articles

Structural analysis of a eukaryotic sliding DNA clamp–clamp loader complex

Gregory D. Bowman[1,2], **Mike O'Donnell**[3] & **John Kuriyan**[1,2]

[1]*Howard Hughes Medical Institute, Department of Molecular and Cell Biology and Department of Chemistry, University of California, Berkeley, California 94720, USA*
[2]*Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*
[3]*Laboratory of DNA Replication, Howard Hughes Medical Institute, The Rockefeller University, New York, New York 10021, USA*

....

Sliding clamps are ring-shaped proteins that encircle DNA and confer high processivity on DNA polymerases. Here we report the crystal structure of the five-protein clamp loader complex (replication factor-C, RFC) of the yeast *Saccharomyces cerevisiae*, bound to the sliding clamp (proliferating cell nuclear antigen, PCNA). Tight interfacial coordination of the ATP analogue ATP-γS by RFC results in a spiral arrangement of the ATPase domains of the clamp loader above the PCNA ring. Placement of a model for primed DNA within the central hole of PCNA reveals a striking correspondence between the RFC spiral and the grooves of the DNA double helix. This model, in which the clamp loader complex locks onto primed DNA in a screw-cap-like arrangement, provides a simple explanation for the process by which the engagement of primer–template junctions by the RFC:PCNA complex results in ATP hydrolysis and release of the sliding clamp on DNA.

Sliding clamps are DNA-tracking platforms that are essential for processive DNA replication in all living organisms[1–5]. Sliding clamps enable DNA polymerases to cope with the considerable torque that results from the production of double-helical DNA, by allowing them to relax and regain their hold on DNA without losing their place at the replication fork. The efficient movement of the replication fork also relies critically on the rapid placement of sliding clamps at newly primed sites on the lagging DNA strand by ATP-dependent clamp loader complexes. This means that the piece-wise generation of Okazaki fragments can keep up with the continuous synthesis of DNA on the leading strand. In addition to their role in DNA replication, sliding clamps are also involved in several other processes that require a mobile contact on DNA[6].

The eukaryotic clamp loader assembly, RFC, consists of five subunits and forms a stable ATP-dependent complex with the eukaryotic sliding clamp, PCNA[7–10]. The RFC:PCNA complex binds specifically to primed DNA, and the recognition of the double-stranded/single-stranded junction stimulates ATP hydrolysis by the clamp loader. This results in the dissociation of RFC from the clamp, leaving PCNA encircling DNA and ready to receive DNA polymerase for the next round of DNA synthesis (Fig. 1a).

The general architecture of the pentameric clamp loader assembly has been revealed by the crystal structure of an inactive form of the *Escherichia coli* clamp loader complex[11], and the structure of one of the subunits of the archaebacterial clamp loader bound to ADP[12]. Each clamp loader subunit contains two structurally conserved domains that together comprise an ATPase module of the AAA+ family[13,14]. AAA+ ATPases are a diverse class of oligomeric (typically hexameric) proteins that couple ATP binding and hydrolysis to the modulation of protein–protein interactions[13]. Many AAA+ ATPases form ring-like assemblies, in which a characteristic feature is the presentation by each subunit of an arginine residue, termed the 'arginine finger', towards ATP bound at the next subunit[13,15,16].

In contrast to the symmetrical organization of the hexameric rings formed by AAA+ proteins such as N-ethylmaleimide-sensitive factor (NSF)[15,16], the structure of the inactive bacterial clamp loader complex is strikingly asymmetric[11], and it has been difficult to infer the molecular details of clamp loader action from this structure[17]. One important unresolved issue concerns the

mechanism by which ATP allows the specific recognition of primed DNA by the clamp loader. Also unknown is the nature of the 'ejection mechanism' by which binding to primed DNA results in the hydrolysis of ATP so that the complex is disassembled, the clamp is loaded onto DNA, and the clamp loader is released for another cycle of clamp loading[10,18].

We now describe the crystal structure of the *S. cerevisiae* RFC complex bound to the PCNA clamp. The structure reveals an unexpected ATP-dependent spiral arrangement of the ATPase domains of RFC that extends above the central channel of PCNA. The geometry of the spiral is such that the ATPase domains track the grooves of primed DNA modelled within the hole of the sliding clamp. The crystal structure suggests that DNA binding would stabilize a catalytically competent configuration of the interfacial ATP-binding sites, thereby providing a hypothesis for how the interaction with the double helix triggers the hydrolysis of ATP.

Structure determination

We have crystallized a variant form of the *S. cerevisiae* RFC complex in which the largest subunit (RFC-A in the notation introduced in Fig. 1) is truncated to a minimal construct (residues 295–785) that is fully functional in clamp loading (data not shown). Because ATP hydrolysis weakens the interaction between clamp and clamp loader, we replaced the arginine finger residues that are present in the highly conserved Ser-Arg-Cys (SRC) motifs of four RFC subunits (RFC-B, RFC-C, RFC-D and RFC-E) by glutamine, which resulted in a significant reduction in ATPase activity (Supplementary Fig. 1). Crystals of this mutated RFC:PCNA complex obtained in the presence of ATP-γS diffract X-rays to only ~5 Å resolution using synchrotron radiation. Substantial improvement in crystal quality was obtained by controlled dehydration of the crystals[19], allowing the structure to be determined to a resolution of 2.85 Å (Supplementary Table 1).

PCNA-bound RFC forms an ATP-dependent spiral

The hetero-pentameric RFC complex is seated on top of a closed PCNA ring, but tipped away from it (Fig. 1b). Domain I of each RFC subunit is a RecA-type ATPase domain, which is followed by a small helical domain (domain II) that is characteristic of AAA+ ATPases.

©2004 **Nature Publishing Group**

Domains I and II together form the AAA+ module, which is connected by a flexible linker to another helical domain (domain III) that is unique to clamp loaders (Fig. 1c, d). The helical third domains of all five RFC subunits pack together to form a stable cylindrical structure, referred to as the 'collar'[11]. The five AAA+ modules of RFC assemble into a right-handed spiral, which results in only three of the five RFC subunits (RFC-A, RFC-B and RFC-C) making contact with PCNA, leaving a wedge-shaped gap between RFC-E and PCNA (Fig. 2).

ATP-γS is bound to each of RFC-A, RFC-B, RFC-C and RFC-D (Supplementary Fig. 2). These nucleotides are key elements that hold the spiral assembly together by anchoring intersubunit interactions through hydrogen bonds to the phosphate groups (Fig. 3b). RFC-E lacks several residues that are conserved at the active sites of functional ATPases, but electron density maps reveal a strongly bound nucleotide (either ADP or ATP-γS) at the RFC-E binding site (Supplementary Fig. 2). This nucleotide packs closely against a unique carboxy-terminal domain (domain IV) of RFC-A, which is situated between the ATPase domains of RFC-A and RFC-E, and provides a physical link between the two ends of the RFC spiral.

PCNA has three conserved hydrophobic grooves for potential interaction with RFC, and two of these sites are engaged by RFC-A and RFC-C. RFC-B, located between RFC-A and RFC-C, makes limited and primarily polar interactions with the intersubunit region of the clamp, and may not contribute significantly to binding energy. All three RFC subunits interact with PCNA primarily through the C-terminal end of the clamp-interacting helix (α4; Fig. 3a) and the loop following it. The interface between RFC-A and PCNA resembles that of other proteins bound to sliding clamps[20,21], and involves a 3_{10} helical turn that follows helix α4 of RFC-A and inserts four hydrophobic residues (A-Val401, A-Val402, A-Tyr404 and A-Phe405) into the hydrophobic binding site on the clamp. Two hydrophobic residues of RFC-C (C-Ile106 and C-Phe107) interact with the corresponding site on the next subunit of the clamp.

The structure of PCNA in the RFC:PCNA complex is not distorted significantly with respect to the structure of free PCNA[5], and we conclude that the crystal structure represents a stable interaction of the RFC complex with a closed PCNA clamp. In the clamp loading cycle, a clamp loader would be bound to a closed clamp immediately before ATP hydrolysis and dissociation of the
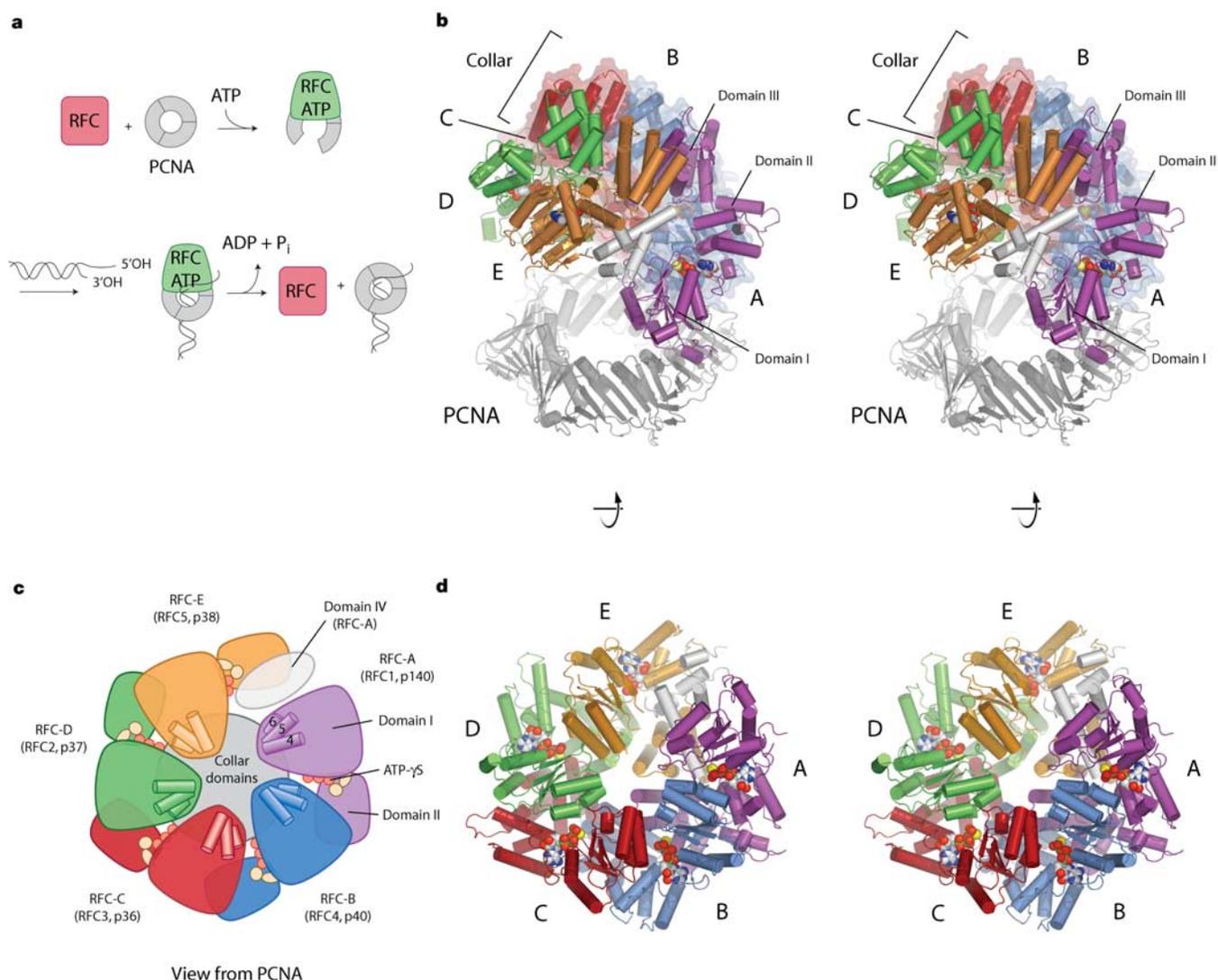

Figure 1 Overview of the RFC:PCNA complex. **a**, The clamp loading cycle. **b**, A stereoview of the RFC:PCNA complex, viewed from the side. **c**, **d**, Schematic (**c**) and backbone (**d**) representations of RFC rotated ~110° about the horizontal with respect to **b**. Viewed from the PCNA-interacting face of the complex, the five AAA+ modules are in the foreground, with the C-terminal collar in the rear. The five subunits of the RFC complex are referred to as RFC-A, RFC-B, RFC-C, RFC-D and RFC-E, respectively, moving in a right-handed sense around the assembly, with the thumb pointing towards the collar. The yeast and human nomenclature for each subunit is shown in parentheses.

complex (Fig. 1a). Consistent with this interpretation, we describe below the organization of the RFC AAA+ modules in what appears to be a catalytically competent state.

Active site configurations in the RFC spiral

The interfaces between AAA+ modules of the RFC complex are related to those seen previously in hexameric AAA+ ATPases such as NSF[15,16], one key difference being the tightness of the interfacial interactions at the ATP-binding sites. There are limited interfacial interactions with the nucleotide in the currently available high-resolution structures of other AAA+ ATPases. Notably, the generation of the spiral seen in the RFC complex has resulted in the distal and proximal faces of adjacent ATPase subunits being significantly closer together in RFC than in the structures of other AAA+ proteins (Supplementary Fig. 3). Some AAA+ proteins do form spirals in the crystal but, in contrast to RFC, the ATP-binding sites are open[22,23].

At the A:B and C:D interfaces, intersubunit contacts result in stabilization of active site residues of RFC-A and RFC-C in configurations that we believe are close to being optimal for catalysis. A striking feature of the A:B interface is the manner in which it locks down the relative orientation of domain II of RFC-A with respect to the nucleotide. For example, B-Ser156 of the SRC motif of RFC-B interacts with A-Asp513 and A-Arg515 in the highly conserved 'sensor 2' element (Fig. 3b). A-Arg515 is also stabilized by interactions with B-Glu132 of helix $\alpha5$ ('central helix') of RFC-B, and is positioned for optimal coordination of the β and γ phosphates of ATP.

The interactions around ATP-γS bound to RFC-A are reminiscent of the organization observed at catalytically competent sites in F_1ATPase, in which two of the three catalytic sites (referred to as β_{DP} and β_{TP}) exhibit tight coordination of the nucleotide[24,25]. The configurations of protein side chains around the nucleotide differ from each other only by fractions of an angstrom in these two sites in F_1ATPase, both of which are expected to be close to the transition-state configuration for ATP hydrolysis. The β_{DP} site of F_1ATPase has a closed interface that excludes bulk water, coordinates the attacking water molecule, and stabilizes the build-up of

negative charge on the transition state[24]. Despite significant differences between the overall architecture of the RFC complex and that of F_1ATPase, several key residues in the β_{DP} site of F_1ATPase seem to have counterparts in the RFC complex in terms of the coordination of ATP (Fig. 3c), suggesting that the ATP-binding site in RFC-A is structurally close to catalytic competency. The RFC-C binding site is similar to that of RFC-A, but an alternating pattern in intersubunit interactions (see below) results in looser ATP coordination at the B and D sites. It is not clear whether this alternating pattern has functional significance.

The RFC spiral matches double-helical DNA

RFC operates on DNA–DNA primers with recessed 3′ ends, formed by the extension of RNA primers by DNA polymerase-α[26]. The RFC complex binds to primer–template junctions with nanomolar affinity, but dissociates rapidly because of ATP hydrolysis[10]. The right-handed spiral arrangement of the five AAA+ domains of RFC displays roughly the same pitch as that of double-stranded B-form DNA. The AAA+ module of RFC-A is related to that of RFC-B by a 61° rotation and 5.5 Å translation. This closely matches the 5.6 Å rise per 60° rotation of double-helical B-form DNA[27]. The RFC spiral is somewhat asymmetric, and the rotations and translations for the B → C, C → D and D → E transformations are 74° and 7.7 Å (that is, 6.2 Å per 60° turn), 60° and 3.5 Å, and 72° and 7.3 Å (6.1 Å per 60° turn), respectively. In each case the axes of rotation are roughly aligned, passing through the centre of the PCNA ring with an angular spread of ~12°.

The non-uniform transformations relating one RFC subunit to the next include compensatory rotations and translations that end up preserving to a remarkable degree the general match with the geometry of duplex DNA over all five RFC subunits. Thus, the transformation that relates RFC-A to RFC-E is a screw rotation of 267° (instead of 240° if a regular 60° rotation had been maintained at each step), and the net translation between RFC-A and RFC-E of 24.4 Å along the screw axis corresponds to a rise of 5.5 Å per 60° turn, matching the helical parameters of B-form DNA quite precisely.

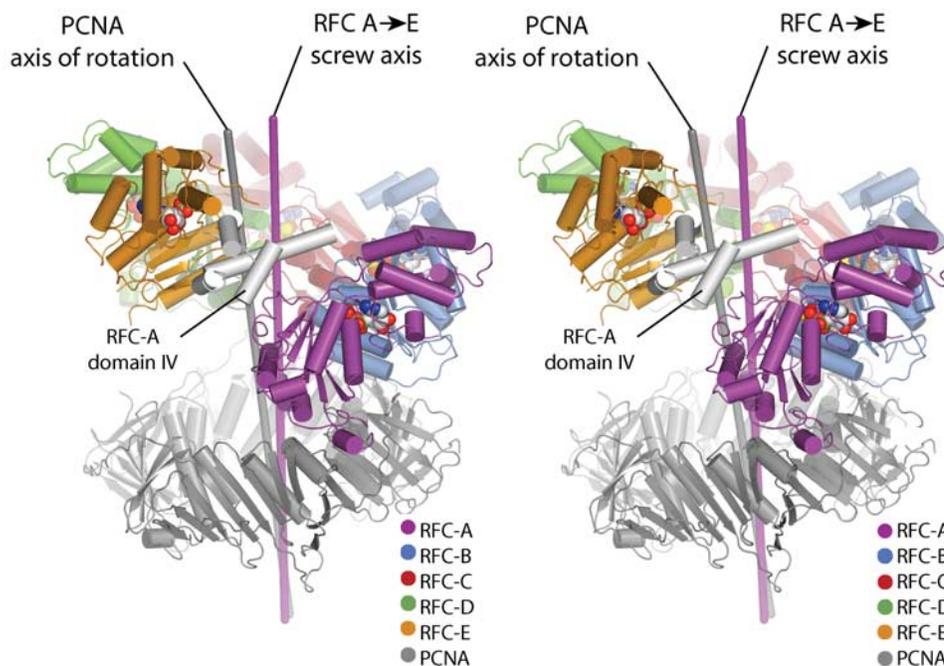The correspondence between the RFC spiral and DNA allows us

Figure 2 The AAA+ domains of the RFC assembly form a right-handed spiral. The screw axis relating RFC-A to RFC-E (magenta rod) is vertical and is tilted by ~9° with respect to the central rotational axis of symmetry of the PCNA ring (grey rod). The helical collar of RFC has been removed for clarity.

to generate a model for an RFC:DNA complex based purely on geometrical considerations. The axis of an ideal B-form DNA duplex was made coincident with the screw axis that relates RFC-A to RFC-E. The sequence-independent recognition of DNA is expected to occur in the minor groove, and so we rotated the DNA about its axis so that RFC-A was positioned alongside the minor groove. Each of the subunits of the RFC complex in this model is seen to track the minor groove of the double helix (Fig. 4a), with a fit that is tight but reasonable. This means that the architecture of the RFC:PCNA complex matches that of DNA not just in terms of the screw rotational operators, but also in terms of preserving the size of the internal chamber, so that it fits tightly around the double helix while allowing the preceding segment of DNA to be positioned comfortably within the PCNA ring (Fig. 4b).

The recognition of RNA–DNA hybrid structures is critical for the function of other clamp loader complexes, such as the *E. coli* γ-complex. Notably, the helical pitch of A-form RNA–DNA duplexes (5.5 Å per 60° turn) closely matches that of B-form DNA–DNA duplexes (5.6 Å per 60° turn)[28], which means that the distinction between the type of primer structure recognized by the clamp loader is not a crucial aspect of the geometrical analysis presented here.

Proposed DNA interaction

The amino-terminal regions of three α-helices of domain I of each RFC subunit (α4, α5, α6; the distal helices, see Fig. 3a), as well as the loops leading into them, are positioned so that they interact with the backbone of the modelled double helix (Fig. 4). The helix dipole effect from the central (α5) and clamp-interacting (α4) helices of each subunit generates a track of positive electrostatic potential within the inner surface of the spiral. In addition, each RFC subunit presents several lysine and arginine residues in a manner that suggests potential interaction with the DNA double helix (Fig. 5). These residues are conserved in all five RFC subunits so that, like the helix dipoles, they track the minor groove of the double helix.

The α-helices and the residues that are potentially important for DNA interaction are conserved in clamp loader subunits from eukaryotes, prokaryotes and archaea (Fig. 5). For example, A-Lys462 in RFC-A is in the loop preceding the SRC helix (α6), and it points directly towards the phosphate backbone of modelled DNA. This residue is either lysine or arginine in all eukaryotic and archaeal RFC sequences, as well as in the sequences of the γ-subunits of bacterial clamp loaders. Likewise, A-Arg383 in RFC-A is invariant in all eukaryotic and archaeal RFC sequences, and is conserved as Arg, Lys or His in 55% of bacterial sequences (bacterial clamp

loaders have conserved basic residues at nearby positions that are not conserved in the RFC sequences). A third highly conserved position lies in the loop preceding the central helix (α5). A-Arg434 is conserved as arginine or lysine in all eukaryotic and about half of archaeal sequences of RFC-A subunits. In subunits other than RFC-A, the residue at this position (for example, B-Thr120 in RFC-B) is either serine or threonine, residues that are commonly found at protein–DNA interfaces[29].

The strong conservation of these residues in the bacterial and archaeal clamp loaders suggests that the DNA-recognition model proposed here may be applicable to all clamp loader complexes. Indeed, comparison of the inactive form of the *E. coli* clamp loader complex[11,30] with the RFC structure shows that the former can be described as a distorted and open form of the RFC spiral, which presumably closes into an RFC-like structure upon ATP and clamp binding (Fig. 6a).

The three distal helices (α4, α5, α6) of the RFC subunits correspond to the E, F and G helices of the canonical RecA fold[31]. In both RecA and the DnaB-type helicase from T7 bacteriophage, the two loops preceding the F and G helices (denoted L1 and L2, see Fig. 5) have been shown by mutagenesis to be involved in DNA recognition[32–35]. In the Rho termination factor, a hexameric RNA helicase, two loops that are implicated in binding to single-stranded RNA within the central region of the hexamer correspond to the L1 and L2 loops of RecA[36]. RecA forms spiral filaments with a significantly larger pitch than the RFC spiral (~14 Å per 60° turn, compared with ~5.6 Å in RFC)[31]. DnaB-type helicases and Rho form closed rings, and also spiral assemblies (but with open ATP-binding sites)[36–38]. Despite these differences in gross subunit organization, it appears that the DNA-recognition mode that we propose for RFC is a variation on a theme that is common to RecA-type ATPases that interact with DNA or RNA.

A mechanism for primer–template recognition

The efficiency of lagging strand synthesis is increased by the ability of clamp loaders to position sliding clamps such that they are oriented appropriately with respect to the 3′ end of the primer strand[18]. Footprinting experiments using human RFC have demonstrated that RFC recognizes DNA structures with recessed 3′ ends specifically, in a sequence-independent manner, and interacts with both double- and single-stranded DNA across the primer–template junction[39]. Our model, which takes primed DNA through the hole in PCNA and into the RFC spiral, fits well with the footprinting data and provides a simple mechanism for distinguishing between
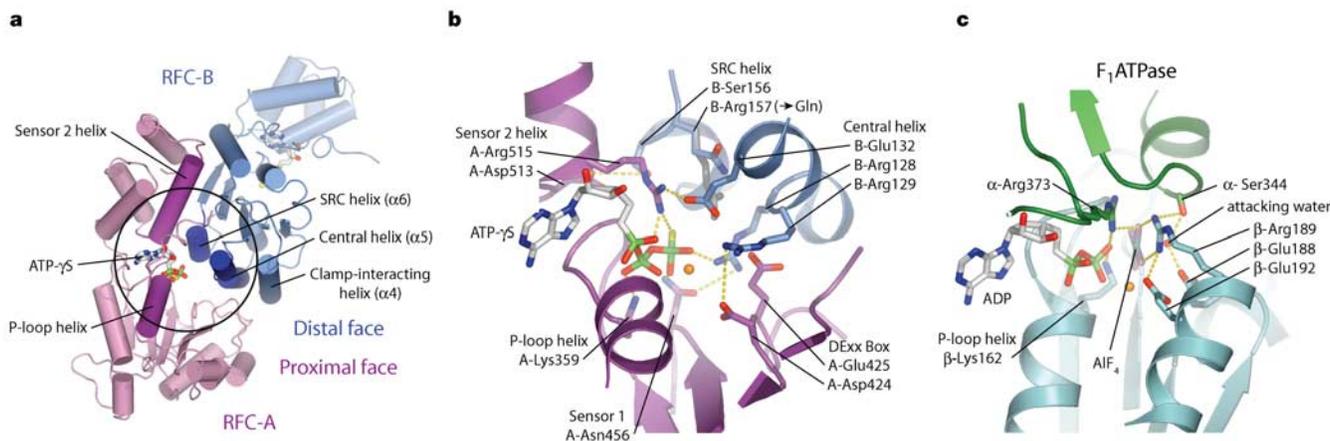

Figure 3 Nucleotide is tightly bound between proximal and distal faces of the RFC-A and RFC-B AAA+ modules. **a**, The RFC-A:RFC-B interface. **b**, Key ATP-interacting residues at the active site of RFC-A. Located near the terminal phosphate group of ATP-γS is the arginine finger residue of the SRC motif of RFC-B (B-Arg157). This position has been mutated to glutamine in this structure, and one possible conformation for an arginine side chain at this position is shown in grey. **c**, A view of the β$_{DP}$ site in F$_1$ATPase[24] (accession code 1H8E).

correctly primed DNA (that is, with a recessed 3′ end) and other forms of DNA. The RFC complex, like a screw-cap, threads onto the last turn of the double helix, further extension of which is blocked by the physical barrier imposed by the C-terminal collar. With the sequence-independent binding of single-stranded DNA extending from the primer–template junction[39], the specificity of RFC for double-stranded/single-stranded junctions is explained by the need to terminate the double helix within the RFC spiral.

If we position RFC-A alongside the minor groove of the ideal DNA duplex (as shown in Fig. 4a), then the strand that runs in the 5′ to 3′ direction along the RFC spiral (the primer strand) ends up running into the interior wall of the RFC complex, near RFC-E. In contrast, the strand that runs in a 3′ to 5′ direction (the template strand) ends up facing the gap between RFC-E and RFC-A, where there appears to be sufficient room for a 5′ extension to snake out of

the interior of the complex (Fig. 4b). DNA polymerases interact with the same face of the sliding clamp as do the clamp loaders[20], and so the clamp is now positioned correctly for DNA synthesis upon departure of RFC.

To illustrate this directional specificity in the recognition of primed DNA, we use the structure of an actual primer–template junction, as seen in the crystal structure of HIV reverse transcriptase trapped to primed DNA[40] (Fig. 6b). The ~8 base pairs of the double helix preceding the 3′-primer terminus, which is bound to the active site of reverse transcriptase, are A-form in this structure and ~13 base pairs of the duplex that are further upstream are in a conformation resembling B-form. As expected, the junction between the B-form and the A-form DNA involves a bend (~40°) in the direction of the helix axis. If we align the A-form portion of the DNA along the axis of the RFC spiral, the preceding B-form
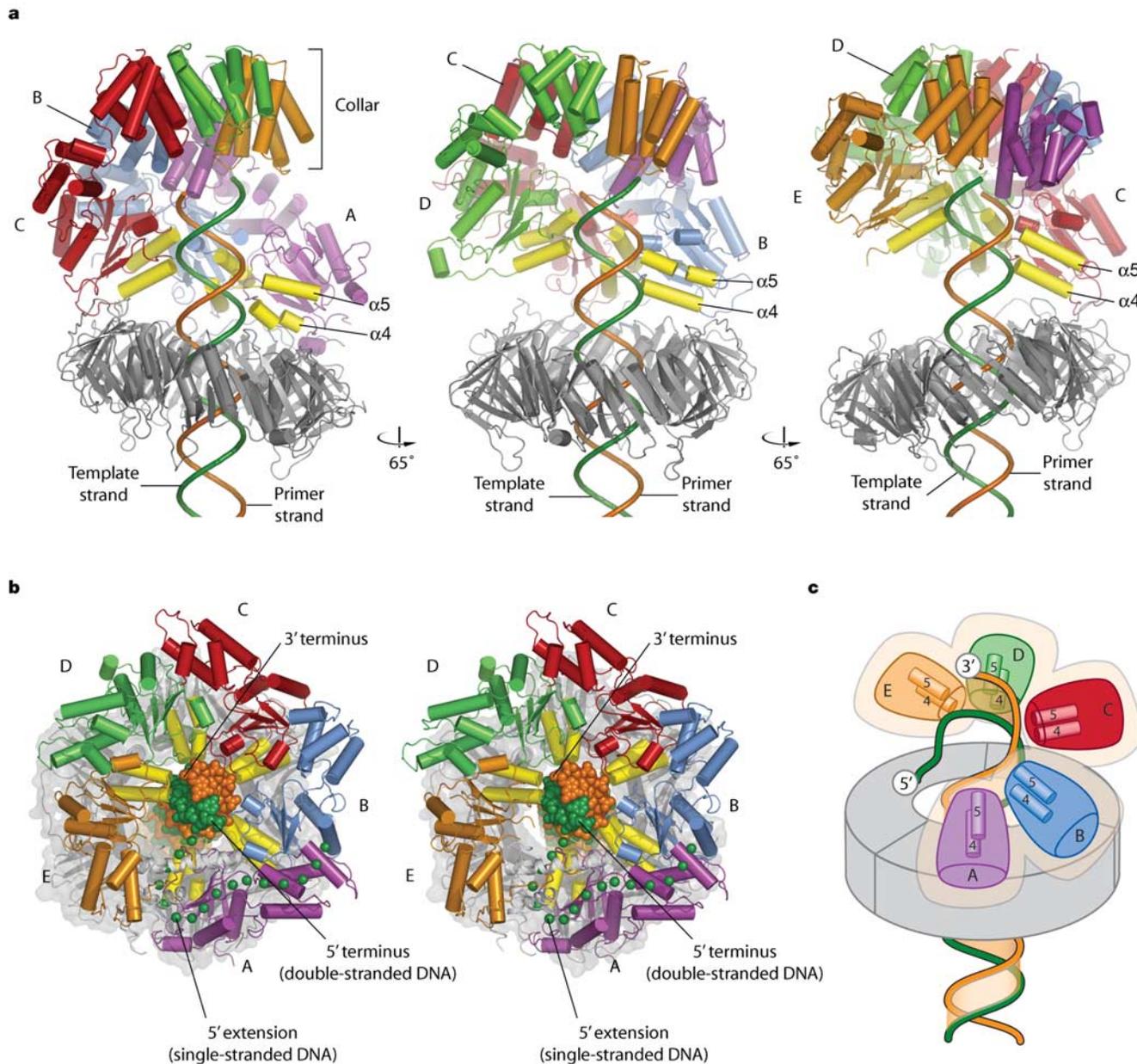

Figure 4 A model for primed DNA interacting with the RFC:PCNA complex. **a**, Three successive 'cut-away' views of the RFC:PCNA complex, showing the tracking of the DNA duplex by the N-terminal ends of the clamp-interacting (α4) and central (α5) helices of each subunit (coloured yellow). In each view, the AAA+ modules in the foreground are removed for clarity. **b**, Stereoview of the DNA:RFC:PCNA model viewed from the collar of

RFC, with the collar removed. A potential exit path for the 5′ end of the template strand is indicated by green spheres. **c**, Schematic representation of the DNA:RFC:PCNA model. Alignment of RFC-A with the minor groove of the double helix positions the 5′ terminus of the template strand near the opening between RFC-E and RFC-A.
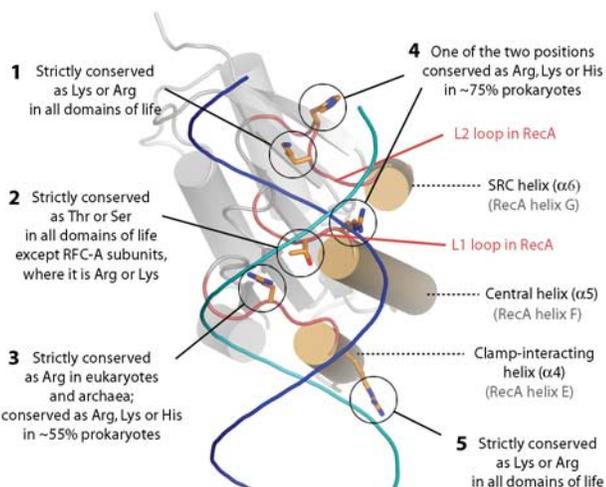
Figure 5 Conserved residues in domain I of clamp loader subunits at the proposed DNA-interacting surface. The loops preceding the central (α5) and SRC (α6) helices are equivalent to the L1 and L2 loops of RecA[31], which have been implicated in DNA binding through extensive mutagenesis of RecA and DnaB-type helicases[32–35]. Conserved positions are indicated on the basis of sequence analysis of 24 eukaryotic (16 small (for example, RFC-B) and 8 large (for example, RFC-A)), 35 archaeal (18 small and 17 large) and 63 prokaryotic (homologous to the *E. coli* γ-subunit) clamp loader subunits. Residue numberings for *S. cerevisiae* RFC-B (RFC4) and the *E. coli* γ-subunit are as follows: (1) B-Lys149, γ-Lys161; (2) B-Thr120, γ-Ser132; (3) B-Arg84, γ-Arg98/γ-Lys100; (4) γ-Arg133 and γ-Gln160; (5) B-Arg90, γ-Arg105.

segment fits comfortably within the PCNA ring, which is large enough to accommodate the bend in the DNA. The single-stranded portion of template extends out of the gap in the RFC spiral, near domain IV of RFC-A. The conservation of positive charge within a surface groove of RFC and the bacterial clamp loaders suggests a possible exit path for the template strand (Fig. 4b).

Conclusions

By introducing mutations into the active sites of the RFC complex, we have captured the clamp loader loaded with an ATP analogue and bound to the closed PCNA sliding clamp. The arrangement of the ATPase domains within the complex suggests that these domains spiral around double-helical primed DNA. The RFC-A subunit is positioned near the mouth of the PCNA ring, enabling it to bind readily to DNA emerging from the clamp. If we assume that this interaction involves the minor groove of the double helix, then the length of the 'readout' of DNA by the RFC spiral ($\sim$11 base pairs) is such that the 3′ end of the double helix must terminate within the RFC complex. This condition provides a simple mechanism for the specific recognition of primer–template junctions by the RFC complex.

The geometry of the double helix appears to match the geometry of an RFC spiral that generates optimally configured active sites, as seen at the RFC-A:RFC-B interface. Just as the central coiled-coil shaft of F$_1$ATPase (the γ-subunit) is coupled to ATP hydrolysis, DNA may serve to stabilize the ATPase domains of RFC in configurations that are optimal for catalysis. Comparison with the structures of RecA, the replicative helicase from T7 bacteriophage and the Rho termination factor suggests that an ancient mechanism for DNA recognition by RecA-type ATPases has been tailored by the clamp loaders for the specific loading of sliding clamps onto primer–template junctions. Our structure of the RFC:PCNA complex is a first view of an AAA+ ATPase assembly engaging one of its substrates, and we look forward to being able to compare the clamp loader mechanism with that of other members of this extensive superfamily.
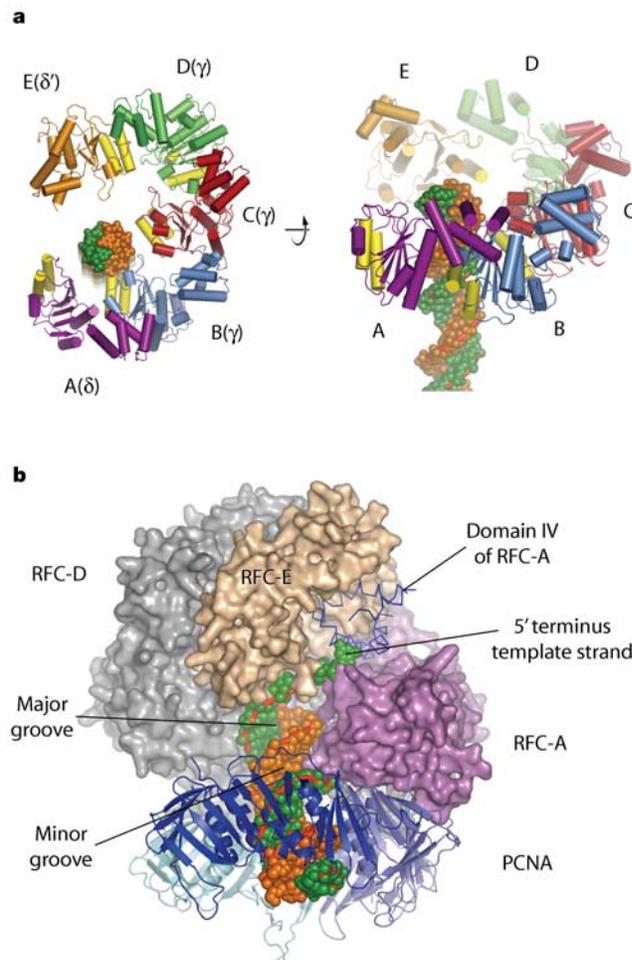

Figure 6 Inactive and active clamp loader complexes. **a**, The nucleotide-free, inactive state of the *E. coli* clamp loader[11] appears to be organized into a proto-spiral. Two views of the clamp loader are shown, oriented with respect to ideal B-form DNA by superimposing the B subunit of the *E. coli* clamp loader onto the B subunit of the RFC:PCNA:DNA model. The five *E. coli* subunits are coloured according to their counterparts in the RFC complex, and the collar has been removed for clarity. **b**, The RFC:PCNA complex modelled with primer–template DNA from the crystal structure of HIV reverse transcriptase[40] (accession code 1RTD). The primer and template strands are in orange and green, respectively. The molecular surface of the RFC complex is shown, except for domain IV of RFC-A.

Note added in proof: A three-dimensional reconstruction of the RFC:PCNA complex from the archaebacterium *Pyrococcus furiosus* has been obtained using cryo-electron microscopy[49]. □

Methods

Protein preparation

S. cerevisiae RFC and PCNA were expressed and purified as described in Supplementary Information. A truncated construct of RFC-A (residues 295–785) and full-length but mutant forms of the other subunits (RFC-B (R157Q), RFC-C (R160Q), RFC-D (R183Q) and RFC-E (R184Q); for nomenclature, see Fig. 2) were combined to form the RFC complex and purified as for wild-type RFC. The RFC complex was mixed with a $\sim$2 molar excess of trimeric PCNA, 10 mM MgCl$_2$ and 0.5 mM ATP-γS, concentrated, and passed over a Superdex-200 (10/30) column (Amersham Biosciences) that was equilibrated in 200 mM NaCl, 20 mM HEPES, pH 7.6, 0.5 mM ATP-γS, 10 mM MgCl$_2$ and 5 mM β-mercaptoethanol (β-ME). Fractions of RFC:PCNA were then concentrated and frozen in small aliquots for crystallization.

Crystal growth and handling

Crystals were grown at 20 °C by the hanging drop method, using equal volumes of protein and crystallization solutions over a 500 μl reservoir. Before mixing RFC:PCNA (30–50 mg ml^{-1}) with the reservoir solution (300 mM NaCl, 14.4–15.2% w/v polyethylene glycol 3350 (PEG 3350), 100 mM 2-(*N*-cyclohexylamino)ethanesulphonic acid (CHES), pH 9.0, 5 mM β-ME), final concentrations of 10 mM ATP-γS, 10 mM MgCl$_2$, 4 mM

©2004 **Nature Publishing Group**

DL-dithiothreitol (DTT), 2 mM Tris (2-carboxyethyl)phosphine hydrochloride (TCEP) were added to the protein solution. Crystals were propagated by streak-seeding at the time of set-up, and were visible within several hours. The diffraction limit for these crystals was no better than ~5 Å, but was significantly improved by controlled crystal dehydration[19]. Two to four days after set-up, crystals measuring approximately $50 \times 100 \times 200\,\mu m$ were harvested into a sitting drop reservoir containing $50\,\mu l$ mother liquor plus 1 mM ATP-γS and 10 mM MgCl$_2$, and the PEG 3350 concentration was steadily increased by slowly replacing the solution over 2 hours with an identical mix containing 33% PEG 3350. Crystals were then plunged into freshly thawed propane slush and stored in liquid nitrogen.

Data collection and structure determination

Diffraction data were measured at the Advanced Light Source, Lawrence Berkeley National Laboratory, on beamline 8.2.2 and processed with HKL2000 Denzo/Scalepack[41]. RFC:PCNA crystallizes in the spacegroup $P2_12_12_1$, with one complex in the asymmetric unit. Dehydration shrinks the unit cell from $a \sim 102.6\,\text{Å}$, $b \sim 116.8\,\text{Å}$, $c \sim 290.1\,\text{Å}$ to $a = 104.2\,\text{Å}$, $b = 110.4\,\text{Å}$, $c = 268.2\,\text{Å}$. Phases were calculated from a two-wavelength anomalous dispersion (MAD) experiment at the selenium inflection and high remote energies[42] using an RFC:PCNA crystal in which only the PCNA component was labelled with selenomethionine, and 17 selenium atoms were located using SOLVE[43]. Electron density maps were calculated using MAD phases to 3.8 Å resolution, and greatly improved with solvent flattening and phase extension to 3.2 Å using SHARP[44] and DM[45]. A further, isomorphous data set was collected to 2.85 Å resolution on a fresh crystal and combined with the MAD phases.

Model building at 2.85 Å resolution was performed with the program O[46], and refinement was carried out using CNS[47]. The R_{work} and R_{free} values are 0.250 and 0.306, respectively. The PCNA subunits contacting RFC and the AAA+ modules of RFC-A, RFC-B and RFC-C are the best defined (average B-values ranging from 56 to 79 Å^2), whereas the RFC-D and RFC-E subunits and domain IV of RFC-A exhibit significantly more disorder (average B-values ranging from 109 to 131 Å^2). In particular, the unique domain IV of RFC-A is built as a polyalanine chain for much of its length. Molecular transformations were calculated using TOP[48] and FIT (http://bioinfo1.mbfys.lu.se/~guoguang).

1. Huang, C. C., Hearst, J. E. & Alberts, B. M. Two types of replication proteins increase the rate at which T4 DNA polymerase traverses the helical regions in a single-stranded DNA template. *J. Biol. Chem.* **256,** 4087–4094 (1981).
2. Prelich, G., Kostura, M., Marshak, D. R., Mathews, M. B. & Stillman, B. The cell-cycle regulated proliferating cell nuclear antigen is required for SV40 DNA replication *in vitro. Nature* **326,** 471–475 (1987).
3. Stukenberg, P. T., Studwell-Vaughan, P. S. & O'Donnell, M. Mechanism of the sliding β-clamp of DNA polymerase III holoenzyme. *J. Biol. Chem.* **266,** 11328–11334 (1991).
4. Kong, X. P., Onrust, R., O'Donnell, M. & Kuriyan, J. Three-dimensional structure of the β subunit of *E. coli* DNA polymerase III holoenzyme: a sliding DNA clamp. *Cell* **69,** 425–437 (1992).
5. Krishna, T. S., Kong, X. P., Gary, S., Burgers, P. M. & Kuriyan, J. Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA. *Cell* **79,** 1233–1243 (1994).
6. Warbrick, E. The puzzle of PCNA's many partners. *Bioessays* **22,** 997–1006 (2000).
7. Tsurimoto, T. & Stillman, B. Purification of a cellular replication factor, RF-C, that is required for coordinated synthesis of leading and lagging strands during simian virus 40 DNA replication *in vitro. Mol. Cell. Biol.* **9,** 609–619 (1989).
8. Tsurimoto, T. & Stillman, B. Functions of replication factor C and proliferating-cell nuclear antigen: functional similarity of DNA polymerase accessory proteins from human cells and bacteriophage T4. *Proc. Natl Acad. Sci. USA* **87,** 1023–1027 (1990).
9. Lee, S. H., Kwong, A. D., Pan, Z. Q. & Hurwitz, J. Studies on the activator 1 protein complex, an accessory factor for proliferating cell nuclear antigen-dependent DNA polymerase δ. *J. Biol. Chem.* **266,** 594–602 (1991).
10. Gomes, X. V. & Burgers, P. M. 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 (2001).
11. Jeruzalmi, D., O'Donnell, M. & Kuriyan, J. Crystal structure of the processivity clamp loader γ complex of *E. coli* DNA polymerase III. *Cell* **106,** 429–441 (2001).
12. Oyama, T., Ishino, Y., Cann, I. K., Ishino, S. & Morikawa, K. Atomic structure of the clamp loader small subunit from *Pyrococcus furiosus. Mol. Cell* **8,** 455–463 (2001).
13. Neuwald, A. F., Aravind, L., Spouge, J. L. & Koonin, E. V. AAA+ : A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. *Genome Res.* **9,** 27–43 (1999).
14. Guenther, B., Onrust, R., Sali, A., O'Donnell, M. & Kuriyan, J. Crystal structure of the δ' subunit of the clamp-loader complex of *E. coli* DNA polymerase III. *Cell* **91,** 335–345 (1997).
15. Yu, R. C., Hanson, P. I., Jahn, R. & Brunger, A. T. Structure of the ATP-dependent oligomerization domain of *N*-ethylmaleimide sensitive factor complexed with ATP. *Nature Struct. Biol.* **5,** 803–811 (1998).
16. Lenzen, C. U., Steinmann, D., Whiteheart, S. W. & Weis, W. I. Crystal structure of the hexamerization domain of *N*-ethylmaleimide-sensitive fusion protein. *Cell* **94,** 525–536 (1998).
17. Goedken, E. R. *et al.* Fluorescence measurements on the *E.coli* DNA polymerase clamp loader: implications for conformational changes during ATP and clamp binding. *J. Mol. Biol.* **336,** 1047–1059 (2004).
18. Ason, B. *et al.* Mechanism of loading the *Escherichia coli* DNA polymerase III β sliding clamp on DNA. Bona fide primer/templates preferentially trigger the γ complex to hydrolyze ATP and load the clamp. *J. Biol. Chem.* **278,** 10033–10040 (2003).
19. Heras, B. *et al.* Dehydration converts DsbG crystal diffraction from low to high resolution. *Structure (Camb.)* **11,** 139–145 (2003).
20. Shamoo, Y. & Steitz, T. A. Building a replisome from interacting pieces: sliding clamp complexed with a peptide from DNA polymerase and a polymerase editing complex. *Cell* **99,** 155–166 (1999).

21. Gulbis, J. M., Kelman, Z., Hurwitz, J., O'Donnell, M. & Kuriyan, J. Structure of the C-terminal region of p21(WAF1/CIP1) complexed with human PCNA. *Cell* **87,** 297–306 (1996).
22. Lee, S. *et al.* The structure of ClpB: a molecular chaperone that rescues proteins from an aggregated state. *Cell* **115,** 229–240 (2003).
23. Niwa, H., Tsuchiya, D., Makyio, H., Yoshida, M. & Morikawa, K. Hexameric ring structure of the ATPase domain of the membrane-integrated metalloprotease FtsH from *Thermus thermophilus* HB8. *Structure (Camb.)* **10,** 1415–1423 (2002).
24. Menz, R. I., Walker, J. E. & Leslie, A. G. Structure of bovine mitochondrial F$_1$-ATPase with nucleotide bound to all three catalytic sites: implications for the mechanism of rotary catalysis. *Cell* **106,** 331–341 (2001).
25. Yang, W., Gao, Y. Q., Cui, Q., Ma, J. & Karplus, M. The missing link between thermodynamics and structure in F1-ATPase. *Proc. Natl Acad. Sci. USA* **100,** 874–879 (2003).
26. Waga, S. & Stillman, B. The DNA replication fork in eukaryotic cells. *Annu. Rev. Biochem.* **67,** 721–751 (1998).
27. Saenger, W. *Principles of Nucleic Acid Structure* (Springer, New York, 1984).
28. Xiong, Y. & Sundaralingam, M. Crystal structure of a DNA RNA hybrid duplex with a polypurine RNA r(gaagaagag) and a complementary polypyrimidine DNA d(CTCTTCTTC). *Nucleic Acids Res.* **28,** 2171–2176 (2000).
29. Jones, S., van Heyningen, P., Berman, H. M. & Thornton, J. M. Protein–DNA interactions: A structural analysis. *J. Mol. Biol.* **287,** 877–896 (1999).
30. Jeruzalmi, D. *et al.* Mechanism of processivity clamp opening by the δ subunit wrench of the clamp loader complex of *E. coli* DNA polymerase III. *Cell* **106,** 417–428 (2001).
31. Story, R. M., Weber, I. T. & Steitz, T. A. The structure of the *E. coli* recA protein monomer and polymer. *Nature* **355,** 318–325 (1992).
32. Hortnagel, K. *et al.* Saturation mutagenesis of the *E. coli* RecA loop L2 homologous DNA pairing region reveals residues essential for recombination and recombinational repair. *J. Mol. Biol.* **286,** 1097–1106 (1999).
33. Mirshad, J. K. & Kowalczykowski, S. C. Biochemical characterization of a mutant RecA protein altered in DNA-binding loop 1. *Biochemistry* **42,** 5945–5954 (2003).
34. Notarnicola, S. M., Park, K., Griffith, J. D. & Richardson, C. C. A domain of the gene 4 helicase/primase of bacteriophage T7 required for the formation of an active hexamer. *J. Biol. Chem.* **270,** 20215–20224 (1995).
35. Washington, M. T., Rosenberg, A. H., Griffin, K., Studier, F. W. & Patel, S. S. Biochemical analysis of mutant T7 primase/helicase proteins defective in DNA binding, nucleotide hydrolysis, and the coupling of hydrolysis with DNA unwinding. *J. Biol. Chem.* **271,** 26825–26834 (1996).
36. Skordalakes, E. & Berger, J. M. Structure of the Rho transcription terminator: mechanism of mRNA recognition and helicase loading. *Cell* **114,** 135–146 (2003).
37. Sawaya, M. R., Guo, S., Tabor, S., Richardson, C. C. & Ellenberger, T. Crystal structure of the helicase domain from the replicative helicase-primase of bacteriophage T7. *Cell* **99,** 167–177 (1999).
38. Singleton, M. R., Sawaya, M. R., Ellenberger, T. & Wigley, D. B. Crystal structure of T7 gene 4 ring helicase indicates a mechanism for sequential hydrolysis of nucleotides. *Cell* **101,** 589–600 (2000).
39. Tsurimoto, T. & Stillman, B. Replication factors required for SV40 DNA replication *in vitro.* I. DNA structure-specific recognition of a primer–template junction by eukaryotic DNA polymerases and their accessory proteins. *J. Biol. Chem.* **266,** 1950–1960 (1991).
40. Huang, H., Chopra, R., Verdine, G. L. & Harrison, S. C. Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. *Science* **282,** 1669–1675 (1998).
41. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276,** 307–326 (1997).
42. Gonzalez, A. Optimizing data collection for structure determination. *Acta Crystallogr. D* **59,** 1935–1942 (2003).
43. Terwilliger, T. C. & Berendzen, J. Automated MAD and MIR structure solution. *Acta Crystallogr. D* **55,** 849–861 (1999).
44. La Fortelle, E. D. & Bricogne, G. Maximum-likelihood heavy-atom parameter refinement in the MIR and MAD methods. *Methods Enzymol.* **276,** 476–494 (1997).
45. Collaborative Computational Project Number 4, The CCP4 suite programs for protein crystallography. *Acta Crystallogr. D* **50,** 760–763 (1994).
46. Kleywegt, G. J. & Jones, T. A. Efficient rebuilding of protein structures. *Acta Crystallogr. D* **50,** 829–832 (1996).
47. Brünger, A. T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54,** 905–921 (1998).
48. Lu, G. TOP: A new method for protein structure comparisons and similarity searches. *J. Appl. Crystallogr.* **33,** 176–183 (2000).
49. Miyata, T. *et al.* Direct view of the clamp-loading complex for processive DNA replication. *Nature Struct. Mol. Biol.* (in the press).

Supplementary Information accompanies the paper on **www.nature.com/nature**.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to J.K. (kuriyan@berkeley.edu). The coordinates and structure factors have been deposited in the Protein Data Bank (accession code 1SXJ).