymerases (Aboussekhra et al., 1995; Li et al., 1994; Shivji et al., 1992). While DNA polymerases δ, and maybe ε appear to be involved in gap repair (Blank et al., 1994; Wang et al., 1993), the process is particularly mutagenic suggesting participation of DNA polymerase(s) with lower fidelity such as Pol β (yPol IV) or error-prone translesion DNA Pols (e.g. yeast Pol ζ-Rev1 and Pol η) which can function with PCNA (McAlear et al., 1996). Another report on the physiological effect of deleting the N-terminal region of Rfc1 that binds 5′ dsDNA showed 2-10 folds reduction in cell growth upon MMS induced damage (Gomes et al., 2000), but these results are smaller compared to earlier studies showing greater differences in Rfc1 mutants (McAlear et al., 1996; Xie et al., 1999). Perhaps the difference may lay in the use of single copy Rfc1 gene mutations in the chromosome in the earlier studies versus mutant RFC genes on a multiple copy centromeric plasmid. Nevertheless, both studies show that the 5' DNA binding of RFC has an effect on base excision repair of short gaps.

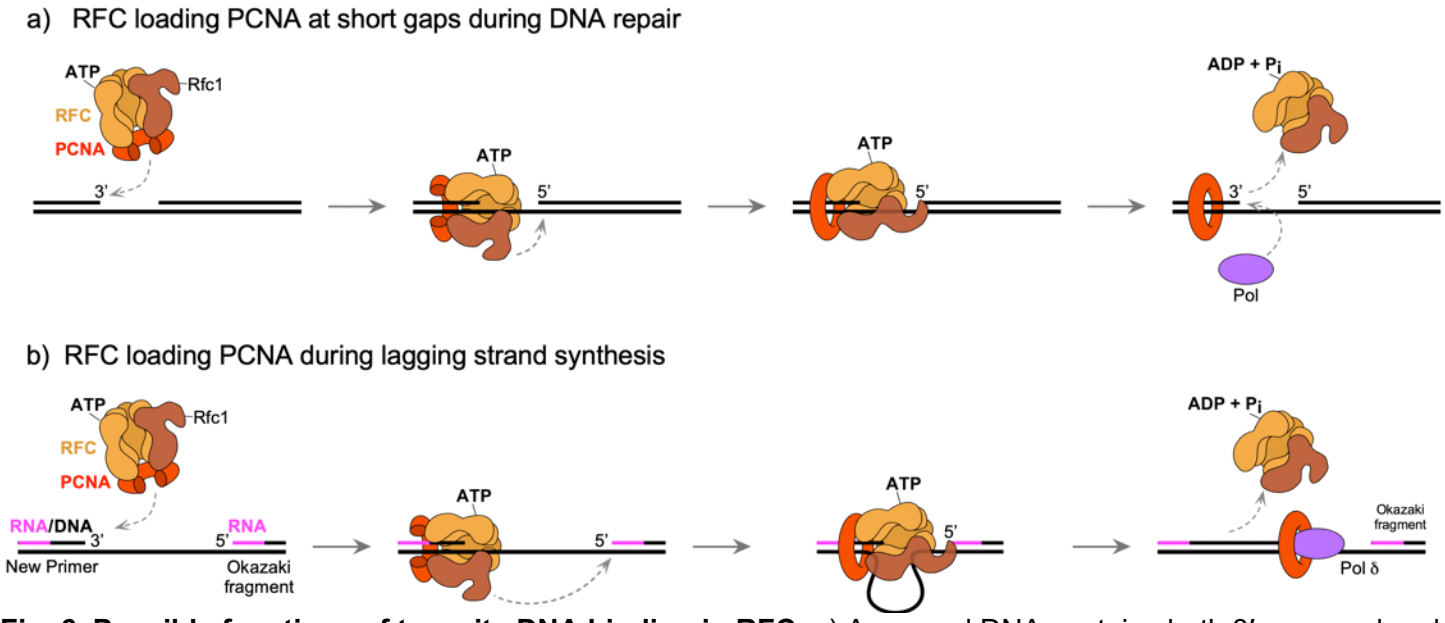


Fig. 6. Possible functions of two-site DNA binding in RFC. a) A gapped DNA contains both 3′-recessed and 5′-recessed DNA ends. RFC, illustrated here as a cut away view, has higher affinity for a gapped DNA. The RFC−PCNA complex likely first binds the 3′-recessed DNA end, which induces a conformation change in Rfc1, leading to RFC binding to the 5′-recessed DNA end. Next, ATP hydrolysis by RFC leads to RFC dissociation from PCNA. Finally, PCNA on the 3′-recessed end attracts a DNA Pol for gap filling. **b**) The lagging strand DNA at a replication fork contains a new primer (i.e., 3′-recessed end) and a 5′-recessed DNA end of the previous Okazaki fragment that is separated by ~150-200-nt from the new primer. Given that RFC appears to bind both 3′ and 5′ ends of two sites separated by 50 nt in our biochemical studies, we propose that RFC might bind both 3′ and 5′ ends of one Okazaki fragment, by looping out ssDNA. The two-site DNA binding by RFC has two potential advantages: one is to keep RFC on DNA without dissociation to facilitate repeated PCNA loading; the other is to protect the 5′-recessed end of the adjacent Okazaki fragment from being utilized by the alternative loader Rad24-RFC, which could unnecessarily load the 911 clamp and trigger the DNA damage response and arrest the replication fork.

2) Is the 5′-DNA2 site a general feature of all DNA clamp loaders? Given that both RFC and Rad24-RFC have a shoulder 5′ DNA binding site located in subunit A, the subunit in the same position as the 5′ DNA binding Rfc1 and Rad24 subunits in their respective clamp loader complexes, it seems possible that other alternative clamp loaders Ctf18-RFC and Elg1-RFC (Bellaoui et al., 2003; Naiki et al., 2001; Stokes et al., 2020; Yao and O'Donnell, 2012), as well as the bacterial β-clamp loader or the T4 phage clamp loader, may also possess the shoulder 5′ DNA binding site? Both Ctf18 and Elg1, in the subunit A position of RFC, likely contain an intact AAA+ module and a similar collar domain that resemble Rfc1 and Rad24, based on AlphaFold structure prediction (**Supplementary Figure 6e-f**). The subunit in the analogous position to Rfc1 and Rad24 in the bacterial β-clamp loader is the δ subunit. The δ subunit maintains an intact AAA+ module and collar domain (Simonetta et al., 2009), and may contain a 5′ DNA binding site. Therefore, the eukaryotic alternative loaders Ctf18-RFC and Elg1 likely possess the shoulder 5′ DNA binding site, and the bacterial loader might contain a 5′ DNA site. However, the first subunit of the T4 phage clamp loader has lost much of the typical AAA+ module, retaining only three α-helices (**Supplementary Fig. 4c**) (Kelch et al., 2011), and is less likely to contain the shoulder DNA binding site. Further biochemical and structural studies are needed to answer this question.

Ideas and Speculation

3) Does the 5′-DNA2 site of RFC also function in lagging strand DNA replication? An Okazaki fragment is about 150-200 nt long (Ohashi and Tsurimoto, 2017; Smith and Whitehouse, 2012). Our biochemical data indicate RFC is stimulated by a 5′ end at a 50-nt gap, and therefore binding to both 3′ and 5′ DNA sites is not confined to an optimal 5-10 nt gap size determined by structural analysis alone. This opens the possibility that RFC maintains some level of contact with the 5’ end at gaps of larger sizes (i.e., 50 nt), and perhaps even

Okazaki fragment gaps. One may speculate that the RFC−5' DNA interaction could prevent premature 5′ exonuclease excision at the 5′ end of Okazaki fragments or may suppress the DNA damage response during replication by preventing the Rad24-RFC (Rad17-RFC in human) from using the 5′ end of Okazaki fragments to load the 9-1-1 clamp that activates the ATR cell cycle control kinase. The two-site interaction of RFC with the 3' and 5' ends of DNA might also aid RFC processivity on the lagging strand, in which RFC might transfer among multiple Okazaki fragments, by walking from one to another during replication. It is also possible that there is some function that in DNA replication and repair that is currently unknown. These possibilities will be a subject of future studies.

4) Does RFC regulate the 9-1-1 response? The reported cellular number of Rad24 is 5-fold less than that of Rfc1, in which the shared Rfc2-5 subunits are far more abundant according to the proteomics in the S.c. database (https://www.yeastgenome.org as of Jan 20, 2022). It seems possible that RFC/PCNA represses the ability of Rad24-RFC to function in the 5′ loading of the 9-1-1 clamp at Okazaki fragments, and even at DNA repair gaps. In this speculation, and given the reported 5-fold higher concentration of Rfc1 over Rad24 protein in the cell, the ATR checkpoint kinase may not be stimulated by the 9-1-1 clamp until significant DNA damage has occurred, overwhelming RFC protection of 5′ ends and PCNA loading for DNA damage gap fill-in. Specifically, at low DNA damage, RFC will load PCNA for gap fill in, but when DNA damage is severe, the RFC will be insufficient to control and fill all the gaps - at which time the Rad24-RFC will load 9-1-1 to turn on the DNA damage checkpoint. Thus, while quite speculative, RFC might act as a buffer against the loading of 9-1-1 for checkpoint activation. At some critical point, RFC will be overwhelmed by the number of repair gaps, and then Rad24-RFC will load the 9-1-1 clamp to upregulate the ATR kinase and induce a cell cycle checkpoint, enabling DNA repair until the genome is safe to continue being duplicated.

Exploring the possible interactions between different clamp loaders for DNA damage control is an exciting area for future research.

Acknowledgements. Cryo-EM micrographs were collected at the David Van Andel Advanced Cryo-Electron Microscopy Suite in Van Andel Institute. We thank G. Zhao and X. Meng for facilitating data collection. This work was supported by the US National Institutes of Health grants GM131754 (to H.L.) and GM115809 (to M.E.O.), Van Andel Institute (to H.L.), and Howard Hughes Medical Institute (to M.E.O).

Author contributions. H.L. and M.E.O. designed research; F.Z., R.E.G. and N.Y.Y. performed research; F.Z., R.E.G., N.Y.Y., H.L. and M.E.O. analyzed the data; F.Z., H.L., and M.E.O. wrote the manuscript with input from all authors.

Competing interests. The authors declare no competing interests.

Data Availability. The 3D cryo-EM maps of *S. cerevisiae* RFC−DNA and RFC−PCNA−DNA complexes have been deposited in the Electron Microscopy Data Bank with accession codes EMD-25872 (RFC−PCNA−DNA1−DNA2), EMD-25873 (RFC−open PCNA−DNA1), EMD-25874 (RFC−closed PCNA−DNA1), EMD-25875 (RFC−DNA1−DNA2), and EMD-25876 (RFC−DNA1). The corresponding atomic models have been deposited in the Protein Data Bank with accession codes 7TFH, 7TFI, 7TFJ, 7TFK and 7TFL.

MATERIALS AND METHODS

Reagents and proteins

Radioactive nucleotides were from PerkinElmer Life Sciences (Waltham, Massachusetts). Unlabeled ATP was from Cytiva (Marlborough, MA). ATPγS was from Roche (Basel, Switzerland). DNA-modification enzymes were from New England BioLabs (Ipswich.Massachusetts). DNA oligonucleotides were from Integrated DNA Technologies (Coralville, Iowa). Protein concentrations were determined with the Bio-Rad Labs (Hercules, California) Bradford Protein stain using bovine serum albumin as a standard. Streptavidin-coated Dynabead M-280 magnetic beads were purchased from Thermo-Fisher Scientific (Waltham, Massachusetts). Anti-Digoxigenin, the anti-Dig Fab, was from Millipore-Sigma (St. Louis, MO).

Cell growth

Expression plasmids were transformed into BLR(DE3) (Novagen, Madison, Wisconsin) *E.* coli-competent cells and selected on LB plates containing ampicillin (100 µg/ml). When needed for use of two expression plasmids, 2 plasmid transformants required selection on both Amp (100 µg/ml) and kanamycin (50 µg/ml). Fresh transformants were grown in 12–24 liters of LB media containing the appropriate antibiotics at 30 °C until reaching an OD_{600} value of 0.6. The cultures were then quickly brought to 15 °C by swirling culture flasks in an ice water bath with a thermometer, then placed into a prechilled shaker incubator at 15 °C before adding 1 mM isopropyl-1-thio-β-d-galactopyranoside. Induced cells were then incubated at 15 °C for ∼18 h. Cells were collected by centrifugation and resuspended in an equal weight of Tris-Sucrose (50mM Tris-HCl pH7.5, 10% sucrose w/v) and frozen until use. Frozen cells were thawed and then lysed by two passages through an Emulsi Flex-C-50 from Avestin (Ottawa, ON, Canada) using pulses of 20,000 psi and 5000 psi at 4 °C. Cell debris was removed by centrifugation (19,000 r.p.m. in a SS-34 rotor for 1h at 4 °C).

PCNA: The expression plasmid for PCNA was generated by PCR of the *pol30* gene from yeast genomic DNA and then cloned into a modified pET16 vector containing a hexahistine tag and 6 residue site for the catalytic subunit of cAMP dependent protein kinase A at the N-terminus of the expressed protein as described (Kelman et al., 1995). 48 L of transformed cells were grown as described above and resuspended in Buffer I (20 mM Tris-HCl (pH 7.9), 500 mM NaCl) plus 5 mM imidazole. The following procedures were performed at 4º C. After cell lysis and clarification by centrifugation, the supernatant (8 g) was loaded onto a 150 ml HiTrap Ni chelating resin (Cytiva, Marlborough, MA) and washed with 300 ml Buffer I containing 60 mM imidazole. Following this, protein was eluted using a 1.6 L linear gradient of Buffer I from 60 mM to 500 mM imidazole, collecting 25 ml fractions. Presence of proteins was followed in 10% SDS-polyacrylamide gels stained with Coomassie Blue. Fractions 10-38 were pooled (3.4 g) and dialyzed against Buffer A. The dialysate was loaded onto a 200 ml DEAE Sepharose column equilibrated in Buffer A (20 mM Tris-HCl, pH 7.5, 5 mM DTT, 0.1 mM EDTA and 10% glycerol). The column was washed with 500 ml Buffer A, then protein was eluted using a 2L linear gradient of Buffer A from zero NaCl to 500 mM NaCl; fractions of 22 ml were collected. Presence of proteins was followed in 10% SDS-polyacrylamide gels stained with Coomassie Blue. Fractions 54-70 were pooled to give 1.8 g >95% pure PCNA. The pool was dialyzed against Buffer A, and a small amount of precipitate was removed by centrifugation, giving 1.4 gram of PCNA which was then aliquoted, snap frozen in liquid N_2, and stored at -80ºC

RFC: Full length *S. cerevisiae* RFC utilized two expression plasmids having different and compatible origins as described previously (Finkelstein et al., 2003). pLANT-2/RIL–RFC[1+5] was co-transformed with pET(11a)-RFC[2+3+4] into BLR(DE3) cells (Novagen, Madison, Wisconsin). The cell lysate was clarified by centrifugation, diluted with Buffer A to ∼150 mM NaCl, and then applied to a 30-ml SP-Sepharose Fast Flow column (Cytiva, Marlborough, MA) equilibrated with Buffer A containing 150 mM NaCl. The column was eluted with a 300-ml gradient of 150–600 mM NaCl in Buffer A. Presence of proteins was followed in 10% SDS-polyacrylamide gels stained with Coomassie Blue. The peak of RFC (which eluted at ∼365 mm NaCl) was pooled and diluted with Buffer A to ∼150 mM NaCl. The protein was then applied to a 40 ml Q-Sepharose Fast Flow column (Cytiva, Marlborough, MA) equilibrated with Buffer A containing 150 mM NaCl and eluted with a 400-ml gradient of 150–600 mM NaCl in Buffer A. The fractions containing RFC complex (which eluted at ∼300 mM NaCl) were pooled (24 mg), aliquoted, flash frozen in liquid N_2 and stored at -80 °C. Truncated $RFC^{\Delta N282}$ was purified by the same procedure.

RPA: RPA was expressed and purified as described (Henricksen et al., 1994).

An SDS-PAGE analysis of these preps is shown in **Supplementary Fig. 11.**

Cryo-EM grids preparation and data collection

The double-tailed DNA substrate with both 3′- and 5′-recessed ends was composed of a 20-nt primer strand (5′-GCA GAC ACT ACG AGT ACA TA -3′) and a 40-nt template strand (5′- TTT TTT TTT TTA TGT ACT CGT AGT GTC TGC TTT TTT TTT T -3′). They were synthesized, HPLC purified, and then annealed by Integrated DNA Technologies Inc (**Supplementary Table 2**). The DNA substrate is also referred to as P/T_DNA in the manuscript. The in vitro assembly of yeast RFC–PCNA–DNA complex followed our previous procedure for assembling the S.c. Pol δ–PCNA–DNA complex (Zheng et al., 2020). Briefly, 1 µl of purified PCNA clamp at 30 µM, 3.0 µl of double tailed DNA at 100 µM were mixed and incubated at 30°C for 10 min, then the mixture along

with 0.75 µl 10 mM ATPγS and 0.75 µl 100 mM Mg-Acetate were added into 10.0 µl of purified RFC protein at 3.3 µM concentration for a final concentration of 2.2 µM RFC, 2.0 µM PCNA clamp, 20.0 µM DNA, 0.5 mM ATPγS, and 5 mM Mg-Acetate, in a total reaction volume of 15 µl. The final molar ratio of RFC: PCNA clamp: DNA was 1.0: 0.9: 9.0. The mixture was then incubated in an ice-water bath for an additional 1 hr. The Quantifoil Cu R2/1 300 mesh grids were glow discharged for 1 min in a Gatan Solarus, then 3 µl of the mixture was applied onto the EM grids. Sample vitrification was carried out in a Vitrobot (Thermo Fisher Mark IV) with the following settings: blot time 2 s, blot force 4, wait time 2 s, inner chamber temperature 6°C, and a 95% relative humidity. The EM grids were flash-frozen in liquid ethane cooled by liquid nitrogen. Cryo-EM data were automatically collected on a 300 kV Titian Krios electron microscope controlled by SerialEM in a multi-hole mode. The micrographs were captured at a scope magnification of 105,000×, with the objective lens under-focus values ranging from 1.3 to 1.9 µm, by a K3 direct electron detector (Gatan) operated in the super-resolution video mode. During a 1.5 s exposure time, a total of 75 frames were recorded with a total dose of 66 $e^-/Å^2$. The calibrated physical pixel size was 0.828 Å for all digital micrographs.

Image processing and 3D reconstruction

The data collection and image quality were monitored by the cryoSPARC Live v3.2 (Punjani et al., 2017) installed in a local workstation. The image preprocessing including patch motion correction, contrast transfer function (CTF) estimation and correction, blob particle picking (70-150 Å diameter) and extraction with a binning factor of 2 were also achieved at the same time, and a total number of 16,401 raw micrographs were recorded during a three-day data collecting session. We performed two rounds of 2-dimensional (2D) image classification, which resulted in ~1.3 million "good" particle images. Then we trained Topaz (Bepler et al., 2019) and used the trained program to pick more particles. We next used the reported "Build and Retrieve" method to avoid missing those less frequently occurring particle views (Su et al., 2021). The cleaned-up particles were combined, and duplicate particles with 40% overlapping (52 Å) or larger than the particle diameter (~130 Å) were removed. The final dataset contained ~2.9 million particles that was used for 3D reconstruction. We first calculated four starting 3D models and obtained two major 3D classes. In the first 3D class, RFC is clearly bound to DNA but not to PCNA. In the second 3D class, RFC is seen bound to both DNA and PCNA. We subjected both classes to non-uniform 3D refinement and 3D viability analysis (3DVA), resulting in eight 3D subclasses for each of the two starting 3D classes.

Among the eight 3DVA-derived subclasses of the RFC–DNA complexes in the absence of PCNA, we found two subclasses bound to one DNA (RFC–DNA1) and two subclasses bound to two DNA molecules (RFC–3′-DNA1–5′-DNA2). The remaining four subclasses contained 3′-DNA1 with partially occupied 5′-DNA2, and they were discarded without further processing. Particles in the two subclasses bound to 3′-DNA1 only were combined into a dataset with 334,876 particles, and they were refined to a 3D map of 3.33 Å average resolution. Particles in the two subclasses bound to 3′-DNA1 and 5′-DNA2 were also combined to a dataset of 315,619 particles, and they were refined to a 3D map at 3.25 Å average resolution.

Among the eight 3DVA-derived subclasses of the RFC–DNA–PCNA complex, 1 subclass bound to 3′-DNA1 only and the bound PCNA was in an open spiral conformation. This subclass contained 118,384 particles and was further refined to a 3D map at 3.41 Å average resolution (RFC–3′-DNA1–open ring PCNA). Two subclasses bound to 3′-DNA1 only with the PCNA in a closed ring conformation. Particles in these two subclasses were combined to a dataset of 166,348 particles, and they were refined to a 3D map at 3.30 Å average resolution (RFC–3′-DNA1–closed ring PCNA). We found 4 subclasses contained 2 DNA molecules, in which, 3 subclasses were combined into a dataset of 432,904 particles and refined to a 3D map (RFC–3′-DNA1–5′-DNA2–PCNA) at 3.09 Å average resolution. The two remaining subclasses only had partial PCNA density, were discarded without further processing. In the 4 subclasses bound to PCNA and 3′-DNA1 and 5′-DNA2, two subclasses did not contain density, one had weak density, and one had strong density for the Rfc1 BRCT domain. We performed focused refinement on 85,932 particles in the subclass with strong density for Rfc1 BRCT. We used a 3D mask that included the Rfc1 NTD (BRCT and AAA+ module) and 5′-DNA2 region for particle subtraction and refined a local 3D map at 3.90 Å average resolution. This local 3D map was aligned to the 3.09-Å 3D map of the full complex and combined in UCSF Chimera (Pettersen et al., 2004) to obtain the final composite map of the RFC–3′-DNA1–5′-DNA2–PCNA complex.

Model building, refinement, and validation

We used the crystal structure of the yeast RFC–PCNA complex (PDB entry 1SXJ) as the initial model for atomic model building of the above described five 3D EM maps. The EM map of the Rfc1 BRCT domain was at a low resolution of 3.90 Å, and there was no available high-resolution structure for the yeast domain. We referenced the crystal structure of the human RFC1 BRCT (PDB entry 2K6G) and the AlphaFold predicted yeast Rfc1 model (AF-P35251-F1) to build the atomic model (Jumper et al., 2021). The DNA from the T4 phage clamp–loader–DNA crystal structure (PDB entry 3U5Z) and the DNA from yeast Rad24-RFC–9-1-1 clamp–DNA complex structure (PDB entry 7SGZ) were adopted as the initial models for chamber (3′-DNA1) and shoulder DNA (5′-DNA2) building, respectively. We also used the *de novo* modeling program Map-to-Model wrapped in PHENIX (Adams et al., 2010) and the module "Automated Nucleic Acid building" function integrated in COOT (Emsley et al., 2010) to help modeling the two DNA molecules. These initial models are fitted into the composite 3D map of RFC–3′-DNA1–5′-DNA2–PCNA and merged into one single coordinate file in the UCSF Chimera (Pettersen et al., 2004) to serve as the starting model for the entire complex. The starting model was refined iteratively between the real space computational refinement in PHENIX and the manual adjustment in COOT. The atomic model of the RFC–3′-DNA1–5′-DNA2–PCNA complex was finally refined to 3.1 Å and went through a comprehensive validation by the MolProbity program (Chen et al., 2010) embedded in PHENIX. This model then served as the initial model of the other four 3D maps. The model for each 3D map was subjected to iteratively real space refinement and manually adjustment as described above. The two PCNA subunits at the gate had weaker EM densities in the spirally open PCNA EM map, and they were modelled by rigid body docking with a PCNA monomer structure. Structure figures were prepared using ChimeraX (Pettersen et al., 2021) and organized in Adobe Illustrator (Adobe Inc, San Jose, CA).

Fluorescent DNA binding assays

The oligonucleotides used to make the DNA substrates (**Supplementary Table 3**) were synthesized and PAGE or HPLC-purified by Integrated DNA Technologies (Coralville, Iowa
). Gapped DNA templates used in the fluorescent binding assays were obtained by annealing the Primer1-Cy3, Primer 2 (or Primer2 5′-P-DNA) with their corresponding complementary DNA template oligos (Tmpl01_40, Tmpl02_41, etc.). The annealed templates were purified from a 10% native PAGE in 0.5x TBE; gel bands were cut out from the gel, extracted in TE buffer and further concentrated and buffer exchanged into standard TE buffer (10 mM Tris-HCl pH 7.5 and 1 mM EDTA).

RFC binding to DNA was measured using each 3′ Cy3-labeled gel purified DNA template (singly primed, nicked, 1-, 2-, 5-, 10- or 50-dT gaps) at constant 10 nM DNA concentration. RFC complex was titrated from 0 to 240 nM into reactions of 50 µL final volume containing 20 mM Tris-HCl (pH 7.5), 100 mM NaCl, 8 mM Mg-Acetate, 4 mM DTT, 0.1 mM EDTA and 100 µM ATPγS. When present, the PCNA trimer was at a concentration ratio of 1:1 with RFC. The reaction mixtures were assembled in a 384-well plate format, then incubated at room temperature for 20 min. The plates were centrifuged for 20 sec to eliminate any bubbles and to assure that the liquid mix is evenly distributed on the well-bottom; the fluorescence signal was recorded using the Synergy Neo2 plate reader fluorimeter (BioTek Instruments, Winooski, VT) by exciting the Cy3 fluorophore at 535 nm and read out of the integrated fluorescence signal between 560-580 nm. The fluorescence recording routine involved a reading every 3-min preceded by an elliptical motion plate stirring for 20 sec. The recording was done for 45 min at constant 23.8°C, for a total 16 readings averaged for each concentration point. Three independent experiments were performed for each titration. Relative intensity (I/I_0) was plotted versus RFC concentration. Apparent K_d measurements were determined using a simple one-site total binding model corrected with a nonspecific binding slope (RFC + DNA ↔ RFC–DNA), using the GraphPad program (GraphPad).

^{32}P-PCNA clamp loading on gapped DNAs.

For magnetic beads assays the DIG/BIO templates and the primer #1 and, when present, primer #2, were mixed with their respective template strand (**Supplementary Table 3**) in a 1:1.2 ratio of template to primer in 50 µl of 5 mM Tris-HCl, 150 mM NaCl, 15 mM sodium citrate (final pH 8.5), then incubated in a 95°C water bath, which was allowed to cool to room temperature over a 30-min interval. Streptavidin-coated Dynabead M-280 magnetic beads were purchased from Thermo-Fisher Scientific (Waltham, MA) and anti-Digoxigenin, the anti-Dig Fab was from Millipore-Sigma (Burlington, MA). Each primed biotinylated DNA template (200 pmol) was incubated with 1 mg Dynabeads M-280 Streptavidin in 5 mM Tris-HCl (pH 7.5), 0.5 mM EDTA and 1M NaCl. An average yield of 130 to 200 pmol DNA/mg of Dynabeads, was determined by releasing DNA from beads by 1% SDS, removing beads, then measuring the absorbance of the supernatant at 260 nm using a nanodrop spectrophotometer. To

block non-specific protein binding the DNA-bead conjugates were pre-incubated with 5 mg/ml BSA in 10 mM Tris-HCl (pH 7.5), 1 mM EDTA, 150 mM NaCl at room temperature for 30 minutes and washed in the same buffer until no BSA remained in the supernatant as detected by Bradford reagent.

The kinase tag on PCNA was used to label it with ^{32}P to a specific activity of 40 cpm/ fmol with [γ-^{32}P] ATP using the recombinant catalytic subunit of cAMP-dependent protein kinase produced in *E. coli* (a gift from Dr. Susan Taylor, University of California at San Diego) as described earlier (Kelman et al., 1995). Before the start of the PCNA-loading assay, 2.5 fold molar excess RPA over DNA was added (for experiments of **Fig. 5c**) and the 5′ Dig of the DNA-bead conjugate was blocked with anti-digoxigenin Fab fragment in a ratio of 1:2 respectively in clamp loading buffer (30 mM HEPES-NaOH (pH 7.5), 1 mM DTT and 1 mM CHAPS) for 10 minutes at room temperature (23°C). The clamp loading assays were then performed in the cold room at 16°C. Each reaction contained 167 nM antibody-blocked DNA-bead conjugate, 376 nM ^{32}P-PCNA in 30 µl of clamp loading buffer, 8 mM $MgCl_2$, 0.5 mM ATP and was initiated with 167 nM RFC for 20 seconds before being quenched with 37 mM EDTA. The beads were then collected in a magnetic concentrator and washed 2X with 200 µl clamp loading buffer containing 100 mM NaCl. The DNA-bound ^{32}P-PCNA was stripped from the beads with 0.5% SDS and 5 minutes of boiling and counted by liquid scintillation. Controls performed the same way but lacking RFC gave less than 5% background ^{32}P-PCNA binding for any given experiment.

References

Aboussekhra, A., Biggerstaff, M., Shivji, M.K., Vilpo, J.A., Moncollin, V., Podust, V.N., Protic, M., Hubscher, U., Egly, J.M., and Wood, R.D. (1995). Mammalian DNA nucleotide excision repair reconstituted with purified protein components. Cell *80*, 859-868.

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W.*, et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr *66*, 213-221.

Allen, B.L., Uhlmann, F., Gaur, L.K., Mulder, B.A., Posey, K.L., Jones, L.B., and Hardin, S.H. (1998). DNA recognition properties of the N-terminal DNA binding domain within the large subunit of replication factor C. Nucleic Acids Res *26*, 3877-3882.

Altieri, A.S., and Kelman, Z. (2018). DNA Sliding Clamps as Therapeutic Targets. Front Mol Biosci *5*, 87.

Bell, S.P., and Labib, K. (2016). Chromosome Duplication in Saccharomyces cerevisiae. Genetics *203*, 1027-1067.

Bellaoui, M., Chang, M., Ou, J., Xu, H., Boone, C., and Brown, G.W. (2003). Elg1 forms an alternative RFC complex important for DNA replication and genome integrity. EMBO J *22*, 4304-4313.

Bepler, T., Morin, A., Rapp, M., Brasch, J., Shapiro, L., Noble, A.J., and Berger, B. (2019). Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. Nat Methods *16*, 1153-1160.

Blank, A., Kim, B., and Loeb, L.A. (1994). DNA polymerase delta is required for base excision repair of DNA methylation damage in Saccharomyces cerevisiae. Proc Natl Acad Sci U S A *91*, 9047-9051.

Bloom, L.B. (2009). Loading clamps for DNA replication and repair. DNA Repair (Amst) *8*, 570-578.

Boiteux, S., and Jinks-Robertson, S. (2013). DNA repair mechanisms and the bypass of DNA damage in Saccharomyces cerevisiae. Genetics *193*, 1025-1064.

Bowman, G.D., Goedken, E.R., Kazmirski, S.L., O'Donnell, M., and Kuriyan, J. (2005). DNA polymerase clamp loaders and DNA recognition. FEBS Lett *579*, 863-867.

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

Burgers, P.M.J., and Kunkel, T.A. (2017). Eukaryotic DNA Replication Fork. Annu Rev Biochem *86*, 417-438.

Cadet, J., Sage, E., and Douki, T. (2005). Ultraviolet radiation-mediated damage to cellular DNA. Mutat Res *571*, 3-17.

Castaneda, J.C., Schrecker, M., Remus, D., and Hite, R.K. (2021). Mechanisms of loading and release of the 9-1-1 checkpoint clamp. bioRxiv. 2021.09.13.460164; doi: https://doi.org/10.1101/2021.09.13.460164

Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr *66*, 12-21.

Chen R, Wold MS. (2014) Replication protein A: single-stranded DNA's first responder: dynamic DNA-interactions allow replication protein A to direct single-strand DNA intermediates into different pathways for synthesis or repair. Bioessay*s. 36,:*1156-1161.

Cullmann, G., Fien, K., Kobayashi, R., and Stillman, B. (1995). Characterization of the five replication factor C genes of Saccharomyces cerevisiae. Mol Cell Biol *15*, 4661-4671.

Davey, M.J., Jeruzalmi, D., Kuriyan, J., and O'Donnell, M. (2002). Motors and switches: AAA+ machines within the replisome. Nat Rev Mol Cell Biol *3*, 826-835.

Ellison, V., and Stillman, B. (2001). Opening of the clamp: an intimate view of an ATP-driven biological machine. Cell *106*, 655-660.

Ellison, V., and Stillman, B. (2003). Biochemical characterization of DNA damage checkpoint complexes: clamp loader and clamp complexes with specificity for 5’ recessed DNA. PLoS Biol *1*, E33.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr D Biol Crystallogr *66*, 486-501.

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

Finkelstein, J., Antony, E., Hingorani, M.M., and O'Donnell, M. (2003). Overproduction and analysis of eukaryotic multiprotein complexes in Escherichia coli using a dual-vector strategy. Anal Biochem *319*, 78-87.

Gaubitz, C., Liu, X., Magrino, J., Stone, N.P., Landeck, J., Hedglin, M., and Kelch, B.A. (2020). Structure of the human clamp loader reveals an autoinhibited conformation of a substrate-bound AAA+ switch. Proc Natl Acad Sci U S A *117*, 23571-23580.

Gaubitz, C., Xingchen, L., Pajak, J., Stone, N.P., Hayes, J.A., Demo, G., and Kelch, B.A. (2021). Cryo-EM structures reveal high-resolution mechanism of a DNA polymerase sliding clamp loader. bioRxiv.

Gomes, X.V., and Burgers, P.M. (2001). ATP utilization by yeast replication factor C. I. ATP-mediated interaction with DNA and with proliferating cell nuclear antigen. J Biol Chem *276*, 34768-34775.

Gomes, X.V., Gary, S.L., and Burgers, P.M. (2000). Overproduction in Escherichia coli and characterization of yeast replication factor C lacking the ligase homology domain. J Biol Chem *275*, 14541-14549.

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

Harrison, J.C., and Haber, J.E. (2006). Surviving the breakup: the DNA damage checkpoint. Annu Rev Genet *40*, 209-235.

Hedglin, M., Kumar, R., and Benkovic, S.J. (2013). Replication clamps and clamp loaders. Cold Spring Harb Perspect Biol *5*, a010165.

Henricksen, L.A., Umbricht, C.B., and Wold, M.S. (1994). Recombinant replication protein A: expression, complex formation, and functional characterization. J Biol Chem *269*, 11121-11132.

Hingorani, M.M., and O'Donnell, M. (2000). A tale of toroids in DNA metabolism. Nat Rev Mol Cell Biol *1*, 22-30.

Indiani, C., and O'Donnell, M. (2006). The replication clamp-loading machine at work in the three domains of life. Nat Rev Mol Cell Biol *7*, 751-761.

Johnson, A., Yao, N.Y., Bowman, G.D., Kuriyan, J., and O'Donnell, M. (2006). The replication factor C clamp loader requires arginine finger sensors to drive DNA binding and proliferating cell nuclear antigen loading. J Biol Chem *281*, 35531-35543.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Zidek, A., Potapenko, A.*, et al.* (2021). Highly accurate protein structure prediction with AlphaFold. Nature *596*, 583-589.

Kelch, B., Makino, D., Simonetta, K., O'Donnell, M., and Kuriyan, J. (2012a). Molecular Mechanisms of DNA Polymerase Clamp Loaders. In Macromolecular Crystallography: Deciphering the Structure, Function and Dynamics of Biological Molecules, M.A. Carrondo, and P. Spandon, eds. (The Netherlands: Springer), pp. 103-114.

Kelch, B.A., Makino, D.L., O'Donnell, M., and Kuriyan, J. (2011). How a DNA polymerase clamp loader opens a sliding clamp. Science *334*, 1675-1680.

Kelch, B.A., Makino, D.L., O'Donnell, M., and Kuriyan, J. (2012b). Clamp loader ATPases and the evolution of DNA replication machinery. BMC Biol *10*, 34.

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

Kelman, Z., Yao, N., and O'Donnell, M. (1995). Escherichia coli expression vectors containing a protein kinase recognition motif, His6-tag and hemagglutinin epitope. Gene *166*, 177-178.

Kobayashi, M., Ab, E., Bonvin, A.M., and Siegal, G. (2010). Structure of the DNA-bound BRCA1 C-terminal region from human replication factor C p140 and model of the protein-DNA complex. J Biol Chem *285*, 10087-10097.

Kobayashi, M., Figaroa, F., Meeuwenoord, N., Jansen, L.E., and Siegal, G. (2006). Characterization of the DNA binding and structural properties of the BRCT region of human replication factor C p140 subunit. J Biol Chem *281*, 4308-4317.

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

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

Kunkel, T.A., and Burgers, P.M. (2014). Delivering nonidentical twins. Nat Struct Mol Biol *21*, 649-651.

Kunkel, T.A., and Burgers, P.M.J. (2017). Arranging eukaryotic nuclear DNA polymerases for replication: Specific interactions with accessory proteins arrange Pols alpha, delta, and in the replisome for leading-strand and lagging-strand DNA replication. Bioessays *39*.

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

Li, H., and O'Donnell, M.E. (2018). The Eukaryotic CMG Helicase at the Replication Fork: Emerging Architecture Reveals an Unexpected Mechanism. Bioessays *40*.

Li, H., Zheng, F., and O'Donnell, M. (2021). Water skating: How polymerase sliding clamps move on DNA. FEBS J *288*, 7256-7262.

Li, R., Waga, S., Hannon, G.J., Beach, D., and Stillman, B. (1994). Differential effects by the p21 CDK inhibitor on PCNA-dependent DNA replication and repair. Nature *371*, 534-537.

Madru, C., Henneke, G., Raia, P., Hugonneau-Beaufet, I., Pehau-Arnaudet, G., England, P., Lindahl, E., Delarue, M., Carroni, M., and Sauguet, L. (2020). Structural basis for the increased processivity of D-family DNA polymerases in complex with PCNA. Nat Commun *11*, 1591.

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

Majka, J., Binz, S.K., Wold, M.S., and Burgers, P.M. (2006). Replication protein A directs loading of the DNA damage checkpoint clamp to 5'-DNA junctions. J Biol Chem *281*, 27855-27861.

Majka, J., and Burgers, P.M. (2003). Yeast Rad17/Mec3/Ddc1: a sliding clamp for the DNA damage checkpoint. Proc Natl Acad Sci U S A *100*, 2249-2254.

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

McAlear, M.A., Tuffo, K.M., and Holm, C. (1996). The large subunit of replication factor C (Rfc1p/Cdc44p) is required for DNA replication and DNA repair in Saccharomyces cerevisiae. Genetics *142*, 65-78.

Miller, J.M., and Enemark, E.J. (2016). Fundamental Characteristics of AAA+ Protein Family Structure and Function. Archaea *2016*, 9294307.

Naiki, T., Kondo, T., Nakada, D., Matsumoto, K., and Sugimoto, K. (2001). Chl12 (Ctf18) forms a novel replication factor C-related complex and functions redundantly with Rad24 in the DNA replication checkpoint pathway. Mol Cell Biol *21*, 5838-5845.

Nick McElhinny, S.A., Gordenin, D.A., Stith, C.M., Burgers, P.M., and Kunkel, T.A. (2008). Division of labor at the eukaryotic replication fork. Mol Cell *30*, 137-144.

Nohmi, T., Kim, S.R., and Yamada, M. (2005). Modulation of oxidative mutagenesis and carcinogenesis by polymorphic forms of human DNA repair enzymes. Mutat Res *591*, 60-73.

O'Donnell, M.E., and Kurth, I. (2013). Machinery of DNA Replication. In Encyclopedia of Biophysics, G.C.K. Roberts, ed. (
Berlin: Springer, Berlin, Heidelberg), pp. 1327-1338.

Ohashi, E., and Tsurimoto, T. (2017). Functions of Multiple Clamp and Clamp-Loader Complexes in Eukaryotic DNA Replication. Adv Exp Med Biol *1042*, 135-162.

Overmeer, R.M., Gourdin, A.M., Giglia-Mari, A., Kool, H., Houtsmuller, A.B., Siegal, G., Fousteri, M.I., Mullenders, L.H., and Vermeulen, W. (2010). Replication factor C recruits DNA polymerase delta to sites of nucleotide excision repair but is not required for PCNA recruitment. Mol Cell Biol *30*, 4828-4839.

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

Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H., and Ferrin, T.E. (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. Protein Sci *30*, 70-82.

Podust, V.N., Tiwari, N., Stephan, S., and Fanning, E. (1998). Replication factor C disengages from proliferating cell nuclear antigen (PCNA) upon sliding clamp formation, and PCNA itself tethers DNA polymerase delta to DNA. J Biol Chem *273*, 31992-31999.

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

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

Pursell, Z.F., Isoz, I., Lundstrom, E.B., Johansson, E., and Kunkel, T.A. (2007). Yeast DNA polymerase epsilon participates in leading-strand DNA replication. Science *317*, 127-130.

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

Shivji, K.K., Kenny, M.K., and Wood, R.D. (1992). Proliferating cell nuclear antigen is required for DNA excision repair. Cell *69*, 367-374.

Simonetta, K.R., Kazmirski, S.L., Goedken, E.R., Cantor, A.J., Kelch, B.A., McNally, R., Seyedin, S.N., Makino, D.L., O'Donnell, M., and Kuriyan, J. (2009). The mechanism of ATP-dependent primer-template recognition by a clamp loader complex. Cell *137*, 659-671.

Smith, D.J., and Whitehouse, I. (2012) Intrinsic coupling of lagging-strand synthesis to chromatin assembly. Nature. *483*,434-438.

Snider, J., Thibault, G., and Houry, W.A. (2008). The AAA+ superfamily of functionally diverse proteins. Genome Biol *9*, 216.

Stokes, K., Winczura, A., Song, B., Piccoli, G., and Grabarczyk, D.B. (2020). Ctf18-RFC and DNA Pol form a stable leading strand polymerase/clamp loader complex required for normal and perturbed DNA replication. Nucleic Acids Res *48*, 8128-8145.

Su, C.C., Lyu, M., Morgan, C.E., Bolla, J.R., Robinson, C.V., and Yu, E.W. (2021). A 'Build and Retrieve' methodology to simultaneously solve cryo-EM structures of membrane proteins. Nat Methods *18*, 69-75.

Su, T.T. (2006). Cellular responses to DNA damage: one signal, multiple choices. Annu Rev Genet *40*, 187-208.

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

Uchiumi, F., Watanabe, M., and Tanuma, S. (1999). Characterization of telomere-binding activity of replication factor C large subunit p140. Biochem Biophys Res Commun *258*, 482-489.

Uhlmann, F., Cai, J., Gibbs, E., O'Donnell, M., and Hurwitz, J. (1997). Deletion analysis of the large subunit p140 in human replication factor C reveals regions required for complex formation and replication activities. J Biol Chem *272*, 10058-10064.

Waga, S., and Stillman, B. (1998). The DNA replication fork in eukaryotic cells. Annu Rev Biochem *67*, 721-751.

Wang, Z., Wu, X., and Friedberg, E.C. (1993). DNA repair synthesis during base excision repair in vitro is catalyzed by DNA polymerase epsilon and is influenced by DNA polymerases alpha and delta in Saccharomyces cerevisiae. Mol Cell Biol *13*, 1051-1058.

Xiao, W., and Samson, L. (1993). In vivo evidence for endogenous DNA alkylation damage as a source of spontaneous mutation in eukaryotic cells. Proc Natl Acad Sci U S A *90*, 2117-2121.

Xie, Y., Counter, C., and Alani, E. (1999). Characterization of the repeat-tract instability and mutator phenotypes conferred by a Tn3 insertion in RFC1, the large subunit of the yeast clamp loader. Genetics *151*, 499-509.

Yao, N.Y., and O'Donnell, M. (2012). The RFC clamp loader: structure and function. Subcell Biochem *62*, 259-279.

Yao, N.Y., and O'Donnell, M.E. (2016a). Comparison of bacterial and eukaryotic replisome components. In Encyclopedia of Cell Biology, R.A. Bradshaw, and P.D. Stahl, eds. (Academic Press, Waltham), pp. 396-417.

Yao, N.Y., and O'Donnell, M.E. (2016b). Evolution of replication machines. Crit Rev Biochem Mol Biol *51*, 135-149.

Yao, N.Y., and O'Donnell, M.E. (2021). The DNA Replication Machine: Structure and Dynamic Function. Subcell Biochem *96*, 233-258.

Zhang, D., and O'Donnell, M. (2016). The Eukaryotic Replication Machine. In The Enzymes, L.S. Kaguni, and M.T. Oliveira, eds. (Academic Press), pp. 191-229.

Zheng, F., Georgescu, R.E., Li, H., and O'Donnell, M.E. (2020). Structure of eukaryotic DNA polymerase delta bound to the PCNA clamp while encircling DNA. Proc Natl Acad Sci U S A *117*, 30344-30353.

Zheng, F., Georgescu, R.E., Yao, N., O'Donnell, M.E., and Li, H. (2021). Rad24-RFC loads the 9-1-1 clamp by inserting DNA from the top of a wide-open ring, opposite the mechanism of RFC/PCNA. bioRxiv. 2021.09.23.461575; doi: https://doi.org/10.1101/2021.09.23.461575

Supplementary information for

A novel 5′ DNA binding site in RFC facilitates PCNA loading for gap DNA repair

Fengwei Zheng[1], Roxana E. Georgescu[2], Nina Y. Yao[2], Huilin Li[1,*], and Michael E. O'Donnell[2,3,*]

Supplementary Tables 1 – 3

Supplementary Figures 1 – 11

Legends for Supplementary Videos 1 and 2

Supplementary Table 1. Cryo-EM data collection, refinement, and atomic model validation

Structures	RFC–3´-DNA1	RFC–3´-DNA1–5´-DNA2	RFC–3´-DNA1–PCNA (open)	RFC–3´-DNA1–PCNA (closed)	RFC–3´-DNA1–5´-DNA2–PCNA
Data collection and processing					
Magnification			105,000		
Voltage (kV)			300		
Electron dose (e^-/Å^2)			66		
Under-focus range (μm)			1.3 - 1.9		
Pixel size (Å)			0.828		
Symmetry imposed			C1		
Initial particle images (no.)	1,039,425			908,744	
Final particle images (no.)	334,876	315,619	118,384	166,348	432,904
Map resolution (Å)	3.33	3.25	3.41	3.30	3.09
FSC threshold			0.143		
Map resolution range (Å)	2.0 – 12.0	2.0 – 13.0	2.0 – 12.0	2.0 – 12.0	2.1 – 13.0
Refinement					
Initial model used (PDB code)	RFC–3´-DNA1–5´-DNA2–PCNA of this study				1SXJ, 7SGZ, 2K6G
Map sharpening B factor (Å^2)	-93.8	-100.8	-97.3	-118.6	-116.6
Map to model CC_{mask}	0.70	0.69	0.82	0.81	0.82
Model composition					
Non-hydrogen atoms	13,690	14,535	20,573	20,607	22,831
Protein and DNA residues	1,653; 24	1,701; 45	2,480; 42	2,486; 42	2,717; 63
Ligands	9	9	9	9	9
R.m.s. deviations					
Bond lengths (Å)	0.003	0.003	0.003	0.003	0.003
Bond angels (º)	0.616	0.634	0.596	0.623	0.562
Validation					
MolProbity score	1.72	1.72	1.66	1.56	1.57
Clashscore	9.60	10.00	8.57	7.59	7.57
Poor rotamers (%)	0.14	0.13	0.05	0.09	0.04
Ramachandran plot					
Favored (%)	96.64	96.78	96.83	97.20	97.14
Allowed (%)	3.36	3.22	3.17	2.80	2.86
Disallowed (%)	0	0	0	0	0

Supplementary Table 2. The double-tailed DNA substrate used for the in vitro assembly of the RFC–PCNA–DNA complex

	Sequence
Primer strand	5′- GCA GAC ACT ACG AGT ACA TA -3′
Template strand	5′- TTT TTT TTT TTA TGT ACT CGT AGT GTC TGC TTT TTT TTT T -3′
Sketch of annealed dsDNA	5' GCAGACACTACGAGTACATA |||||||||||||||||||| 3' TTTTTTTTTTCGTCTGTGATGCTCATGTATTTTTTTTTTT

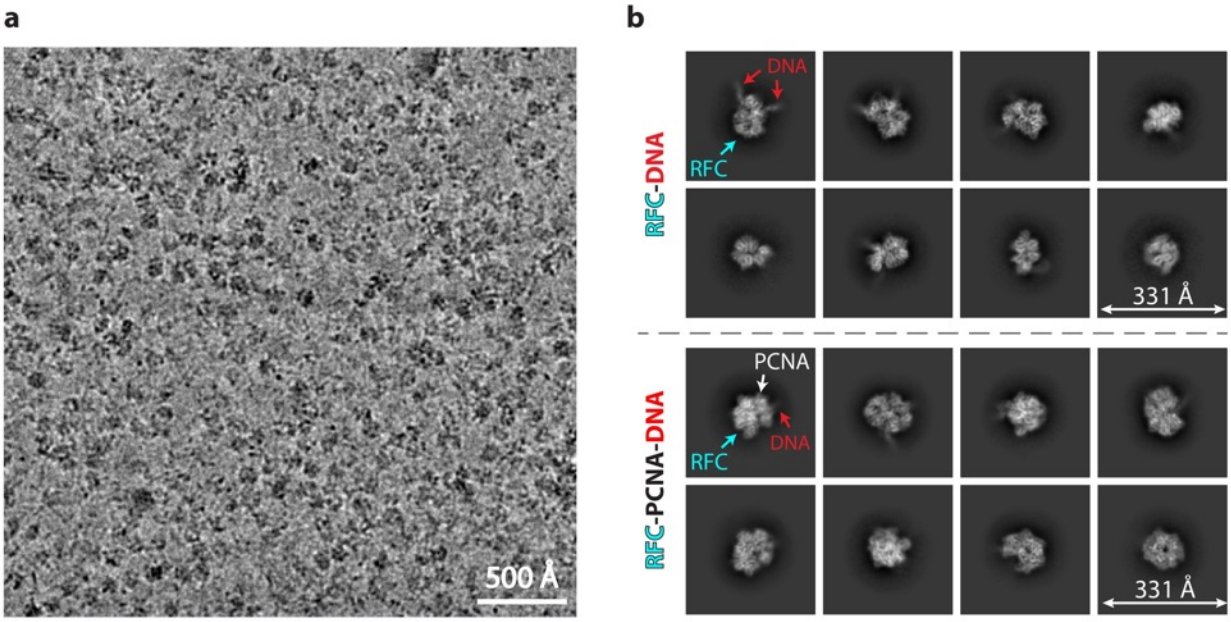
* The stably bound and modeled regions of DNA1 and DNA2 in the quaternary RFC–DNA1–DNA2–PCNA complex are marked by a dashed blue and red box, respectively.

19 **Supplementary Table 3. Sequence of DNA oligonucleotides used for in vitro assays**

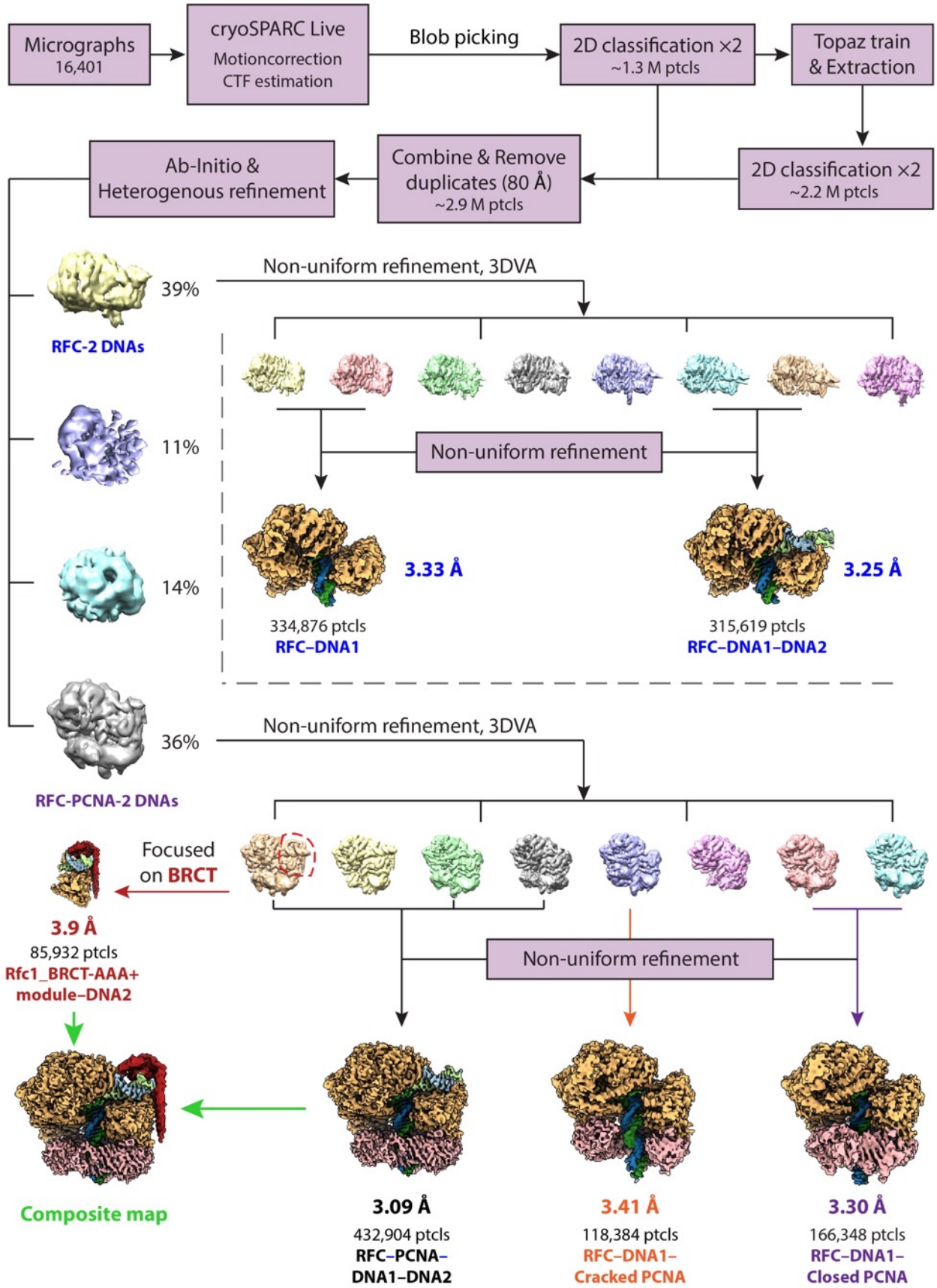
20

Primer 1	5′ GGTATGGTCGGGCTCGCTACGGCTATACGG 3′
Primer1-Cy3	5′- GGCTCGCTACGGCTATACGG-3′-Cy3′
Primer2	5′- GATTCGTATCGCCTATACCG -3′
Primer2 - 5′Phos	5′-P-GATTCGTATCGCCTATACCG-3′
Tmpl01_40	5′- CGGTATAGGCGATACGAATCCCGTATAGCCGTAGCGAGCC -3′
Tmpl02_41	5′- CGGTATAGGCGATACGAATCTCCGTATAGCCGTAGCGAGCC -3′
Tmpl03_42	5′- CGGTATAGGCGATACGAATC$(T)_2$CCGTATAGCCGTAGCGAGCC -3′
Tmpl04_45	5′- CGGTATAGGCGATACGAATC$(T)_5$CCGTATAGCCGTAGCGAGCC -3′
Tmpl05_50	5′- CGGTATAGGCGATACGAATC$(T)_{10}$CCGTATAGCCGTAGCGAGCC -3′
Tmpl06_90	5′- CGGTATAGGCGATACGAATC$(T)_{50}$CCGTATAGCCGTAGCGAGCC -3′
Bead-Pr#1-template	5′-Dig-CGGTATAGGCGATACGAATC$(T)_{10}$CCGTATAGCCGTAGCGAGCCCGACCATACC-3′-Bio
Bead-gap50-template	5′-Dig-CGGTATAGGCGATACGAATC$(T)_{50}$CCGTATAGCCGTAGCGAGCCCGACCATACC-3′-Bio

21

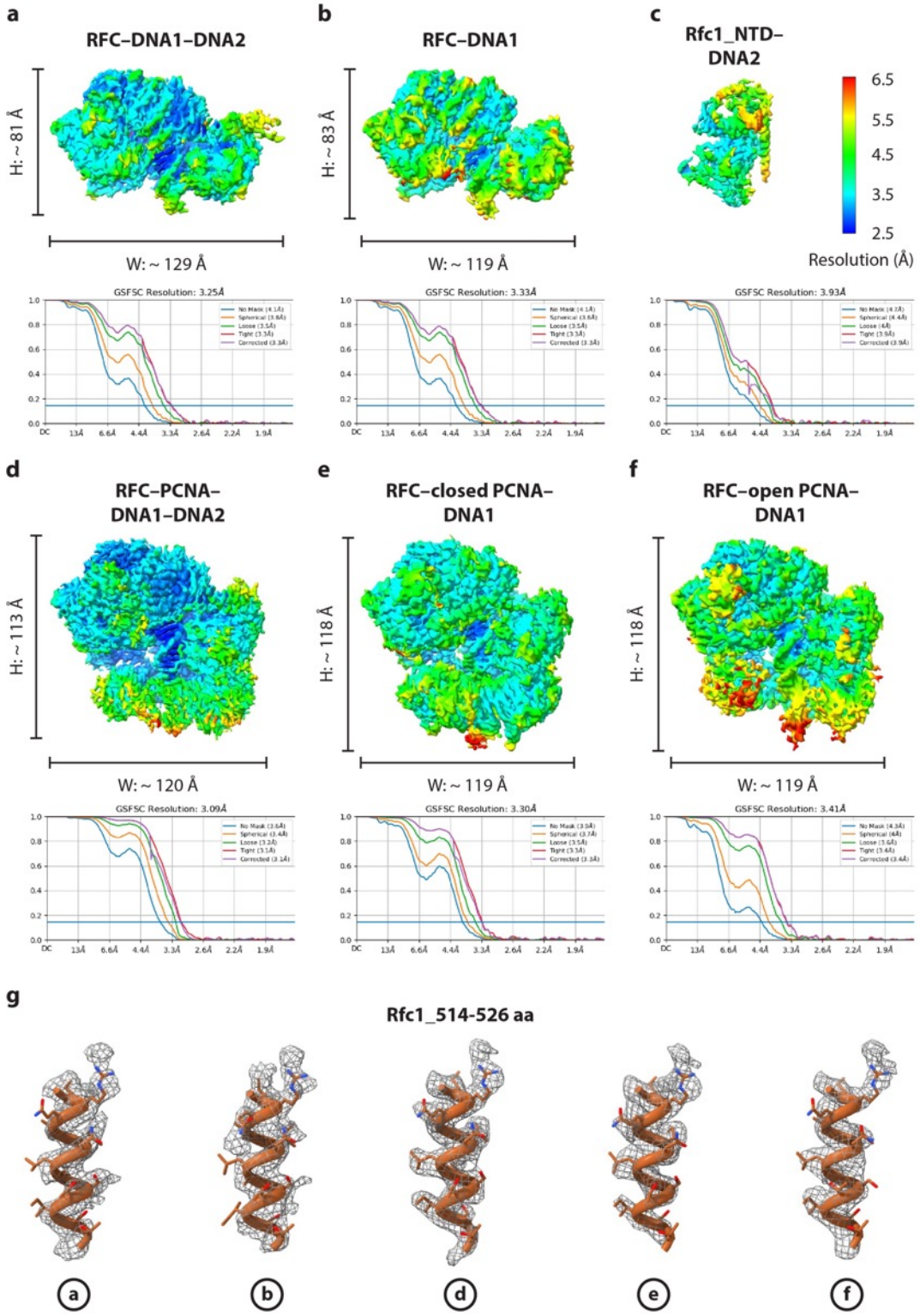


Supplementary Figure 1. Cryo-EM of in vitro assembled *S. cerevisiae* RFC–PCNA–DNA complexes. a) A typical raw micrograph of the mixture of RFC with PCNA and DNA. **b**) Selected 2D class averages with different views. Upper panel shows 2D averages of particles derived from mixture of RFC and DNA. Lower panel shows 2D averages of particles derived from mixture of RFC, DNA, and PCNA.



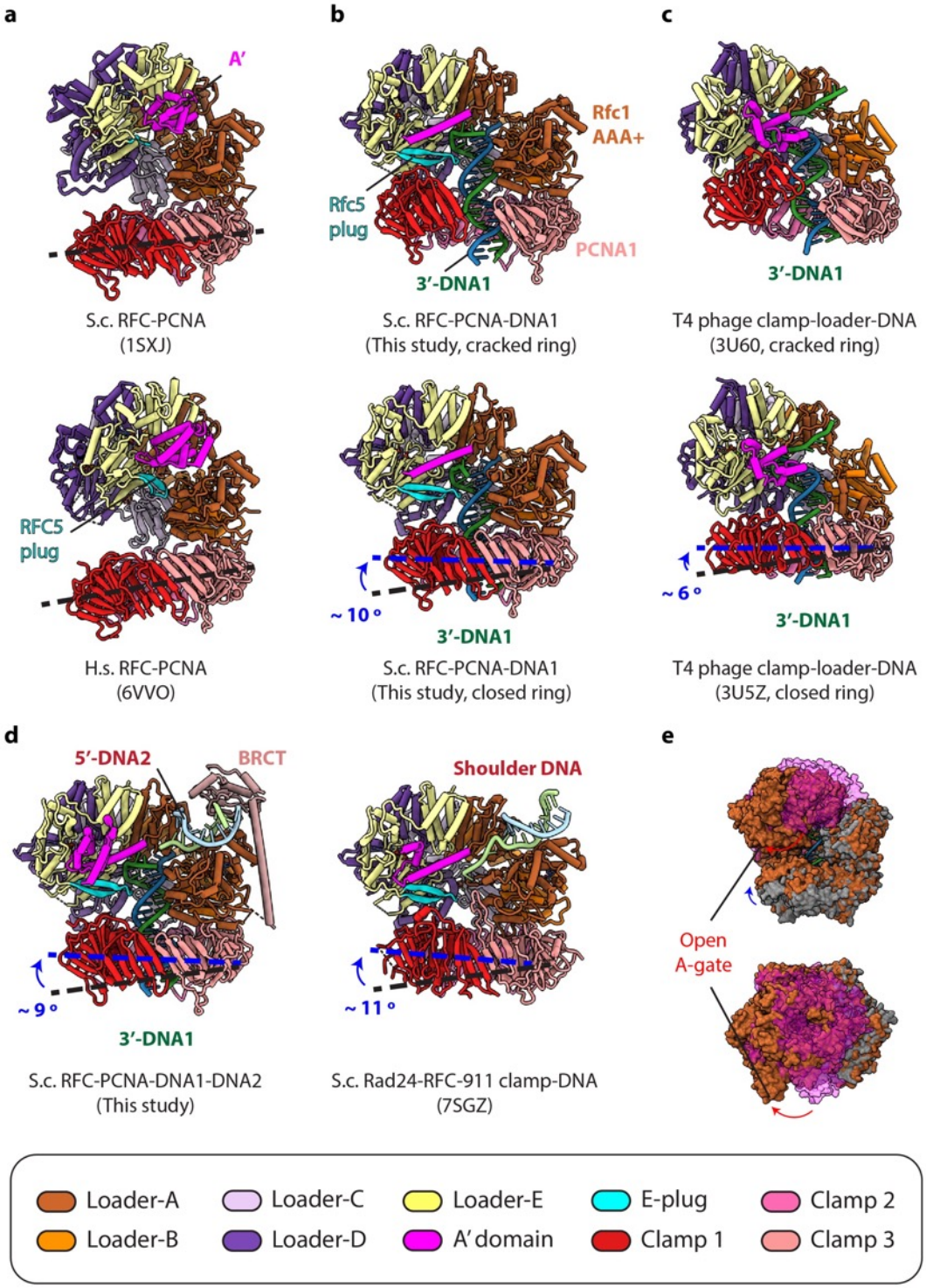
Supplementary Figure 2. Cryo-EM image processing workflow. CryoSPARC Live v3.2 was used to monitor the data collection and to process data in real time. We first used blob particle picking, then the machine learning-based Topaz to pick out additional particles. 3D variability analysis (3DVA) was

performed to segregate different structures, resulting in the five final 3D maps. Focused refinement on Rfc1_BRCT-AAA+ module−5′-DNA2 region improved the EM density of BRCT domain, which was combined with the RFC−3′-DNA1−5′-DNA2−PCNA map in Chimera to obtain the composite map.



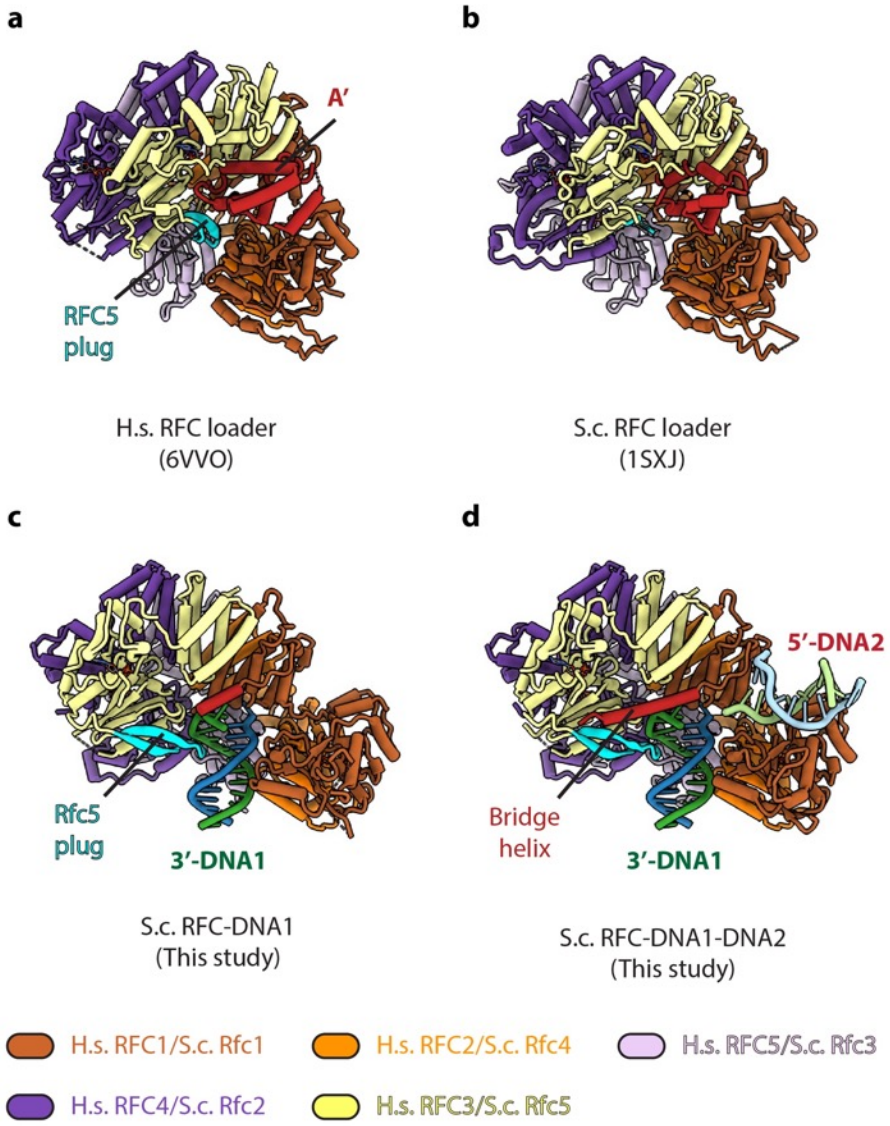
Supplementary Figure 3. Resolution estimation of the six 3D EM maps of the RFC loading complexes. a-f) Color-coded local resolution maps of the six 3D EM maps are shown on top, with the dimensions for each structure as labeled. The lower panels show the gold standard Fourier shell correlation (GSFSC) of the corresponding maps. **g)** Fitting of the structural model (cartoon with side

chains in sticks) with the EM density (grey meshes) is shown (as an example) for Rfc1 α-helix (aa 514-526) in each of the five full 3D maps.

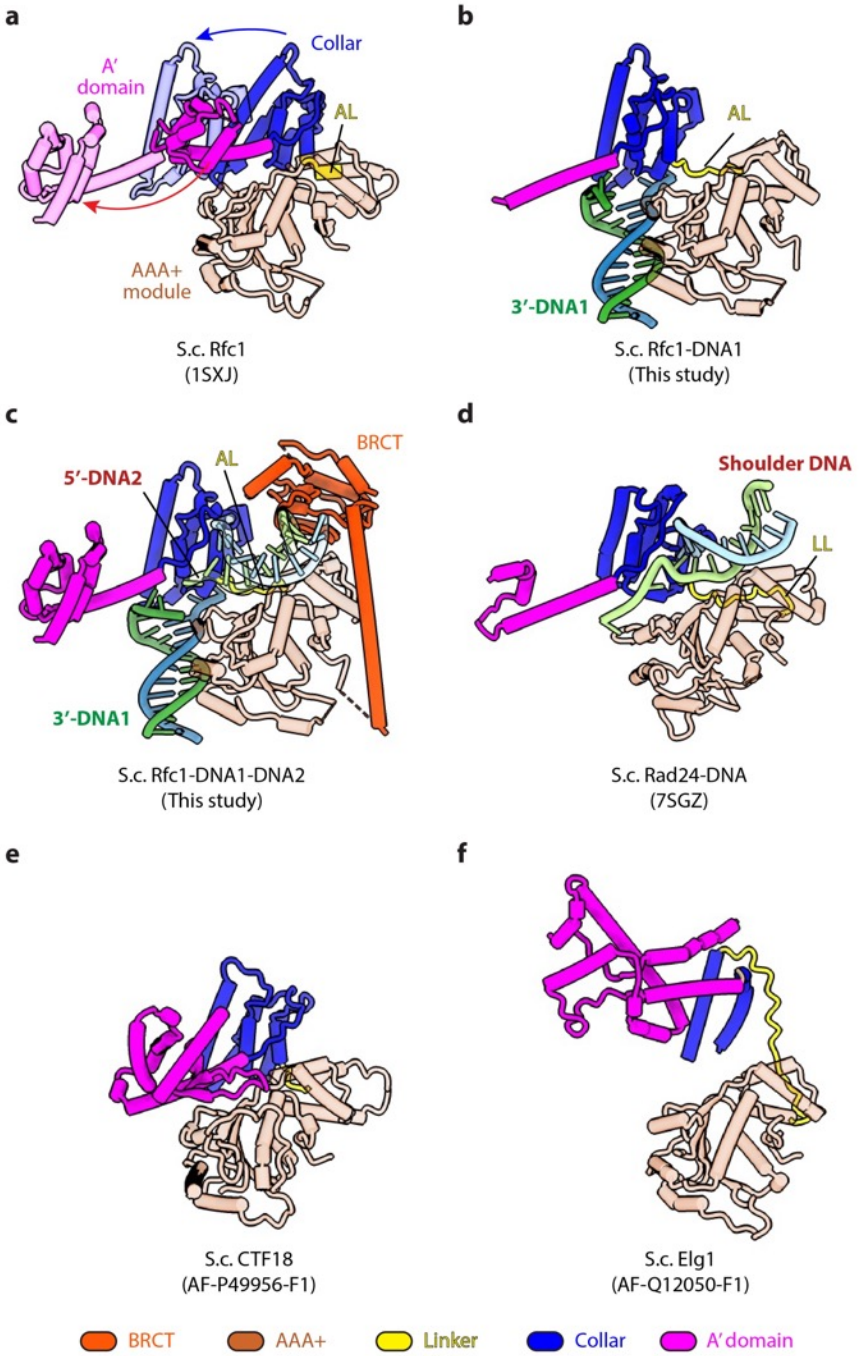


Supplementary Figure 4. Comparison of the PCNA poses in various RFC−(DNA)−PCNA structures. a) PCNA is tilted away from RFC in the yeast (upper) and human (lower) RFC−PCNA structures in the absence of DNA. **b-c)** The PCNA clamp is either cracked open (upper) or closed (lower) when bound to the RFC−3′-DNA1 (b), resembling the two T4 phage clamp conformations (c). **d)** The 5′-DNA2 binding site in Rfc1 (left) resembles the DNA binding site in Rad24 of the 911 clamp loader Rad24-RFC (right). **e)** Superimposition of RFC−PCNA (grey surface) with RFC−3′-DNA1−Closed ring PCNA (brown surface) in a front side (upper) and top view (lower). Rfc2, 3, 5 along

with the entire collar ring that will move upon DNA binding are shown in magenta. The relative movement that opens the A-gate is indicated by curved arrows.

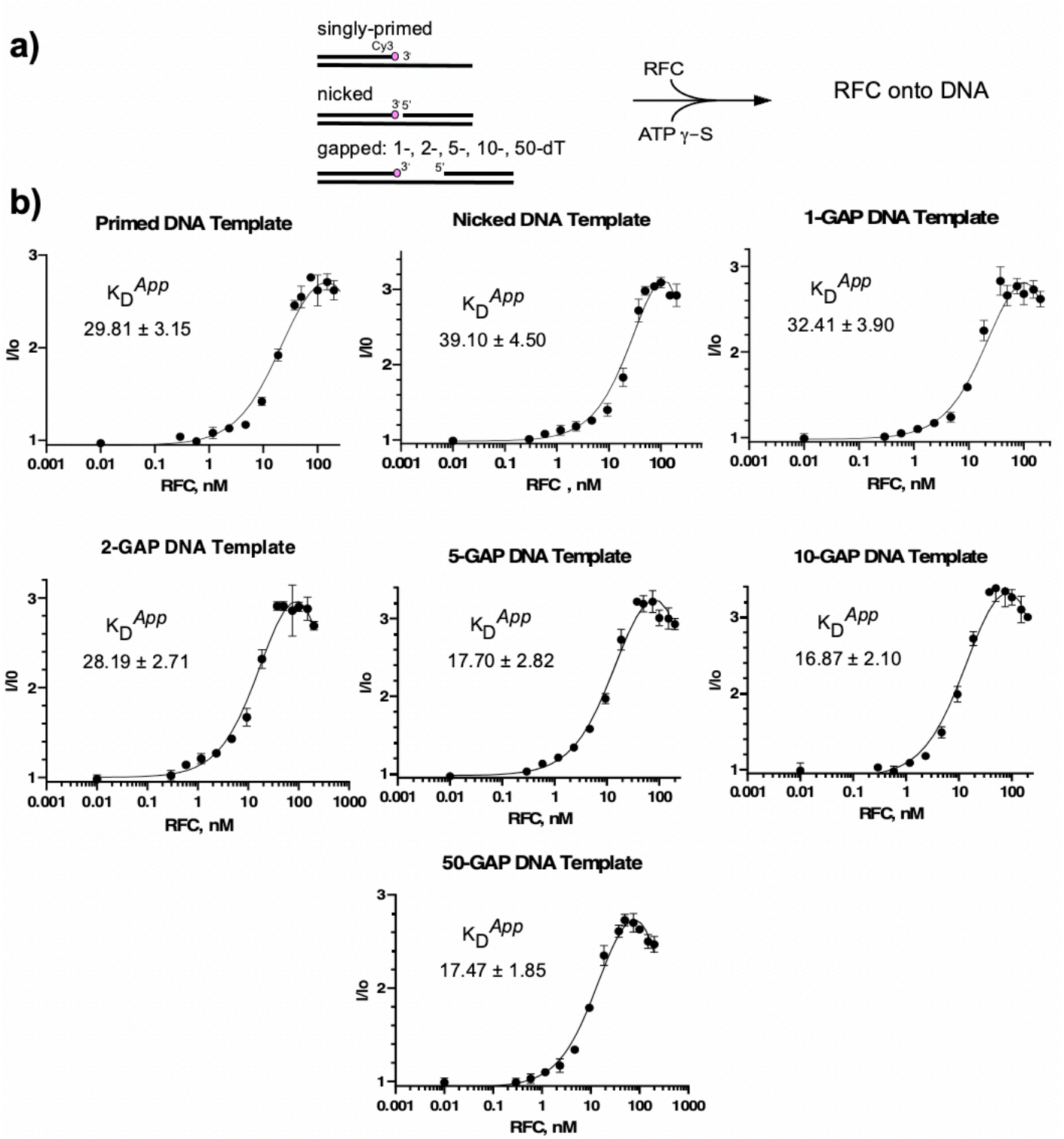


Supplementary Figure 5. Structural comparison of the RFC loaders in different functional states. a-b) The reported human RFC (a) and yeast RFC (b) structures in the absence of DNA. c-d) The structure of RFC bound to 3′-DNA1 (c) or to 3′-DNA1 and 5′-DNA2 (d). The RFC structures without DNA are extracted from published RFC−PCNA structures. 5′-DNA2 binding does not markedly change the RFC structure.



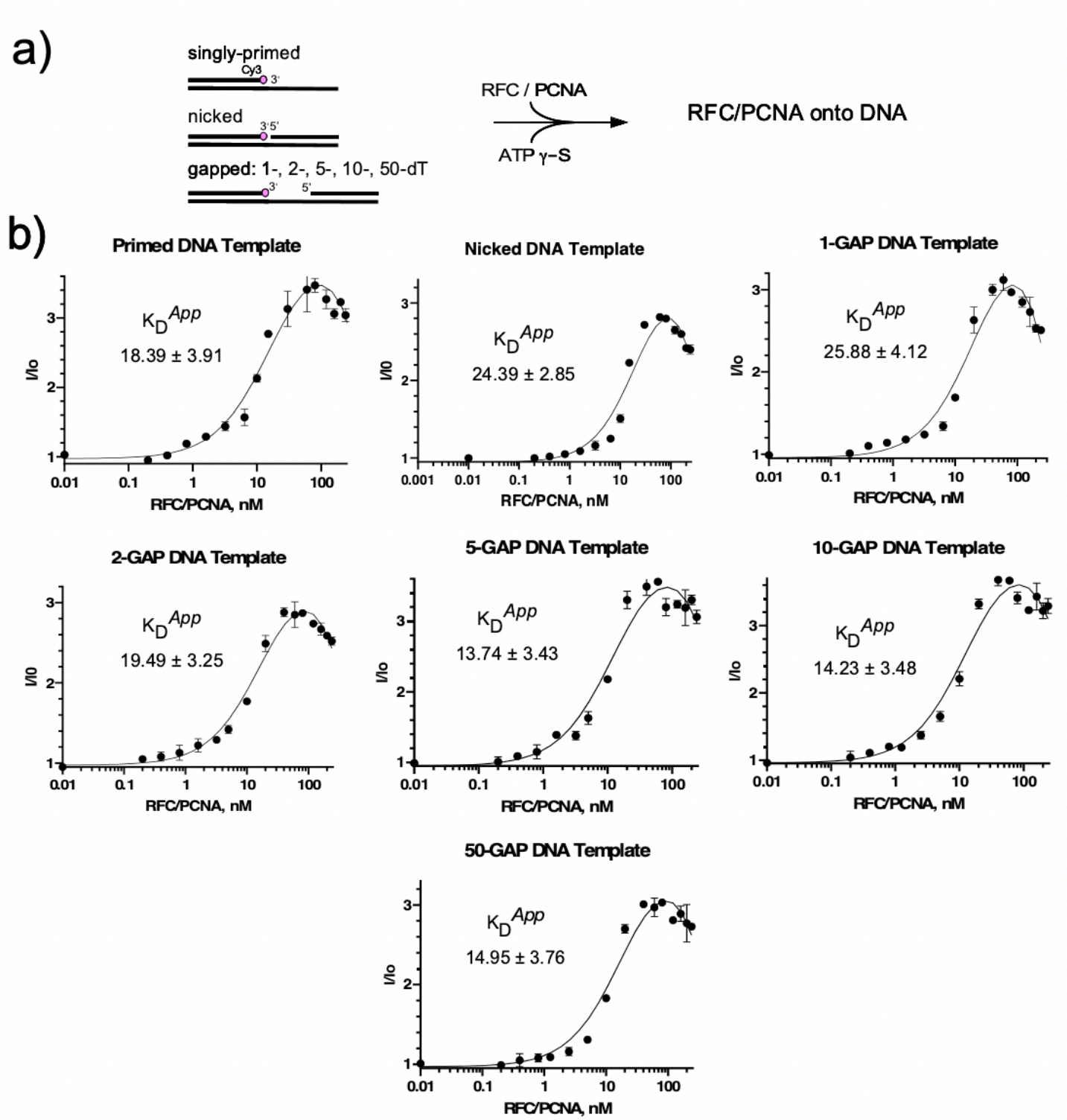
Supplementary Figure 6. 3′-DNA1 binding induces major conformation changes in Rfc1 to form the 5′-DNA2-binding groove. a) Rfc1 from the S.c. RFC–PCNA crystal structure in the absence of DNA (PDB ID 1SXJ) is superimposed to that from the RFC–3′-DNA1–5′-DNA2–PCNA quaternary complex, revealing large scale leftward movements of the two domains (indicated by the two curved arrows), similar with that described in **Fig. 2e**. **b)** 3′-DNA1 binding is sufficient to induce the formation of the 5′-DNA2 binding groove as observed in the RFC–3′-DNA1 structure even without the binding of PCNA. The same 5′-DNA2-binding groove is formed in the RFC–3′-DNA1–PCNA (closed ring) and RFC–3′-DNA1–PCNA (cracked open ring) structure. **c)** Rfc1 binding to 3′-DNA1 and 5′-DNA2 simultaneously in RFC–3′-DNA1–5′-DNA2–PCNA complex. The Rfc1 BRCT domain is in dark orange. **d)** Rad24-RFC binds 5′ DNA (PDB ID 7SGZ) at a site equivalent to the Rfc1 5′-DNA2 binding site. The Rad24 long linker LL is equivalent to the Rfc1 alternative linker AL. Panels containing superimpositions are aligned by the AAA+ module of Rfc1 or Rad24. **e-f)** AlphaFold predicts Ctf18 (e) and Elg1 (f)

structures appear to be compatible with 5′-DNA binding at the shoulder. Some loop regions (with low predicting confidence) by AlphaFold are omitted for clarity.



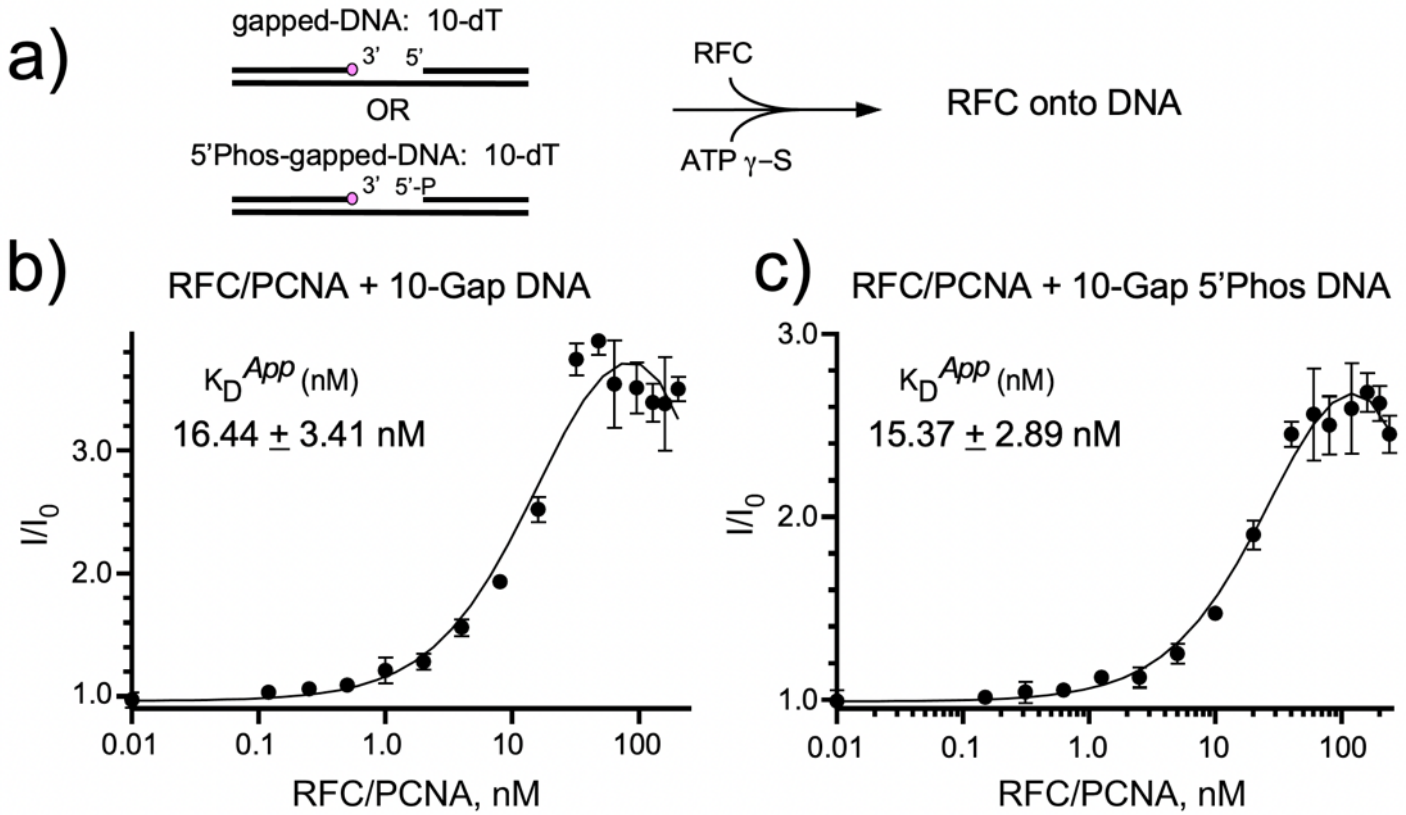
Supplementary Fig. 7. Binding of RFC to gapped DNA. a) Reaction scheme as described in Materials and Methods. **b)** The plots represent fluorescence titrations of RFC (0.01-200 nM) with different DNA templates (10 nM), as noted above each plot. Fluorescence intensity was normalized (I/I_0) to compare each fluorescent DNA template in the presence of DNA to that in the absence of RFC (see Methods). Experiments were performed in triplicate and the mean values (filled circles), as well as

the standard deviation (error bars) are shown. Data was analyzed using GraphPad Prism software and the apparent K_d values are shown for each graph.

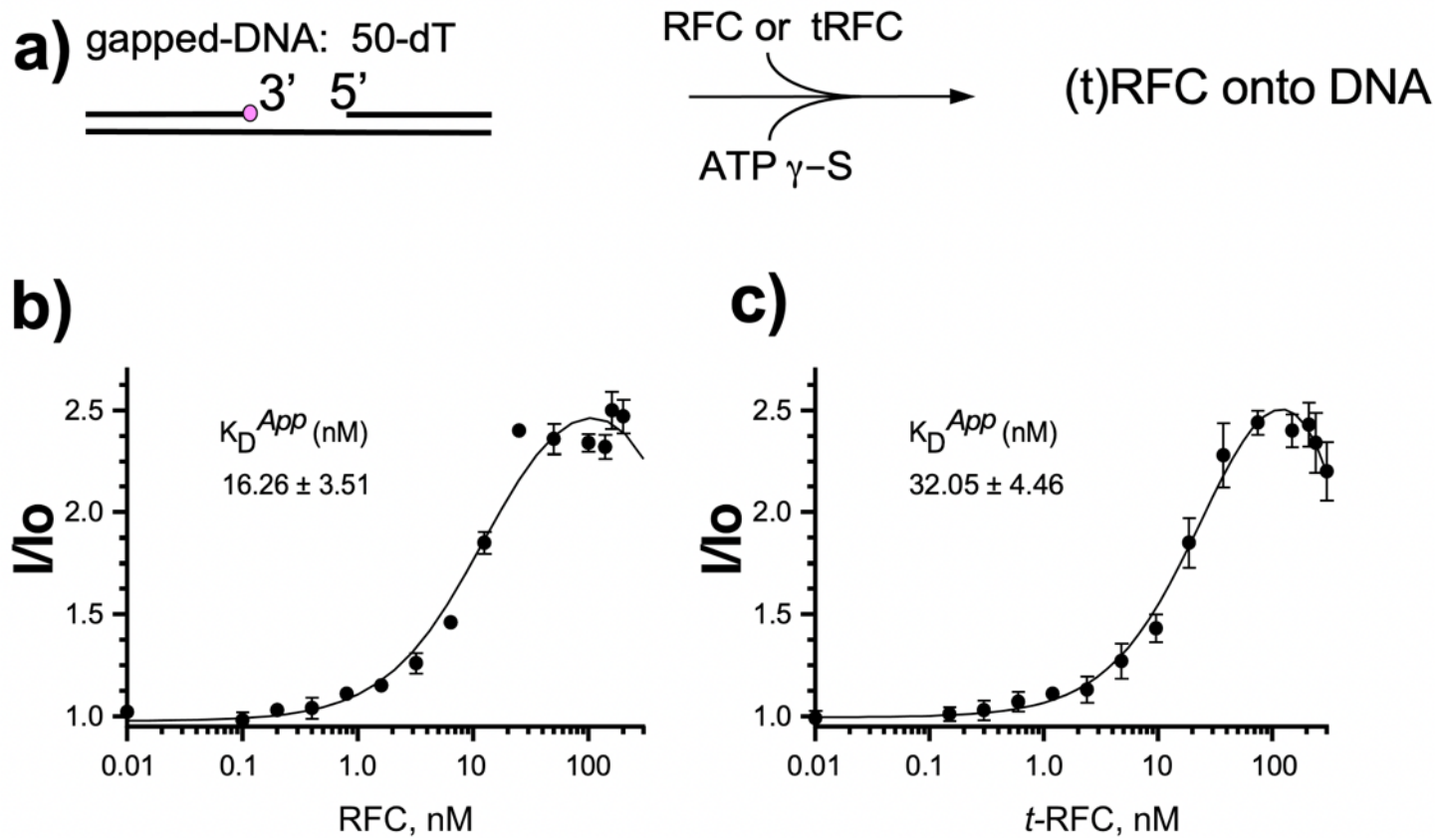


Supplementary Fig. 8. Binding of RFC/PCNA to gapped DNA. a) Reaction scheme as described in Methods. **b)** The plots represent fluorescence titrations of RFC/PCNA at equimolar ratio (between 0.01-200 nM) with different DNA templates (10 nM), as noted above each plot. Fluorescence intensity was normalized (I/I_0) to compare each fluorescent DNA template in the presence of DNA to that in the

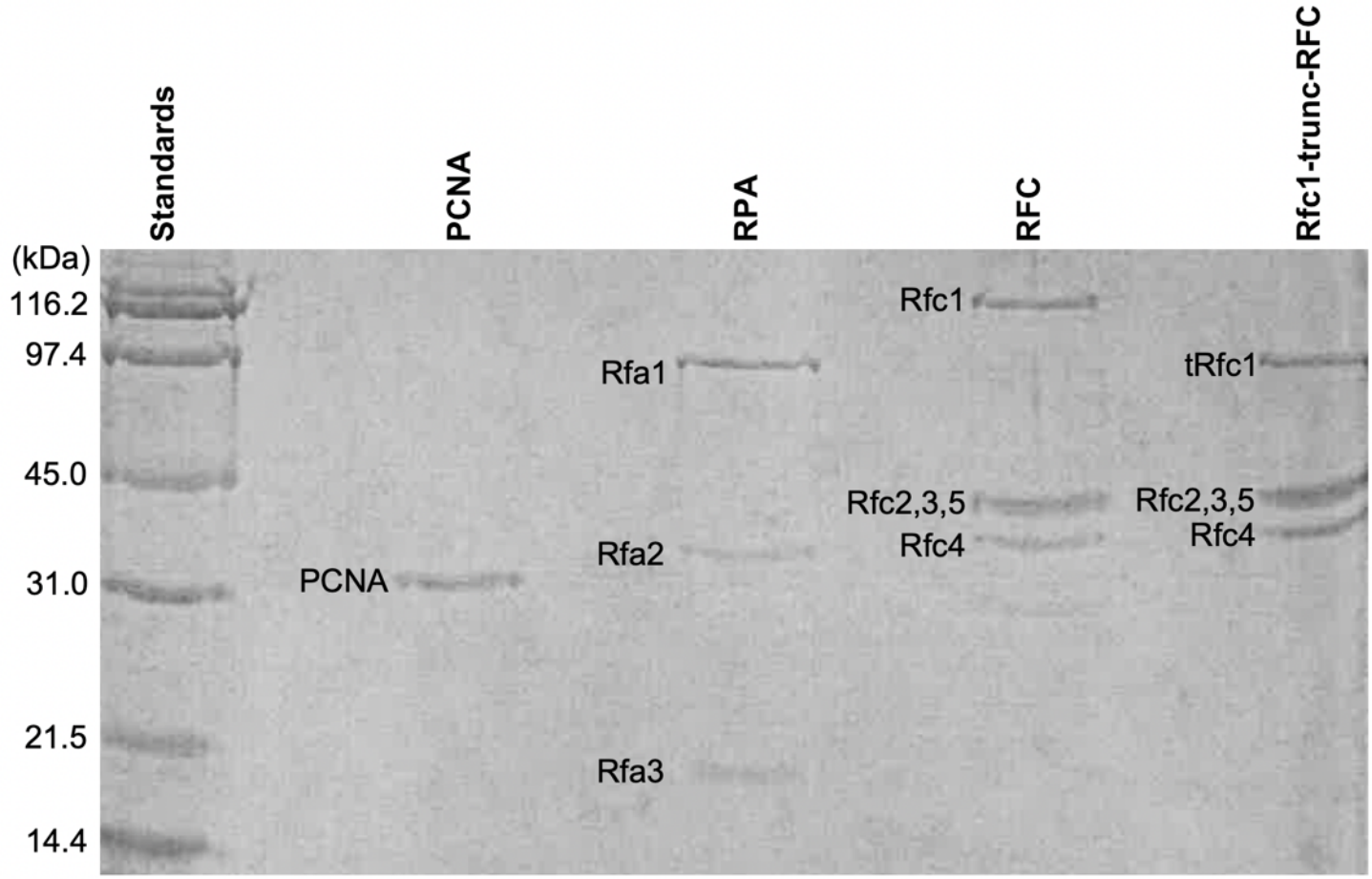
absence of RFC/PCNA (see Methods). Experiments were performed in triplicate and the mean values (filled circles), as well as the standard deviation (error bars) are shown. Data was analyzed using GraphPad Prism software and the apparent K_d values are shown for each graph.



Supplementary Figure 9. A 5' P at a 10-nt gap has no effect on RFC or RFC–PCNA binding. The experiments were performed as described in Materials and Methods. **a)** Scheme to examine the effect on RFC binding to DNA containing a 10mer gap having either a 5' OH or 5' P. **b)** Results of a titration of RFC–PCNA binding to DNA having a 5′ OH at the gap. **c)** Results of a similar titration but having 5′ phosphate group at the gap. Experiments were performed in triplicate and the mean values (filled circles), as well as the standard deviation (error bars) are shown. Data was analyzed using GraphPad Prism software and the apparent K_d values are shown for each graph.



Supplementary Figure 10. RFC binds a gapped DNA tighter than a mutant RFC lacking the 5′-DNA domain in Rfc1. **a)** Reaction scheme as described in Material and Methods. **b)** The plots represent fluorescence titration reactions (0.01-200 nM) of RFC, or **c)** truncated RFC (t-RFC) lacking the N-terminal 275 residues in Rfc1 of the RFC pentamer. Fluorescence intensity was normalized (I/I_0) to compare each fluorescent DNA template in the presence of DNA to that in the absence of RFC or t-RFC. Experiments were performed in triplicate and the mean values (filled circles), as well as the standard deviation (error bars) are shown. Data was analyzed using GraphPad Prism software and the apparent K_d values are shown for each graph.



Supplemental Figure 11. SDS-PAGE of proteins used in this report. The mw markers are to the left, and each protein factor is indicated above its corresponding lane.

LEGENDS OF SUPPLEMENTARY VIDEOS

Supplementary Video 1. A side-by-side comparison of the five structures obtained in this study. The five atomic models are rotated around the Y axis by 360º followed by another 360º rotation around the X axis. These models are first aligned to the RFC–3′-DNA1–5′-DNA2–PCNA model (lower right) and shown separately as cartoons in the same color scheme as in **Fig.3a**.

Supplementary Video 2. Morph of Rfc1 from in the RFC–PCNA structure to in the structure of RFC– 3′-DNA1–PCNA. The two models are aligned and shown first side by side in cartoon view. Then, the RFC–PCNA is changed to pale green, and RFC–PCNA–3′-DNA1 to rosy brown with DNA1 in cyan. Next, the two structures are superimposed and the PCNA clamp and DNA are removed for clarity. The remaining structures rotate around the X axis by 35º for better visualization of the 5′-DNA2 binding groove. The subunits Rfc3, 2 and 5, along with the entire collar region, move towards lower left to open the gate between the AAA+ module and the A' domain of Rfc1 allowing DNA1 entry into the RFC chamber (step 1). Concomitantly, a large groove is introduced between Rfc1 collar domain and AAA+ module to accommodate 5′-DNA2 at its shoulder (step 2). The Rfc1 BRCT domain participating in the shoulder 5′-DNA2 binding is also shown.