Published in final edited form as:

J Mol Biol. 2008 October 17; 382(4): 859–869. doi:10.1016/j.jmb.2008.07.058.

Insights into the Replisome from the Structure of a Ternary Complex of the DNA Polymerase III α -subunit

Richard A. Wing^{1,*}, Scott Bailey^{1,3,*}, and Thomas A. Steitz^{1,2,3,†}

¹Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06520, USA.

²Department of Chemistry, Yale University, New Haven, CT 06520, USA.

³Howard Hughes Medical Institute, New Haven, CT 06520, USA.

Summary

The crystal structure of the catalytic subunit of the DNA Polymerase III holoenzyme (PolIII α) bound to primer-template DNA and an incoming deoxynucleoside triphosphate has been determined at 4.6 Å resolution. The polymerase interacts with the sugar phosphate backbone of the DNA across its minor groove, which is made possible by significant movements of the thumb, fingers and β -binding domains relative to their orientations in the unliganded polymerase structure. Additionally, the DNA and incoming nucleotide are bound to the active site of PolIII α nearly identically as they are in their complex with DNA polymerase β (Pol β), thereby proving that the eubacterial replicating polymerase is homologous to Pol β , but not the eukaryotic replicating polymerases. Finally, superimposing a recent structure of the clamp bound to DNA on this PolIII α complex with DNA places a loop of the β -binding domain into the appropriate clamp cleft and supports a mechanism of polymerase switching.

Introduction

The DNA polymerase III holoenzyme replicates the genome of eubacteria. This fifteen-subunit complex functions as a tripartite assembly consisting of the core polymerase, the clamp loader, and the clamp¹. In *Escherichia coli* (Eco), the core polymerase contains the catalytic α -subunit (PolIII α), the 3'–5' exonuclease ϵ -subunit, and the θ -subunit whose function is essentially unknown². The ring-shaped clamp, which encircles duplex DNA^{3,4}, imparts processivity to the core polymerase through its interactions with PolIII α ^{5–7}. The clamp loader operates as a heptameric adenosine 5'-triphosphate hydrolase that loads the clamp onto RNA-primed sites prior to binding of clamp by DNA Polymerase III⁸.

DNA-dependent DNA polymerases have been classified into six families (A, B, C, D, X and Y) based upon analysis of their sequences^{9–12}. Members of families A, B, C and D participate directly in DNA replication. The main eubacterial replicating polymerase, PolIII α , is a C-

© 2008 Elsevier Ltd. All rights reserved.

[†]To whom correspondence should be addressed. E-mail: thomas.steitz@yale.edu.

*These authors contributed equally to this work.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Accession Numbers

The coordinates and structure factor amplitudes of the Taq PolIII α ternary complex have been deposited under the accession code #####.

family polymerase, whereas many archaeal and all eukaryotic replicative polymerases are B-family polymerases⁹. X and Y family polymerases, such as the eukaryotic DNA Polymerase β (Pol β) and the archaeal DNA Polymerase IV (Dpo4), respectively, are involved in DNA repair^{9; 10}. Members of all six polymerase families accomplish polymerization utilizing three domains, termed the fingers, palm, and thumb¹³. The fingers interact with the incoming nucleotide, the palm positions the catalytic Mg²⁺ ions that perform catalysis and the thumb grips the DNA substrate, typically through backbone interactions with the minor groove¹⁴. In addition to these domains PolIII α also includes a N-terminal polymerase and histidinol phosphatase (PHP) domain^{15; 16}, a β -binding domain that contains an “internal” clamp binding site^{17; 18} and a C-terminal domain (CTD) that includes both an oligonucleotide binding-fold (OB-fold)¹⁹ and an “external” clamp binding site located at the extreme C-terminus of the polymerase²⁰.

The six polymerase families can be grouped into two superfamilies based on the fold of their palm domains. The palm domains of members of families A, B and Y are formed by a four-stranded antiparallel β -sheet^{21–24} whereas members of families C and X have a palm domain formed by a five-stranded mixed β -sheet^{25–27} that belongs to the Pol β -like nucleotidyl-transferase (β NT) superfamily²⁸. Although both types of palm domain folds utilize a two-metal ion mechanism for catalysis, the two palm folds are considered to be cases of convergent evolution²⁹. As such, this distinct difference between eukaryotic and eubacterial replicative polymerases may potentially be exploitable for the design of novel antibiotics.

A key difference between the conclusions concerning the functional implications of the two structures of unliganded PolIII α recently reported is whether the palm domain of PolIII α is homologous to the palm domain of Pol β and thus binds its DNA substrate and incoming dNTP as Pol β does^{25; 27}. Bailey et al. (2006) determined the structure of the full-length protein from *Thermus aquaticus* (Taq) and Lamers et al. (2006) that of a truncated mutant lacking the CTD from Eco (throughout this manuscript Eco and Taq residue numbers will have the prefix “Eco” and “Taq”, respectively). Lamers et al. (2006) conclude that Eco PolIII α is not homologous to Pol β and cannot bind DNA as Pol β does because when they homology-modeled the DNA, by superimposing the structures of PolIII α and the ternary complex of Pol β using their catalytic aspartates and the β -sheets of their palms, the homology modeled DNA directly clashes with a short α -helix (residues Eco543 – Eco549 and Taq605 – Taq612) adjacent to the active site aspartates of the palm domain (the active site helix). Therefore, to model a complex of Eco PolIII α with DNA the authors placed a segment of standard B-form DNA into the PolIII α active site and rotated its orientation so that it did not overlap significantly with the protein and such that it could also interact with the β -binding domain. Their resulting model places the plane of the base pairs of the active site DNA in an orientation not parallel with the direction of the β -strands of the palm domain, whereas the hallmark of DNA binding in the X family is that the plane of the base pairs is parallel with the direction of the β -strands of the palm domain^{30; 31}. In contrast, Bailey et al. (2006) concluded that the palm domains of PolIII α and Pol β are homologous and found that homology modeling the DNA from the ternary complex of Pol β onto the structure of Taq PolIII α resulted in no main chain clashes between the DNA and the palm domain, including the active site helix. Consequently, they were able to produce a model of the DNA complex in which the plane of the base pairs is parallel to the direction of the β -strands, as is the case in the structure of the ternary complex of Pol β . In this model the β -binding domain was also hypothesized to rotate towards and interact with the DNA duplex.

Clearly, at least one of these two models of the DNA complex and the associated conclusions had to be incorrect. The crystal structure of Taq PolIII α bound to primer template DNA with incoming nucleotide presented here establishes that in contrast to the conclusion of Lamers et al. (2006) PolIII α does indeed bind DNA and incoming nucleotide as Pol β does. Consequently, PolIII α is homologous to Pol β and not to the replication DNA polymerases of eukaryotes.

Results

Mutagenesis and Inactivation of the N-terminal PHP Nuclease Domain

The N-terminal PHP domain of *Thermus thermophilus* PolIII α possesses a novel Zn²⁺-dependent exonuclease activity¹⁶, which had to be inactivated in order to co-crystallize a complex between Taq PolIII α and its DNA substrate. The crystal structure of the unliganded Taq PolIII α exhibited two Zn²⁺ ions bound in a deep cleft within the PHP nuclease domain²⁵. Site-directed mutagenesis was employed to eliminate two of the observed side-chain ligands of the Zn²⁺ ions by mutating two aspartates (Taq20 and Taq212) to asparagines (Supplementary Figure 1). The nuclease activity of this double mutant on a substrate of primer-template DNA was significantly attenuated when compared with wild-type protein of equivalent purity (Supplementary Figure 2). However, the residual nuclease activity could be completely inhibited by the addition of 2',3'-dideoxycytidine 5'-triphosphate (ddCTP) and 2'-deoxyadenosine 5'-triphosphate (dATP) to the reaction buffer. Under these conditions, the primer strand was terminated by the incorporation of ddCMP and subsequently the dATP bound stably to the incoming nucleotide site, which resulted in a stable ternary complex of polymerase, primer-template DNA and incoming nucleotide. It is presumably the stability of this complex that restricts the PHP nuclease domain's access to the DNA, thereby further inhibiting the nuclease activity.

Structure Determination

The ternary complex of the D20N, D212N double mutant of Taq PolIII α with the ddCTP terminated primer-template DNA and dATP was crystallized. The structures of several other ternary complexes of DNA polymerases have been determined by employing an analogous strategy^{31; 32}. The resulting crystals diffracted X-rays to beyond 4.9 Å resolution, where the I/ σ is 2.0. Although these crystals were merohedrally twinned, as could be demonstrated by analyzing the distribution of diffraction intensities³³, phases were readily obtained by molecular replacement using the unliganded structure of Taq PolIII α as a search model²⁵. Two copies (A and B) of the polymerase ternary complex were found in the asymmetric unit. Given the resolution of the X-ray data, optimization of the model was restricted to rigid-body refinement of the positions of the individual domains (see Experimental Procedures) and the subsequent analysis of the interactions between the polymerase and its substrates was limited to whether alternate side-chain rotamers could bring the side chains into contact distance with the substrates. The final model was refined at 4.6 Å resolution, where the I/ σ is 1.4. This resolution was chosen due to the low value of the twinned R-factor and R-free (36.9% and 37.6%, respectively) obtained in the high resolution shell (Table 1). The ultimate testament to the accuracy of the model was seen when inspecting the two-fold averaged electron density map calculated using phases from the protein coordinates alone before the DNA and incoming deoxynucleoside triphosphate (dNTP) were incorporated into the model (twinned R-factor and R-free of 31.4% and 30.4%, respectively). Readily apparent in this unbiased electron density map were features corresponding to the major and minor grooves of the double-stranded primer-template DNA as well as the triphosphate moiety of the incoming dNTP (Figures 1 and Supplementary Figure 3). The two copies of the protein components of the ternary complex found in the asymmetric unit are essentially identical with a root mean squared deviation (rmsd) of 1.1 Å over 1148 C α atoms. The DNA substrates are effectively identical apart from a bend in the DNA after the -12 position in copy A that is imposed by crystal packing (Supplementary Figure 4). However, after this position there are no observed contacts between the polymerase and the DNA in either copy. Furthermore, the positions of the incoming nucleotide and the 3' primer terminus are identical in both copies. We have chosen to discuss copy B here because its DNA substrate appears unperturbed by crystal packing. The final model including the DNA and incoming dNTP has a twinned R-factor and R-free of 28.7% and 27.1% (Figure 2A), respectively. Data collection and refinement statistics are presented in Table 1.

The Overall Structure of PolIII α and Substrate Interactions

Upon binding the primer-template DNA, the fingers, thumb and β -binding domains of PolIII α are seen to move towards the DNA bound to the palm and grip its backbone at the minor groove (Supplemental Movie 1). The thumb domain uses several positively charged residues within two α -helices (residues Taq500 – Taq511 and Taq515 – Taq526) that lie parallel to the DNA helical axis and loops emanating from the palm domain to make electrostatic interactions with the sugar-phosphate backbone of the minor groove (Figure 3A). A loop (Taq527 – Taq535 and Eco465 – Eco471) at the tip of the thumb extends towards the polymerase active site. More distant from the active site, the β -binding domain interacts with the DNA substrate using its helix-hairpin-helix (HhH) motif and flanking loops (Figure 3B). The HhH motif presumably utilizes a bound metal ion to mediate interactions between the protein and the phosphates of the backbone, as is seen in the structures of other complexes of HhH motifs with DNA³⁴. The two flanking loops appear to contact the sugar-phosphate backbone of the DNA directly using positively charged residues (Taq848, Taq932, Taq933 and Taq895). The fingers domain completes the substrate binding cleft by forming a wall towards which the 3' end of the primer is oriented. In addition, the fingers domain acts in concert with the palm and the 3' terminus of the primer to create a pocket that coordinates the incoming dATP molecule. The PHP nuclease domain does not contact the DNA substrate except possibly for a loop (residues Taq232 – Taq241) situated below the thumb domain (Supplementary Figure 5).

The 3' end of the primer strand and the incoming dNTP appear to be coordinated by several highly-conserved residues in the active site (Figure 2B). The phosphate group of the 3' primer terminus is positioned by electrostatic interactions with a lysine residue (Taq616) of the palm domain. Although at the current resolution the precise interactions are not clear, the binding pocket for the base and sugar moieties of the incoming dNTP appear to be primarily formed by residues from the fingers domain (Taq698 and Taq813 – 821). However, clear electron density corresponding to the triphosphate moiety of the incoming dNTP (Figure 1) is seen above the three aspartate residues (Taq463, Taq465 and Taq618) that have been previously seen by site-directed mutagenesis to be essential for catalysis³⁵. The γ -phosphate of the incoming dNTP appears to contact the side chain of the serine of the glycine-serine (GS) motif (Taq425 – Taq426), a motif found in all nucleotidyl transferases²⁸, as well as an arginine residue (Taq767) from the fingers domain. All polymerases utilize a two metal ion mechanism of catalysis²⁹. One metal ion (metal B) sits between the triphosphate moiety of the incoming nucleotide and the catalytic aspartates while the second metal ion (metal A) is coordinated by the α -phosphate of the incoming nucleotide and the primer terminal 3' hydroxyl. At the current resolution it is difficult to assign the position of metal ions into our electron density maps. However, we assume that at a minimum metal B has to be present in order to bridge the interaction between the negatively charged triphosphate moiety of the dNTP and the negative charge of the catalytic carboxylates. It is probable that any metal ions present at the polymerase active site are calcium ions rather than magnesium ions because the crystals were obtained in the presence of 10mM magnesium chloride and 100mM calcium chloride.

Conformation of the Primer-Template DNA

The electron density corresponding to the upstream duplex is consistent with the conformation of B-form DNA (Figure 2A), as observed previously in the structures of the ternary complexes of Pol β ^{31, 36} and RB69 polymerase³⁷. However, a 30° bend, whose midpoint is located at the –2 position (Figure 1), is imposed in the DNA by the loops that connect the thumb and palm domains (Figure 3C). The current electron density maps are too ambiguous to model the base moieties of the downstream template, however, the path of the sugar-phosphate backbone is clear (Figure 2A). The downstream template exits the active site with a sharp kink that is imposed by the fingers domains as has also been seen in other ternary complex structures³¹;

^{32; 37–39}. In contrast to other DNA polymerases, PolIII α appears to interact via the OB-fold of its CTD more extensively with the downstream ssDNA template (Figure 2A), consistent with the suggested role of the CTD in the lagging strand processivity switch^{25; 40}.

Conformational Changes Induced by Substrate Binding

A comparison of the unliganded and ternary complex structures of Taq PolIII α reveals numerous conformational changes in the polymerase that take place upon binding DNA, some of which are similar to those that were previously predicted²⁵ and some of which were unanticipated but are functionally important. As predicted, the β -binding domain rotates by $\sim 20^\circ$ towards the palm, thereby allowing it to interact with the DNA substrate (Figure 4A· Supplemental Movie 1) and positioning the internal clamp binding site of the polymerase adjacent to the upstream duplex DNA (see below). The CTD rotates in conjunction with the β -binding domain but, as was not predicted, also undergoes an additional $\sim 10^\circ$ rotation that appears to bring its OB-fold into a position to bind the downstream ssDNA template strand (Figure 4B· Supplemental Movie 2). Movement of the thumb away from the PHP domain (Figure 4A) optimally positions the thumb to grip the duplex DNA from the -3 to the -10 position (Figure 3A). Only a minor displacement of the PHP nuclease domain is observed. A 15° rotation of a region within the fingers domain (residues Taq699–Taq810), denoted as the index finger by Lamers et al., (2006), inwards toward the palm (Figure 4A) appears to help form the binding pocket for the bound incoming nucleotide. In particular, this movement positions an absolutely conserved arginine residue (Taq767) within contact distance of the γ -phosphate of the incoming nucleotide. A similar closing of the fingers domain has been observed in the structures of other polymerase ternary complexes^{32; 37–39}.

Discussion

PolIII α binds active site DNA in the same manner as Pol β

Although the previously determined structures of the unliganded Eco fragment of PolIII α ²⁷ and the intact enzyme from Taq were interpreted differently with respect to whether PolIII α is homologous to Pol β and hence binds its DNA substrate similarly, the structure of the presented ternary complex of Taq PolIII α both resolves and explains these differences. Not only are the structures of the palm domains of PolIII α and Pol β the same as noted by Bailey et al (2006), but additionally the structures of the ternary complexes of PolIII α and Pol β show the same orientation of their DNA substrates relative to their palm domains. In both cases, the plane of the base pairs of the DNA substrate is parallel to the strands of the palm domain. Furthermore, the positions of active site residues and, 3' primer terminus and incoming nucleotide are likewise similar. These elements are positioned nearly identically to accomplish catalysis in the ternary complex structures of PolIII α and Pol β (Figure 4C and 4D). Once in the active site, two Mg²⁺ ions, the triphosphate moiety of the incoming nucleotide and the primer-terminal 3' hydroxyl must be properly positioned in order for the phosphoryl transfer reaction to occur. The catalytic magnesium ions will be coordinated by three conserved aspartate residues (Taq463, Taq465 and Taq618). A conserved GS motif, two arginine residues (Taq452 and Taq458) and a magnesium ion (metal B) coordinate the triphosphate moiety of the incoming nucleotide. Positioning of the primer-terminal 3' hydroxyl relies on coordination of the primer-terminal phosphate by Taq616 and one of the two magnesium ions (metal A). Indeed, a least-squares superposition of the 3' primer terminus, the incoming nucleotide, the conserved Mg²⁺ binding catalytic aspartates, the GS motif and lysine (Taq616) observed in the structure of the ternary complex of PolIII α onto the equivalent residues of Pol β (D190, D192, D256, G179, S180 and R254) results in an RMSD of 1.36 Å over the C α atoms of the protein residues and all the atoms of the incoming nucleotide.

These similarities between the structures of the PolIII α and Pol β ternary complexes not only confirm that the palm domains share the same fold, but importantly establish that they bind DNA in the same manner, thereby providing further evidence that they share a common ancestor. It is not surprising however that the pathway of the DNA outside the active site diverges between the two enzymes, since they do not share any common protein folds outside of the palm domain.

Comparison of the structures *E. coli* and *Taq* PolIII α

Comparison of the overall structures of unliganded Eco PolIII α and the ternary complex of *Taq* PolIII α explains the different previous conclusions concerning its homology to Pol β and DNA binding. Although in the unliganded Eco structure the β -binding domain and fingers are in a more open conformation and the thumb and PHP nuclease domains are rotated away from the palm domain, the movement of the thumb, fingers and β -binding domain that would be required for the Eco PolIII α to form a ternary complex are consistent with the conformational changes that are observed between the structures of unliganded *Taq* PolIII α and its ternary complex. However, the critical difference between the Eco and *Taq* PolIII α structures^{25, 27} is the conformation of β -sheet of the palm domain, particularly the 6th strand of their palm domains that contains the third catalytic carboxylate residue (*Taq*618 and Eco555) (Figure 5A). The distance between the C α atoms of the second (Eco403) and third (Eco555) catalytic carboxylates in the structure of Eco PolIII α is 7.6 Å, compared with a distance of 4.57 Å in the structure of *Taq* PolIII α (Figure 5A). The average distance between these C α atoms among the 16 structures of enzymes that possess the NT fold is 5 Å with a standard deviation of 0.7 Å (Supplementary Figure 6). Indeed, a least-squares fit of the C α atoms of the three catalytic carboxylates of Eco PolIII α onto Pol β results in an RMSD of 1.8 Å versus 0.3 Å for the equivalent fit between *Taq* PolIII α and Pol β .

The conclusion of Lamers et al. (2006) that “PolIII cannot bind DNA in a Pol β -like manner” resulted from the differences in the conformations of the structure of their palm domains, specifically the C-termini of their β -strands which contain the third catalytic carboxylate (Figure 5A), complicating their ability to align the Eco PolIII α structure with other structures containing the β NT fold. Lamers et al. (2006) included the third carboxylate in their alignment of the ternary complex of Pol β ³¹ onto their Eco PolIII α structure, which resulted in main-chain clashes between the DNA substrate of Pol β and the active site helix of Eco PolIII α . Likewise, inclusion of the third catalytic carboxylate in an alignment of the palm domains of the structures of unliganded Eco PolIII α and the ternary complex of *Taq* PolIII α results in similar clashes between the aligned DNA substrate of *Taq* PolIII α and the active site helix of Eco PolIII α (Figure 5B). However, if superposition is performed by excluding the third catalytic carboxylate, then no clashes are observed between the modeled DNA and the active site helix of Eco PolIII α (Figure 5C). Main-chain clashes are, however, observed between the homology-modeled DNA and the C-terminus of the 6th strand that contains the third catalytic carboxylate. In order to alleviate these clashes such that Eco PolIII α can bind DNA in the same manner as *Taq* PolIII α and Pol β , this strand must assume the same conformation that is observed in all other enzymes that utilize the β NT fold.

Thus, it appears that Lamers et al. (2006) concluded that Eco PolIII α does not bind DNA as Pol β does, because for some reason the tertiary structure of its palm domain is distorted. It is very unusual to find the ternary structure of a domain distorted, but may be a result of crystal packing, removal of the C-terminal domain or the lack of bound DNA. We expect that when the structure of a complex of Eco PolIII α with DNA and dNTP is determined, this distorted palm domain conformation will change to that observed and the resulting complex structure will closely resemble those of *Taq* PolIII α and Pol β .

A Structural Model of the complex between PolIII α , the clamp and DNA

PolIII α accomplishes processive replication by partnering with the β -sliding clamp protein^{7; 41; 42}. In the absence of the clamp, the polymerase only replicates 10 – 20 nucleotides before falling off of the DNA substrate^{41; 43}. Upon binding of the clamp, this processivity is increased to thousands of nucleotides per binding event^{7; 41; 42}. The clamp is a dimeric protein that forms a ring shaped molecule that contains two well studied binding sites for polymerase on its N-terminal face^{18; 20}.

The recent crystal structure of the β -sliding clamp bound to DNA³ allows us to construct a model of the interaction between PolIII α and clamp bound to primer-template DNA by superimposing the DNA's in the two co-crystal structures. Initially, the structure of the last base pair of the duplex DNA emanating from the N-terminal face of the clamp was superimposed on the structure of the last base pair of the upstream duplex DNA of the polymerase ternary complex, and then the structure of clamp and DNA was moved one base pair at a time towards the polymerase until one of the two polymerase binding sites (using PDB ID: 1JQL from Jeruzalmi et al., 2001)⁴⁴ on the N-terminal surface of the clamp aligned with the internal clamp binding site of the polymerase. To optimize the fit between the two proteins, the DNA was allowed to bend $\sim 5^\circ$ (Supplementary Figure 7). Such a bend appears to be reasonable given that the DNA substrate is flexible past the -12 position as observed in copy A of the polymerase ternary complex (Supplementary Figure 4). The resulting model (Figure 6A) produced no steric clashes between the two proteins. Further movement of the clamp towards the polymerase by just one base pair results in clashes between the clamp and the PHP domain as well as the β -binding domain of the polymerase, and movement of the clamp away from the polymerase by one base pair rotates the polymerase binding site of the clamp by 36° and out of contact distance with the polymerase. A striking finding in the clamp co-crystal structure with DNA is the 22° tilt of the clamp plane from being perpendicular to the DNA and its specific interactions with the DNA backbone. The excellent docking of the two complexes supports the validity of both the clamp-DNA interaction and the polymerase-DNA interaction. Additionally, the modeled distance of ~ 20 base pairs between the N-terminal surface of the clamp and 3' terminus of the primer (Figure 6A) is consistent with previous fluorescence⁴⁵, crosslinking⁴⁶ and DNA footprinting experiments⁴⁷.

The clamp is a homodimer and thus has two identical protein binding sites. Indeed, using fluorescence energy transfer experiments, it has been shown that both PolIII and PolIV are able to bind clamp simultaneously⁴⁸. It is believed that on encountering DNA damage, PolIII is displaced from the primer-template by PolIV, and that once PolIV has replicated past the damage, PolIII regains control of the primer-template DNA⁴⁸. The tilted configuration of the DNA observed in the clamp-DNA structure may aid in this polymerase switching mechanism³. We have been able to dock the structure of PolIV onto our model of the structure of the complex between PolIII α , clamp and DNA by superimposing the structure of full-length PolIV²¹ onto the structure of the clamp bound the little finger domain of PolIV⁴⁹. Consequently, these three co-crystal structures have allowed us to construct a model (Figure 6B) of a complex between PolIII α , PolIV, clamp and DNA that supports the proposed mechanism of polymerase switching at points of DNA damage^{3; 48}.

Experimental Procedure

Mutagenesis and Protein Purification

The exonuclease-deficient D20N, D212N double mutant of *Taq* PolIII α was generated using the QuikChange method (Qiagen). The expression and purification of the mutant polymerase was performed as previously described for the wild-type protein²⁵, except for a 200mM to 50mM reduction in the concentration of sodium chloride in the gel filtration buffer.

Nuclease Assays

The double mutant of *Taq* PolIII α (70 μ M) was incubated with primer-template DNA (140 μ M) either alone or in the presence of ddCTP (1 mM) and dATP (1 mM) at 37°C for 1 hour or overnight. Reactions were quenched using a mixture of formamide and bromophenolblue and xylene cyanol dyes at a ratio of 4 volumes dye mixture to 1 volume reaction mixture. Subsequent to quenching, reactions were incubated at 90°C for 10 min prior to gel loading. Nuclease reactions were then loaded onto a 25% acrylamide and 7M urea denaturing gel run in Tris-Borate-EDTA buffer. The gel was stained with Ethidium Bromide for visualization.

Formation of the Ternary Complex

Ternary complex was formed by incubating D20N, D212N mutant polymerase (70 μ M) with MgCl₂ (10 mM), ddCTP (1 mM), dATP (1 mM) and primer-template DNA (primer sequence CGAAACGACGGCCAGTGCCA; template sequence TTTTTTGTGGCACTGGCCGTCGTT TCG) (140 μ M) for 1 hr at 37°C.

Crystallization and Data Collection

Crystals were grown by the sitting-drop method of vapor diffusion, using well solutions consisting of 16 – 21% (w/v) polyethylene glycol 400, 100 mM calcium chloride and 100 mM HEPES (pH 6.9 – 7.2). Prior to flash freezing in liquid propane, crystals were transferred stepwise into a solution of 35% polyethylene glycol 400, 100 mM calcium chloride, and 100 mM HEPES pH 6.9. Data were collected at beamline X-29 at the National Synchrotron Light Source (NSLS) at Brookhaven National Laboratory or at beamline 24-ID at the Advanced Photon Source (APS). Data were integrated and scaled to 4.6 Å (to an $\langle I \rangle / \sigma I$ of 1.4) using the HKL suite of programs⁵⁰. Merohedral twinning of the crystals was detected in the program TRUNCATE⁵¹ and the initial estimate of the twin fraction was determined in the CNS software suite⁵². Crystallographic statistics are shown in Table 1.

Structure Determination and Refinement

In principle, crystal twinning should not hinder structure determination by molecular replacement (MR)³³. Therefore, MR was performed with the twinned data in the program PHASER⁵³ using the unliganded *Taq* PolIII α structure²⁵ as a search model. Two copies of the polymerase were found using this procedure, which were then broken into rigid bodies (body 1: residues 1 to 285; body 2: residues 286 to 492 and residues 576 to 522; body 3: residues 493 to 575; body 4: residues 523 to 835; body 5 residues 836 to 1012; and body 6; residues 1013 to 1220) and the positions of these rigid bodies were refined against the twinned amplitudes in the CNS suite⁵² including the twinning operator (k, h, -l), the twin fraction and non-crystallographic symmetry (NCS) restraints. Test reflections were selected such that all the twin-related pairs belonged to either the test set or the working set. The initial model output by phaser yielded a twinned R-factor and R-free of 32.47% and 31.94%, respectively. Phases derived from this model were then improved by solvent flattening and two-fold NCS averaging in the program DM⁵¹ using amplitudes that had been de-twinned by the program DETWIN⁵¹. Individual rigid bodies were then fit to omit maps, that to further reduce model bias, were generated using phases that had been solvent flattened and two-fold averaged. All model building was performed in the program COOT⁵⁴. Omit maps of each of the polymerase domains can be seen in Supplementary Figure 8. Once refinement and model building of the protein had converged the DNA and incoming nucleotide were fit to their unbiased electron density (Figure 1). The resulting ternary complex model was then subjected to a final round of rigid body refinement. The final twin fraction of 45% was determined by comparing Rfree values after refinement.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the staff at NSLS beamline X-29 and APS beamline 24-ID for help with data collection. This research was supported by the NIH grant GM57510 to T.A.S. The project described was supported by Grant Number T32GM007223 from the National Institute of General Medical Sciences to R.A.W. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute Of General Medical Sciences or the National Institutes of Health.

References

1. Kornberg, A.; Baker, T. DNA Replication. New York: Freeman; 1992.
2. Johnson A, O'Donnell M. Cellular DNA replicases: components and dynamics at the replication fork. *Annu Rev Biochem* 2005;74:283–315. [PubMed: 15952889]
3. Georgescu RE, Kim SS, Yurieva O, Kuriyan J, Kong XP, O'Donnell M. Structure of a sliding clamp on DNA. *Cell* 2008;132:43–54. [PubMed: 18191219]
4. Kong XP, Onrust R, O'Donnell M, Kuriyan J. Three-dimensional structure of the beta subunit of E. coli DNA polymerase III holoenzyme: a sliding DNA clamp. *Cell* 1992;69:425–437. [PubMed: 1349852]
5. Kuwabara N, Uchida H. Functional cooperation of the dnaE and dnaN gene products in Escherichia coli. *Proc Natl Acad Sci U S A* 1981;78:5764–5767. [PubMed: 6458043]
6. LaDuca RJ, Crute JJ, McHenry CS, Bambara RA. The beta subunit of the Escherichia coli DNA polymerase III holoenzyme interacts functionally with the catalytic core in the absence of other subunits. *J Biol Chem* 1986;261:7550–7557. [PubMed: 3519609]
7. Stukenberg PT, Studwell-Vaughan PS, O'Donnell M. Mechanism of the sliding beta-clamp of DNA polymerase III holoenzyme. *J Biol Chem* 1991;266:11328–11334. [PubMed: 2040637]
8. O'Donnell M, Jeruzalmi D, Kuriyan J. Clamp loader structure predicts the architecture of DNA polymerase III holoenzyme and RFC. *Curr Biol* 2001;11:R935–R946. [PubMed: 11719243]
9. Filee J, Forterre P, Sen-Lin T, Laurent J. Evolution of DNA polymerase families: evidences for multiple gene exchange between cellular and viral proteins. *J Mol Evol* 2002;54:763–773. [PubMed: 12029358]
10. Ito J, Braithwaite DK. Compilation and alignment of DNA polymerase sequences. *Nucleic Acids Res* 1991;19:4045–4057. [PubMed: 1870963]
11. Ohmori H, Friedberg EC, Fuchs RP, Goodman MF, Hanaoka F, Hinkle D, Kunkel TA, Lawrence CW, Livneh Z, Nohmi T, Prakash L, Prakash S, Todo T, Walker GC, Wang Z, Woodgate R. The Y-family of DNA polymerases. *Mol Cell* 2001;8:7–8. [PubMed: 11515498]
12. Patel PH, Loeb LA. Getting a grip on how DNA polymerases function. *Nat Struct Biol* 2001;8:656–659. [PubMed: 11473246]
13. Kohlstaedt LA, Wang J, Friedman JM, Rice PA, Steitz TA. Crystal structure at 3.5 Å resolution of HIV-1 reverse transcriptase complexed with an inhibitor. *Science* 1992;256:1783–1790. [PubMed: 1377403]
14. Steitz TA. DNA polymerases: structural diversity and common mechanisms. *J Biol Chem* 1999;274:17395–17398. [PubMed: 10364165]
15. Aravind L, Koonin EV. Phosphoesterase domains associated with DNA polymerases of diverse origins. *Nucleic Acids Res* 1998;26:3746–3752. [PubMed: 9685491]
16. Stano NM, Chen J, McHenry CS. A coproofreading Zn(2+)-dependent exonuclease within a bacterial replicase. *Nat Struct Mol Biol* 2006;13:458–459. [PubMed: 16604084]
17. Dalrymple BP, Kongsuwan K, Wijffels G, Dixon NE, Jennings PA. A universal protein-protein interaction motif in the eubacterial DNA replication and repair systems. *Proc Natl Acad Sci U S A* 2001;98:11627–11632. [PubMed: 11573000]

18. Dohrmann PR, McHenry CS. A bipartite polymerase-processivity factor interaction: only the internal beta binding site of the alpha subunit is required for processive replication by the DNA polymerase III holoenzyme. *J Mol Biol* 2005;350:228–239. [PubMed: 15923012]
19. Murzin AG. OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *EMBO J* 1993;12:861–867. [PubMed: 8458342]
20. Lopez de Saro FJ, Georgescu RE, O'Donnell M. A peptide switch regulates DNA polymerase processivity. *Proc Natl Acad Sci U S A* 2003;100:14689–14694. [PubMed: 14630952]
21. Ling H, Boudsocq F, Woodgate R, Yang W. Crystal structure of a Y-family DNA polymerase in action: a mechanism for error-prone and lesion-bypass replication. *Cell* 2001;107:91–102. [PubMed: 11595188]
22. Ollis DL, Brick P, Hamlin R, Xuong NG, Steitz TA. Structure of large fragment of *Escherichia coli* DNA polymerase I complexed with dTMP. *Nature* 1985;313:762–766. [PubMed: 3883192]
23. Wang J, Sattar AK, Wang CC, Karam JD, Konigsberg WH, Steitz TA. Crystal structure of a pol alpha family replication DNA polymerase from bacteriophage RB69. *Cell* 1997;89:1087–1099. [PubMed: 9215631]
24. Zhou BL, Pata JD, Steitz TA. Crystal structure of a DinB lesion bypass DNA polymerase catalytic fragment reveals a classic polymerase catalytic domain. *Mol Cell* 2001;8:427–437. [PubMed: 11545744]
25. Bailey S, Wing RA, Steitz TA. The structure of *T. aquaticus* DNA polymerase III is distinct from eukaryotic replicative DNA polymerases. *Cell* 2006;126:893–904. [PubMed: 16959569]
26. Davies JF 2nd, Almasy RJ, Hostomska Z, Ferre RA, Hostomsky Z. 2.3 A crystal structure of the catalytic domain of DNA polymerase beta. *Cell* 1994;76:1123–1133. [PubMed: 8137427]
27. Lamers MH, Georgescu RE, Lee SG, O'Donnell M, Kuriyan J. Crystal structure of the catalytic alpha subunit of *E. coli* replicative DNA polymerase III. *Cell* 2006;126:881–892. [PubMed: 16959568]
28. Aravind L, Koonin EV. DNA polymerase beta-like nucleotidyltransferase superfamily: identification of three new families, classification and evolutionary history. *Nucleic Acids Res* 1999;27:1609–1618. [PubMed: 10075991]
29. Steitz TA, Smerdon SJ, Jager J, Joyce CM. A unified polymerase mechanism for nonhomologous DNA and RNA polymerases. *Science* 1994;266:2022–2025. [PubMed: 7528445]
30. Garcia-Diaz M, Bebenek K, Krahn JM, Kunkel TA, Pedersen LC. A closed conformation for the Pol lambda catalytic cycle. *Nat Struct Mol Biol* 2005;12:97–98. [PubMed: 15608652]
31. Pelletier H, Sawaya MR, Kumar A, Wilson SH, Kraut J. Structures of ternary complexes of rat DNA polymerase beta, a DNA template-primer, and ddCTP. *Science* 1994;264:1891–1903. [PubMed: 7516580]
32. Doubie S, Tabor S, Long AM, Richardson CC, Ellenberger T. Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Å resolution. *Nature* 1998;391:251–258. [PubMed: 9440688]
33. Yeates TO. Detecting and overcoming crystal twinning. *Methods Enzymol* 1997;276:344–358. [PubMed: 9048378]
34. Doherty AJ, Serpell LC, Ponting CP. The helix-hairpin-helix DNA-binding motif: a structural basis for non-sequence-specific recognition of DNA. *Nucleic Acids Res* 1996;24:2488–2497. [PubMed: 8692686]
35. Pritchard AE, McHenry CS. Identification of the acidic residues in the active site of DNA polymerase III. *J Mol Biol* 1999;285:1067–1080. [PubMed: 9887268]
36. Sawaya MR, Prasad R, Wilson SH, Kraut J, Pelletier H. Crystal structures of human DNA polymerase beta complexed with gapped and nicked DNA: evidence for an induced fit mechanism. *Biochemistry* 1997;36:11205–11215. [PubMed: 9287163]
37. Franklin MC, Wang J, Steitz TA. Structure of the replicating complex of a pol alpha family DNA polymerase. *Cell* 2001;105:657–667. [PubMed: 11389835]
38. Huang H, Chopra R, Verdine GL, Harrison SC. Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. *Science* 1998;282:1669–1675. [PubMed: 9831551]
39. Li Y, Korolev S, Waksman G. Crystal structures of open and closed forms of binary and ternary complexes of the large fragment of *Thermus aquaticus* DNA polymerase I: structural basis for nucleotide incorporation. *EMBO J* 1998;17:7514–7525. [PubMed: 9857206]

40. Leu FP, Georgescu R, O'Donnell M. Mechanism of the E. coli tau processivity switch during lagging-strand synthesis. *Mol Cell* 2003;11:315–327. [PubMed: 12620221]
41. Bloom LB, Chen X, Fygenon DK, Turner J, O'Donnell M, Goodman MF. Fidelity of Escherichia coli DNA polymerase III holoenzyme. The effects of beta, gamma complex processivity proteins and epsilon proofreading exonuclease on nucleotide misincorporation efficiencies. *J Biol Chem* 1997;272:27919–27930. [PubMed: 9346941]
42. Maki S, Kornberg A. DNA polymerase III holoenzyme of Escherichia coli. III. Distinctive processive polymerases reconstituted from purified subunits. *J Biol Chem* 1988;263:6561–6569. [PubMed: 3283127]
43. Fay PJ, Johanson KO, McHenry CS, Bambara RA. Size classes of products synthesized processively by DNA polymerase III and DNA polymerase III holoenzyme of Escherichia coli. *J Biol Chem* 1981;256:976–983. [PubMed: 7005228]
44. Jeruzalmi D, Yurieva O, Zhao Y, Young M, Stewart J, Hingorani M, O'Donnell M, Kuriyan J. Mechanism of processivity clamp opening by the delta subunit wrench of the clamp loader complex of E. coli DNA polymerase III. *Cell* 2001;106:417–428. [PubMed: 11525728]
45. Griep MA, McHenry CS. Fluorescence energy transfer between the primer and the beta subunit of the DNA polymerase III holoenzyme. *J Biol Chem* 1992;267:3052–3059. [PubMed: 1737760]
46. Reems JA, Wood S, McHenry CS. Escherichia coli DNA polymerase III holoenzyme subunits alpha, beta, and gamma directly contact the primer-template. *J Biol Chem* 1995;270:5606–5613. [PubMed: 7890680]
47. Reems JA, McHenry CS. Escherichia coli DNA polymerase III holoenzyme footprints three helical turns of its primer. *J Biol Chem* 1994;269:33091–33096. [PubMed: 7806536]
48. Indiani C, McInerney P, Georgescu R, Goodman MF, O'Donnell M. A sliding-clamp toolbelt binds high- and low-fidelity DNA polymerases simultaneously. *Mol Cell* 2005;19:805–815. [PubMed: 16168375]
49. Bunting KA, Roe SM, Pearl LH. Structural basis for recruitment of translesion DNA polymerase Pol IV/DinB to the beta-clamp. *EMBO J* 2003;22:5883–5892. [PubMed: 14592985]
50. Otwinowski, Z.; Minor, W. Processing of X-ray diffraction data collected in oscillation mode. In: Sweet, JCWCARM., editor. *Macromolecular Crystallography, Part A*. New York, New York: Academic Press; 1997. p. 307–326.
51. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* 1994;50:760–763. [PubMed: 15299374]
52. Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* 1998;54:905–921. [PubMed: 9757107]
53. Storoni LC, McCoy AJ, Read RJ. Likelihood-enhanced fast rotation functions. *Acta Crystallogr D Biol Crystallogr* 2004;60:432–438. [PubMed: 14993666]
54. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 2004;60:2126–2132. [PubMed: 15572765]

Figure 1. DNA substrate density and schematic

Above: Electron density map calculated using observed amplitudes and phases calculated from the protein coordinates alone and contoured at 1σ . The observed DNA substrate is depicted using a stick representation. Below: The primer and template strands are colored grey and wheat, respectively. The bases are represented as rounded rectangles with the phosphate backbone represented as yellow circles. The 3' primer terminal ddC base is colored blue and the incoming nucleotide is colored red. The numbering used to identify specific bases is indicated in black above the template strand. The two bases of the substrate that are not observed in the crystal structure at the -20 and -21 positions are rendered transparent.

Figure 2. Crystal Structure of a Replicating Complex of Taq PolIII α

(A) A view of the crystal structure of the replicating complex of Taq PolIII α . Domains are labeled as follows: PHP Nuclease (yellow), Palm (pink), Thumb (light green), Index Finger (light blue), Middle Finger (dark blue), β -binding (orange) and C-terminal (red). The template and primer strand are represented as spheres and colored wheat and white, respectively. The 3' primer terminal base is depicted with blue spheres and the incoming nucleotide is depicted with red spheres. PolIII α is depicted in a cartoon representation. The density for the downstream DNA template strand is represented in mesh.

(B): A close-up of the active site of PolIII α . PolIII α is depicted using a surface representation. The DNA substrate is depicted as sticks and colored as in the left panel. The incoming nucleotide and 3' prime terminal base are colored red and blue, respectively. Key active site residues are colored as follows: D463, D465 and D618 are colored red, K616 is colored yellow, S426 is colored green, and R767 is colored blue. The Active Site Helix is colored magenta.

Figure 3. Interactions with the DNA substrate

(A) Surface representation of interactions between the Thumb domain of PolIIIα with the DNA substrate in cartoon representation. Green surfaces indicate the Thumb domain with blue surfaces representing residues proposed to interact with the DNA substrate. Yellow surfaces represent the PHP Nuclease domain. The polymerase active site is seen at left as demarcated by the blue stick representation of the 3' Primer Terminal Base (ddC) and the red stick representation of the incoming nucleotide (dATP).

(B) Surface representation of interactions between the β-binding domain of PolIIIα with the DNA substrate in cartoon representation. Orange surfaces indicate the β-binding domain with three key loops implicated in DNA binding colored as follows: Taq927 – Taq923 (blue), HhH motif Taq892 – 910 (cyan), and Taq846 – Taq852 (purple). Positively-charged residues that are within proximity to interact with the sugar-phosphate backbone and labeled and their positions are indicated with white arrows. An additional arrow points to the location of the polymerase active site.

(C) The observed DNA substrate of PolIIIα is depicted using a cartoon representation with the primer strand colored white and the template strand colored wheat. Idealized B-form DNA aligned with the DNA substrate prior to bending is displayed in a surface representation. Black lines indicate the helical axes of the DNA substrate (outside of the active site) and the idealized B-form DNA. The 3' primer terminal base is represented in blue sticks. The loops of the thumb domain (Taq492 – Taq499 and Taq569 – 575) that induce DNA bending are colored green and represented as a surface.

Figure 4. Comparison of Polβ and PolIIIα and conformational changes

(A) Cartoon Depiction of domain changes in PolIIIα associated with binding of DNA substrate. Domains are colored as in figure 2. Colored arrows indicate the direction of domain movement. Domain positions from the unliganded PolIIIα structures are presented in a cartoon representation and colored grey.

(B) Changes in the position of the OB-fold associated with binding of DNA substrate. The positions of the OB-fold of PolIIIα in unliganded and ternary complex conformations are shown in a cartoon presentation and colored grey and red, respectively. Here, the very C-terminus of the C-terminal domain and the β-binding domain are omitted for clarity. The density for the downstream DNA template strand is represented in mesh.

(C) The polymerase active site in the structure of the ternary complex of PolIIIα. The DNA substrate, the 3' primer terminal base, metals and the incoming nucleotide are colored as in Figure 2. The 3' primer terminal base and incoming nucleotide are represented as sticks. The

five β -sheets of the palm domain and the DNA substrates are displayed in a cartoon representation and colored pink and wheat (template) and white (primer), respectively. (D) The polymerase active site in the structure of the ternary complex of Pol β .

Figure 5. Comparison of *E. coli* and *T. aquaticus* PolIII α Structures

(A) The two sheets of the palm domain which contain the catalytic aspartate residues are depicted in a cartoon representation. The catalytic residues and sheets of the palm domains of Eco PolIII α , Taq PolIII α , and Pol β are colored grey, pink, and green, respectively. The catalytic aspartates are depicted as sticks and are labeled as follows: “1” represents Eco401, Taq463 and D190 of Pol β ; “2” represents Eco403, Taq465 and D192 of Pol β ; and “3” represents Eco555, Taq618 and D256 of Pol β .

(B) The palm domain of Eco PolIII α in relation to the DNA substrate of Taq PolIII α when including the third catalytic carboxylate as a reference for superposition of the palm domain of Eco PolIII α onto that of Taq PolIII α . The DNA substrate is depicted as in Figure 1. The Eco palm is colored grey. The active site helix is colored magenta. A mainchain clash between the active site helix of Eco PolIII α and the DNA substrate of Taq PolIII α is indicated by a white arrow. Sidechains closest to the DNA substrate are depicted as spheres.

(C) A superposition of the palm domains of Eco and Taq PolIII α when the third catalytic carboxylate is omitted as a reference.

Figure 6. Models of Clamp Binding to PolIII α and polymerase switching

(A) The polymerase and duplex DNA substrates are represented and colored as in Figure 1. The clamp is shown in a cartoon representation and colored purple. The polymerase binding sites on the clamp are colored cyan and depicted in a sphere representation.

(B) Here, the model of two polymerases simultaneously bound to the sliding clamp is depicted with PolIII and PolIV in surface representation and colored as in Figure 2 and green, respectively and rotated about the helical axis of the DNA into the page by about 60°. The clamp is depicted in a cartoon representation and colored purple. The position of PolIV was achieved using a model of the little finger domain of *E. coli* Pol IV bound to clamp (PDB ID: 1UNN) (Bunting et al., 2003) and the structure of the ternary complex of Pol IV (PDB ID: 1JXL) (Ling et al., 2001) and the same modeling procedure as in (Indiani et al., 2005).

Table 1Crystallographic Data and Refinement Statistics of the ternary complex of PolIII α

Space group	P4 ₁
Unit Cell Dimensions (Å)	149.546 × 149.546 × 163.338
Resolution (Å)	50 – 4.6
Unique Reflections	20,097
R _{sym} (%) ^{a,b}	7.8 (70.0)
I/σ ^a	15.3 (1.4)
Completeness (%) ^a	99.7 (97.6)
Redundancy ^a	3.6 (3.0)
Copies in AU	2
R _{cryst} (%) ^c	28.7 (37.4)
R _{free} (%) ^c	27.1 (36.7)
Twin Fraction (%)	45
# of protein atoms in AU	18,338
# of DNA atoms in AU	1,583
# of dNTP atoms in AU	58
# of metal atoms in AU	2
PDB ID	XXXX

^aNumbers in parentheses correspond to the highest resolution shell.^bR_{sym} = Σ|I – <I>|/ΣI where I = observed intensity and <I> = average intensity of multiple observations of symmetry-related reflections.^cAfter rigid body refinement.