Three-Dimensional Structure of the β Subunit of E. coli DNA Polymerase III Holoenzyme: A Sliding DNA Clamp

Xiang-Peng Kong,* Rene Onrust,† Mike O'Donnell,†‡
and John Kuriyan*§
*Laboratory of Molecular Biophysics
§Howard Hughes Medical Institute
The Rockefeller University
New York, New York 10021
†Microbiology Department
Hearst Research Center
‡Howard Hughes Medical Institute
Cornell University Medical College
New York, New York 10021

Summary

The crystal structure of the β subunit (processivity factor) of DNA polymerase III holoenzyme has been determined at 2.5 Å resolution. A dimer of the β subunit ($M_r = 2 \times 40.6$ kd, 2×366 amino acid residues) forms a ring-shaped structure lined by 12 α helices that can encircle duplex DNA. The structure is highly symmetrical, with each monomer containing three domains of identical topology. The charge distribution and orientation of the helices indicate that the molecule functions by forming a tight clamp that can slide on DNA, as shown biochemically. A potential structural relationship is suggested between the β subunit and proliferating cell nuclear antigen (PCNA, the eukaryotic polymerase δ [and ε] processivity factor), and the gene 45 protein of the bacteriophage T4 DNA polymerase.

Introduction

DNA polymerases are enzymes that duplicate the information content of DNA by catalyzing the template-directed polymerization of nucleic acids. A distinction can be made between polymerases that are primarily involved in the replication of chromosomal DNA during cell division and those that normally operate on shorter stretches of template during, for example, the repair of damaged DNA. Polymerases in the latter class are generally nonprocessive, i.e., they polymerize only a few nucleotides before dissociating from the template (Kornberg and Baker, 1991). In contrast, the chromosomal replicative polymerase of Escherichia coli, DNA polymerase III (PolIII) holoenzyme, is distinguished by its ability to perform rapid replication (750 bases per second) of very long stretches of DNA without dissociation (Fay et al., 1981; Burgers and Kornberg, 1982; O'Donnell and Kornberg, 1985; Kornberg and Baker, 1991). This property is conferred upon the enzyme by the presence of associated proteins that clamp the polymerase onto primed DNA, in a process that expends ATP energy. This mechanism, first worked out in detail for E. coli PolIII holoenzyme and the bacteriophage T4 DNA polymerase system, appears to operate analogously in the eukaryotic DNA polymerases δ and ε (Kornberg and Baker, 1991). The three-dimensional structure of one poly-

merase has been determined by X-ray crystallography, that of the Klenow fragment of E. coli PolI (Ollis et al., 1985). Although this is not a highly processive polymerase, the structure has general relevance for understanding the mechanism of the enzymatic subunits of DNA polymerases. No structural information has yet been available, however, for any of the accessory proteins of the processive polymerases.

Intact PolIII holoenzyme is a complex of at least 10 different protein subunits (α, ε, θ, τ, γ, δ, δ', χ, ψ, and β) (Maki and Kornberg, 1988). The α subunit performs the catalytic polymerase function, and the ε subunit is the 3'-5' exonuclease. A three-subunit core polymerase subassembly of the holoenzyme, containing α, ε, and θ, is unable to act processively on its own, although it can fill in short single-stranded regions. The highly processive character of the holoenzyme can be reconstituted upon mixing the core polymerase with both the β subunit and the five-protein γ complex (γ, δ, δ', χ, and ψ) (Wickner, 1976; O'Donnell, 1987; Maki and Kornberg, 1988). The reconstitution of the processive polymerase proceeds in two distinct stages. In the first stage, the γ complex hydrolyzes ATP to transfer the β subunit to the primed template. In the second stage, the core polymerase assembles with the β subunit on DNA to form the processive polymerase. Thus, it is the β subunit that confers the remarkable processivity onto the core polymerase. Study of the minimal number of subunits required to assemble the processive polymerase showed that only the γ and δ subunits of the γ complex are needed to transfer β from solution to the primed template (O'Donnell and Studwell, 1990). Two subunits of the core polymerase, α and ε (as an αε complex), are needed for processive polymerization (Studwell and O'Donnell, 1990). Once on DNA, the β subunit confers complete processivity onto the αε polymerase, even upon subsequent removal of the γ complex (Stukenberg et al., 1991).

Once the γ complex has performed the operation of clamping the β subunit onto DNA, the β subunit is very strongly bound in that it cannot be easily separated from circular DNA. It has, however, been shown to slide freely along duplex DNA, consistent with its role as a clamp that tethers the polymerase core to the template and moves along with the polymerase during replication (Stukenberg et al., 1991). 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 DNA by sliding to the site of the break and falling off (Stukenberg et al., 1991). These and related experiments in the same study suggest that, in contrast to site-specific DNA-binding proteins such as transcription factors or nucleases, which make specific hydrogen-bonding or other stabilizing interactions with DNA, the β subunit is bound to DNA mainly by virtue of its topology rather than by stabilizing interactions. It was proposed that the β subunit might form a closed ring, and that one role of the γ complex might be to open and then close the ring around DNA,
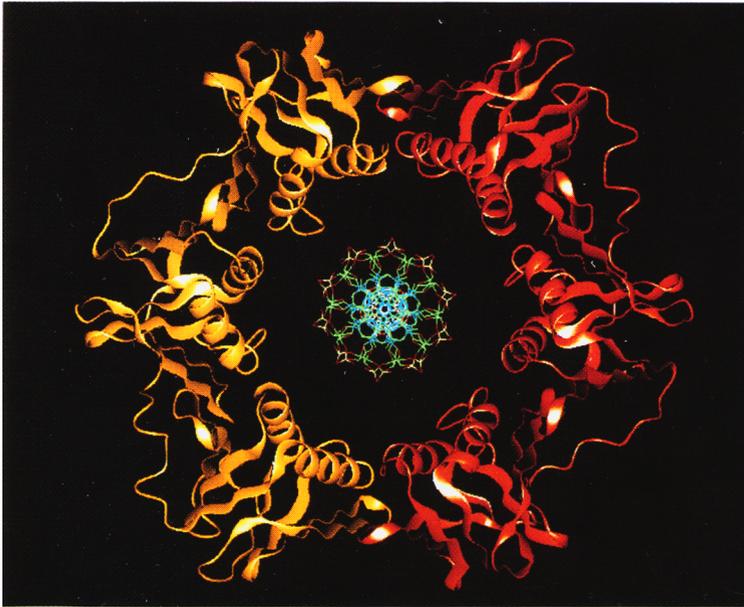
Figure 1. Ribbon Representation of the Polypeptide Chain of a β Subunit Dimer, Looking Down the 2-Fold Axis of the Ring

The α helices are shown as spirals and the β sheets as flat ribbons. The two monomers are colored yellow and red. A standard model of B-form DNA (Saenger, 1984) is in the middle of the structure, represented in stick form with phosphorus, oxygen, nitrogen, and carbon atoms colored yellow, red, blue, and green, respectively. The DNA structure is hypothetical, and is placed in the geometric center of the β subunit ring with the helix axis aligned along the 2-fold rotation axis of the ring. Figures 1, 2, 3, 6, and 8 were generated using the program QUANTA (Polygen Corp.). Refined atomic coordinates and X-ray structure factors are being deposited in the Protein Databank.
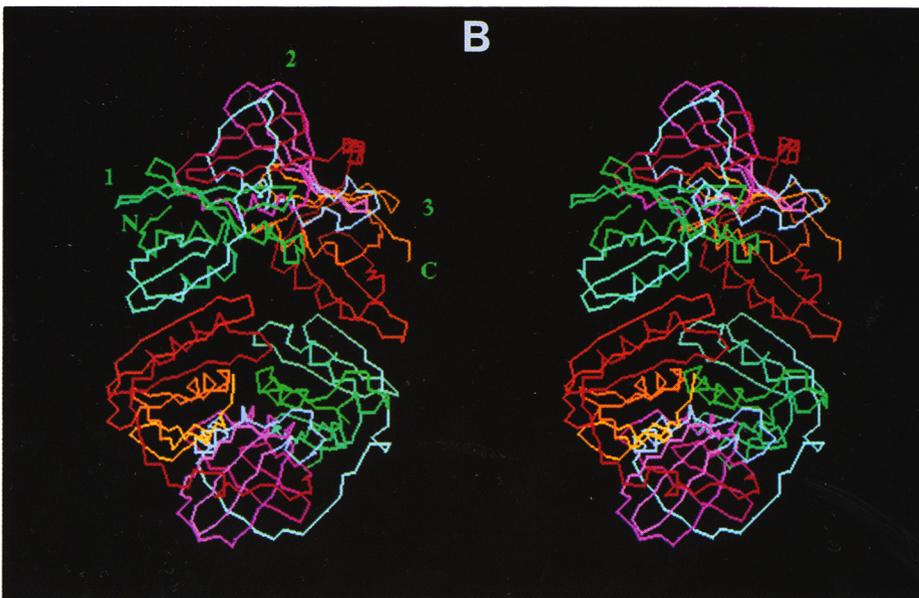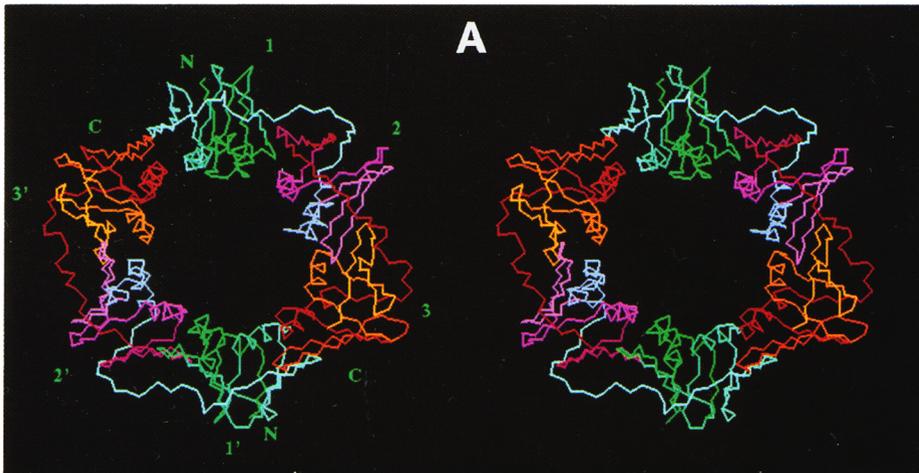
effectively trapping the β subunit on DNA (Stukenberg et al., 1991).

As a step toward a detailed understanding of the molecular basis of the processive properties of PolIII holoenzyme, we have crystalized the β subunit and determined its three-dimensional structure by X-ray diffraction at 2.5 Å resolution. In pleasing congruence with previous speculation, we find that the structure is indeed that of a closed ring, with overall shape similar to that of a donut or toroid. In this paper we discuss the consequences of this structure for the function of the β subunit, and suggest a possible structural relationship between this protein and its functional equivalent in the eukaryotic polymerase δ (and ε) replicases, the processivity factor PCNA (proliferating cell nuclear antigen), and the bacteriophage T4 DNA polymerase gene 45 protein. This represents the first atomic resolution view of any of the accessory proteins of processive polymerases involved in DNA replication, and complements the previously reported structure of the polymerase I Klenow fragment (Ollis et al., 1985).

Results and Discussion

Architecture of the β Subunit Dimer

The β subunit forms a head-to-tail dimer in the crystal, consistent with previous observations that the isolated protein is a dimer in solution (Johanson and McHenry, 1980). Representations of the polypeptide backbone of the dimer are shown in Figures 1 and 2, and a space-filling representation of all atoms is in Figure 3. The overall structure is that of a star-shaped ring, of approximate diameter 80 Å, with a hole of diameter ~35 Å in the middle (Figure 1). 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 (~34 Å).

The starlike shape of the ring is due to an unexpected feature of the structure: it is internally symmetric, with each monomer of the β subunit consisting of three structural domains of identical chain topology and very similar three-dimensional structure. Each domain is roughly 2-fold symmetric in its architecture, with an outer layer of two β sheets providing a scaffold that supports two α helices. Replication of this motif around a circle results in a rigid molecule with 12 α helices lining the inner surface of the ring, and with 6 β sheets forming the outer surface (Figure 1). Surprisingly, this structure is reminiscent of the ring-shaped pentameric assembly of the B subunits of cholera toxin-related heat-labile enterotoxin from E. coli (Sixma et al., 1991). The toxin structure has 5 α helices that line the inner surface of the ring with antiparallel β sheets forming the outer surface. The hole in the ring is plugged by an extended polypeptide strand from the A subunit. The hole

Figure 3. Space-Filling Model of the β Subunit Dimer with B-Form DNA

One monomer is colored red and the other yellow. The radius of the spheres corresponds to the van der Waals radius of the corresponding atom. Hydrogen atoms are not explicitly displayed, but manifest themselves as increased radii for atoms that they are bonded to. The hypothetical model of B-form DNA is as in Figures 1 and 6, and is shown with one strand colored white and the other green. The double helix passes through the hole in the β subunit dimer with no steric repulsions.

in the toxin structure (diameter 11 Å) is much smaller than that observed in the β subunit dimer, and the detailed topology of each of the B subunits of the toxin is different from that of the β subunit modules.

The secondary structure and chain topology of a domain is particularly simple and is shared by all three domains in one monomer: two adjacent antiparallel α helices are flanked by two 4-stranded antiparallel β sheets (see schematic diagram in Figure 4A). One of the β sheets forms the commonly found "greek key" motif (Branden and Tooze, 1991). The other β sheet contains the N- and C-termini of the chain. If an imaginary connection is drawn between the termini, this β sheet also forms a greek key, and a striking 2-fold symmetry in the chain topology is revealed (Figure 4A). Although there is insufficient sequence similarity to draw conclusions about the evolution of this fold, we note that one domain can be generated by duplication of a βαββ motif (Figure 4A). The chain topology diagram also reveals the simple principle underlying the architecture of the entire ring. The two outer strands of β sheets in one domain form hydrogen-bonding interactions with corresponding strands in two adjacent domains, contin-

Figure 2. Cα Connectivity of the β Subunit Dimer

The stereo diagrams are colored based on the sequence number of the residues, with the colors smoothly varying in the order green, light blue, purple, red, and yellow, with increasing sequence number within each monomer, from the N-terminus to the C-terminus. The dimer interfaces are between the domains colored green/blue and red/yellow. (A) The 2-fold axis is approximately perpendicular to the page. The approximate locations of the chain termini are marked with N and C. The domains are numbered 1, 2, 3 and 1', 2', 3' in the two monomers. (B) Edge view of the ring.
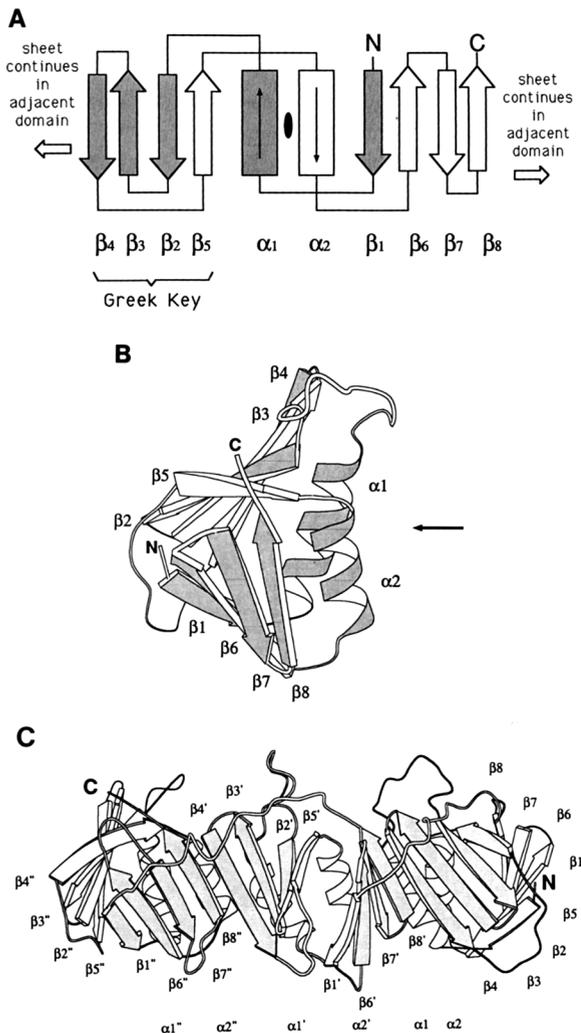
Figure 4. Secondary and Tertiary Structure of the β Subunit

The secondary structure elements were defined using the program DSSP (Kabsch and Sander, 1983).

(A) Schematic diagram of the secondary structure of one domain of the β subunit. α helices are shown as rectangles and β sheets as arrows. The topology is 2-fold symmetric about the point indicated by the ellipse. The structure can be generated by duplication of a βαββ unit, shaded in gray.

(B) Ribbon diagram (Priestle, 1988) of a domain. The approximate 2-fold symmetry axis is indicated by the arrow. The secondary structural elements are labeled as in (A).

(C) Ribbon diagram of one monomer. The β sheets that are continued across domain boundaries are shaded gray. These constitute two of the six β sheets in the dimer. The other two are the corresponding ones in the other molecule, and two others are formed by continuing the sheets (shown unshaded) across the molecular boundaries. Note the presence of three protruding loops on the top half of the molecule. The secondary structural elements are labeled as in (A), with the unprimed, singly primed, and doubly primed labels referring to the first, second, and third domain, respectively.

ued around the circle (Figure 4C). No distinctions are apparent between such β sheet extensions across internal domain boundaries as opposed to intermolecular contacts, i.e., the two dimer interfaces also form continuous antiparallel β sheets. These interactions lead to a com-

pletely closed circle with six "seamless" β sheets on the outer surface (Figures 1, 3, and 4).

Each domain consists of about 110 residues, and forms a compact and well-folded structure (Figure 4B). The 2-fold symmetry apparent in the topology diagram manifests itself as a very approximate 2-fold axis between the two helices (Figure 4B). Despite the simple architecture, each domain is quite clearly an independent folding unit, with a well-defined hydrophobic core consisting of about 20 residues. The symmetry axis relating the two molecules in the dimer is noncrystallographic, i.e., the two monomers are packed into different crystal environments and are crystallographically independent. The transformation that optimally superimposes the two molecules is a 180° rotation about an axis perpendicular to the plane of the ring, and results in a root mean square (rms) deviation of 0.40 Å between equivalent C_α positions (not including residues in the loops formed by residues 18–28 and 210–213, and the last three residues at the C-terminus). This axis is inclined by 12° with respect to the b-axis of the $P2_1$ crystal form, and therefore the holes of translationally related dimers do not line up to generate a linear tunnel in the crystal. Rotations about the dimer axis also superimpose the three different domains that constitute each monomer, and this basic structural unit repeats after every 60° rotation about the axis. Although the amino acid sequences of each domain are quite different, superimposition of the C_α positions of the three domains reveals that approximately 80% of these positions can be considered to be structurally analogous (Figure 5). The rms deviation in the positions of these C_α positions is 1.2 Å.

Potential Mode of Interaction with DNA

In the replication of a chromosome, the initial clamping of a β subunit dimer on DNA occurs at a primer terminus which is RNA/DNA hybrid duplex, presumably A-form (Saenger, 1984; Kornberg and Baker, 1991). Subsequent to this event, the β subunit can freely move along duplex DNA, presumably B-form (Stukenberg et al., 1991). The interactions of the β subunit with both the A and B forms of the double helix are therefore of interest. A-form DNA is similar to the B form in terms of the width of its cross section, and the local direction of the phosphate backbone with respect to the helix axis is very similar in both forms (Saenger, 1984). All subsequent analysis is focused on just these two gross features of the double helix, and the results are applicable to both the A and B forms.

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 locally specific interactions. The high symmetry of the structure is well suited to interact with the cylindrically symmetric DNA duplex, and the hole in the middle of the ring (of diameter ~35 Å, not including extended sidechains) is large enough to easily accommodate either the A or B forms of DNA (diameter ~25 Å) with no steric repulsion. Insertion of a model of either form of DNA into the ring results in a precise relationship between the common tilt in the orientations of all 12 α helices and the tilt of the phosphate backbone (see below). Finally,
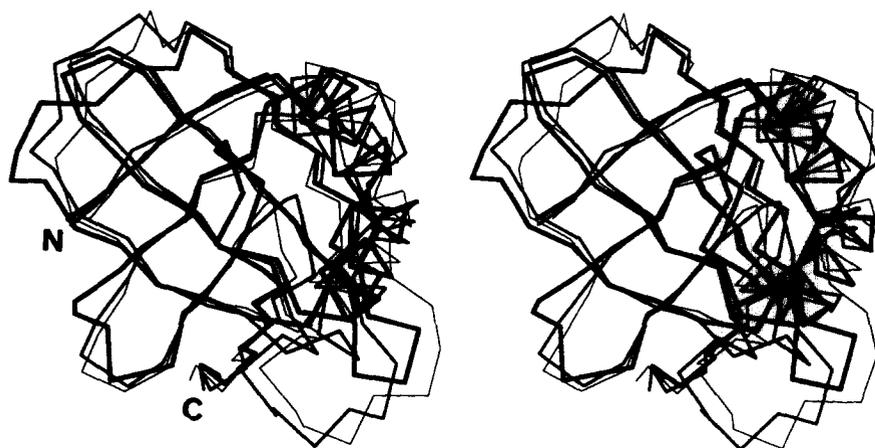
Figure 5. Structural Similarity between Three Domains of the Monomer

The C_α connectivity of each monomer is displayed in a stereo diagram after the three are optimally superimposed. The domains were considered in pairs, and a least-squares algorithm was used to align the structures (QUANTA, Polygen Corp.). This process was started by matching the central portion of one of the β sheets, then increasing the list of atoms included in the matching list and repeating the least-squares optimization. The process was repeated until the list included approximately 80% of the C_α atoms in each domain, with an rms deviation of 1.2 Å between corresponding C_α atoms. The remaining atoms are in the loops and in one of the helices, and are judged to have somewhat altered structure in the three domains. Domains 1 and 3 have very similar structure for both helices. In domain 2, one of the helices (α2) is displaced by about half a helical turn along the helix axis. This does not significantly alter its orientation with respect to the proposed model of DNA.

although the protein is strongly negatively charged, calculation of the electrostatic field generated by the molecule reveals a focusing of positive electrostatic field in the center of the ring, precisely where the negatively charged phosphate backbone of DNA is expected to be.

One consequence of the symmetrical arrangement of the six domains is that each of the 12 α helices has a similar tilt with respect to the axis of the ring. Assuming that duplex DNA passes through the middle of the ring, we generated a model of standard B-form DNA and placed it in the center of the β subunit (Figures 1, 3, and 6). The simple assumption that the duplex is perpendicular to the plane of the ring results in an intriguing relationship between the axes of all 12 helices and the phosphate backbone lining the major and minor grooves of DNA (Figure 6). The axis of each helix is almost precisely perpendicular to the local direction of the phosphate backbone, i.e., the helices span the major and minor grooves. This feature seems designed to prevent entry of the protein into either groove, and should facilitate rapid motion along the duplex. An additional consequence of aligning the DNA perpendicular to the ring is that each helix interacts with a different phase of DNA: if one helix spans the major groove, the one directly across the ring spans the minor groove. This is likely to lead to a damping out of the variation in interaction energy with the phosphate backbone as the protein moves across the grooves of DNA. The α helices are maintained in this precise orientation due to packing interactions with each other and with the underlying β sheet. Despite the intrinsic curvature of the β sheets, the strands, by and large, run in directions parallel to the helices, and therefore perpendicular to the DNA backbone in this model (Figures 2 and 4).

When viewed down the double helix, the phosphate backbone of DNA is 10-fold and 11-fold symmetric in projection for the A and B forms of DNA, respectively (such a projection is shown for the B form in Figure 1). Thus, there is no specific correspondence between the number of helices (12) that line the hole of β subunit and repetitive periods in DNA (10 or 11). Rather, the important features of the structure are the circular symmetry and the conservation of the helix–DNA backbone interaction. Given the dimensions of the ring and the tilt of the helices, the 12 helices pack against each other with no additional space remaining.

A dimer of β subunit contains 38 aspartate, 58 glutamate, 24 lysine, 50 arginine, and 14 histidine residues. The protein thus has a net charge of -22 if all histidines are assumed neutral, and -15 if half of them are charged. This negative charge is consistent with the inability of the β subunit to bind DNA without ATP activation by the γ complex. The electrostatic charge is not, however, uniformly distributed over the protein. We have found it useful to visualize the effects of the asymmetric charge distribution by calculating the electrostatic field generated by the protein, using a continuum electrostatic model that treats the protein as a low dielectric medium with embedded charges, immersed in a high dielectric solvent (water) of variable ionic strength (Gilson et al., 1988). While a quantitative analysis of charge effects would require a careful consideration of approximations introduced in this model, we focus solely on qualitative features that are reproducibly obtained with various calculational parameters (see legend to Figure 7).

Two important qualitative features consistently emerge from analysis of the computed electrostatic field. The outer edge and both faces of the ring are in regions of strongly negative electrostatic potential (Figures 7A and 7B). How-
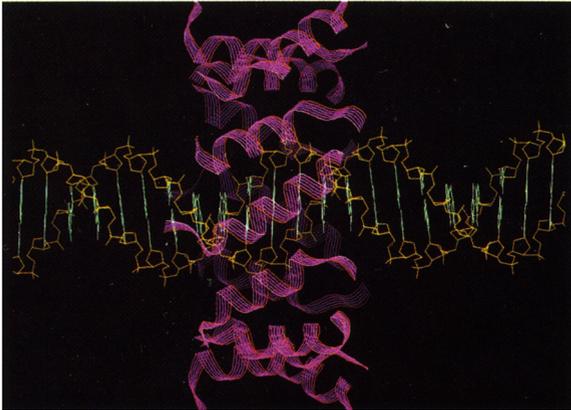
Figure 6. The Helices of the β Subunit Are Perpendicular to the Phosphate Backbone of the DNA

The 12 α helices of the β subunit dimer and a hypothetical model of B-form DNA are shown. The relative orientation of the double helix and the β subunit is exactly as in Figures 1 and 3. For clarity, the α helices are shown schematically in a ribbon representation, with the brighter ones being closer to the viewer. The central helix in front of the DNA is perpendicular to the phosphate backbone and spans the major groove of the double helix, while the one furthest behind it spans the minor groove. Each of the 12 helices has similar disposition with respect to the DNA backbone, but faces a different combination of the major and minor grooves. In this projection, the major and minor grooves are superimposed, with the major groove in the center of the figure being nearer the viewer.

ever, the surface of the hole has strongly positive electrostatic potential, and would be expected to interact favorably with the negatively charged backbone of DNA. Thus, the β subunit dimer can be described as a ring of negative charge surrounding a positively charged core (Figure 7). This focusing of positive charge is probably necessary to stabilize the dimer around DNA, as the dimer interface is itself mainly electrostatic in nature (see discussion below) and is unlikely to withstand repulsive interactions with DNA.

The other important feature is that the two faces of the ring are quite dissimilar in their properties, due to the head-to-tail dimer formation (Figure 7B). Although both are negatively charged, the negative electrostatic potential is clearly more extended on one face than the other. The other face has six prominent loops that extend away from the ring and appear to be well suited for interactions with another protein subunit (Figure 4C). We presume that only one of these faces interacts with the αε polymerase, which is known to bind strongly to the β subunit (Stukenberg et al., 1991). Although the faces of the β subunit are asymmetric, the symmetry of duplex DNA makes it unlikely that one orientation will be preferred over the other. It is likely that the γ complex interacts with the primer template junction in a specific orientation, and may thereby correctly orient the β subunit with respect to the primer terminus for productive interaction with the αε polymerase.
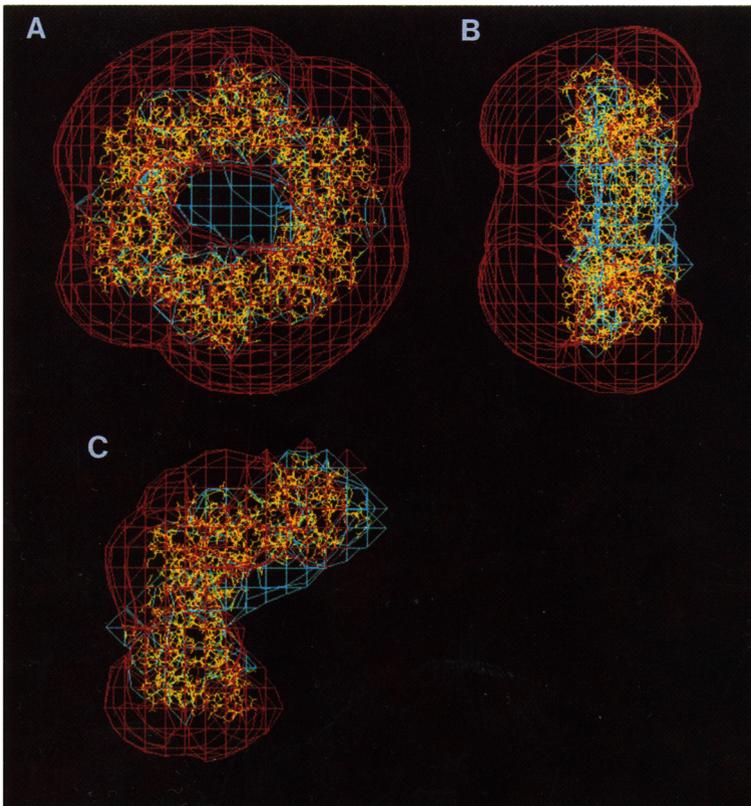
Figure 7. Electrostatic Potential Maps for the β Subunit Dimer and an Isolated Monomer

The maps are calculated using the programs DELPHI (Gilson et al., 1988) and INSIGHTII (Biosym Technologies). Lys and Arg residues have a single positive charge localized on the terminal nitrogen atoms of the sidechains. Asp and Glu residues have a single negative charge, localized on the terminal oxygen atoms of the sidechains. His sidechains have a 1/2 positive charge each. All other atoms in the molecule are considered neutral. Qualitatively similar results are obtained upon changing the charges on the protein to make all histidines neutral. The calculation was done assuming a uniform dielectric of 80 for the solvent and 2 for the protein interior. The ionic strength was set to zero. The red and blue mesh contours represent negative electrostatic potential (energy of $-2.5 \, k_BT/e$) and positive electrostatic field (energy of $+2.5 \, k_BT/e$), respectively. k_BT is the product of the Boltzmann constant and the temperature, and e is the charge of the electron. Two orthogonal views of the electrostatic potential of the dimer are shown in (A) and (B), and (C) is the potential for an isolated monomer.
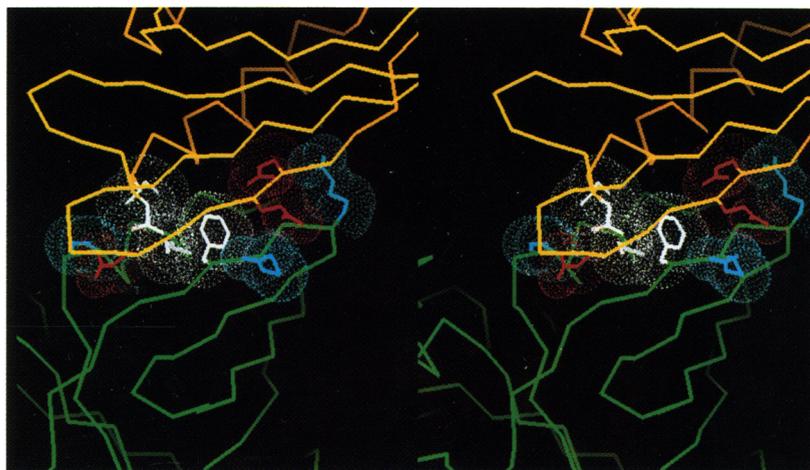
Figure 8. Stereo Diagram of the Dimer Interface

Bonds linking C$_\alpha$ atoms in one monomer are yellow, and those in the other are green. Note the continuation of the β sheet across the interface. Ten amino acid sidechains are shown at the interface, with the dot surfaces representing the van der Waals spheres of the atoms. Four of these are hydrophobic residues and are colored white (Phe-106 and Ile-108 from one monomer and Leu-273 and Ile-272 from the other). Arg-103, Arg-105, and Lys-74 are contributed by one monomer and are colored blue. Glu-300, Glu-303, and Glu-304 are from the other monomer and are colored red. For clarity, Arg-96 and Glu-303 are not shown.

The spacing between the hypothetical phosphate backbone of DNA and the protein sidechains is such that direct contacts are not likely to be made. The distance of closest approach between the terminal atoms of fully extended arginine residues lining the hole and the phosphate groups is expected to be no less than 3.5 Å. This suggests that water molecules will play an important role in mediating the protein–DNA interactions, which will also increase the ability of the protein to move along DNA without becoming attracted to any particular region. This lack of specificity makes it unlikely that useful crystals of a β subunit–DNA complex would ever be obtained. Molecular dynamics simulations of a solvated model of β subunit and DNA may be a useful approach toward understanding the details of this interaction.

The Dimer Interface and the Formation of the Head-to-Tail Dimer

The nature of the dimer interface is relevant to an understanding of the role played by the γ complex and ATP in assembling the β subunit on duplex DNA. The main feature of the interface is a continuation, across the molecular boundary, of β sheet structure (Figures 2, 8, and 10). This appears indistinguishable from β sheet continuation at interdomain boundaries within a monomer, and contributes at least four strong hydrogen bonds at each of the two interfaces. Relatively little surface area is buried upon formation of the β subunit dimer. The exposed surface area calculated using a water-sized probe of radius 1.5 Å (Lee and Richards, 1971) is 33,218 Å^2 for two isolated monomers, and decreases by only 8% (to 30,525 Å^2) upon dimerization.

In addition to the hydrogen bonds contributed by the β sheet, further stabilization at the dimer interface is provided by two distinct sets of interactions between amino acid sidechains from neighboring domains (Figure 8). At the center of the interface, the sidechains of Phe-106 and Ile-278 from one monomer pack against Ile-272 and Leu-273 from the other and form a small hydrophobic core. Surrounding these residues are six potential intermolecular ion pairs: Lys-74 and Glu-298' (closest observed distance between charged groups of 2.9 Å, with the prime indicating a residue in the other monomer), Lys-74 and Glu-300' (2.5 Å), Arg-96 and Glu-300' (4.7 Å), Arg-103 and Glu-304' (3.0 Å), Arg-105 and Glu-301' (3.3 Å), and Arg-105 and Glu-303' (3.0 Å). Two of the ion pairs (Arg-96–Glu-300' and Arg-103–Glu-304') involve charged groups that are both inaccessible to solvent, as determined using a water-sized probe (Lee and Richards, 1971), which is expected to lead to particularly strong ionic interactions. A feature of these interactions is that all the positively charged residues are contributed by one monomer, with the other one contributing the negatively charged residues. The computed electrostatic potential for an isolated monomer reflects this charge asymmetry, with the N- and C-terminal parts of the monomer being in regions of positive and negative electrostatic potential, respectively (Figure 7C). This extensive electrostatic complementarity is unique to the dimer interface, and is not observed at the interdomain boundaries within a monomer.

The dimer interface is thus seen to have a number of specific and potentially strong interactions, and a plausible explanation for the ATP requirement for β subunit assembly on DNA is that energy is required to break the interfacial hydrogen bonds and buried ion pairs. Nevertheless, the relatively small interaction surface suggests that a monomer–dimer equilibrium is possible in solution. In this regard, it is interesting to contrast the observed head-to-tail dimer (with the N-terminal domain of one monomer interacting with the C-terminal domain of the other) with a

β1   α1   β2   β3   β4   β5   α2

```
BETA-1   -MKFTVERE-HLLKPLQQVSGPLGGRPTLPILGNLLLQVAD-----GTLSLTGIDL-EMEMVARVALVQPHEHGATTVHARKFFDICRGL   82
BETA-2   -VEFTLPQA-TMKRLIEATQFSMAHQDVRYYLNGMLFETEG-----EELRTVATDG-HRLAVCSMPIGQSLPSHSVIVPRKGVIELMRML  207
BETA-3   -KHLEAGCD-LLKQAFARAAILSNEKF-----RGVRLYVSE-----NQLKITANNPEQEEAEEILDVTYSGAEMEIGENVSYVLDVLNAL  331

human-1  MFEARLVQGSILKKVLEALKDLLNEACWDISSSGVNLQSMDS----SHVSLVQLTLRSEGFDTYRCDRNLAMGVNLTSMSKILKCAGNED   86
yeast-1  MLEAKFEEASLFKRIIDGFKDCVQLVNFQCKEDGIIAQAVDD----SRVLLVSLEIGVEAFQEYRCDHPVTLGMDLTSLSKILRCGNNTD   86
gene45-1 -MKLSKDTTALLKNFATINSGIMLKSG-----QFIMTRAVNGT---TYAEANISD--VIDFDVAIYDLN--GFLGILSLVNDAEISQSED   77

human-2  -GEFARICR-DLSHIGDAVVISCAK-------DGVKFSASGEL---GNGNIKLSQTSNVDKEEEAVTIEMNEPVQLTFALRYLNFFTK--  217
yeast-2  -SEFSKIVR-DLSQLSDSINIMITK-------ETIKFVADGDI---GSGSVIIKPFVDMEHPETSIKLEMDQPVDLTFGAKYLLDIIK--  217
gene45-2 -TEIKAEDLQQLLRLSRGLQIDTIAITVKEGKIVINGFNKVEDSALTRVKYSLTL--GDYDGENTFN-----FIINMANMKMQPGNYKLL  186
            **  ***   ******  *  *** *       ** **   *     *  ***      *    *        *     ** ***      *
```

β6   β7   β8

```
BETA-1   -------PEGAEIAVQLEG----ERMLVRSG---RSRFSLSTLPAADFPNLDDWQSE   125
BETA-2   ------DGGDNPLRVQIGS----NNIRAHVG---DFIFTSKLVDGRFPDYRRVLPKNPD 253
BETA-3   --------KCENVRMMITD--SVSSVQIEDAASQSAAYVVMPMRL   366

human-1  IITLRAEDNADTLALVFEAPNQEKVSDYEMKLM-DLDVEQLGIPEQEYSCVVKMPS   141
yeast-1  TLTLIADNTPDSIILLFEDTKKDRIAEYSLKLM-DIDADFLKIEELQYDSTLSLPS   141
gene45-1 -------GNIKIADARSTIFWRAADPSTDVAPNKPIPFPVASAV   114

human-2  ATPLS-----STVTLSMSADV-PLVVEYKIADMGHLKYYLAPKIEDEEGS   261
yeast-2  GSSLS-----DRVGIRLSSFA-PALFQFDLKS-GFLQFFLAPKFNDEE   258
gene45-2 -----LWAKGKQGAAKFEG----EHANYVWA----LEADSTHDF   227
              **    **         **    *  *  *          **
```

Figure 9. Alignment of the Sequences of the Domains of the β Subunit with Human PCNA, Yeast PCNA, and Gene 45 Protein

The three domains of the β subunit are labeled BETA-1, BETA-2, and BETA-3, and are aligned based on three-dimensional structural correspondence alone. The human, yeast, and gene 45 sequences have been split into two domains labeled 1 and 2. The secondary structural elements within the β subunit domains are boxed and labeled. The meanings of the shaded bars for the β subunit sequences are different from those for the PCNA and gene 45 sequences. Within the 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 (Lee and Richards, 1971). No criterion of sequence similarity was used in the selection of these residues. For the PCNA and gene 45 sequences, the shaded bars indicate residues that are deemed to be similar in property to those that are buried in at least two of the β subunit domains. The residues Asp, Glu, Gln, Asn, His, Arg, and Lys have been excluded from the gray bars in PCNA and gene 45 unless a charged residue is present in the corresponding β subunit. The asterisks underneath the sequences indicate positions where at least two of the PCNA and gene 45 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.

head-to-head assembly. The latter would have identical faces on both sides of the ring and is not forbidden from considerations of chain topology alone. Model building shows that by rotating one of the monomers, a structure can be constructed which maintains similar β sheet and α helix packing at the interface, but in which the N- and C-terminal domains interact with the corresponding domains from the other monomer (this would correspond to rotating one of the monomers by 180° about a vertical axis in Figure 2B). Such an arrangement would, however, lead to unfavorable electrostatic interactions, as the N- and C-terminal regions of the monomers are positively and negatively charged, respectively (Figure 7C). Electrostatic complementarity is therefore likely to play an important role in favoring an initial head-to-tail association that could then lead to the formation of the complete interface and the burial of ion pairs.

Sequence Comparison between β Subunit Domains, PCNA, and Gene 45 Protein

The internal symmetry and domain structure of the β subunit was not suspected earlier because the three domains share very little sequence identity. Pairwise sequence alignments, based on the three-dimensional structure, result in only 16%, 9%, and 9% amino acid identity between domains 1 and 2, 2 and 3, and 1 and 3, respectively (Figure 9). This is well below the threshold of 20%–25% identity required for the predication of similar three-dimensional

structure (Sander and Schneider, 1991). Likewise, conventional sequence-matching algorithms fail to detect any relationship between the sequences of the three domains. However, knowledge of the three-dimensional structure makes possible an unambiguous internal sequence alignment, based simply on structural correspondence, and this reveals two conserved features (Figure 9). First, there are 17 positions in each domain that are completely buried and where only Phe, Leu, Ile, Met, Val, or Ala is present. These constitute the conserved hydrophobic cores of the domains. There are another 10 or so buried positions where charged residues are not tolerated (with the exception of residues involved in dimer formation; see below). Second, although the protein is overall negative in charge, the two α helices in each domain have net positive charge (the 12 helices have a total of 22 Arg and Lys residues and 8 Glu and Asp residues).

The three domains presumably diverged from an ancestral module by gene duplication events. That the present level of sequence identity is so low probably reflects the fact that the structural constraints on the domain are fairly loose compared to the geometric requirements at, for example, an enzyme active site or the recognition site of a site-specific transcription factor. Even in systems that do have such geometric requirements, it appears that large numbers of different sequences can lead to the same functional fold as demonstrated by random mutagenesis experiments on λ repressor (Lim and Sauer, 1991). The di-

vergence in sequence between the three domains of the β subunit is not surprising in this context.

Given the lack of significant sequence similarity between the three domains of the β subunit, it seems unlikely that conventional sequence-matching schemes will be successful in recognizing protein sequences that adopt this three-dimensional fold. We have, instead, focused our attention on proteins that are known to be functionally similar to the β subunit, and looked to see if they are likely to have the same overall architecture. These proteins are the various "processivity factors" of other replicative polymerases (Kornberg and Baker, 1991), the best characterized of which are the bacteriophage T4 DNA polymerase system (Nossal and Alberts, 1984) and the yeast and mammalian chromosomal replication polymerase, Pol δ (Tsurimoto and Stillman, 1990; Burgers, 1991; Lee et al., 1991a, 1991b). For highly processive DNA replication, each of these polymerases requires a set of ATP-dependent accessory factors that appear to act to tether the core polymerase enzyme onto DNA, in a manner analogous to the γ complex and β subunit of E. coli PolIII holoenzyme. The accessory factor that has functional correspondence with the β subunit is the gene 45 protein for the T4 DNA polymerase system, and PCNA for eukaryotic Pol δ (Kornberg and Baker, 1991). T4 gene 45 protein and PCNA are approximately two-thirds the size of the β subunit (T4 gene 45 protein: 227 residues; Saccharomyces cerevisiae PCNA: 258 residues; human PCNA: 261 residues). Given their smaller size and knowing the three-domain structure of the β subunit, we reasoned 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. There is no obvious amino acid sequence homology between the β subunit and T4 gene 45 protein or any PCNA. However, measurements of the molecular weight of T4 gene 45 protein and S. cerevisiae PCNA in solution indicate that they are likely to be trimeric (Jarvis et al., 1989; Bauer and Burgers, 1988).

The sequences of T4 gene 45 protein and of yeast and human PCNA were compared with those of the three domains of the β subunit, using the following criteria. Insertions and deletions were allowed only 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. The yeast and human sequences are closely related (30% identity) (Almendral et al., 1987; Bauer and Burgers, 1990), and these conditions led to a plausible, albeit weak, alignment for both PCNA sequences with the β subunit (Figure 9). A similar alignment is obtained for gene 45 protein with the β subunit (Figure 9). This alignment differs from one previously reported between yeast PCNA and T4 gene 45 protein, which included several very long insertions and deletions (Tsurimoto and Stillman, 1990).

The resulting alignment between the β subunit domains and other sequences (two domains each) is shown in Figure 9, which also indicates the buried residues in the β subunit. The alignment preserves in PCNA the hydrophobic core of the β subunit, and also reveals a suggestive

level of sequence identity between PCNA and the domains of the β subunit. At 37 positions out of 110, there are at least two amino acid identities between the domains of the β subunit and PCNA (Figure 9). Some of these sequence identities are intriguing. For example, sheet 4 in the third domain of the β subunit has two buried Glu residues that are involved in dimer formation. These are conserved in corresponding positions in the PCNA sequences. Two of the positively charged residues involved in ion pairing at the dimer interface (Lys-74 and Arg-96) have corresponding Lys or Arg residues at the same position or one amino acid removed in the PCNA sequences. Eight positions in the domains of the β subunit have at least one buried aromatic residue. Five of these have corresponding aromatic residues in one of the PCNA sequences. Finally, although both PCNA sequences have a net excess of negatively charged residues, the regions corresponding to the helices have net positive charge, qualitatively similar to what is observed in the β subunit. Normalized to 12 α helices, the yeast sequences have 18 positive and 12 negative residues on the helices (the human sequence has an additional negative charge). Similar conclusions can be drawn from the alignment of gene 45 protein with the β subunit (Figure 9). As in the β subunit and PCNA, although the overall charge of the gene 45 protein is negative, the net charge for the α helices is positive (15 positively charged and 9 negatively charged when normalized to 12 helices). Ion pairs can also be formed in the trimer interface of the gene 45 protein. The negative charged Asp-49 and Asp-51 (which are aligned one residue away from the buried Glu residues) can be paired with the positively charged Lys-177 and -184, or alternatively, the positively charged Arg-127 and -130 can be paired with Asp-69, Glu-71, and Glu-76.

This alignment suggests that if PCNA and gene 45 form a ring around DNA, then they must do so as trimers, with each monomer of the protein contributing two domains that are roughly similar in architecture to the β subunit domains. However, it must be stressed that the alignment is weak and should only be taken as a hypothesis for further experimental testing. For example, the recently developed profile method of sequence comparison (Bowie et al., 1991), which relies on knowledge of the three-dimensional structure of one of the structures being compared, is unable to detect any significant similarity between β subunit domains and PCNA or gene 45 protein when provided with the β subunit structure. It is, however, able to detect some, but not all, of the internal symmetry of the β subunit. A conclusive understanding of PCNA and gene 45 protein architecture must await determination of the three-dimensional structures.

Conclusion

A satisfying feature of the structure of the β subunit is the beautiful way in which the circular symmetry of the double helix has been reflected in the ring-shaped and highly symmetric structure of the protein clamp. This symmetry manifests itself on a smaller scale than the domain boundaries, and the entire structure can be considered to be generated by replication of a simple βαββ motif 12 times around a

circle, with the β strands and α helices being oriented perpendicular to the direction of the grooves of DNA. This simple and elegant structure carries out a simple function and stands in sharp contrast to the structure of the actual catalytic subunit of the polymerase, as probably exemplified by the Klenow fragment of PolI. That enzyme, which must recognize the information content of DNA and then replicate one strand, has none of the symmetry of DNA and resembles rather the palm of a hand gripping the double helix (Ollis et al., 1985). For processive polymerases involved in DNA replication, it appears that the proper combination of symmetry and asymmetry is a requirement for functionality.

Experimental Procedures

Crystalization

The β subunit was purified to >99% purity as described (Onrust et al., 1991), with the following modifications. Chromatography on AH–Sepharose was performed in place of SP–sephadex, and the chromatofocusing step was replaced by chromatography on a fast flow Q column and a mono-Q column (Pharmacia). The purified β subunit was concentrated to 16 mg/ml, as determined by Bradford assay (Bradford, 1976), by ultrafiltration in a buffer containing 20% glycerol, 20 mM Tris buffer (pH 7.5), and 0.5 mM EDTA. The protein solution in glycerol was used as such for all further crystalization trials, using a sparse matrix method (Jancarick and Kim, 1991). Crystalization was rapid, with many different conditions yielding crystals within a day or two of the first experiments. Optimization of several of these conditions led to three different crystal forms (Forms I, II, and III) that are suitable for high-resolution X-ray diffraction analysis.

All crystals were grown using the hanging drop method (McPherson, 1990). The drops were initially set up by mixing 2 μl of protein and 2 μl of a "reservoir" solution, and equilibrated against the reservoir at room temperature or 4°C. The unit cell constants and space groups of the various crystal forms were determined by a combination of oscillation data collection (see below), single counter diffractometry (using a Rigaku AFC5 diffractometer), and precession photography. Form I crystals are in space group P1 (a = 86.8 Å, b = 73.9 Å, c = 65.7 Å, α = 75.2°, β = 86.8°, γ = 81.8°) and are grown at 4°C from reservoirs containing 13%–15% isopropanol, 100 mM CaCl₂, and 100 mM MES buffer (pH 7.5). Form II crystals are obtained at room temperature (21°C) from reservoirs containing 30% polyethylene glycol (average M_r = 400), 100 mM CaCl₂, and 100 mM MES (pH 6.5). The space group is P1 (a = 41.7 Å, b = 72.9 Å, c = 65.5 Å, α = 74.6°, β = 85.1°, γ = 82.2°).

Forms I and II are related, with the asymmetric unit in Form I being a doubling of that in Form II. Finally, Form III crystals are obtained at room temperature (21°C) from conditions very similar to those that yield Form II crystals, except that the pH is lowered to 6.0. This crystal form is in space group P2₁ (a = 80.6 Å, b = 68.3 Å, c = 82.3 Å, β = 114.2°).

All three crystal forms grow to large sizes (0.3 × 0.4 × 0.7 mm³) and diffract strongly to 2 Å resolution. The presence of the rotational symmetry axis in Form III crystals makes them much more suitable for oscillation data collection than the lower-symmetry P1 forms, particularly for the simultaneous measurement of Bijvoet pairs, and thus the structure determination was carried out using this form alone. Based on molecular volume calculations, Forms II and III contain a dimer of the β subunit in the asymmetric unit, while Form I contains two dimers. Assuming this stoichiometry, the volume per unit mass, V_m, as defined by Matthews, is 2.5, 2.3, and 2.6 Å³ for the three forms, respectively, within the range typical for protein crystals (Matthews, 1968).

Data Collection and Structure Determination

X-ray intensity data collection was carried out by the oscillation method (Arndt and Wonacott, 1977), using a Rigaku R-AXIS IIC imaging phosphor area detector, mounted on a Rigaku RU200 rotating anode X-ray generator (Molecular Structure Corp., Houston). Typical crystal-to-detector distances and exposure times were 153 mm and 15 min, respectively, for 2° oscillations. Data processing and reduction were done entirely by software provided by Rigaku (Table 1).

The structure determination was carried out by the multiple isomorphous replacement (MIR) method. The amino acid sequence of the β subunit shows that the molecule contains four cysteines (Ohmori et al., 1984), at least one of which is reactive to N-ethylmaleimide (Johanson et al., 1986). The search for isomorphous heavy atom derivatives was therefore focused on mercury compounds, and two good derivatives were obtained using mercuric chloride and ethyl mercury phosphate. The binding sites for mercury are similar in both cases, with mercuric chloride and ethylmercury phosphate reacting with three and four cysteines per monomer, respectively (Table 2). Although crystals of the native protein have reasonable lifetimes in the X-ray beam, allowing complete data sets to be collected on single crystals, the derivative crystals exhibit much earlier radiation decay. For the ethylmercury phosphate derivative, this was overcome by measuring data at −5°C. The mercuric chloride–treated crystals were not stable at the lower temperature, and a total of four crystals were used for data collection at room temperature (Table 2). For both derivatives, measurement of anomalous differences was optimized by aligning crystals so that the 2-fold rotation axis (b*) was precisely along the oscillation axis, leading to simultaneous measurement of Bijvoet pairs.

Difference Patterson maps for the mercuric chloride derivative showed a small number of strong peaks (10–16 standard deviations above the mean density value) in the y = 1/2 Harker section, and heavy

Table 1. Refined Heavy Atom Positions

Derivative		Relative Occupancy	x	y	z	Derivatized Cysteine	Distance to Sulphur (Å)
HgCl₂	1	1.0	0.26	0.40	0.44	79	1.88
	2	1.0	0.41	0.80	0.99	79'	2.14
	3	0.6	0.00	0.86	0.09	180	1.60
	4	0.6	0.84	0.50	0.52	180'	1.69
	5	0.6	0.83	0.39	0.88	333	1.95
	6	0.6	0.00	0.54	0.53	333'	1.51
EMP	1	0.6	0.26	0.40	0.43	79	1.37
	2	0.4	0.41	0.80	0.98	79'	1.68
	3	0.8	0.01	0.86	0.09	180	1.52
	4	0.9	0.84	0.50	0.52	180'	1.81
	5	0.3	0.83	0.39	0.88	333	1.90
	6	0.3	0.00	0.54	0.53	333'	1.59
Sites not present in HgCl₂ derivative:							
	7	0.9	0.81	0.38	0.93	260	1.60
	8	1.0	0.02	0.50	0.49	260'	2.16

The positions are fractional coordinates in the unit cell, and the prime on the cysteine number indicates that it is in the second monomer.

Table 2. Statistics for Data and Derivatives

	Native	HgCl₂ I	HgCl₂ II	EMP
Number of crystals	1	3	1	1
Concentration (mM)		2	2	2
Soaking time (hr)		12	12	6
Resolution (Å)	2.4	2.78	3.0	2.79
Measured reflections	101,590	86,408	42,097	61,458
Unique reflections	30,957	19,660	16,148	19,568
Completion (%)	94.4	93.3	97.3	94.1
R-merge[a] (%)	7.29	7.87	7.41	6.78
Mean isomorphous difference[b] (%)		24.4	25.6	17.8
Phasing power[c]		2.1	2.1	1.4
Mean figure of merit	0.67			

The two HgCl₂ data sets are for differently oriented crystals (see text). [a] $\Sigma_h \Sigma_i |I_{h,i} - I_h|/\Sigma_h \Sigma_i I_{h,i}$, where $I_{h,i}$ is the scaled intensity of the ith observation of reflection h, and I_h is the mean value. [b] $\Sigma |F_{PH} - F_P|/\Sigma F_P$, where F_{PH} and F_P are the scaled derivative and native structure factor amplitudes, respectively. [c] Phasing power: rms heavy-atom structure factor / rms lack of closure.

atom positions were readily determined by manual inspection. Additional sites were found by difference Fourier techniques. Heavy atom parameters were refined and initial phases were calculated using the program HEAVY (Terwilliger and Eisenberg, 1983), including anomalous differences for only those reflections that were fully recorded on

one image (71.7% and 70.2% of the data had anomalous measurements included for mercuric chloride and ethylmercury phosphate, respectively). The quality of the anomalous data was judged to be good, since the anomalous difference Patterson map recapitulated the major peaks in the difference Patterson map. For calculating and refining the MIR phases, the mercuric chloride data were divided into two sets. The first set consists of data (including anomalous differences) from three crystals, each of which had been oriented with the b*-axis parallel to the oscillation axis. The second set consists of data from a crystal mounted in an arbitrary orientation and does not include anomalous differences, and better results were obtained by treating these data separately in the phase calculations. Summary statistics from the heavy atom phasing procedure are given in Table 1.

The MIR phases were further improved by solvent flattening (Wang, 1985) using a program written by W. Kabsch (COMBINE). These phases were then used to generate a 3 Å electron density map. Predominant features of the structure, such as the donut shape, the two-layer architecture with helices in the middle and β sheets outside, and the approximate 6-fold symmetry, were immediately recognized. The program BONES (Jones and Thirup, 1986) was used to generate a skeleton representation of the density, which made evident the fact that the structure is made up of six copies of very similar modules. An additional feature that emerged at this stage was that the 2-fold noncrystallographic symmetry apparent in the mercury positions correlated extremely well with a 2-fold noncrystallographic symmetry in the electron density. This 2-fold axis is only 12° away from the crystallographic 2-fold screw axis, and was therefore not identified in self-rotation functions.

A partial model of poly-Ala chain was built for 80% of the structure fairly readily, using the program FRODO and a data base of protein structures (Jones and Thirup, 1986). During this process, partial atomic models were repeatedly used to improve the phases by phase combination with the MIR and solvent-flattened phases, using COMBINE. Although the noncrystallographic symmetry was never explicitly used in the phase refinement process, it provided a useful check on the accuracy of the phases, as did the unexpected but obvious 3-fold symmetry within each monomer. The topology of the protein fold and
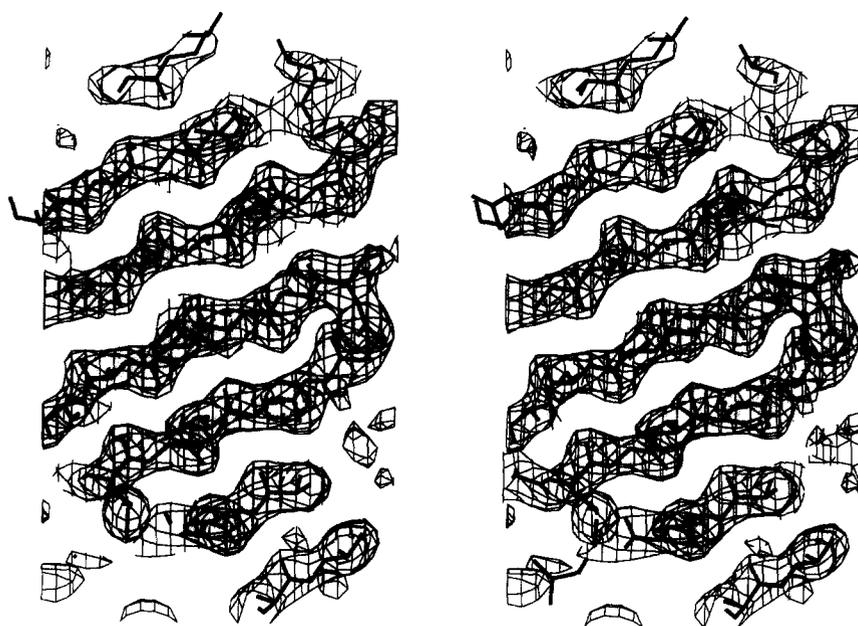
Figure 10. Stereo View of the Electron Density

Electron density at 2.5 Å resolution, at one of the two dimer interfaces of the β subunit. The three β strands at the top are from one molecule and the other four from the second molecule. Electron density contours are displayed at 1 σ above the mean value of the density. Note that the directions of the carbonyl oxygens are apparent as bumps in the backbone density, which makes the direction of the chain unambiguous. The quality of the electron density map apparent here is typical of the rest of the map, which has no breaks in backbone density anywhere in the molecule. The displayed map is a $(2|F_o|-|F_c|)\exp\{i\alpha_c\}$ Fourier synthesis, where F_c and F_o are calculated and observed structure factors, respectively, and α_c is the calculated phase. The model used is the final refined one.

the locations of the N- and C-termini of the monomers were clearly recognized at this stage. A complete model of one monomer with the correct sequence (Ohmori et al., 1984) was readily built (except for a loop of 10 residues that was left as poly-Ala) by assuming that the mercury atoms had bound to cysteine residues, and tracing the sequence out from these points. The second monomer was generated from this model by the noncrystallographic symmetry operation.

This atomic model was refined against 3.0 Å native X-ray data by least-squares refinement with both monomers treated independently, using the program X-PLOR (Brünger, 1988). Two hundred steps of Powell optimization and 40 steps of individual B factor refinement smoothly reduced the R factor from 46.2% to 24.9% with no manual intervention, indicating the accuracy of the initial model. Repeated model building using FRODO and least-squares refinement using X-PLOR were carried out. A total of five rounds of simulated annealing refinement were then carried out, using initial and final temperatures of 1000 K and 300 K, respectively (Brünger et al., 1987; Weis et al., 1990). The resolution of the native data included was extended to 2.5 Å, and 150 well-resolved solvent molecules (interpreted as water) were included. The current model has the entire sequence built in, and has unbroken backbone electron density from the N- to the C-terminus in difference Fourier maps. The R factor is 18.9% (for 27,614 reflections with $|F|>2\sigma(F)$) and the rms deviation from ideal geometry is 0.017 Å for bond lengths and 3.5° for bond angles, with good stereochemistry for the backbone torsion angles. Representative electron density for the final refined model is shown in Figure 10, at the dimer interface. Although the electron density is weak for a few surface sidechains (arginines and lysines in particular), the backbone density is strong throughout the structure, including all surface loops. The orientation of the carbonyl groups is clear for almost all residues (Figure 10).

Acknowledgments

References

Almendral, J. M., Huebsch, D., Blundell, P. A., Macdonald-Brown, H., and Bravo, R. (1987). Cloning and sequence of the human nuclear protein cyclin: homology with DNA binding proteins. Proc. Natl. Acad. Sci. USA 84, 1575–1579.

Arndt, V. W., and Wonacott, A. J. (1977). The Rotation Method in Crystallography (Amsterdam: North-Holland).

Bauer, G. A., and Burgers, P. M. J. (1988). The yeast analog of mammalian cyclin/proliferating-cell nuclear antigen interacts with mammalian DNA polymerase δ. Proc. Natl. Acad. Sci. USA 85, 7506–7510.

Bauer, G. A., and Burgers, P. M. J. (1990). Molecular cloning, structure and expression of the yeast proliferating cell nuclear antigen gene. Nucl. Acids Res. 18, 261–265.

Bowie, J. U., Lüthy, R., and Eisenberg, D. (1991). A method to identify protein sequences that fold into a known three-dimensional structure. Science 253, 164–170.

Bradford, M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein–dye binding. Anal. Biochem. 72, 248–254.

Branden, C., and Tooze, J. (1991). Introduction to Protein Structure (New York: Garland).

Brünger, A. T. (1988). X-PLOR (Version 2.2) Manual. Howard Hughes Medical Institute and Dept. of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut.

Brünger, A. T., Kuriyan, J., and Karplus, M. (1987). Crystallographic R-factor refinement by molecular dynamics. Science 235, 458–460.

Burgers, P. M. J. (1991). Saccharomyces cerevisiae replication factor C. II. Formation and activity of complexes with proliferating cell nuclear antigen and with DNA polymerases δ and ε. J. Biol. Chem. 266, 22698–22706.

Burgers, P. M. J., and Kornberg, A. (1982). ATP activation of DNA polymerase III holoenzyme of Escherichia coli. J. Biol. Chem. 257, 11468–11473.

Fay, P. J., Johanson, K. O., McHenry, C. S., and Bambara, R. A. (1981). Size classes of products synthesized processively by DNA polymerase III and DNA polymerase III holoenzyme of Escherichia coli. J. Biol. Chem. 256, 976–983.

Gilson, M. K., Sharp, K. A., and Honig, B. H. (1988). Calculating the electrostatic potential of molecules in solution: method and error assessment. J. Comp. Chem. 9(4), 327–335.

Jancarick, J., and Kim, S.-H. (1991). Sparse matrix sampling: a screening method for crystallization of proteins. J. Appl. Crystallogr. 24, 409–411.

Jarvis, T. C., Paul, L. S., and von Hippel, P. H. (1989). Structural and enzymatic studies of the T4 DNA replication system. I. Physical characterization of the polymerase accessory protein complex. J. Biol. Chem. 264, 12709–12716.

Johanson, K. O., and McHenry, C. S. (1980). Purification and characterization of the β subunit of DNA polymerase III holoenzyme of Escherichia coli. J. Biol. Chem. 255, 10984–10990.

Johanson, K. O., Haynes, T. E., and McHenry, C. S. (1986). Chemical characterization and purification of the β subunit of the DNA polymerase III holoenzyme from an overproducing strain. J. Biol. Chem. 261, 11460–11465.

Jones, T. A., and Thirup, S. (1986). Using known substructures in protein model building and crystallography. EMBO J. 5, 819–822.

Kabsch, W., and Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers 22, 2577–2637.

Kornberg, A., and Baker, T. A. (1991). DNA Replication (New York: W. H. Freeman).

Lee, B. K., and Richards, F. M. (1971). The interpretation of protein structures: estimation of static accessibility. J. Mol. Biol. 55, 379–400.

Lee, S.-H., Kwong, A. D., Pan, Z.-H., and Hurwitz, J. (1991a). 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.

Lee, S.-H., Pan, Z.-Q., Kwong, A. D., Burgers, P. M. J., and Hurwitz, J. (1991b). Synthesis of DNA by DNA polymerase ε in vitro. J. Biol. Chem. 266, 22707–22717.

Lim, W. A., and Sauer, R. T. (1991). The role of internal packing interactions in determining the structure and stability of a protein. J. Mol. Biol. 219, 359–376.

Maki, S., and Kornberg, A. (1988). DNA polymerase III holoenzyme of Escherichia coli. III. Distinctive processive polymerases reconstituted from purified subunits. J. Biol. Chem. 263, 6561–6569.

Matthews, B. W. (1968). Solvent content of protein crystals. J. Mol. Biol. 33, 491–497.

McPherson, A. (1990). Current approaches to macromolecular crystallization. Eur. J. Biochem. 189, 1–23.

Nossal, N. G., and Alberts, B. M. (1984). Mechanism of DNA replication catalyzed by purified T4 replication proteins. In Bacteriophage T4 (Washington, D.C.: American Society for Microbiology).

O'Donnell, M. E. (1987). Accessory proteins bind a primed template and mediate rapid cycling of DNA polymerase III holoenzyme from Escherichia coli. J. Biol. Chem. 262, 16558–16565.

O'Donnell, M. E., and Kornberg, A. (1985). Complete replication of templates by Escherichia coli DNA polymerase III holoenzyme. J. Biol. Chem. 260, 12884–12889.

O'Donnell, M., and Studwell, P. S. (1990). Total reconstitution of DNA polymerase III holoenzyme reveals dual accessory protein clamps. J.

Biol. Chem. *265*, 1179–1187.

Ohmori, H., Kimura, M., Nagata, T., and Sakakibara, Y. (1984). Structural analysis of the *dnaA* and *dnaN* genes of *Escherichia coli*. Gene *28*, 159–170.

Ollis, D. L., Brick, P., Hamlin, R., Xuong, N. G., and Steitz, T. A. (1985). Structure of large fragment of *Escherichia coli* polymerase I complexed with dTMP. Nature *313*, 762–766.

Onrust, R., Stukenberg, P. T., and O'Donnell, M. (1991). Analysis of the ATPase subassembly which initiates processive DNA synthesis by DNA polymerase. J. Biol. Chem. *266*, 21681–21686.

Priestle, J. P. (1988). RIBBON—stereo cartoon drawing program for proteins. J. Appl. Crystallogr. *21*, 572–576.

Saenger, W. (1984). Principles of Nucleic Acid Structure (New York: Springer-Verlag).

Sander, C., and Schneider, R. (1991). Database of homology-derived protein structures and the structural meaning of sequence alignment. Proteins *9*, 56–68.

Sixma, T. K., Pronk, S. E., Kalk, K. H., Wartna, E. S., van Zanten, B. A. M., Witholt, B., and Hol, W. G. J. (1991). Crystal structure of the cholera toxin-related heat-labile enterotoxin from *E. coli*. Nature *351*, 371–377.

Studwell, P. S., and O'Donnell, M. (1990). Processive replication is contingent on the exonuclease subunit of DNA polymerase III holoenzyme. J. Biol. Chem. *265*, 1171–1178.

Stukenberg, P. T., Studwell-Vaughan, P. S., and O'Donnell, M. (1991). Mechanism of the sliding β-clamp of DNA polymerase III holoenzyme. J. Biol. Chem. *266*, 11328–11334.

Terwilliger, T. C., and Eisenberg, D. (1983). Unbiased three-dimensional refinement of heavy-atom parameters by correlation of origin-removed Patterson functions. Acta Crystallogr. *A39*, 813–817.

Tsurimoto, T., and Stillman, B. (1990). 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.

Wang, B. C. (1985). Resolution of phase ambiguity in macromolecular crystallography. Meth. Enzymol. *115*, 90–112.

Weis, W. I., Brünger, A. T., Skehel, J. J., and Wiley, D. C. (1990). Refinement of the influenza virus hemagglutinin by simulated annealing. J. Mol. Biol. *212*, 737–761.

Wickner, S. (1976). Mechanism of DNA elongation catalysed by *Escherichia coli* DNA polymerase III, dna Z protein, and DNA elongation factors I and III. Proc. Natl. Acad. Sci. USA *73*, 3511–3515.