Structural and functional similarities of prokaryotic and eukaryotic DNA polymerase sliding clamps

Zvi Kelman* and Mike O'Donnell1

Microbiology Department Hearst Research Foundation and 1Howard Hughes Medical Institute, Cornell University Medical College, 1300 York Avenue, New York, NY 10021, USA

Received June 9, 1995; Revised and Accepted August 15, 1995

ABSTRACT

The remarkable processivity of cellular replicative DNA polymerases derive their tight grip to DNA from a ring-shaped protein that encircles DNA and tethers the polymerase to the chromosome. The crystal structures of prototypical 'sliding clamps' of prokaryotes (β subunit) and eukaryotes (PCNA) are ring shaped proteins for encircling DNA. Although β is a dimer and PCNA is a trimer, their structures are nearly superimposable. Even though they are not hexamers, the sliding clamps have a pseudo 6-fold symmetry resulting from three globular domains comprising each β monomer and two domains comprising each PCNA monomer. These domains have the same chain fold and are nearly identical in three-dimensions. The amino acid sequences of 11 β and 13 PCNA proteins from different organisms have been aligned and studied to gain further insight into the relation between the structure and function of these sliding clamps. Furthermore, a putative embryonic form of PCNA is the size of β and thus may encircle DNA as a dimer like the prokaryotic clamps.

INTRODUCTION

Chromosomal replicases are multiprotein assemblies that polymerize thousands of nucleotides without dissociating from DNA. In three well characterized systems (E.coli, eukaryotes and bacteriophage-T4) this remarkable processivity is achieved by a ring-shaped protein (‘sliding clamp’) that encircles DNA and anchors the polymerase to the template (reviewed in 1,2). The sliding clamp is assembled on DNA by a multiprotein ‘clamp loader’ apparatus that recognizes a primer terminus and couples ATP hydrolysis to assembly the clamp around DNA (3). In eukaryotes the clamp loader is the 5-subunit RF-C complex (also known as activator-1) and the sliding clamp is the PCNA protein which confers processivity to DNA polymerase δ (Pol δ). The clamp loader of the prokaryotic DNA polymerase III holoenzyme (Pol III) is the 5-subunit γ-complex and the clamp is the 3 subunit. In the T4 system the clamp loader is the gene 44/62 protein complex (g44/62p) and the clamp is the product of gene 45.

The first indication of the processivity factor’s circular shape came from the observation that the β subunit was tightly fastened to circular DNA but easily dissociated from linear DNA, suggesting that β slides off the ends of DNA (4). This hypothesis was enforced by the observation that the exit of β from linear DNA was blocked by proteins bound near the ends of the DNA. The three dimensional structure of the E.coli β subunit showed it to be a dimer in the shape of a ring with a central cavity large enough to accommodate duplex DNA (5) (Fig. 1).

The sliding clamp in the T4 system has been observed on DNA by cryoelectron microscopy (6). The clamps not only appeared to encircle the DNA, but their appearance also indicated that they may slide along the DNA. These clamps are presumably gene 45 protein trimers encircling the DNA in a fashion similar to that of the β dimer. Using a linear template it was shown that gene 45 protein can support processive replication in the absence of the clamp loader (7) and that the T4 sliding clamp consists solely of gene 45 protein (8). Interestingly, the gene 45 protein can bind to a circular template in the absence of the clamp loader, provided a macromolecular crowding agent is present (9).

Evidence for the topological binding of PCNA to DNA came from the observation that PCNA alone can support replication by Pol δ, in the absence of RF-C complex, only when the DNA is linear and has a double-stranded end (10). This result suggests that PCNA can slide along duplex DNA until it reaches the 3' terminus where it interacts with the polymerase to initiate processive replication. Supporting evidence for PCNA sliding along duplex DNA comes from photocrosslinking experiments in which PCNA can be crosslinked to DNA following assembly around DNA (11). However, crosslinking of PCNA to DNA is not observed upon linearization of the duplex (11). Furthermore, using tritiated human PCNA it has been demonstrated that the protein can be loaded onto nicked circular DNA, but upon linearization of the DNA it slides off the ends, similar to the E.coli β subunit (N. Yao, Z. Kelman and M. O'Donnell, unpublished). Proof that PCNA and β are structurally similar was obtained upon solution of the three dimensional structure of the yeast Saccharomyces cerevisiae PCNA which showed it to be a trimer with an overall shape very similar to that of E.coli β (12) (Fig. 1).

The sliding clamps (β and PCNA) have a 6-fold appearance, yet they are not hexamers. In β the 6-fold appearance results from three globular domains which comprise each monomer, while in PCNA each monomer contains two domains. In both prokaryotes

*To whom correspondence should be addressed at present address: Department of Molecular Biology and Genetics, The Johns Hopkins Medical School, Baltimore, MD 21205, USA
These sequences were obtained from the databases by searching for known DNA sequences of dmN, which is a superimposition of the PCNA and dimer of β subunits and 13 PCNAs (Fig. 3). The alignments were made such that no gaps were introduced between elements of secondary structure and that even though these proteins have high net negative charges, the residues aligning with the α helices of the central cavity have a net positive charge. These alignments generally follow a third criteria, that positions corresponding to the hydrophobic amino acid residues (Phe, Leu, Ile, Met, Val or Ala) in E.coli β and yeast PCNA that are inaccessible to water, are also occupied by hydrophobic residues.

The central cavity

In general these clamps are acidic proteins (Tables 1 and 2). The pI of the PCNAs are slightly lower then the pI of β subunits and gene 45 protein has an intermediate pI (pI = 4.8). The distribution of charges on the ring is asymmetric. The outer surface has a strong negative electrostatic potential and the inside of the central cavity has a net positive electrostatic potential. Hence, these clamps would be repelled by DNA, but after assembly onto DNA, the ring may even interact favorably with it. A possible function of the negative outside surface may be to increase the specificity of assembly on DNA by preventing the ring from associating with DNA by itself, thus constraining it to enlist the help of the clamp loader and ATP. Alternatively, the negative surface may help destabilize local interaction with DNA. A notable exception to this rule is the β subunit from B.aphidicola with a pI of 9.0. Hence, providing this β subunit functions like β of E.coli, an overall acidic charge would appear not to be essential to clamp function.

Both β and PCNA have 12 α helices in the central cavity all of which are perpendicular to the DNA. The α helices are long enough to span the major and minor grooves and thus may act as molecular crossbars to prevent the clamp from entering the grooves of DNA during sliding. The α helices have an overall positive potential as a consequence of eight (PCNA) or 12 (β) Lys or Arg residues within the ring. However, there is at least one acidic amino acid (Asp or Glu) in every α helix of β, and in six out of the 12 α helices in PCNA. These negatively charged residues may be needed for a balance to prevent the local interaction between the clamp and DNA. Although the overall structure of the rings are similar, the central cavity in β (35 Å) is larger than the one in PCNA (34 Å) and has a slightly more oval shape (Fig. 1).

The dimer interface

The interfaces between the protomers of β and PCNA are formed by β sheet. The β sheet is part of the continuous layer of sheet structure on the outside of these rings. Three interactions may contribute to stabilization of the interface: (i) hydrogen bonds between the β sheets; (ii) a small hydrophobic core; and (iii) putative ion pairs. In the following sections the amino acids that contribute to the formation of the interface close to the N-terminus will be referred to as the ‘head interface’ and those near the C-terminus will be referred to as the ‘tail interface’ (Fig. 4).
Figure 2. Sequence alignment of β subunits from different prokaryotes. The proteins are aligned to show maximum similarity with proteins directly above and below. Identical residues between neighboring proteins have a vertical dash. Gaps were introduced to maximize homology. The sequences are split into three parts which correspond to the three globular domains of the E.coli β subunit. The Glu and Asp residues are shown in red and the Lys and Arg residues are blue. To the right are the amino acid numbers at the end of each domain of the E.coli β. The secondary structures corresponding to E.coli β are indicated above each domain. Amino acids that are conserved in all species are indicated in the bottom row of each domain. The sequences are from the following references: Buchnera aphidicola [BUCAP (24)], Bacillus subtilis [BACSU (25)], Micrococcus luteus [MICLU (26)], Streptomyces coelicolor [STRO (27)], Pseudomonas putida [PSEP (28)], Serratia marcescens [SERMA (29)], Salmonella typhimurium [SALTY (29)], Escherichia coli [ECOLI (30)], Proteus mirabilis [PROMI (31)], Mycoplasma capricolum [MYCCA (32)], Actinobacillus pleuropneumoniae [ACTPAL (33)].

Hydrogen bonds. Hydrogen bonds between the two antiparallel β sheets may contribute to stabilization of the interface in both β and PCNA. The difference, however, is that within the PCNA interface there are eight potential hydrogen bonds between the sheets compared to only four in the β subunit. Perhaps the greater entropic force against stable association of a trimer versus a dimer requires more interactions among the interfaces of the protomers.

Hydrophobic forces. Hydrophobic interactions may also be important in the strength of the interface. The PCNA interface contains four pairs of hydrophobic amino acids which form a hydrophobic core (Fig. 4). In β, however, the interface hydrophobic core consists of only two pairs of hydrophobic residues (Fig. 4). This may suggest that hydrophobic interactions play an important role in the stabilization of a trimeric structure and again support the notion that more interactions are needed to stabilize a trimer relative to a dimer.

In β the two hydrophobic pairs are: Phe106 to Leu273 and Leu108 to Ile272 ([5]; Fig. 4). The position of the four hydrophobic residues of the interface have been highly conserved in all the bacterial species examined (Fig. 2). For example, Phe106 and Leu108 have been conserved in every species except B. aphidicola (Phe is changed to Tyr) and M. capricolum (Leu is changed to Ile). The Ile272 and Leu273 are conserved, though not identical and are replaced by either LeuVal or LeuLeu.

In PCNA the four hydrophobic pairs are: Ile78 to Leu154, Ala112 to Val180, Tyr114 to Ile181, and Leu116 to Ile182 ([12]). The eight hydrophobic residues in PCNA are not as conserved as in the prokaryotes. In the head interface the four hydrophobic amino acids in different species range from being identical to those in
**S. cerevisiae** or have been replaced by other hydrophobic amino acids. In the tail interface the picture is more complex. In all species the hydrophobic residue at position 154 is conserved (either Leu, Ile or Phe). However, in several species the amino acids in the triplet Val/Ile/Ile (residues 180-183) (Fig. 3) are replaced by either polar amino acids (Thr or Cys) or even a charged residue (Lys) in mammalian and *D. melanogaster* PCNAs (Fig 3).

**Ion pairs.** Ionic interactions between amino acids with opposing charges may also contribute to the stability of the interface. There are six putative ion pairs in the interface of *E. coli* β, but only one in yeast PCNA (Fig. 4). Interestingly, this is the opposite situation than that of the putative hydrophobic forces in the interface where there are four pairs of hydrophobic amino acids in PCNA and only two in β. This suggests that ion pairs may play a significant role in the strength of the interface in prokaryotic clamps.

In the primary sequence of the β subunit, the tail interface is the most tightly focused region of negative charge; five of seven residues are glutamates (residues 298–304; Figs 2 and 4). The putative tail interfaces of the other bacterial species also show this focused region of negative charge. Conversely, the head interface contains the positively charged partners, but these residues are scattered over a longer region.

The β sequences from other bacteria each have residues in positions that correspond to at least three of the ion pairs of *E. coli* β and they have additional charged residues in their interface region which could possibly form other ion pairs. The only putative ion pair conserved among all prokaryotes studied is the Lys74–Glu298 pair. The Arg105–Glu301 pair is reversed in *B. subtilis* (Glu105–Lys301). Residues corresponding to the buried Arg96–Glu300 of *E. coli* β is present in all the species except for *B. subtilis* where both members of the pair are missing, suggesting intolerance at burying a lone member of this pair. In PCNA from all 13 species studied, there is an Asp at position 150 and either an Arg or Lys at position 110.
Figure 4. Residues that comprise the dimer interface of *E. coli* β and of *S. cerevisiae* PCNA. (A) The interface of the *E. coli* β and (B) the interface of the yeast PCNA. Amino acids that form the sheet structure at the dimer interface are underlined. Lines drawn between residues of the two interfaces indicate putative interactions which stabilize the structure. The N-terminal domain interface (head interface) is shown above the C-terminal domain interface (tail interface).

Table 1. Molecular mass, net charge and calculated pI of β subunits from different species (using the DNASTAR program)

<table>
<thead>
<tr>
<th>Species</th>
<th>Amino acids</th>
<th>MW (Da)</th>
<th>Net charge (pH = 7)</th>
<th>pI</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. coli</em></td>
<td>366</td>
<td>40589.9</td>
<td>-9.982</td>
<td>5.218</td>
</tr>
<tr>
<td><em>B. subtilis</em></td>
<td>378</td>
<td>42106.6</td>
<td>-14.388</td>
<td>4.819</td>
</tr>
<tr>
<td><em>S. coelicolor</em></td>
<td>376</td>
<td>39959.4</td>
<td>-18.877</td>
<td>4.487</td>
</tr>
<tr>
<td><em>P. putida</em></td>
<td>366</td>
<td>40593.8</td>
<td>-10.279</td>
<td>5.095</td>
</tr>
<tr>
<td><em>P. mirabilis</em></td>
<td>367</td>
<td>40751.3</td>
<td>-9.286</td>
<td>5.137</td>
</tr>
<tr>
<td><em>B. aphidicola</em></td>
<td>365</td>
<td>42057.2</td>
<td>8.207</td>
<td>9.063</td>
</tr>
</tbody>
</table>

Only β subunits for which the entire sequence is known are listed.

Table 2. Molecular mass, net charge and calculated pI of PCNA from different species (using the DNASTAR program)

<table>
<thead>
<tr>
<th>Species</th>
<th>Amino acids</th>
<th>MW (Da)</th>
<th>Net charge (pH = 7)</th>
<th>pI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human</td>
<td>262</td>
<td>28899.5</td>
<td>-16.728</td>
<td>4.457</td>
</tr>
<tr>
<td>Murine</td>
<td>261</td>
<td>28787.5</td>
<td>-15.563</td>
<td>4.553</td>
</tr>
<tr>
<td>Rat</td>
<td>261</td>
<td>28751.4</td>
<td>-16.728</td>
<td>4.457</td>
</tr>
<tr>
<td><em>X. laevis</em></td>
<td>261</td>
<td>28899.3</td>
<td>-17.896</td>
<td>4.363</td>
</tr>
<tr>
<td><em>D. melanogaster</em></td>
<td>260</td>
<td>28832.5</td>
<td>-13.827</td>
<td>4.543</td>
</tr>
<tr>
<td><em>C. roseus</em></td>
<td>268</td>
<td>29767.6</td>
<td>-17.531</td>
<td>4.460</td>
</tr>
<tr>
<td><em>D. carota</em> (short)</td>
<td>264</td>
<td>29226.1</td>
<td>-15.532</td>
<td>4.556</td>
</tr>
<tr>
<td><em>D. carota</em> (long)</td>
<td>365</td>
<td>40105.1</td>
<td>-28.670</td>
<td>4.343</td>
</tr>
<tr>
<td><em>O. sativa</em></td>
<td>263</td>
<td>29274.2</td>
<td>-15.697</td>
<td>4.503</td>
</tr>
<tr>
<td><em>S. cerevisiae</em></td>
<td>258</td>
<td>28919.3</td>
<td>-19.837</td>
<td>4.261</td>
</tr>
<tr>
<td><em>S. pombe</em></td>
<td>260</td>
<td>28971.5</td>
<td>-17.671</td>
<td>4.367</td>
</tr>
<tr>
<td>AcNPV</td>
<td>256</td>
<td>28637.5</td>
<td>-9.072</td>
<td>5.219</td>
</tr>
</tbody>
</table>

Only PCNA for which the entire sequence is known are listed.
The evolutionary conservation of charged pairs in these interfaces suggest they are important to the function of the clamps. The most obvious role is to stabilize the dimer interface. However, it is possible that the need to exclude water from the buried charges could require energy and thus destabilize the interface, adding to a balance between stability and instability in this region that may be needed to open and close the interface. Another possible role for charged residues at the interface derives from the high degree of symmetry in the ring. Computer modeling of a head to head β dimer results in a ring which looks like the head to tail ring, having tight interfaces with an antiparallel sheet as in the head to tail interface. Therefore, another possible function of the charged residues may be to provide specificity to form the head to tail dimer rather than the head to head dimer.

The interface may be a 'busy' region of protein interaction as this is the site at which work may be performed by the clamp loader to open and close the ring. Hence, another possible role of the charged residues at the interface may be to function in the recognition of other subunits of the polymerase. Also, if the polymerase (or another subunit) were to bind the clamp by spanning the interface it could act as a brace to further stabilize the interface thereby increasing the stability of the clamp on DNA and leading to greater processivity.

**Sliding clamp interactions with other proteins**

The sliding clamps must interact with at least two components of their corresponding DNA polymerase holoenzyme, the clamp loader and the DNA polymerase. However, these sliding clamps also interact with cellular components involved in other processes (Table 3). In the T4 system, gene 45 protein interacts with RNA polymerase to activate late gene transcription (13). Similarly, the PCNA homologue from the baculovirus AcNPV is important for late gene transcription (14). Beside its interactions with Pol III holoenzyme, the β subunit also interacts with DNA polymerase II an enzyme implicated in DNA repair. Human PCNA interacts with D-type cyclins (15), the cell cycle dependent kinase (CDK) inhibitor p21 (16,17) and the DNA damage induced gene, Gadd45 (18) (Table 3). In vitro replication studies have shown that the binding of p21 to PCNA inhibits replication (16,17). Gadd45, like p21, is induced by p53 upon DNA damage, presumably to block DNA replication. Consistent with this notion both p21 (16,17) and...
Gadd45 (Jerard Hurwitz, pers. comm.) have been shown to inhibit DNA replication in vitro.

**Table 3. Multiple enzymes interact with sliding clamps of prokaryotic and eukaryotic DNA polymerases**

<table>
<thead>
<tr>
<th>Clamp</th>
<th>Interacts with:</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. coli</em> β</td>
<td>DNA polymerase II</td>
</tr>
<tr>
<td></td>
<td>DNA polymerase III</td>
</tr>
<tr>
<td>T4 gene 45 protein</td>
<td>T4 DNA polymerase (gene 43 protein)</td>
</tr>
<tr>
<td></td>
<td><em>E. coli</em> RNA polymerase (modified by gene 33 protein and gene 55 protein)</td>
</tr>
<tr>
<td>Human PCNA</td>
<td>DNA polymerase δ</td>
</tr>
<tr>
<td></td>
<td>DNA polymerase ε</td>
</tr>
<tr>
<td></td>
<td>D-type cyclins</td>
</tr>
<tr>
<td></td>
<td>p21 CDK inhibitor</td>
</tr>
<tr>
<td></td>
<td>Gadd45 protein</td>
</tr>
</tbody>
</table>

The sequences of the different β subunits in Figure 2 are much more divergent than the PCNA genes in Figure 3, due to the greater evolutionary distance between the prokaryotic organisms relative to the eukaryotes. Hence, identical residues in the β subunits from all species may be expected to be functionally important. Among the β subunits analyzed only 15 residues are identical in all species (Fig. 2). These conserved amino acids are either in loops or in the α-helices lining the central cavity (Fig. 5). Interestingly, the conserved amino acids on the loops are all on only one face of the ring. The conserved residues are good candidates for interactions with other proteins. Consistent with this, the C-termini of β which protrude from this face have been shown to be a touch point for interaction of β with both the clamp loader and the DNA polymerase, consistent with the recognition of β by both the polymerase and the clamp loader (V. Naktinis and M. O'Donnell, unpublished observation).

The C-terminal amino acids of the *E. coli* β subunit are important in the binding of both the γ complex and the polymerase (V. Naktinis and M. O'Donnell, unpublished observation). Examination of the C-terminus of PCNA shows a notable high negative charge. In all species there are 3–6 Glu or Asp residues adjacent to the C-terminus (Fig. 3). The three dimensional structure of the yeast PCNA shows that the three acidic amino acids form a ‘hook’ with the side chains pointing to the solvent (12). This highly acidic region is unique to PCNA and is not present in β or gene 45 protein. It is tempting to speculate that the acidic C-terminus of PCNA, lacking in β, may be involved in interaction with cellular regulators.

In the region preceding the acidic C-terminus of PCNA there are several sequence similarities to β. In all PCNAs, there is a conserved Lys before the Glu/Asp stretch and in all β subunits there is a positively charged amino acid in the C-terminus (Fig. 2). Additionally, both PCNA and β have a hydrophobic residue (Ile or Leu) following the positively charged amino acid and both have one Pro within the terminal four amino acids. The amino acid sequence similarities at the C-terminal region of both β and PCNA, together with the observation that the subunits of the *E. coli* γ complex share sequence similarities with the RF-C complex (19), suggest that this region in PCNA may be responsible for interaction with RF-C and/or with Pol δ.

**Dimer and trimer sliding clamps are found in both prokaryotes and eukaryotes**

All of the prokaryotic dnaN genes examined here encode proteins of approximately the same size (length of 365–378 residues; Table 1). Hence, they presumably form dimers like *E. coli* β. The 13 PCNA genes encode proteins that are ~2/3 the size of β (lengths of 257–269 residues; Table 2) and presumably form trimers like yeast PCNA.

Why did eukaryotes evolve trimers? The identical chain topology of the domains suggests that both PCNA and β evolved from a single domain that either formed a ring by hexamerising or that fused. The three domains of a β monomer required two fusions; the two domains of a PCNA monomer required only one fusion. Hence, it is possible that eukaryotes (and phage-T4) have yet to undergo the second fusion to form the three domain monomer like prokaryotes. Alternatively, the trimeric form may have evolved from the prokaryote three domain protein by deletion of one domain.

A trimeric ring composed of three proteomers may not be as stable as a ring composed of two proteomers, having a greater probability for proteomer disassembly and a lower probability for three monomers to associate relative to only two. It is conceivable that a less stable ring is used to an advantage during the replication of the lagging strand. The lagging strand is synthesized as a series of fragments. Although the clamps exist in excess over the polymerase inside the cell, there are not enough clamps to provide one for each Okazaki fragment. Hence, the clamps must be recycled. The β subunit is very stable on DNA and a clamp unloading activity was identified to maintain a pool of available clamps within the cell (20). A less stable clamp on DNA may dissociate more easily from the DNA when the polymerase abandons it and thus a specific unloading mechanism may not be needed to maintain an intracellular pool of clamps.

It appears that both prokaryotes and eukaryotes make use of both type of rings, those composed of a trimer and those composed of a dimer. In the case of *E. coli*, it was recently demonstrated that a shorter version of β (called β*) is induced upon exposure of cells to UV light. β* is expressed from an internal promoter within the dnaN gene. This β* is comprised of the C-terminal 2/3 of normal β and hence each monomer contains two domains instead of three. Characterization of β* showed it behaves as a trimer on gel-filtration, presumably forming a 6-domain ring and it stimulates DNA synthesis by Pol III (Z. Livneh, pers. comm.).

A possible dimeric ring in eukaryotes was observed in an embryonic form of PCNA from *D. carota* (21). Two distinct PCNA genes were isolated from *D. carota*; one encodes a PCNA of typical length for a trimer (264 residues, 29.2 kDa) and the other encodes a larger protein of 365 amino acids (40.1 kDa). This is the size of a typical prokaryotic dnaN gene suggesting that the embryonic long form of PCNA may act as a dimer. Further, during oogenesis of the amphibian, *Xenopus laevis*, the presence of a 43 kDa protein that cross-reacts with anti-PCNA antibodies correlates with production of two differently sized transcripts (22).

Thus it appears that eukaryotes and prokaryotes alike may utilize rings made of dimers and trimers. Speculation as to why the chromosomal replicase of prokaryotes use the dimer form of...
the clamp and eukaryotes use the trimer has been treated (23), but experimentally derived evidence pertaining to this question remains for further study.

ACKNOWLEDGEMENTS

We are grateful to Jerard Hurwitz and Zvi Livneh for sharing their data prior to publication. We also wish to thank Talluru Krishna and Jennifer Turner for their contributions to Figures 1 and 5. This work was supported by a grant from the National Institutes of Health (GM38839).

REFERENCES