Structural and functional insights provided by crystal structures of DNA polymerases and their substrate complexes

Chad A Brautigam* and Thomas A Steitz*†

New levels in the understanding of DNA replication have been achieved from recent crystal structure determinations of several DNA polymerases and their substrate complexes. The structure of an α family DNA polymerase from bacteriophage RB69 shows some similarities, but also considerable differences in structure and organization from the pol I family DNA polymerases. Also, the functions of three polymerase domains and their conserved residues have been clarified by studying structures of pol I family DNA polymerases complexed to their substrates. These structures also confirm that an identical two-metal ion catalytic mechanism proposed previously is used by both the nonhomologous pol I and pol β family DNA polymerases.

Addresses
*Department of Molecular Biophysics and Biochemistry, †Department of Chemistry, *Howard Hughes Medical Institute, Yale University, PO Box 208114, New Haven, CT 06520-8114, USA

Current Opinion in Structural Biology 1998, 8:54–63

http://biomednet.com/elecref/0959440X00800054

Abbreviations

ddNTP	dideoxynucleoside triphosphate
gp43	gene 43 protein
KF	Klenow fragment
pol I	DNA polymerase I
pol α	DNA polymerase α
pol β	DNA polymerase β
RT	reverse transcriptase
T7 pol	T7 DNA polymerase
***Taq* pol**	*Taq* DNA polymerase

Introduction

Significant advances in the structural biology of DNA polymerases have been made on two fronts. First, the library of known polymerase structures has expanded to include the gene 43 protein (gp43) from bacteriophage RB69 [1••], a member of the human DNA polymerase α family, offering insights into eukaryotic replicative polymerases. This expanded library also includes structures of polymerases from bacteriophage T7 [2••] and a *Bacillus stearothermophilus* strain [3•]—examples of polymerases homologous to DNA polymerase I from *Escherichia coli*. On the second front, a combination of structural and biochemical techniques has expanded our understanding of several properties of DNA polymerases, such as substrate binding [4••,5•,6••], processivity [2••,6••], fidelity [2••,7•], and nucleotidyl transfer [2••]. Although DNA polymerases often occur as multifunctional polypeptides, we shall restrict the scope of this review to the polymerase domains.

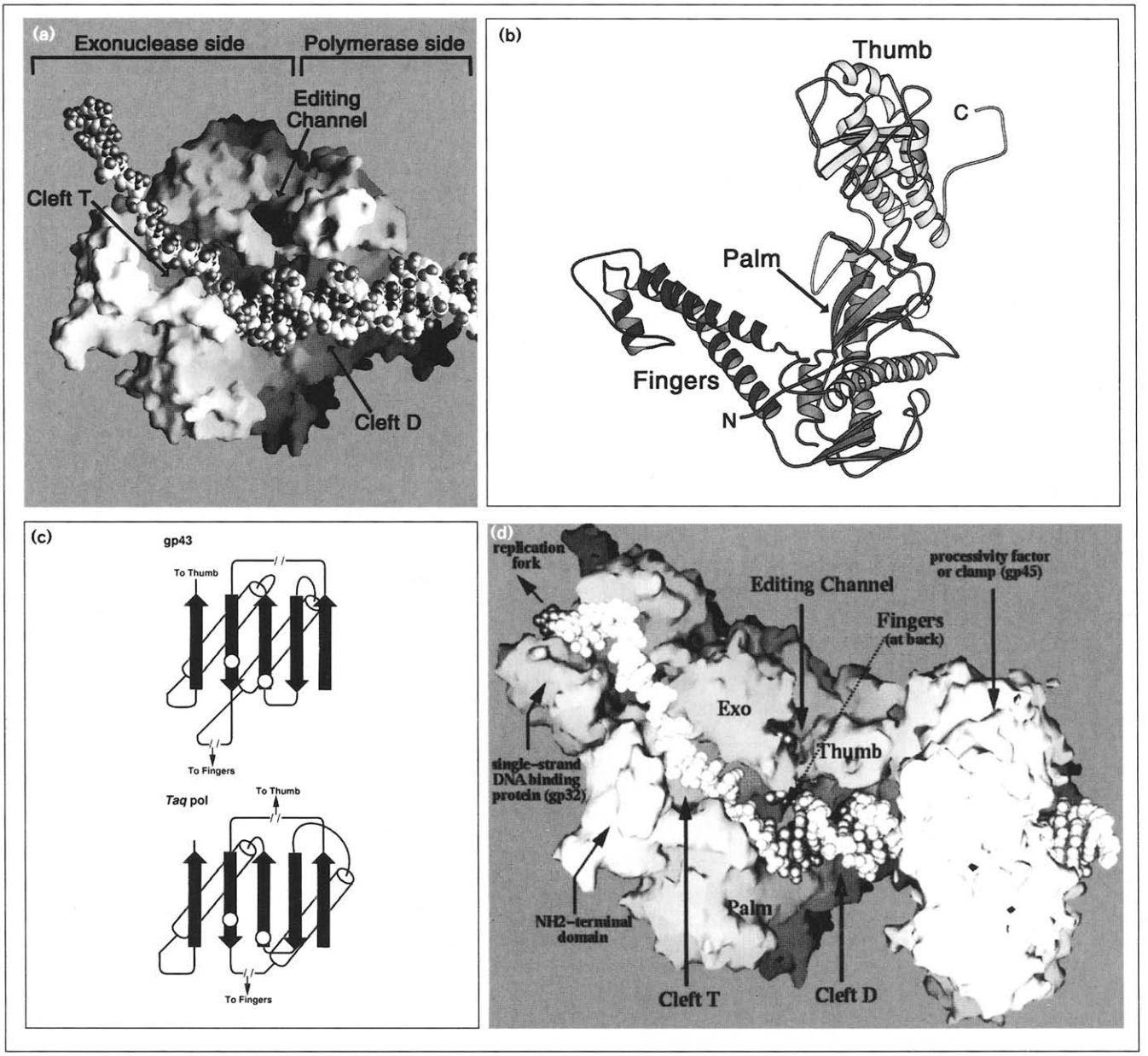
The structure of gp43 from bacteriophage RB69

The gp43 protein from bacteriophage RB69, which is a close homologue of bacteriophage T4, is primarily responsible for the replication of the viral genome [8]. It shares both sequence and functional homology with human DNA polymerase α (pol α), an enzyme implicated the replication of chromosomal DNA. The gp43 protein contains the six conserved sequence hallmarks, known as regions I–VI, that are common to all pol α-like or 'B family' polymerases [9]. Functionally, it shares several attributes with these polymerases: it is responsible for both leading- and lagging-strand synthesis, it is tethered to the substrate DNA by a protein processivity factor and it is stimulated by a single-stranded DNA binding protein [8]. Thus, knowledge gained from the study of gp43 will be applicable to eukaryotic replicative DNA polymerases.

The crystal structure of RB69 gp43 at 2.8 Å resolution [1••] reveals some surprising features along with some known structural motifs. The overall structure of the 103 kDa polypeptide is that of a disc with a small hole in the center (Figure 1). Arrayed around this central cavity are five structurally and functionally distinct domains that form three clefts, termed cleft D, cleft T, and the editing channel (Figure 1a). Approximately one-half of the disc is formed by the polymerase portion of gp43, which adopts the right hand shape common to all known polymerase structures. It comprises three domains: the fingers, palm and thumb (Figure 1b). The catalytic center of the polymerase is located on the palm domain, near the central cavity. The disc is completed by an N-terminal domain and the 3′–5′ exonuclease domain, which is homologous to the Klenow fragment (KF) editing domain. The structure of this half of the disc is nearly identical to that of the homologous KF domain from bacteriophage T4 gp43 determined by Wang *et al.* [10].

One unexpected feature of the RB69 gp43 structure is the position of the proof-reading domain with respect to the polymerase domain, which is opposite to their arrangement in the DNA polymerase I (pol I) family polymerases. Also, unlike the pol I family, a single-stranded tetranucleotide binds in the editing channel so that no translation of the primer–template duplex is required for the primer terminus to bind either in this channel or at the polymerase active site. In *Taq* DNA polymerase (*Taq* pol), the primer–template duplex bound to the polymerase active site is translocated relative to the frayed duplex DNA bound with its primer terminus in the KF-like proof-reading site [4••,11]. As is generally the case for the pol α family, the editing activity of this phage polymerase

Figure 1



The structure of gp43, the α family DNA polymerase from phage RB69. **(a)** A surface representation of the protein with homology modeled DNA. The polymerase half of the enzyme is on the right, whereas the exonuclease portion is on the left. The three clefts are labeled. The DNA in the structure has been modeled [1••] by comparison with the *Taq* pol protein–DNA co-crystal structure. The duplex product fits into cleft D while the modeled 5′ template was placed in cleft T. The central cavity is obscured by the DNA model. **(b)** A ribbon representation of the polymerase domain of gp43. The structure is color-coded by domain, with the fingers in the darkest gray, the thumb in the lightest gray, and the palm in an intermediate shade. The palm is dominated by an antiparallel β sheet. The thumb is mainly α-helical, but also contains some β strands. The unusual fingers feature two very long α helices. **(c)** Schematic representation showing the similar overall topologies of the palm domains of gp43 and *Taq* pol. Both comprise an antiparallel β sheet flanked on the noncatalytic side by two α helices. The positions of the conserved catalytic aspartates are shown as open circles. **(d)** A model of DNA and accessory proteins bound to gp43 [1••]. The DNA is modeled as in (a). The clamp protein, gp45, is proposed to interact with both the DNA and gp43 at the exit point of cleft D. The single-stranded DNA-binding protein, gp32, is proposed to interact with the single-stranded template after it leaves cleft T. Reproduced with permission from [1••].

is about 1000-fold higher than the exonucleolytic activity of the *E. coli* pol I enzyme [12–16]. The structure of a single-stranded DNA substrate bound to the exonuclease active site of gp43 shows the ε-amino group of a lysine residue in close proximity to the DNA scissile phosphate [1••,10]. This lysine is conserved among the pol α family, but is not present in the pol I family. The ε-amino group of this residue may enhance the exonuclease activity of gp43

by stabilizing the pentacovalent transition state through an interaction with a negatively charged, non-bridging phosphate oxygen.

A comparison of the polymerase motif of gp43 with the analogous motif in a binary complex between *Taq* pol [4••] and double-stranded DNA shows that the topology of the palm domains is identical and thus, by homology modeling, suggests the direction of DNA replication when bound to gp43. The structures of the β-sheet portions of the gp43 and *Taq* pol palm domains are very similar (Figure 1c). By superimposing the homologous palm domains of the binary complex of *Taq* pol and gp43 and by assuming that the two polymerases bind DNA similarly, a model for primer–template bound to gp43 was built [1••]. It positions the double-stranded portion of the DNA substrate at cleft D (Figure 1d). Furthermore, cleft T is in a suitable position to bind the single-stranded DNA template. This model of DNA bound to these clefts allowed Wang *et al.* [1••] to surmise the positions of the accessory proteins of gp43 within a functioning replisome (Figure 1d). The processivity factor of RB69, gp45, like that of the T4 phage, is a sliding-clamp protein whose donut-like structure is homologous to that of the β subunit of DNA polymerase III from *E. coli* (J Kuriyan, personal communication; see also [17]). It is proposed that gp45 interacts with the long C-terminal tail of gp43 as well as the double-stranded DNA on the polymerase side of the disc, at the exit point of cleft D [1••]. This model is supported by the fact that removal of the C-terminal tail of gp43 eliminates gp45-generated processivity [18]. The single-stranded DNA-binding protein from RB69, gp32, probably associates with gp43 on the exonuclease side of the disc, since cleft T exits the polymerase there. Thus, taken in the context of other polymerase structures, the structure of gp43 allows a reasonable model of a portion of the eukaryotic replication apparatus to be inferred.

The structure of gp43 also clarifies some aspects of the interrelatedness of pol α family polymerases and the relationships between this family and other polymerase families [1••]. Using the gp43 structure, the sequence alignments of all pol α family polymerases can be extended beyond regions I–VI to a span of nearly 200 residues encompassing all six of the previously identified sequence hallmarks (Figure 2a). An earlier attempt [19] to align the sequences of all DNA and RNA polymerases resulted in the conclusion that the active-site motif DTD (residues 621–623) of pol α was equivalent to a DD motif (residues 184–185) in HIV-1 reverse transcriptase (RT) or a DE motif (residues 882–883) in KF. The threonine of the pol α sequence was aligned as an insertion between the two conserved acidic residues. The structures of gp43, KF and HIV-1 RT [20], however, show that these two acidic residues are in similar positions only in KF and HIV-1 RT. Across all polymerases, only the second D of the DTD motif and the first D of the DD(E) motif are structurally conserved [1••] (Figure 2b).

Other polymerase structures

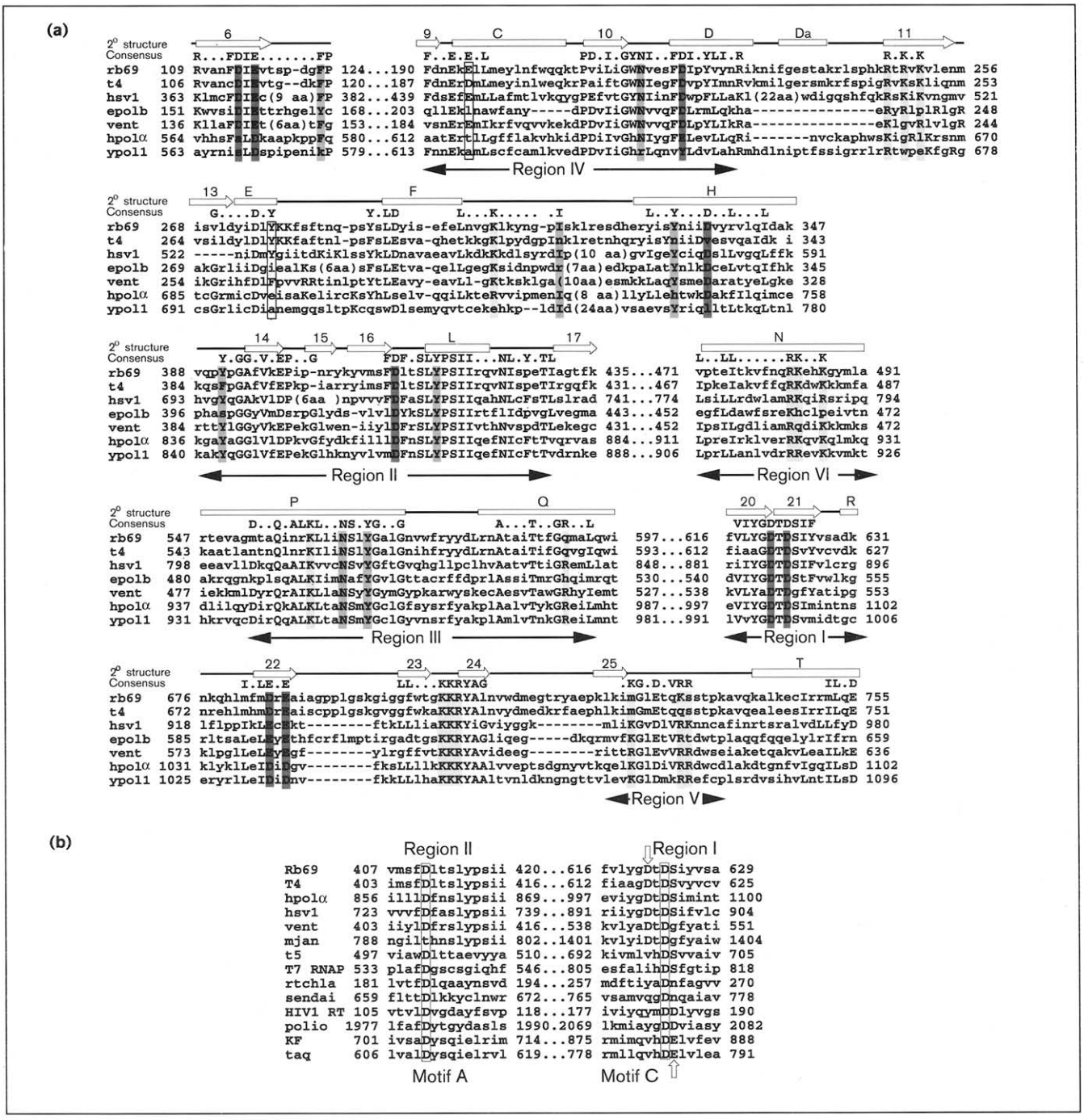
The crystal structures of two homologues of DNA polymerase I from *E. coli* have been solved recently. The first, a Klenow-like fragment of the DNA polymerase I of a thermophilic *Bacillus* species [3•], has been solved at 2.1 Å resolution. The structure of this enzyme is very similar to Klenow fragment [21], especially in its polymerase domain. The second structure, a quaternary complex of T7 DNA polymerase (T7 pol), thioredoxin (a processivity factor), DNA, and incoming dideoxynucleoside triphosphate was solved to 2.2 Å [2••]. The complex shows that the three-dimensional structure of T7 DNA polymerase is also analogous to that of KF.

Structural comparisons among DNA polymerases

A comparison of all DNA polymerase structures elucidated to date reveals some interesting commonalities and some striking differences. Presently, the structures of enzymes from four DNA polymerase families [9] have been established: RB69 gp43 from the pol α or 'B family' [1••]; KF [21], *Taq* pol [22], T7 pol [2••], and both the *Bacillus* and *Thermus aquaticus* KF analogues [3•,23] from the pol I family; rat and human DNA polymerase β (pol β) [5•,24,25] from the terminal transferase family [26]; and HIV-1 RT [20,27–29] and a fragment of Moloney murine leukemia virus RT [30] from the reverse transcriptase family. The palm domains of all DNA polymerases consist of a four- to six-stranded β sheet that is flanked on one side by two α helices. In the pol I, pol α and reverse transcriptase families, the topologies of the palm domain are the same. Even though the palm domain of pol β is mostly β sheet, its topology differs from that of the other polymerase families and thus was not derived from a common evolutionary ancestor. Beyond the palm domains, structural comparisons break down. The structures of the fingers domains vary widely from family to family, and, although the thumb domains are mainly α-helical, the detailed structures of these domains are also not related. Perhaps surprisingly, the fingers and thumb domains in all four families have arisen from different ancestors.

Although there are structural differences among the families of DNA polymerases, complexes of these enzymes with primer–template DNA show that the anatomical features of these enzymes play analogous roles across family boundaries, often using similarly oriented secondary structural elements. This can be illustrated by viewing a binary complex of blunt-ended double-stranded DNA and *Taq* pol, an enzyme whose polymerase motif is a close homologue of the polymerase motif of KF (Figure 3a) [4••,22]. The primer terminus of the DNA abuts the fingers domain of *Taq* pol and several acidic residues from the palm domain that are responsible for binding the catalytically essential metal ions are located close to the primer. The DNA leads away from the fingers and has several contacts with the palm domain. It then encounters the thumb domain, which has extensive contacts with

Figure 2



Sequence comparisons among DNA polymerases. **(a)** A comparison within the pol α family. Included in this comparison are polymerases from RB69 and T4 (rb69 and t4), herpes simplex virus 1 (hsv1), *E. coli* (pol II, [epolb]), *Pyrococcus furiosis* (vent), humans (pol α, [hpolα]), and yeast (pol I, [ypol1]). Arrows and cylinders above the alignment denote secondary structure, and regions I–VI are shown below. Capital letters are consensus residues. A consensus sequence is shown above the alignment. Residues shown in the darkest shade are conserved carboxylates, medium gray residues are solvent-exposed residues that may be involved in substrate binding, and lightest gray are positively charged residues. Boxed residues are carboxylates whose roles are unclear from the gp43 structure. **(b)** Sequence conservation among all DNA polymerases. In addition to some of the polymerases enumerated in (a), those from phage T5 (t5), *Methanococcus jannaschii* (mjan), mitochondrial *Chlamydomonas rheinhardtii* DNA intron (rtchla), sendai virus L (sendai), HIV-1 (HIV-1 RT), polio virus (polio), *E. coli* (KF), *Thermus aquaticus* (taq), and T7 RNA polymerase (T7 RNAP) are shown. Region II of the pol α family polymerases corresponds to motif A of [19], while region I corresponds to motif B. This sequence alignment is based on the conserved positions of carboxylates in the palm domain. Only two carboxylates are completely conserved across all of these polymerases (shown in boxes). The first D of the DTD motif and the second carboxylate of the DD(E) motif (arrows) are not structurally equivalent. Reproduced with permission from [1••].

the DNA across its minor groove. Whereas the above statements specifically concern *Taq* pol, they also hold true for T7 pol ([2••]; Figure 4a), HIV-1 RT ([27]; Figure 3b) and also for the homology model-built DNA complex with RB69 gp43 ([1••]; Figure 3c). The ternary complex of pol β with DNA and incoming dideoxynucleoside triphosphate (ddNTP) is shown in Figure 3d. Here, the primer–template DNA abuts domain D, and domain B contacts the DNA across the minor groove. Based on this functional analogy, it is clear that domains D and B would most usefully be called fingers and thumb, respectively. We shall follow this nomenclature throughout this review, even though a previous inappropriate alignment of pol β with non-homologous polymerases has led to an opposite, functionally less useful naming of these domains [25]. Viewed with this assignment in mind, all of the statements made above regarding *Taq* pol are true for pol β and an additional similarity among the pol α, pol I, and terminal transferase families becomes evident. They all have an α helix located in the fingers domain near the primer terminus that houses residues responsible for nucleotide binding. No analogous α helix is present in the reverse transcriptases, but rather there is an anti-parallel β hairpin located at the primer terminus.

The enzyme–substrate complexes of *Taq* pol [4••] and T7 pol [2••], along with the changed assignment of pol β, allow us to resolve a controversy that exists in the field [31–33]. A structure of an editing complex between KF and DNA shows the double-stranded portion of the substrate bound between the thumb and the 3′–5′ exonuclease domains [11]. This led Steitz and colleagues to propose that when the DNA was bound at the polymerase active site, it occupies roughly the same position as in *Taq* pol (Figure 3a). However, the alignment of KF and pol β resulting from the previous inappropriate superposition of their non-homologous palm regions spawned a proposal that the DNA should be bound to KF in the opposite orientation [25]. The structures of *Taq* pol and T7 pol demonstrate unequivocally that the proposal of Steitz and coworkers is correct and that the latter proposal should be disregarded.

Structural and functional insights into the activities of DNA polymerases

The quaternary complex of T7 pol [2••], the structure of *Taq* pol bound to DNA [4••] and several new structures of pol β [5•,7•] are remarkable for the insight they give into the processivity, fidelity, catalytic mechanism and conformational flexibility of DNA polymerases.

Processivity

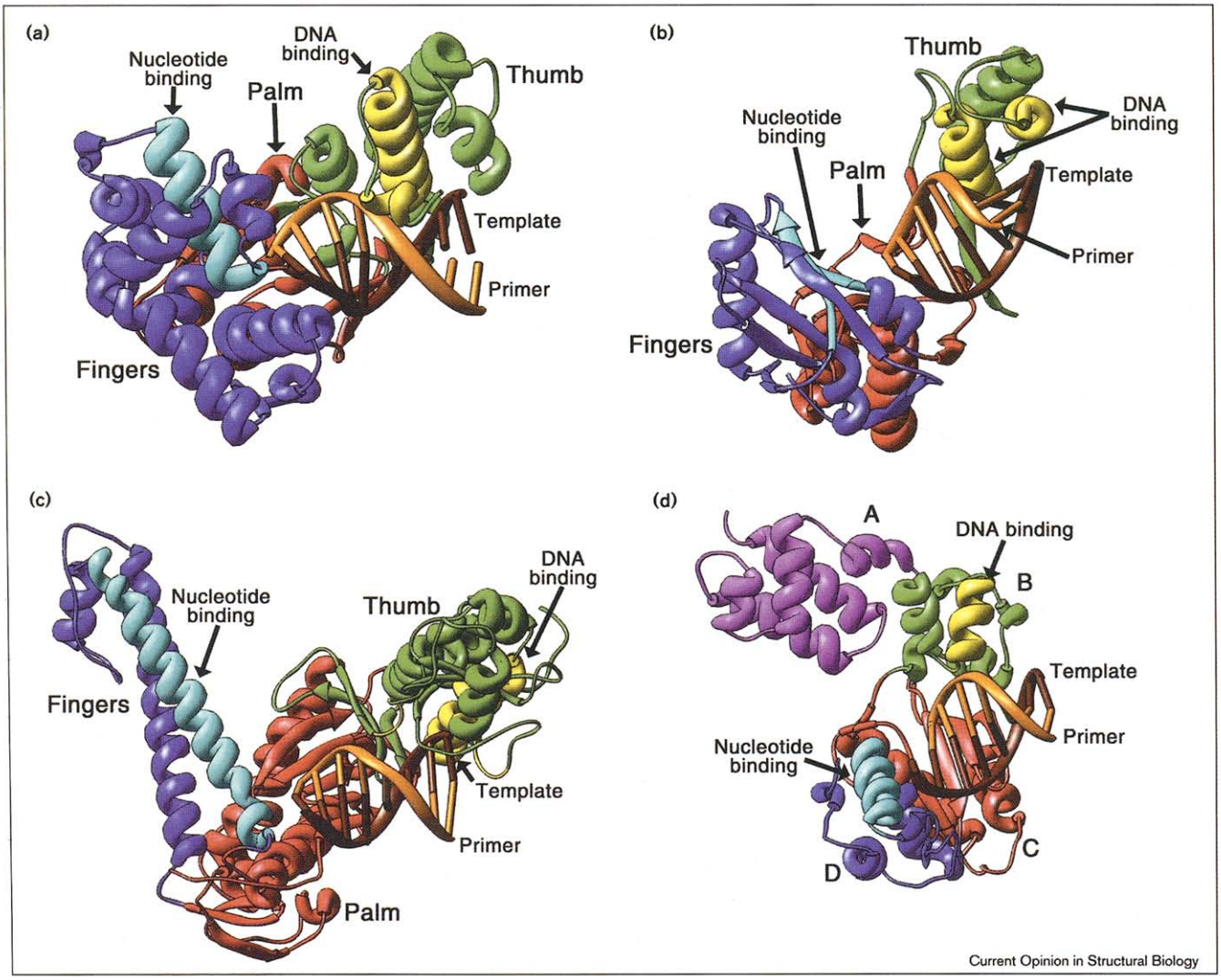
DNA polymerases have the unique feature that they may remain associated with the substrate primer–template over many catalytic cycles, that is, they are processive [8]. Replicative DNA polymerases need the aid of a protein factor in order to remain processive over the many thousands of nucleotide additions required for genome replication. T7 pol, a viral replicative DNA polymerase, commandeers the *E. coli* protein thioredoxin for this purpose [34]. The crystal structure of the quaternary complex of T7 pol and its substrates includes thioredoxin ([2••]; Figure 4a). The factor binds to a loop that protrudes from the tip of the polymerase thumb. Although electron density for DNA does not appear near thioredoxin due to disorder, extrapolation of the visible DNA helix positions the primer–template for a potential interaction with the factor (Figure 4a). Since all kinetic schemes for DNA polymerases include a step in the catalytic cycle during which the DNA is susceptible to dissociation from the enzyme [13,16,35,36], thioredoxin or other processivity factors act at this stage, either sterically or electrostatically hindering the dissociation of the DNA.

The thumb domain has been implicated in processivity in two other polymerases. When viewed along the bound duplex DNA, KF is seen to surround the DNA on three sides with the tip of the thumb on top [11]. Deleting the tip of the KF thumb reduces its processivity about fourfold [6••]. Furthermore, a thioredoxin-sensitive increase in processivity can be conferred on KF by grafting the thioredoxin-binding loop from T7 pol onto the tip of the KF thumb [37••]. In pol β, deletion of an 8 kDa extension from the tip of its thumb likewise reduces its processivity [38]. Also, a metal ion mediates interactions between the pol β thumb and the substrate DNA, possibly discouraging the DNA from dissociating during the enzyme's catalytic cycle [5•]. Thus, in these three examples (and perhaps in all polymerases), the thumb plays an important role in maintaining the polymerase's grip on its DNA substrates.

Fidelity

An enzyme's fidelity, or its ability to incorporate the correct substrate from among a sea of potential imitators, is a general problem in enzymology. The problem is magnified with DNA polymerases. These enzymes must choose from a pool of four deoxynucleoside triphosphates (dNTPs) and this choice is directed by yet another substrate (the single-stranded template). Furthermore, the consequences of an incorrect choice may be catastrophic to the organism. Thus, fidelity is of central importance in the reaction catalyzed by DNA polymerases. It is therefore surprising that in the T7 DNA polymerase quaternary structure [2••] there are no hydrogen bonds between the protein and the bases of the incipient base pair formed by the template and incoming nucleotide. Instead, it seems that the basis for error discrimination at this juncture is the steric complementarity between the protein and a correctly formed Watson–Crick base pair. This is in contrast to the pol β ternary structure [25], in which polymerase residues hydrogen bond to the nascent base pair. Amino acids from the palm of the T7 pol do, however, interact with base pairs near the 3′ terminus of the primer, suggesting that these residues are scanning for newly-made errors in polymerization. Since mutation of these residues results in a lower affinity of the enzyme for DNA [39,40], they could

Figure 3

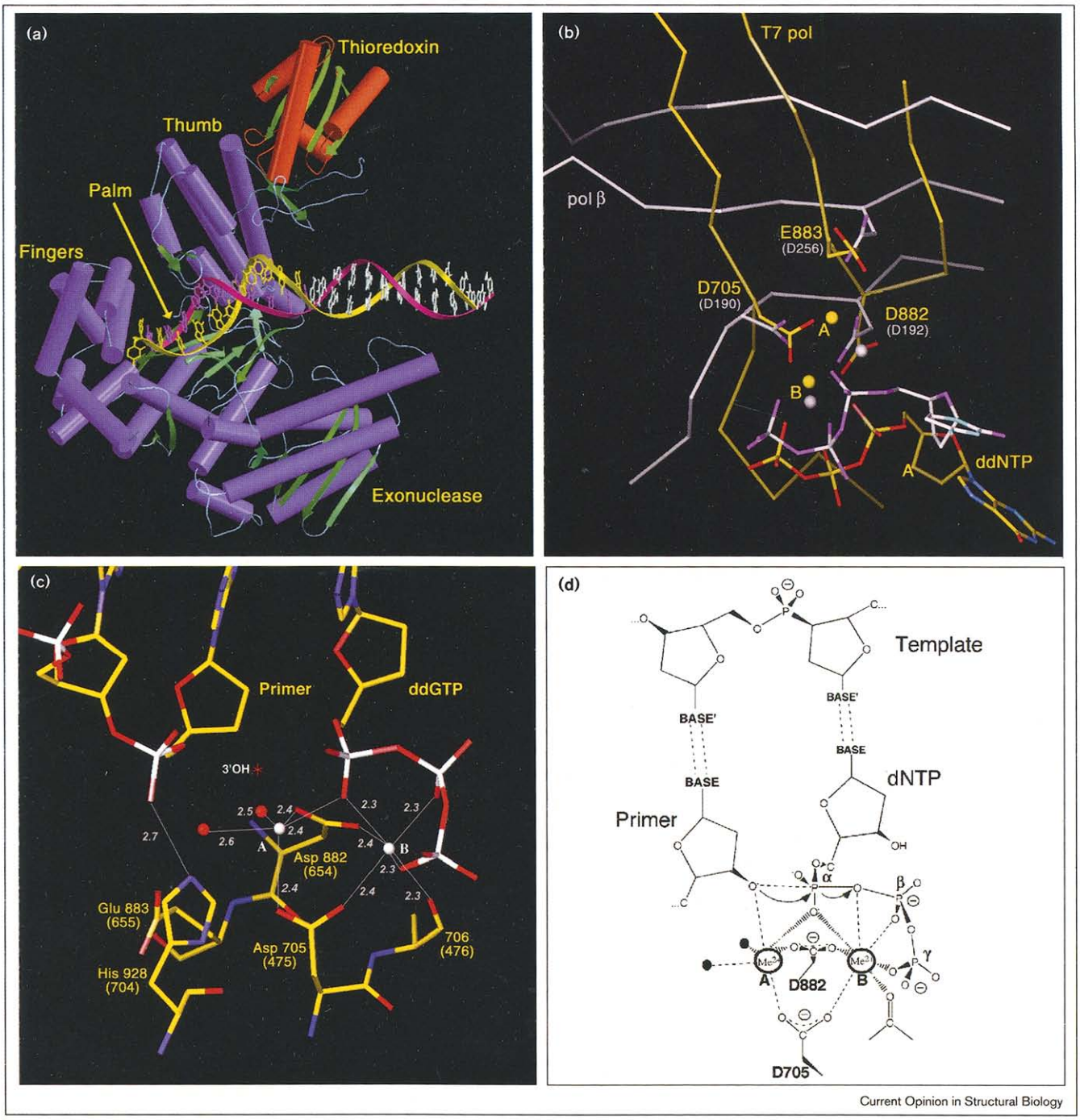


A comparison of primer–template DNA bound to four DNA polymerases. The complexes shown in (a), (b), and (d) are co-crystal structures, whereas the complex in part (c) is a homology model [1••]. These four structures have been aligned by the first two base pairs at the primer terminus. The fingers of these structures are colored blue, the palms are red, and the thumbs are green. Secondary structural elements at the primer terminus that help bind nucleotide are shown in cyan and those that contact DNA are yellow. The primer strand is shown in orange, while the template strand is brown. **(a)** *Taq* pol bound to DNA [4••]. As with the other three structures shown in this figure, the DNA stacks against the fingers and is contacted across the minor groove by the thumb domain. **(b)** The binary complex of HIV-1 RT and DNA [27]. This structure does not have a nucleotide-binding α helix in the fingers domains. Instead, a β hairpin probably performs this function. **(c)** The model of DNA bound to RB69 gp43 [1••]. A likely DNA-binding α helix has been highlighted. It appears that the thumb domain would have to move toward the primer terminus in order to bind DNA analogously to the other polymerases. **(d)** The ternary complex of rat pol β with DNA and ddNTP [25]. Domain D (purple) plays the role of the fingers and presents an α helix at the primer terminus. Domain B is analogous to other polymerase thumb domains and binds the minor groove of the duplex substrate.

be involved in shuttling newly incorporated mismatched nucleotides to the 3′–5′ exonucleolytic site of T7 pol [2••].

Several recently determined structures of human pol β complexed with substrate DNA also address the issue of fidelity. First, polymerases are more error-prone when the native, catalytic Mg^{2+} is replaced by Mn^{2+} [41–43]. Crystalline pol β is capable of adding a nucleotide, untemplated, onto a blunt-ended DNA in the presence of Mn^{2+} (or other mutagenic metal ions), but not in the presence of Mg^{2+} [7•]. This implies that mutagenic cations, perhaps owing to their increased affinity for carboxylates, cause mistakes by stabilizing non-native primer–template–nucleotide combinations. Also, observations of the interactions of pol β with DNA in the crystal have led to a proposal for how this polymerase binds at the single-stranded DNA gap [5•] that is presumably its native substrate. The position of the primer–template in

Figure 4



Nucleotidyl transfer in the T7 pol co-crystal structure. **(a)** Structure of T7 DNA polymerase complexed with thioredoxin, DNA, and incoming ddNTP. The primer is depicted in magenta and the template is in yellow. White bases denote the portion of the DNA that has been modeled. Purple cylinders represent polymerase α helices and the red cylinders are α helices from thioredoxin. The DNA abuts the fingers at the primer terminus and the thumb contacts DNA across the minor groove. **(b)** Superposition of the active sites of T7 pol and pol β using only the DNAs from the T7 quaternary structure and the pol β ternary structure. For T7 pol, KF numbering (yellow) is used and the pol β residues are labeled in white. The figure shows that the essential carboxylate residues and metal ions are in a similar 3D arrangement, despite the obvious disagreement in the direction of the β strands that make up the palm domains. **(c)** Active site of T7 pol shown using KF numbering, with the T7 pol numbers shown below in parentheses. Metal ions A and B are contacted by two protein residues [D882(654) and D705(475)], the phosphates of the ddNTP and two waters. The putative position of the primer's 3′-OH, which would contact metal ion A, is shown by a red star. Figures provided by Tom Ellenberger. **(d)** Mechanism of nucleotidyl transfer in T7 pol, using the KF numbering system. The active site features two metal ions that stabilize the resulting pentacoordinated transition state. Metal ion A activates the primer's 3′-OH for attack on the α-phosphate of the dNTP. Metal ion B plays the dual role of stabilzing the negative charge that builds up on the leaving oxygen, and chelating the β- and γ-phosphates.

this model is the same as in the ternary complexes, with the single-stranded template making a sharp turn so that the next double-stranded portion of the DNA

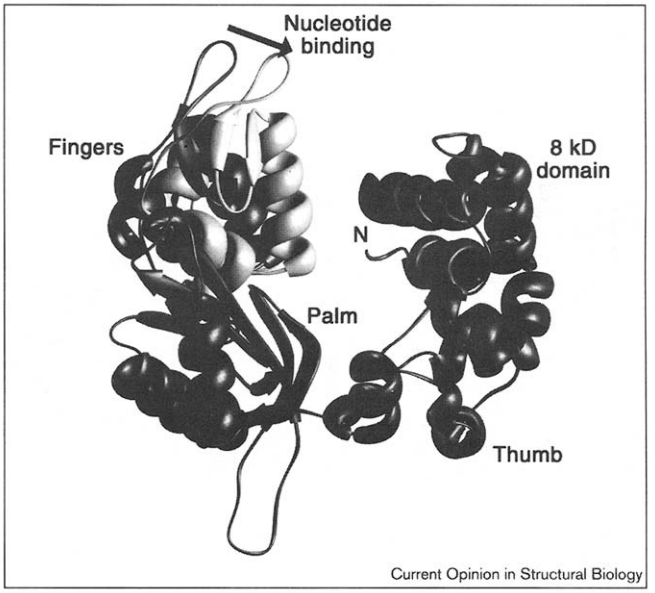
may interact with the 8 kDa domain at the tip of the thumb. Kraut and colleagues [5•] postulate that such a kink in the template has advantages for fidelity, as it would minimize non-specific stacking contacts between the incoming dNTP and the primer; the template strand is also sharply bent in the T7 quaternary complex [2••].

Nucleotidyl transfer

The crystal structures of T7 pol [2••] and of pol β [25] complexed with DNA and next correct dNTP show that these non-homologous enzymes use an identical catalytic mechanism to transfer a nucleotide onto the 3′ end of a DNA primer. The active sites of these unrelated polymerases share many essential and common features. When Ellenberger and colleagues [2••] aligned the two structures by superimposing their DNA substrates, the catalytically essential carboxylates from these disparate structures also superimposed (Figure 4b), as do the two metal ions bound to the triphosphate moiety in both structures. Although the terminal transferase family (to which pol β belongs) has evolved separately from other polymerase families and has a different catalytic domain topology, it shares a common three-dimensional arrangement of critical catalytic components. The structure and positions of the primer–template, dNTP and divalent metal ions are nearly identical in T7 pol and pol β.

The structure of the T7 pol quaternary complex provides a basis for understanding the roles of certain residues known from other studies to be essential for nucleotidyl transfer ([2••]; Figure 4c,4d). For the following discussion, the KF numbering scheme is used. Two conserved carboxylates, D705 and D882, are bridging ligands to the two catalytic metal ions. Significantly, these are the only two carboxylate residues that are absolutely conserved between the pol α, pol I and reverse transcriptase families ([1••]; Figure 2b). E883, which is conserved in the pol I family, does not contact the metal ions, as is consistent with its absence in the pol α family. The two cations have been dubbed A and B. The ligation environment of metal ion A is completed by two water molecules and the *pro*-R oxygen of the nucleotide's α-phosphate. This cation would also be contacted by the 3′-OH of the primer, if it were present (a dideoxy-terminated primer was utilized in the crystallization experiment). This hydroxyl group would be perfectly positioned to undertake an in-line attack on the nucleotide's α-phosphate. Metal ion B is contacted by oxygens from all three of the nucleotide's phosphates, as well as a main-chain oxygen. Curiously, the metal ion–ligand distances in the active site of T7 pol are somewhat longer than expected (Figure 4c). Nevertheless, the features of this Michaelis complex make it clear that a proposed two metal ion mechanism for DNA polymerases [44], which was based on an analogy to the 3′–5′ exonucleolytic two metal ion mechanism of KF [45], is essentially correct (Figure 4d). In this polymerase mechanism, metal ion A activates the 3′-OH of the primer for attack by lowering its pK_a. This nucleophilic attack at the α-phosphate generates a pentacoordinated intermediate, whose structure would be stabilized by both metal ions. A build-up of negative charge on the leaving oxygen of the β-phosphate could be stabilized by metal ion B. Once the reaction is complete, the pyrophosphate product and the metal ions dissociate, and the DNA must translate and rotate relative to the polymerase so that the newly formed primer terminus is correctly positioned in the primer-binding site for another catalytic cycle. The only significant deviation from the earlier proposal is that metal ion B appears to be playing the dual role of transition-state stabilization and chelation of both the β- and γ-phosphates.

Figure 5



Finger domain movement in DNA polymerases. The palm domains of the binary [5•] and ternary [27] complexes of pol β were used to superimpose the two protein structures. The binary (protein and DNA) form of the polymerase is shown in the darker shade and only the fingers domain of the ternary (protein and DNA and dNTP) form is shown in the lighter shade. The fingers close down in the presence of the nucleotide.

Conformational changes

The binding of dNTP and duplex DNA to both pol I family and pol β family DNA polymerases produces large changes in the orientations of the fingers and the thumb domains relative to the palm domain. A comparison of the binary complexes of polymerases and DNA with ternary complexes containing an incoming dNTP shows that the fingers close down on the DNA upon binding of the nucleotide in both pol β ([5•]; Figure 5) and pol-I-like enzymes [2••,4••]. A comparison of the *Taq* and

T7 pol complexes shows a few angstroms difference in the position of the primer terminus in these two complexes. Although it is unclear whether these conformational differences are part of the polymerase catalytic cycle or are influenced by crystallographic artifacts, they provide a rational explanation for the kinetic observation [36,46–48] that the rate-limiting step in the polymerase reaction pathway occurs after the binding of dNTP [2••,5•]. The thumb domains of polymerases also move in response to binding DNA [4••,11,25,49], and this mobility may also play a role in the catalytic cycle.

Conclusions

Principal among the many new structural and functional insights into DNA polymerases obtained during the last year have been the structure determinations of a pol α family DNA polymerase and a T7 DNA polymerase complexed with DNA and nucleotide substrates. The structure of RB69 gp43 strongly suggests how the replisomes of this and related polymerases are organized. The structures of *Taq* pol with DNA, pol β with various metal ions and substrates, and the quaternary structure of T7 DNA polymerase all give a much needed insight into the workings of DNA polymerases. These studies, taken together with the results of mutagenic experiments, suggest the functions of the domains that make up polymerases: the fingers are involved in correctly positioning the template and the complementary NTP relative to the catalytic metal ions, and undergo a substrate-induced conformational change that may be important to catalysis; the palm domain harbors two completely conserved carboxylate residues that bind the essential metal ions and also contains residues that may detect replication mistakes; and the thumb domain is important for DNA positioning and processivity. Additionally, the catalytic mechanism of nucleotidyl transfer in the polymerase reaction may be universal, involves two metal ions and is very similar to that of the proof-reading exonucleases. Though the model of RB69 gp43 with DNA and accessory proteins is instructive, it is essential to determine the crystal structures of this and similar polymerases with their substrates and associated factors in order to confirm and extend this model.

Perhaps the most striking findings from these polymerase structures, as anticipated by the limited sequence similarity among families, are the differences in the structures of the fingers and thumb domains of polymerases from different families. The thumb and fingers domain structures are completely unrelated among the four DNA polymerase families now known: pol I, pol α, pol β and reverse transcriptase. Nevertheless they function similarly in all families, using in most cases analogous secondary structural elements. Why are these domains, so essential to the polymerase reaction, derived from different ancestors? It would of interest to establish the structure of the bacterial DNA polymerase III, which shows no recognizable sequence similarity to other polymerases. It may possess yet another polymerase fold and could yield clues to understanding the surprising differences among these vital enzymes.

Acknowledgements

The authors wish to thank S Doublié and T Ellenberger for providing data and figures in advance of publication, J Wang for providing figures and molecular models, and J Jäger and J Pata for access to molecular models. This work was funded in part by the American Cancer Society grant #BE-52K to TAS.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. •• Wang J, Sattar AKMA, Wang CC, Karam JD, Konigsberg WH, Steitz TA: **Crystal structure of a pol α replication DNA polymerase from bacteriophage RB69.** *Cell* 1997, **89**:1087-1089.

The authors describe the first crystal structure of either a pol α or 'B family' polymerase. This structure has wide-ranging implications for the relatedness of DNA polymerases, proof-reading, the organization of the RB69 replisome and eukaryotic DNA polymerases that could not be modeled on the pol I enzyme.

2. •• Doublié 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.

The elucidation of the structure of the ternary complex of T7 DNA polymerase with its substrates is described. This 2.2 Å resolution structure shows how substrates bind to this KF-like polymerase and suggests a mechanism by which thioredoxin allows the DNA to stay associated with the polymerase. A detailed view of the active site shows that the catalytic mechanism involves two metal ions.

3. • Kiefer JR, Mao C, Hansen CJ, Basehore SL, Hogrefe HH, Braman JC, Beese LS: **Crystal structure of a thermostable *Bacillus* DNA polymerase I large fragment at 2.1 Å resolution.** *Structure* 1997, **5**:95-108.

The high-resolution structure of the large fragment of this enzyme shows that the polymerase domain of pol I from *B. stearothermophilus* is shown to be highly homologous to the same domain of KF. The structural basis for the lack of 3′–5′ exonuclease activity of the *Bacillus* enzyme is also discussed.

4. •• Eom SH, Wang J, Steitz TA: **Structure of *Taq* polymerase with DNA at the polymerase active site.** *Nature* 1996, **382**:278-281.

The structure of a binary complex of *Taq* pol and blunt-ended DNA shows that the DNA binds with the blunt end stacked against the fingers and the double-stranded portion leaving the active site between the thumb and the vestigial exonuclease domain.

5. • Pelletier H, Sawaya MR, Wolfle W, Wilson S, Kraut J: **Crystal structures of human DNA polymerase β complexed with DNA: implications for catalytic mechanism, processivity, and fidelity.** *Biochemistry* 1996, **35**:12742-12761.

The authors describe several binary complexes of pol β and DNA. These complexes reveal facts about the processivity, fidelity and the conformational changes upon nucleotide binding. In addition, a model for the binding of pol β to gapped DNA is proposed and discussed in these contexts.

6. •• Minnick DT, Astatke M, Joyce CM, Kunkel TA: **A thumb subdomain mutant of the large fragment of *Escherichia coli* DNA polymerase I with reduced DNA binding affinity, processivity, and frameshift fidelity.** *J Biol Chem* 1996, **271**:24954-24961.

The results in this paper show that deleting 24 residues from the tip of the thumb of KF reduce the processivity of this enzyme. Further, such a polymerase mutant makes far more frameshift mutations than wild-type KF.

7. • Pelletier H, Sawaya MR, Wolfle W, Wilson SH, Kraut J: **A structural basis for metal ion mutagenicity and nucleotide selectivity in human DNA polymerase β.** *Biochemistry* 1996, **35**:12762-12777.

A variety of structures are presented that illustrate the importance of the identity of the metal ions and the base, ribose and triphosphate moieties on the binding and utilization of nucleotides by human pol β.

8. Kornberg A, Baker TA: *DNA Replication*, edn 2. New York: WH Freeman; 1992.

9. Ito J, Braithwaite DK: **Compilation and alignment of DNA polymerase sequences.** *Nucleic Acids Res* 1991, **19**:4045-4057.

10. Wang J, Yu P, Lin TC, Konigsberg WH, Steitz TA: **Crystal structures of an NH_2- terminal fragment of T4 DNA polymerase and its complexes with single-stranded DNA and with divalent metal ions.** *Biochemistry* 1996, **35**:8110-8119.

11. Beese LS, Derbyshire V, Steitz TA: **Structure of DNA polymerase I Klenow fragment bound to duplex DNA.** *Science* 1993, **260**:352-355.

12. Huang WM, Lehman IR: **On the exonuclease activity of phage T4 deoxyribonucleic acid polymerase.** *J Biol Chem* 1972, **247**:3139-3146.

13. Capson TL, Peliska JA, Kaboord BF, Frey MW, Lively C, Dahlberg M, Benkovic SJ: **Kinetic characterization of the polymerase and exonuclease activities of the gene 43 protein of bacteriophage T4.** *Biochemistry* 1992, **31**:10984-10994.

14. Lin TC, Karam G, Konigsberg WH: **Isolation, characterization, and kinetic properties of truncated forms of T4 DNA polymerase that exhibit 3′–5′ exonuclease activity.** *J Biol Chem* 1994, **269**:19286-19294.

15. Sattar AKMA, Lin TC, Jones C, Konigsberg WH: **Functional consequences and exonuclease kinetic parameters of point mutations in bacteriophage T4 DNA polymerase.** *Biochemistry* 1996, **35**:16621-16629.

16. Kuchta RD, Benkovic P, Benkovic SJ: **Kinetic mechanism whereby DNA polymerase I (Klenow) replicates DNA with high fidelity.** *Biochemistry* 1988, **27**:6716-6725.

17. Kong X-P, Onrust R, ODonnell M, Kuriyan J: **Three-dimensional structure of the β- subunit of *E. coli* DNA polymerase III holoenzyme: a sliding clamp.** *Cell* 1992, **69**:425-437.

18. Berdis LS, Soumillion P, Benkovic SJ: **The carboxyl terminus of the bacteriophage T4 DNA polymerase is required for holoenzyme complex formation.** *Proc Natl Acad Sci USA* 1996, **93**:12822-12827.

19. Delarue M, Poch O, Tordo N, Moras D, Argos P: **An attempt to unify the structure of polymerases.** *Protein Eng* 1990, **3**:461-467.

20. Kohlstaedt LA, Wang J, Friedman JM, Rice PA, Steitz TA: **Crystal structure at 3.5 Å of HIV-1 reverse transcriptase complexed with an inhibitor.** *Science* 1992, **256**:1781-1790.

21. 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.

22. Kim Y, Eom SH, Wang J, Lee DS, Suh SW, Steitz TA: **Crystal structure of *Thermus aquaticus* DNA polymerase.** *Nature* 1995, **376**:612-616.

23. Korolev S, Nayal M, Barnes WM, DiCera E, Waksman G: **Crystal structure of the large fragment of *Thermus aquaticus* DNA polymerase I at 2.5 Å resolution: structural basis for thermostability.** *Proc Natl Acad Sci USA* 1995, **92**:9264-9268.

24. Davies JF Jr, Almassy RJ, Hostomska Z, Ferre RA, Hostomsky Z: **2.3 Å crystal structure of the catalytic domain of DNA polymerase β.** *Cell* 1994, **76**:1123-1133.

25. Pelletier H, Sawaya MR, Kumar A, Wilson SH, Kraut J: **Structures of ternary complexes of rat DNA polymerase β, a DNA template-primer, and ddCTP.** *Science* 1994, **264**:1891-1903.

26. Holm L, Sander C: **DNA polymerase β belongs to an ancient nucleotidyl transferase superfamily.** *Trends Biochem Sci* 1995, **20**:345-347.

27. Jacobo-Molina A, Ding J, Nanni RG, Clark AD, Jr, Lu X, Tantillo C, Williams RL, Kamer G, Ferris AL, Clark P *et al.*: **Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 Å resolution shows bent DNA.** *Proc Natl Acad Sci USA* 1993, **90**:6320-6324.

28. Ren J, Esnouf R, Hopkins A, Ross C, Jones Y, Stammers D, Stuart D: **The structure of HIV-1 reverse transcriptase complexed with 9-chloro-TIBO: lessons for inhibitor design.** *Structure* 1995, **3**:915-926.

29. Rodgers DW, Gamblin SJ, Harris BA, Ray S, Culp JS, Hellmig B, Woolf DJ, Debouck C, Harrison, SC. **The structure of unliganded reverse transcriptase from the human immunodeficiency virus type 1.** *Proc Natl Acad Sci USA* 1995, **92**:1222-1226.

30. Georgiadis MM, Jessen SM, Ogata CM, Telesnitsky A, Goff SP, Hendrickson WA: **Mechanistic implications from the structure of a catalytic fragment of Moloney murine leukemia virus reverse transcriptase.** *Structure* 1995, **3**:879-892.

31. Steitz, TA, Smerdon SJ, Jäger J, Joyce CM: **A unified mechanism for nonhomologous DNA and RNA polymerases.** *Science* 1994, **266**:2022-2025.

32. Pelletier H: **Polymerase structures and mechanism.** *Science* 1994, **266**:2025-2026.

33. Hughes SH, Hostomsky Z, LeGrice SFJ, Lentz K, Arnold E: **What is the orientation of DNA polymerases on their templates?** *J Virol* 1996, **70**:2679-2683.

34. Modrich P, Richardson CC: **Bacteriophage T7 deoxyribonucleic acid replication *in vitro*. A protein of *Escherichia coli* required for bacteriophage T7 DNA polymerase activity.** *J Biol Chem* 1975, **250**:5508-5514.

35. Kati WM, Johnson KA, Jerva LF, Anderson KS: **Mechanism and fidelity of HIV reverse transcriptase.** *J Biol Chem* 1992, **267**:25988-25997.

36. Werneberg BG, Ahn J, Zhong X, Hondal RJ, Kraynov VS, Tsai M-D: **DNA polymerase β: pre-steady-state kinetic analysis and roles of arginine-283 in catalysis and fidelity.** *Biochemistry* 1996, **35**:7041-7050.

37. •• Bedford E, Tabor S, Richardson CC: **The thioredoxin binding domain of bacteriophage T7 DNA polymerase confers processivity on *Escherichia coli* DNA polymerase I.** *Proc Natl Acad Sci USA* 1997, **94**:479-484.

By simply grafting the thioredoxin-binding domain of T7 DNA polymerase onto the tip of the thumb of a KF, the authors show that the KF can be rendered into a thioredoxin-sensitive processive DNA polymerase. Though the overall activity of the mutant KF is lower than the wild-type, the gains in processivity are impressive.

38. Singhal RK, Wilson SH: **Short gap-filling synthesis by DNA polymerase β is processive.** *J Biol Chem* 1993, **268**:15906-15911.

39. Polesky AH, Steitz TA, Grindley NDF, Joyce CM: **Identification of residues critical for the polymerase activity of the Klenow fragment of DNA polymerase I from *Escherichia coli*.** *J Biol Chem* 1990, **265**:14579-14591.

40. Polesky AH, Dahlberg ME, Benkovic SJ, Grindley NDF, Joyce CM: **Side chains involved in catalysis of the polymerase reaction of DNA polymerase I from *Escherichia coli*.** *J Biol Chem* 1992, **267**:8417-8428.

41. El-Deiry WS, Downey KM, So AG: **Molecular mechanisms of manganese mutagenesis.** *Proc Natl Acad Sci USA* 1984, **81**:7378-7382.

42. Orgel A, Orgel LE: **Induction of mutations in bacteriophage T4 with divalent manganese.** *J Mol Biol* 1965, **14**:453-457.

43. Beckman RA, Mildvan AS, Loeb LA: **On the fidelity of DNA replication: manganese mutagenesis *in vitro*.** *Biochemistry* 1985, **24**:5810-5817.

44. Steitz TA: **DNA- and RNA-dependent DNA polymerases.** *Curr Opin Struct Biol* 1993, **3**:31-38.

45. Beese LS, Steitz TA: **Structural basis for the 3′–5′ exonuclease activity of *Escherichia coli* DNA polymerase I: a two metal ion mechanism.** *EMBO J* 1991, **10**:25-33.

46. Bryant FR, Johnson KA, Benkovic SJ: **Elementary steps in the DNA polymerase I reaction pathway.** *Biochemistry* 1983, **22**:3537-3546.

47. Patel SS, Wong I, Johnson KA: **Pre-steady-state kinetic analysis of processive DNA replication including complete characterization of an exonuclease-deficient mutant.** *Biochemistry* 1991, **30**:511-525.

48. Spence RA, Kati WM, Anderson KS, Johnson KA: **Mechanism of inhibition of HIV-1 reverse transcriptase by non-nucleoside inhibitors.** *Science* 1995, **267**:988-993.

49. Jäger J, Smerdon SJ, Wang J, Boisvert DC, Steitz TA: **Comparison of three different crystal forms shows HIV-1 reverse transcriptase displays an internal swivel motion.** *Structure* 1994, **2**:869-876.