REVIEW ARTICLE

Sliding Clamps of DNA Polymerases

John Kuriyan†

Laboratories of Molecular Biophysics
Howard Hughes Medical Institute
The Rockefeller University, 1230 York Avenue
New York, NY 10021, U.S.A.

and Mike O'Donnell

Microbiology Department, Hearst Research Center
Howard Hughes Medical Institute
Cornell University Medical College, 1300 York Avenue
New York, NY 10021, U.S.A.

(Received 22 July 1993; accepted 14 September 1993)

The determination of the structure of the processivity factor (β subunit) of *Escherichia coli* DNA polymerase III holoenzyme showed that this protein acts to clamp the polymerase onto DNA by forming a closed circular structure that can encircle duplex DNA (X.-P. Kong, R. Onrust, M. O'Donnell & J. Kuriyan. (1992). *Cell*, 69, 425-437). In this review we describe the features of the β subunit that allow it to be linked tightly but non-specifically to DNA, and discuss the surprisingly symmetrical architecture of the molecule. The simple repeating pattern of the chain fold allows a connection to be made to the as yet unknown structures of eukaryotic proliferating cell nuclear antigen and the gene 45 protein of bacteriophage T4, which are the processivity factors of the corresponding DNA polymerases.

*Keywords:* DNA polymerase; sliding clamps; processivity factor

DNA polymerases that are involved in chromosomal replication during cell division can polymerize thousands of nucleotides in one continuous process, without dissociating from the template (Kornberg & Baker, 1991). This remarkable property of high processivity is achieved in many cases by the attachment of the catalytic machinery to a "sliding clamp" that allows the polymerase to move rapidly along DNA while remaining topologically bound to it (Huang et al., 1981; Burgers & Kornberg, 1982; Tsunemoto & Stillman, 1990; Lee et al., 1991a; Jarvis et al., 1991; Stukenberg et al., 1991). The proteins that form the clamps are distinct from the enzymatic subunits and are unusual in that they do not have any intrinsic affinity for DNA. The formation of the tight DNA clamp requires the presence of certain accessory proteins, or "clamp loader" molecules, that bind to single strand/duplex junctions and then assemble the clamps onto DNA in an ATP-dependent process. These highly processive DNA polymerases can therefore be thought of as having three components: a core catalytic machinery (including polymerase and exonuclease), a protein that forms the sliding DNA clamp, and accessory proteins that load the clamp onto DNA (see Table 1).

The best-understood of these highly processive polymerases are DNA polymerase III holoenzyme (Pol III) of *Escherichia coli* and the DNA polymerase of T4 bacteriophage. In addition a great deal of information is now emerging for the eukaryotic DNA polymerases δ and ε. The mechanism for achieving high processivity was first described for the T4 DNA polymerase, which has only four component proteins: a polymerase/exonuclease (gene 43), a clamp (gene 45) and two other proteins (genes 44, 62) that act to load the clamp onto DNA (Newport et al., 1980; Huang et al., 1981; Nossal & Alberts, 1984). The early studies on the processivity of T4 DNA polymerase introduced the descriptive phrase "sliding clamp" for the collective action of

† Author to whom all correspondence should be addressed.
all three accessory proteins. The E. coli Pol III is more complicated, and the holoenzyme contains a polymerase (α subunit), an exonuclease (ε subunit), a sliding clamp (β subunit), accessory proteins that assemble the clamp at the primer, which are also required for maximal efficiency of lagging strand synthesis (γ complex, containing the γ, δ, δ', χ, θ subunits), the τ subunit, which dimerizes the polymerase for replication of duplex DNA, and the 9 subunits (for review, see Kornberg & Baker, 1991; McHenry, 1991; O’Donnell, 1992). The eukaryotic system contains polymerase subunits (δ or ε), a clamp protein (proliferating cell nuclear antigen, or PCNA) and 8 accessory proteins (replication factor C, RF-C) (for review, see So & Downey, 1992). In this review we describe the crystal structure of the DNA clamp (β subunit) of the chromosomal replicase of Escherichia coli and discuss briefly how high processivity is achieved by this polymerase as well as those of eukaryotic systems and the bacteriophage T4.

E. coli DNA polymerase III holoenzyme

Pol III, which is responsible for the replication of the E. coli chromosome, can be purified from bacteria as an intact holoenzyme that is a complex of at least 10 different polypeptide subunits (Mak & Kornberg, 1988). These subunits together enable the holoenzyme to be bound very stably to DNA, and carry out highly processive replication of the leading and lagging strands. The high processivity of Pol III is readily demonstrated (Fay et al., 1981; Burgers & Kornberg, 1982; Maki & Kornberg, 1988; O’Donnell & Kornberg, 1985; Stadwell & O’Donnell, 1990). For example, model in vitro reactions utilize primed single-stranded circular DNA as a template that is replicated using Pol III and labeled nucleotides. When the time course of the reaction is monitored by autoradiography, it is seen that partially completed duplex circles do not accumulate before the entire circle is replicated. The lack of intermediates in the reaction, even with sub-stoichiometric amounts of Pol III, means that essentially all holoenzyme molecules that initiate the process go on to complete the replication of the entire circle without dissociation (O’Donnell & Kornberg, 1985).

The β subunit of Pol III

Pol III holoenzyme can stay bound to a template for longer than the time required to replicate the entire E. coli chromosome (30–40 min) (McHenry, 1991). However, the catalytic core of the enzyme (α, ε, ϊ) does not have this property and its ability to replicate DNA resembles that of E. coli Pol I, which functions to fill in relatively short gaps and incorporates just a few nucleotides at a time before dissociation (Fay et al., 1981). The high processivity of Pol III can be reconstituted by mixing the core with the β subunit and the γ complex (γ, δ, δ', χ, θ) in a two-stage process (Wickner, 1976; O’Donnell, 1987; Maki & Kornberg, 1988). First, the γ complex hydrolyzes ATP to transfer the β subunit to the primed template. The core polymerase then assembles with the β subunit on DNA to form the processive polymerase. Once on DNA, the β subunit confers complete processivity onto the polymerase, even upon removal of the γ complex (Stukenberg et al., 1991; Onrust et al., 1991). The role of the γ complex has been described as that of a “molecular matchmaker” (Sancar & Hearst, 1993) in the sense that it acts to bring DNA and the β subunit together, but is not itself an intrinsic part of the clamp. Once the loading process is completed, the resulting complex between DNA and the β subunit is very stable. However, experiments using restriction endonucleases have revealed that if circular DNA is cut after the β subunit is clamped on, the β subunit completely separates from the DNA by sliding to the site of the break and falling off (Stukenberg et al., 1991). In contrast to site-specific DNA-binding proteins that make specific hydrogen-bonding or other stabilizing interactions with DNA, the β subunit is bound non-specifically, but very tightly, suggesting a topological rather than a chemical interaction with DNA. These biochemical experiments led to the proposal, made in advance of three-dimensional structural information, that the β subunit would be a ring-shaped molecule that encircles DNA (Stukenberg et al., 1991).

We have determined the structure of the β subunit by X-ray crystallography at 2.5 Å resolution (Kong et al., 1992). The overall structure is indeed that of a ring, of approximate diameter 80 Å, with a hole of diameter ~35 Å in the middle (Figs 1, 2). The protein forms a head-to-tail dimer, consistent

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Subunit composition of chromosomal replicases</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Polymerase</td>
</tr>
<tr>
<td>E. coli</td>
<td>Pol III core</td>
</tr>
<tr>
<td></td>
<td>(α, ε, ϊ)</td>
</tr>
<tr>
<td>T4 phage</td>
<td>Gene 43</td>
</tr>
<tr>
<td></td>
<td>Yeast/human</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Taken from O’Donnell et al. (1993).
with previous observations that the isolated protein is a dimer in solution (Johanson & McHenry, 1980). The 2-fold dimer axis is perpendicular to the face of the ring, the thickness of which is about that of one full turn of duplex B-form DNA (approx. 34 Å). Although the structure has been determined in the absence of DNA, several obvious features indicate that the protein is designed to wrap around the double helix with a minimum of specific interactions. The hole in the middle is large enough to accommodate A- or B-form DNA with no steric repulsions. Although the protein is strongly negatively charged (−22, if all the histidines are assumed to be neutral and −15 if they are assumed to be half charged), calculation of the electrostatic field generated by the molecule (Gilson et al., 1988) reveals a positive electrostatic potential in the middle of the ring, precisely where the negatively charged phosphate backbone of the DNA is expected to be (Fig. 2).

The crystal structure explains immediately the observed biochemical properties of the molecule. Figure 2a shows the molecular surface of the dimer (Nicholls et al., 1991), emphasizing the continuity of the ring-like structure, with no break evident at the dimer interfaces. The β subunit does not normally bind to DNA at low concentrations in the absence of γ complex and ATP, which presumably are required to open and close the ring. Once the ring is on closed circular DNA it does not readily fall off (Stukenberg et al., 1991), suggesting a stable interaction. Indeed, the dimer interface of the molecule appears to be particularly tight: it includes hydrogen bonds that are part of an inter-molecular β sheet, a small hydrophobic core as well as a number of ion pairs.

A completely unexpected feature of the structure is its symmetry, since the amino acid sequence has no internal homology (Fig. 1). There are three levels of symmetry. First, each monomer of the β subunit consists of three domains of identical chain topology and very similar three-dimensional structure (Fig. 3a). Second, each domain is roughly 2-fold symmetric in its architecture, with an outer layer of two β sheets providing a scaffolding that supports two helices (Figs 1 and 3). Third, the entire structure has circular symmetry: rotation of the structure results in a repeating pattern every 60°. Thus, the molecule has 12 helices lining the inner surface of the ring, and six β sheets forming the outer surface. Each of the domains consists of about 110 residues, and forms a compact and well-formed structure (Fig. 3b). Despite the simple architecture, each domain is quite clearly an independent folding unit, with a well-defined hydrophobic core.
Figure 2. Molecular surfaces of the β subunit. The solvent-accessible surface of the molecule is displayed using the program GRASP (Nicholls et al., 1991). The surface is colored on the basis of the local electrostatic potential, calculated using GRASP with a dielectric constant of 80 for the solvent and 2 for the protein interior (Gislon et al., 1988). The extreme ranges of red and blue represent electrostatic potentials of \(-4\) to \(+4\) \(k_B T\), where \(k_B\) is the Boltzmann constant and \(T\) is the temperature. Negative charges were assigned to the terminal atoms of Asp and Glu side-chains, and positive charges to those of Arg and Lys residues. All other residues (including His) were uncharged, and no partial charges were used. The molecular surface and electrostatic potential on the dimer is shown in (a), and that of an isolated monomer is shown in (b). A schematic representation of the chain fold of a monomer is shown in (c), in the same orientation as in (b). Helices are represented as cylinders, and strands of β sheet as arrows, displayed using the “O” program of Jones et al. (1991).
Figure 3. The modular structure of the $\beta$ subunit (from Kong et al., 1992). (a) Schematic diagram of the secondary structure of one domain of the $\beta$ subunit. $\alpha$ Helices are shown as rectangles, and $\beta$ strands as arrows. The topology is 2-fold symmetric about the point represented by the ellipse. The structure can be generated by repeating the pattern shaded in gray. (b) Ribbon diagram (Priestle, 1988) of a domain. The approximate 2-fold axis is indicated by the arrow. (c) Ribbon diagram of one monomer. The long loops that connect the domains are highlighted in black. The $\beta$ sheets that are continued across domain boundaries are shaded in gray.

consisting of about 20 residues. This is an important point, which allowed us to make a connection to PCNA$^+$ and gene 45 protein (see below).

The circular architecture of the $\beta$ subunit is quite rigidly maintained by the network of hydrogen bonds in the $\beta$ sheets, and by a curious feature of the inter-domain connections. Instead of the domains being connected by short loops in a head-to-tail fashion, the topology of each domain is such that the polypeptide chain connecting two domains

$\mathbf{\dagger}$ Abbreviation used: PCNA, proliferating cell nuclear antigen.

runs forward along the surface of the second domain before connecting with it (Fig. 3c). These long loops make a number of interactions with the outer surface of the ring, and appear to act like bands on a barrel. Each loop spans one of the two inter-domain interfaces in each monomer, and a number of hydrophobic interactions between the loops and residues at each interface appear to buttress the inter-domain connections.

Models of standard A- or B-form DNA can be inserted into the center of the hole in the $\beta$ subunit dimer with no steric hindrance: the closest distance between extended arginine side-chains and the phosphate backbone of DNA is more than 5 Å. The 12 helices that line the inner surface of the ring are oriented perpendicular to the local directions of the phosphate backbone of B-form DNA, and are long enough to span the major groove. Each of the helices of the protein interacts with a different combination of the major and minor grooves, a feature that would assist motion along the DNA by preventing any particular local interaction from dominating. The non-specific nature of these interactions is highlighted by comparing the structure of the $\beta$ subunit with that of a site-specific DNA-binding protein in complex with DNA (Fig. 4). Two of the helices of the engrailed homeodomain (Kissinger et al., 1990) can be aligned approximately with two helices of the $\beta$ subunit. Although this structural alignment is fortuitous, it has an interesting illustrative consequence: the DNA of the homeodomain co-crystal structure inserts into the middle of the $\beta$ subunit ring (Fig. 4). Note that the "recognition helix" of the homeodomain, which engages the major groove for site-specific recognition of base-pairs, has no corresponding element in the $\beta$ subunit. Rather, the inner surface of the $\beta$ subunit stays away from the grooves of DNA, and is likely instead to form loose water-mediated interactions with the phosphate backbone (Kong et al., 1992).

The $\beta$ subunit dimer is formed by a head-to-tail interaction of the monomers. Although the symmetry of the structure is consistent with formation of a head-to-head dimer (the relative orientation of the sheets and helices is approximately the same in both cases), only the head-to-tail dimer would form at reasonable ionic strengths due to electrostatic complementarity of the two ends of the monomer (Fig. 2b). This has an important consequence: the two faces of the ring have quite different properties. Interestingly, one face has a number of protruding loops, which may be important in binding to the polymerase or the $\gamma$ complex (Fig. 2c). This asymmetry in its faces may allow the $\beta$ subunit to provide directionality for the core polymerase once it is correctly oriented on DNA by the $\gamma$ complex.

The ATPase accessory complex of Pol III

The picture of a polymerase core riding along DNA by virtue of being bound to the ring-shaped $\beta$
subunit suffices to explain continuous synthesis of a single DNA template. In the cell, however, the lagging strand is synthesized in many discontinuous pieces known as Okazaki fragments (Kornberg & Baker, 1991). Pol III holoenzyme is thought to contain two active polymerase subunits, one for each strand (Johanson & McHenry, 1984; Maki & Kornberg, 1988; Studwell & O’Donnell, 1993). However, the number of polymerase molecules in the cell is much smaller than the number of Okazaki fragments, and so the catalytic units must move from one fragment to the initiation site of the other.

The accessory proteins of the polymerase play a critical role in cycling the lagging strand polymerase between sites of discontinuous DNA synthesis (O’Donnell, 1987). Substantial advances have been made in the last year regarding the protein–protein interactions and activities of these accessory proteins. The genes encoding δ, δ', θ, γ and ψ have been cloned and the proteins have been over-expressed, purified and characterized (Carter et al., 1992; Xiao et al., 1993a; Dong et al., 1993; Onrust & O’Donnell, 1993; Studwell-Vaughan & O’Donnell, 1993; Xiao et al., 1993b). The genes encoding γ and τ have been cloned previously, and it had been shown that these two proteins arise from the same gene, and that γ corresponds to the N-terminal two-thirds of τ, due to a frame shift event that leads to a premature stop codon (Tsuechihashi & Kornberg, 1990; Flower & McHenry, 1990; Binkawa & Walker, 1990). With all the components of the holoenzyme available as purified individual proteins, the subunit–subunit and subunit–DNA interactions have been probed by gel filtration and reconstitution analyses.

Based on these results, a simple hypothetical scheme for the action of Pol III holoenzyme at a replication fork is proposed (Fig. 5), in which one catalytic core is clamped by a β subunit ring, and carries out uninterrupted synthesis of the leading strand. The lagging strand polymerase must rapidly dissociate from DNA upon completion of an Okazaki fragment, and reassociate with a new RNA primer for the next fragment. However, the lagging strand polymerase is also clamped to DNA by a β ring. Discontinuous replication of the lagging strand is made possible by the dynamics of Pol III action.

Figure 4. Comparison of the structure of the β subunit and the engrailed homeodomain. The structure of the DNA complex of the homeodomain (Kissing & et al., 1990) is shown as a red ribbon (for the protein) and as yellow and orange bonds (for the DNA). The G' atoms of the β subunit are drawn in blue. Two of the helices of the homeodomain were aligned with 2 of the helices of the β subunit, resulting in the structure shown, with the DNA of the homeodomain passing through the hole in the β subunit dimer. This comparison is for illustrative purposes only.

Figure 5. Highly simplified model for the action of Pol III holoenzyme at a replication fork. The parent double-stranded DNA is at the top in each of the 3 pictures, and is separated into single stranded templates by a helicase. A primase continues synthesizing primers on the lagging strand (the strand on the right). Pol III holoenzyme is shown as a dimer that is held together mainly by the τ subunit. The γ complex is depicted as an elongated jaw, and the β subunit as a ring on DNA. The leading strand (on the left) is replicated by one polymerase core that is clamped by one β ring. The γ complex loads a β clamp onto the next primer, in a step that consumes ATP (diagram A → diagram B). The polymerase core expels its β clamp upon completing an Okazaki fragment and then binds to the β clamp on the next primer (diagram B → diagram C). In this fashion the core jumps from one β clamp to the next one. Biochemical experiments have shown that such rapid cycling is possible (O’Donnell, 1987; Studwell et al., 1990).
in which the β clamp separates from Pol III only after a primer is extended to the very last nucleotide (Fig. 5B), thus freeing the polymerase to jump rapidly to the next primer (Fig. 5C) provided that the γ complex has endowed the primer with a β clamp. The β clamps that are left behind on the DNA product could, in principle, be used by other enzymes to couple their action to DNA replication (as in activation of phage T4 late genes discussed below).

The genes for some of the equivalent accessory proteins of eukaryotic DNA polymerases (RFC complex) have also recently been cloned (Chen et al., 1992a,b). Analysis of the sequences of the Pol III accessory proteins, and comparison with those for the eukaryotic RFC complex and T4 gene 44 protein has revealed a surprising redundancy in the sequences as well as similarity between the bacterial, phage and human sequences (O'Donnell et al., 1993). The results can be summarized as follows: E. coli proteins δ, γ and τ are homologous to the four known sequences of the human RFC (activator 1) complex, and to the gene 44 protein of T4 phage. δ is 27% identical in sequence with γ, which strongly suggests that they will have very similar three-dimensional structures (Sander & Schneider, 1991). The sequence of τ is identical to the N-terminal two-thirds of τ (except for the last amino acid). The homology of the four human proteins is weaker, but clearly recognizable (O'Donnell et al., 1993). This structural redundancy appears to be an intrinsic feature of these "clamp loader" molecules: the T4 system has only two different proteins (genes 44 and 62), but these are present in a 4:1 ratio (Jarvis et al., 1989). No structural information is available as yet for any of these accessory proteins. The δ subunit of Pol III, which is required for efficient action of the γ complex, has been crystallized, and its structure determination is underway (B.G. Guenther, X.-P. Kong, R. Onrust, M.O'D. & J.K., unpublished results).

The only DNA-dependent DNA polymerase for which a three-dimensional structure has been determined is the Kleenow fragment of E. coli Pol I (Ollis et al., 1985). Unfortunately, Pol I does not share sequence similarity with the polymerase subunit (α) of Pol III (Tomasiewicz & McHenry, 1987), and it has not been possible to draw any conclusions about how α subunit might interact with the rather complicated machinery for achieving processivity. The DNA polymerase of T7 bacteriophage (T7 Pol) is an interesting contrast to Pol III in this respect, because it uses just one small protein to achieve a comparable level of processivity. T7 Pol is by itself not a processive polymerase. However, the host protein thioredoxin binds tightly to T7 Pol in a 1:1 ratio and converts it to a highly processive polymerase (huber et al., 1988). The structure of Pol I, determined by Steitz and co-workers (Ollis et al., 1985), is expected to resemble that of T7 Pol in its general aspects. This structure resembles a hand-gripping DNA between its palm and fingers, and thioredoxin may act by closing off one of the open grooves on the surface of the protein, thereby trapping the DNA within a channel.

PCNA and T4 phage gene 45 protein

PCNA (originally named cyclin) was discovered, as a nuclear protein that correlates with the proliferative phase of the cell cycle (Bravo & Celis, 1980). It was also characterized independently as a nuclear antigen in patients suffering from a particular subtype of the autoimmune disease, systemic lupus erythematosus, and was named "proliferating cell nuclear antigen" (Miyachi et al., 1978; Mathews et al., 1984). The name "cyclin" is no longer used, because the amount of PCNA in the cell is not controlled by specific degradation (Morris & Mathews, 1989). This is in contrast to the cell cycle proteins called cyclins that regulate mitosis and are rapidly destroyed by proteolysis after cell division (Evans et al., 1983).

PCNA was found to stimulate the activity of DNA polymerase δ (the leading strand polymerase in eukaryotic cells) by several hundredfold, and subsequent work identified it as a processivity factor similar to β subunit of E. coli and gene 45 protein (Tan et al., 1986; Bravo et al., 1987; Prelich et al., 1987; Bauer & Burgers, 1988; Tsurimoto & Stillman, 1990; Burgers, 1991; Lee et al., 1991a; Ng et al., 1991; Yoder & Burgers, 1991). The protein has been cloned from a variety of sources including human, rat, yeast and plants (Almendral et al., 1987; Suzuki et al., 1989; Travali et al., 1989; Bauer & Burgers, 1990).

The mechanism of PCNA is less well understood than that of the E. coli β subunit, but there appears to be considerable functional correspondence between the two proteins. There are at least three different polymerases working at the replication fork in mammalian cells. Of these, Pol α has primase function and is also thought to replicate the lagging strand (Downey et al., 1990). The leading strand is replicated by Pol δ, which is made processive by PCNA. Recent results show that PCNA interacts with Pol ε, a proof-reading DNA polymerase essential for yeast growth that might also participate in lagging strand synthesis (Burgers, 1991; Lee et al., 1991b). Apart from its role in DNA replication, PCNA is also involved in excision repair (Shivji et al., 1992). Intriguingly, Beach and co-workers (Xiong et al., 1993) have shown that PCNA interacts with D-type cyclins, proteins that are implicated in controlling the transition from the G1 phase of the cell cycle to the S phase where DNA synthesis occurs. This study shows that PCNA can be co-precipitated with cyclin D1 and with protein kinases involved in cell cycle regulation (edk2, edk4, edk5). The functional relevance of these interactions are not yet understood.

Gene 45 protein is the processivity factor for T4 phage DNA polymerase, but is unrelated in sequence to β subunit or PCNA (Huang et al., 1981; Jarvis et al., 1991). Like PCNA, it is implicated in at least one other function other than DNA replica-
tion: the transcription of bacteriophage late genes is enhanced by interactions between RNA polymerase and the three proteins that are required for DNA polymerase processivity (Herendeen et al., 1989). Most interestingly, cryoelectron microscopy has shown that at least one of these accessory proteins (presumably gene 45 protein) forms a ring around DNA (see below) (Gogol et al., 1992).

Possible structural similarity between β subunit, PCNA and gene 45 protein

Given the functional similarity between these three proteins, a natural question is whether they are structurally related despite a lack of sequence similarity. T4 gene 45 protein and PCNA are approximately two-thirds the size of β subunit (β subunit, 306 residues; gene 45 protein, 220 residues; Saccharomyces cerevisiae PCNA, 258 residues; human PCNA, 261 residues). Given their smaller size and knowing the three-domain structure of β subunit, one can speculate that these proteins may be composed of two domains each and thereby function as trimeric molecules that assemble six domains around DNA in sets of two per monomer (Kong et al., 1992). A trimeric functional form for these proteins is consistent with previous measurements of the molecular weights of T4 gene 45 protein and S. cerevisiae PCNA in solution (Bauer & Burgers, 1988; Jarvis et al., 1989).

The sequences of T4 gene 45 protein and yeast and human PCNA were compared with those of the three domains of β subunit, using the following criteria. Insertions and deletions were only allowed between the secondary structural elements of the β subunit, and the hydrophobic core and other buried residues of β were matched with hydrophobic or neutral polar residues in the other sequences (Kong et al., 1992). A plausible alignment results for the PCNA and gene 45 sequences with the β subunit (Fig. 6). A different alignment between yeast PCNA and T4 gene 45 protein has previously been reported, but this included several very long insertions and deletions (Tsurimoto & Stillman, 1990).

Our alignment suggests that PCNA and gene 45 might indeed form trimeric rings that are roughly similar in architecture to the β subunit. Intriguingly, low-resolution cryoelectron microscopic studies of the three accessory proteins of T4 phage (genes 44, 62 and 45) show features that bear a striking resemblance in shape and size to our structure of the β subunit (Gogol et al., 1992). The electron micrographs show the presence of multiple "hash marks" on DNA, consistent with ring-like structures through which DNA is threaded. The structures are disk-shaped, with a diameter of approximately 80 to 90 Å and a thickness of about 35 Å (Gogol et al., 1992), which matches closely the dimensions of the β subunit. However, these rings have not been identified specifically as being formed by gene 45 protein, since all three accessory factors are present.

It must be stressed that the sequence alignment between the various processivity factors is extremely weak. Comparison of the yeast and E. coli sequences reveals less than 5% sequence identity per domain, which is well below the level required to

---

**Figure 6.** From Kong et al. (1992). Alignment of the sequences of the domains of β subunit with human and yeast PCNA. The 3 domains of β subunit are labeled BETA 1, BETA 2, BETA 3, and are aligned based on 3-dimensional structural correspondence alone. The human and yeast sequences have been split into 2 domains, labeled 1 and 2, respectively. The secondary structural elements within the β subunit domains are boxed and labeled. The shaded bars have different meanings within the β subunit sequences and the PCNA sequences. Within β subunit, the bars indicate all the residues that are completely buried, as judged by solvent accessibility calculations using a water-sized probe and the intact dimer. No criteria of sequence similarity were used in the selection of these residues. For the PCNA sequences, the shaded bars indicate residues that are deemed to be similar in property to those that are buried in the β subunit sequence. Unless a charged residue is present in the corresponding β subunit position, the residues Asp, Glu, Gln, Asn, His, Arg and Lys have been excluded from the gray bars in PCNA. The asterisks underneath the sequences indicate positions where at least 2 of the PCNA domains have sequence identity with amino acids in the β subunit domains. The numbers to the right are the last residue numbers in each row.
predict structural correspondence with certainty (Sander & Schneider, 1991). The distance relationship between the sequences is emphasized by the fact that the recently developed profile method of sequence comparison (Bowie et al., 1991) fails to detect any similarity between these proteins. A conclusive understanding of PCNA and gene 45 architecture must await determination of the three-dimensional structures. Crystallization of gene 45 protein was reported 10 years ago (Thaller et al., 1984), but no three-dimensional structure has resulted, due to the limited diffraction quality of the crystals. Crystals of yeast PCNA that diffract to 2.4 Å have been obtained recently (T. S. R. Krishna, X.-P. Kong, P. Burgers & J. K. Kuriyan, unpublished results).

**Conclusion**

The structure of the β subunit explains how a protein that is highly negatively charged can be bound tightly, yet non-specifically, to DNA and is a particularly clear example of the connection between structure and function. The structure does, however, raise a number of intriguing questions. It provides no clues about the mechanism for the assembly of the clamp on DNA. Since a closed ring cannot be loaded onto DNA, what is the mechanism by which the γ complex puts β onto DNA? How does the clamp hold onto the polymerase core, and how does it let go for discontinuous lagging strand synthesis? Understanding these mechanisms will require a sustained effort at obtaining the three-dimensional structures of the components of these polymerases by crystallography and NMR and, ultimately, putting all the pieces together by electron microscopy.

This work was supported in part by grants from the National Institutes of Health (GM43094 to J.K. and GM39830 to M.O.D.). J.K. is a Pew Scholar in the Biomedical Sciences. We acknowledge the central contributions of Xiang Peng Kong and Rene Onrust in determining the structure of the β subunit, and thank them for making possible a stimulating collaboration. We thank Charles H. Williams, Jr. and Peter Model for bringing us together initially, and Stephen K. Burley and Greg Petsko for numerous suggestions. We are grateful to Xiang Peng Kong for Figure 4 and to P. Todd Stulkenberg for Figure 5.

**References**


Edited by W. Hendrickson