Crystal Structure of a Y-Family DNA Polymerase in Action: A Mechanism for Error-Prone and Lesion-Bypass Replication

Hong Ling,[1] François Boudsocq,[2]
Roger Woodgate,[2] and Wei Yang[1,3]
[1] Laboratory of Molecular Biology
National Institute of Diabetes and Digestive
   and Kidney Diseases
[2] Section on DNA Replication, Repair,
   and Mutagenesis
National Institute of Child Health
   and Human Development
National Institutes of Health
Bethesda, Maryland 20892

Summary

Sulfolobus solfataricus P2 DNA polymerase IV (Dpo4) is a DinB homolog that belongs to the recently described Y-family of DNA polymerases, which are best characterized by their low-fidelity synthesis on undamaged DNA templates and propensity to traverse normally replication-blocking lesions. Crystal structures of Dpo4 in ternary complexes with DNA and an incoming nucleotide, either correct or incorrect, have been solved at 1.7 Å and 2.1 Å resolution, respectively. Despite a conserved active site and a hand-like configuration similar to all known polymerases, Dpo4 makes limited and nonspecific contacts with the replicating base pair, thus relaxing base selection. Dpo4 is also captured in the crystal translocating two template bases to the active site at once, suggesting a possible mechanism for bypassing thymine dimers.

Introduction

DNA polymerases replicate millions of nucleotides efficiently and accurately during each cycle of cell division. In *Escherichia coli*, DNA polymerase III, which is responsible for the bulk of DNA replication, possesses high fidelity in base selection and only makes one error in $\sim10^5$ base pairs. With the intrinsic exonucleolytic proofreading and postreplicative mismatch correction, the overall error rate of DNA replication is as low as $\sim10^{-10}$ (Schaaper, 1993). High-fidelity DNA polymerases are, however, often sensitive to minor distortions in the template DNA and incoming nucleoside triphosphate (Kunkel and Bebenek, 2000). Coincidentally, many organisms have equipped themselves with a multitude of repair enzymes to remove damaged DNA and maintain the integrity of the nucleotide pools prior to replication (Lindahl and Wood, 1999; Wood et al., 2001). Despite the plethora of repair enzymes, DNA lesions, such as cyclobutane pyrimidine dimers resulting from UV damage, often persist in DNA and impede the progress of a growing replication fork. Recent studies suggest that specialized DNA polymerases are recruited to traverse DNA lesions, where they temporarily substitute for the high-

fidelity polymerase (Woodgate, 1999; Friedberg and Gerlach, 1999; Goodman and Tippin, 2000).

Many of the polymerases that replicate damaged DNA are phylogenetically related to each other. Originally called the UmuC/DinB/Rev1/Rad30 superfamily, after the prototype of each branch of the family, proteins belonging to this superfamily are now identified as the Y-family of DNA polymerases (Ohmori et al., 2001). The Y-family polymerases share five highly conserved sequence motifs, which were identified long before the enzymatic activities were characterized (Kulaeva et al., 1996), yet they do not show any significant sequence identity to any previously characterized polymerases from the A-, B-, C-, D-, and X-families. The Y-family polymerases replicate DNA in a distributive manner and lack any intrinsic exonucleolytic activity for proofreading (Goodman and Tippin, 2000). In addition to bypassing damaged template or bulky DNA adducts, these polymerases replicate undamaged DNA with a 10- to 100-fold increased error rate compared to the A-, B-, C- and X-family polymerases (Johnson et al., 2000a, 2000b; Ohashi et al., 2000a; Tang et al., 2000). The most extreme example is human polι, which preferentially inserts G opposite T, rather than the canonical Watson-Crick base A, by a factor of 3- to 11-fold, depending upon the template sequence context (Johnson et al., 2000c; Tissier et al., 2000; Vaisman et al., 2001; Zhang et al., 2000a).

Structures of several high-fidelity DNA polymerases from the A-family (Klenow fragment of *E. coli* and *bacillus* polymerase I, T7, and *T. aquaticus* DNA polymerases), B-family (RB69 gp43), X-family (polβ), and HIV *reverse transcriptase* (RT) have been reported (Doublie et al., 1998; Eom et al., 1996; Franklin et al., 2001; Huang et al., 1998; Kiefer et al., 1998; Li et al., 1998; Ollis et al., 1985; Pelletier et al., 1996). Despite being phylogenetically unrelated, these polymerases have a similar overall shape, which has been likened to a right hand composed of palm, finger, and thumb domains (Steitz, 1999). The palm domain is the most conserved structurally and contains three invariant carboxylates forming a catalytic center near the interface with the finger domain. Although structurally varied, the finger domain tightly envelops the Watson-Crick base pair formed between a template base and an incoming nucleotide, and plays a critical role in the induced-fit mechanism to discriminate against mismatches (Doublie et al., 1999; Li et al., 1998; Johnson, 1993; Sawaya et al., 1997). When the fit is perfect, the polymerase efficiently catalyzes the phosphatidyl transfer reaction.

In the first step toward understanding the Y-family polymerases, Zhou and colleagues have recently solved the crystal structure of a truncated but catalytically active fragment of the Dbh (*DinB homolog*) from the crenarchaeon *Sulfolobus solfataricus* P1 (Zhou et al., 2001). The P1 Dbh protein was first identified in a search for UmuC-DinB orthologs in archaea (Kulaeva et al., 1996), and was recently shown to bypass an abasic site in a template strand at high enzyme and nucleotide concentrations (Zhou et al., 2001). Despite the lack of detectable sequence conservation, the Dbh catalytic fragment is

[3] Correspondence: wei.yang@nih.gov

structurally similar to other DNA polymerases with a right-handed configuration of finger, thumb, and palm domains and an invariant tricarboxylate active site (Zhou et al., 2001).

Another DinB ortholog has recently been identified in the *S. solfataricus* P2 genome (She et al., 2001). *S. solfataricus* P2 has three B-family polymerases (Dpo1, Dpo2, and Dpo3) and consequently, the DinB-like Y-family polymerase has been called DNA polymerase IV (Dpo4) (Boudsocq et al., 2001). We have overproduced, purified, and characterized the enzymatic properties of *S. solfataricus* P2 Dpo4, and find that it exhibits misincorporation rates in the range of 8×10^{-3} to 4×10^{-4} (Boudsocq et al., 2001), which is consistent with its phylogenetic placement in the DinB branch of the Y-family tree. Interestingly, unlike *E. coli* DinB and the related human polκ, which are unable to bypass a *cis-syn* cyclobutane thymine dimer (Johnson et al., 2000b; Ohashi et al., 2000b; Tang et al., 2000; Zhang et al., 2000b), Dpo4 can bypass this lesion relatively efficiently (Boudsocq et al., 2001). With respect to lesion bypass, the enzymatic properties of Dpo4 thus bear similarity to that of the distantly related eukaryotic polη (Ishikawa et al., 2001; Johnson et al., 1999; Masutani et al., 1999a, 1999b).

Toward the goal of understanding the molecular basis for the Y-family's low-fidelity DNA synthesis on undamaged DNA templates and their ability to traverse normally replication-blocking lesions in DNA, we report here the biochemical characterization of Dpo4 and the crystal structures of two ternary complexes of Dpo4 and primer-template DNA with either a correct or incorrect incoming nucleotide at 1.7 Å and 2.1 Å resolution, respectively.

Results

Structure Determination

Crystals of the Dpo4 ternary complexes were grown from a mixture of purified full-length protein, preannealed primer-template DNA, and an appropriate dideoxynucleoside triphosphate (Experimental Procedures). In order to trap an active ternary complex, the 3′ end of the primer was extended with the dideoxynucleotide by incubating the enzyme-substrate mixture with Mg^{2+} for 10 min at room temperature prior to crystallization. Depending upon the sequence of the template and the type of dideoxynucleoside triphosphate supplied, we expected to generate one complex containing only Watson-Crick base pairs and another containing a potential mismatch.

The perfectly base-paired ternary complex with $^{3'}TT^{5'}$ as the template and ddATP as the incoming nucleotide produced crystals diffracting to better than 1.7 Å Bragg spacings. These crystals, which are hereafter referred as type I, contain a single ternary complex in each asymmetric unit. The structure was solved by the multiwavelength anomalous diffraction (MAD) method using selenomethionine substituted Dpo4 (see Experimental Procedures). The final model, which consists of residues 1 to 341 of Dpo4, 17 nt template, 13 nt primer, ddADP, and a divalent cation, is refined to 1.7 Å with an R value

of 21.3%, R free of 22.8%, and excellent geometry (Table 1 and Experimental Procedures).

The ternary complex generated from a template with a sequence of $^{3'}GC^{5'}$ and ddGTP as the incoming nucleotide, which potentially makes a G-G mismatch, produced crystals (type II) isomorphous to those of type I, but diffracted X-rays only to $\sim$2.0 Å Bragg spacings. The structure solution was readily obtained with the refined structure of type I complex (Experimental Procedures). Interestingly, the expected G:G mismatch did not occur, but instead, the incoming nucleotide base paired with the next template base, C.

Overall Structure of the Dpo4-DNA Ternary Complexes

We use the higher resolution type I structure to describe the general features of the ternary complex since the type II complex is essentially identical except for the active site region. The Dpo4-DNA ternary complex is analogous to a right hand grasping a rope made of double-stranded DNA, and as such is similar in structure to all known polymerase-DNA complexes (Figure 1). The protein contains four structural domains. The palm, finger, and thumb domains are defined by the conserved function universal to all polymerases. Dpo4 also contains an additional fourth domain at the C terminus, which is not included in the Dbh catalytic core structure (Zhou et al., 2001). This domain is tethered to the thumb domain in sequence but physically located next to the finger domain and interacts with DNA in the major groove (Figures 1 and 2). It is hereafter referred to as the "little finger" domain based on its spatial location relative to the thumb, palm, and finger domains in the anatomy of a right hand. The single-stranded 5′ end of the template enters Dpo4 between the finger and little finger domains. The 13 base pair template and primer duplex is in a straight B-form conformation, except for the exposed end distal from the polymerase, which is slightly distorted due to the packing of the DNA end against the back of the palm domain of a neighboring Dpo4 molecule. Such lattice contact is essential for growing crystals of the ternary complexes, yet limits DNA duplex to 14–15 base pairs. The incoming nucleotide makes a perfect 14th Watson-Crick base pair with the template base, T. A Ca^{2+} ion, chelated by two active-site carboxylates, stabilizes the phosphate moiety of the incoming nucleotide. Surprisingly, the ddATP was hydrolyzed to ddADP in the crystal. We confirmed that Dpo4 does contain a weak phosphatase activity in solution. A mutation at the active site (D105A/E106A), which abolishes the polymerase activity (data shown later), eliminates the phosphatase activity (F. Boudsocq, unpublished data).

Structural Features of Dpo4

The finger (residues 11–77), palm (1–10 and 78–166), thumb (167–233), and little finger (244–341) domains of Dpo4 occur sequentially from the N to C terminus except for the very first β strand, which belongs to the palm instead of the finger. A similar domain arrangement (finger-palm-thumb) is found in HIV RT and RB69 polymerase (Franklin et al., 2001; Huang et al., 1998). In the A-family of DNA polymerases, the thumb domain occurs

Table 1. Data Collection and Refinement Statistics

MAD data statistics (15.0–2.4 Å)

λ (Å)[1]	0.9832	0.9794	0.9790	0.9686
	0.017	0.025	0.024	0.027
		0.030	0.017	0.029
			0.041	0.019
				0.033
R merge[2]	0.026	0.032	0.037	0.032
Completeness	96.3%	95.2%	95.7%	95.8%

Figure-of-merit: 0.63 & Score: 51.02 (determined by SOLVE)

Refinement

Crystal (space group)	Type I ($P2_12_12$)	Type II ($P2_12_12$)
Unit cell (a, b, c) (Å)	97.1, 102.8, 52.3	97.1, 101.9, 52.5
Nonhydrogen atoms	3948	3585
Resolution range[3]	26.6–1.70 Å (1.73–1.70 Å)	29.2–2.10 Å (2.14–2.10 Å)
R merge[2,3]	0.040 (0.554)	0.033 (0.331)
Unique reflection	57,010	30,292
Completeness[3]	97.6 (98.4)	97.3 (98.1)
R value[4]	0.213	0.221
R_{free}[5]	0.228 (1122 reflections)	0.278 (1452 reflections)
R.m.s.d. bond length (Å)	0.017	0.016
R.m.s.d. bond angle (°)	1.943	1.74
Ave. B-value (Wilson) (Å²)	34.9 (27.2)	41.7 (37.9)

[1] Values represent the Bijvoet difference ratios at each wavelength (diagonal elements) and dispersive difference ratios between two wavelengths (off-diagonal elements).
[2] $R_{merge} = \Sigma_h\Sigma_i|I_{hi} - \langle I_h\rangle|/\Sigma\langle I_h\rangle$, where I_{hi} is the intensity of the ith observation of reflection h, and $\langle I_h\rangle$ is the average intensity of redundant measurements of the h reflections.
[3] Data of the highest resolution shell is shown in parentheses.
[4] R value $= \Sigma||F_o| - |F_c||/\Sigma|F_o|$, where F_o and F_c are the observed and calculated structure-factor amplitudes.
[5] R_{free} is monitored with the number of reflections in parentheses excluded from refinement.

first and is followed by the finger and palm domains (Steitz, 1999). Eukaryotic polβ is unique in that its domains occur in the order of thumb, palm, and finger (Pelletier et al., 1996). Regardless of the topological sequence and difference in secondary structures, the three domains form a functionally equivalent U-shaped struc-
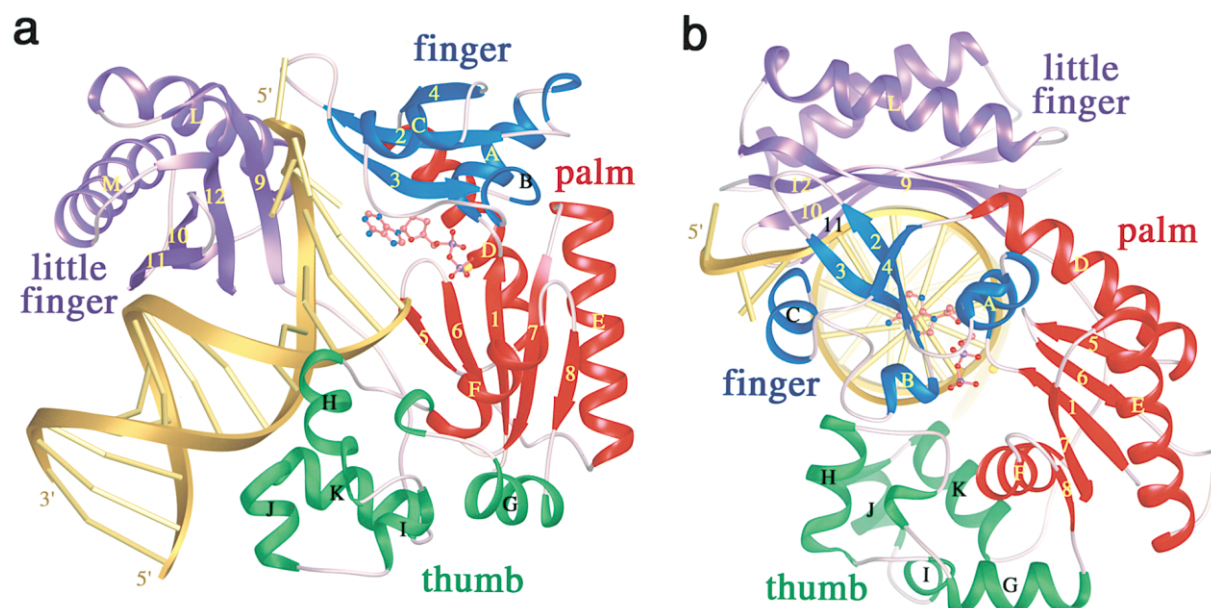

Figure 1. Crystal Structure of the Dpo4 Ternary Complex

Both the protein and DNA backbones are shown in ribbon diagrams, DNA bases as rods, and the incoming nucleotide and the Ca²⁺ ion in a ball-and-stick model. The Dpo4 structural domains are shown in red (palm), green (thumb), blue (finger), and purple (little finger); the DNA is in gold. Secondary structures are named in alphabetic or numeric order for α helices and β strands, respectively, following their order in the primary sequence.
(a) A view looking into the active site with the palm, finger, thumb, and little finger domains well separated. (b) A view down the DNA-helical axis.
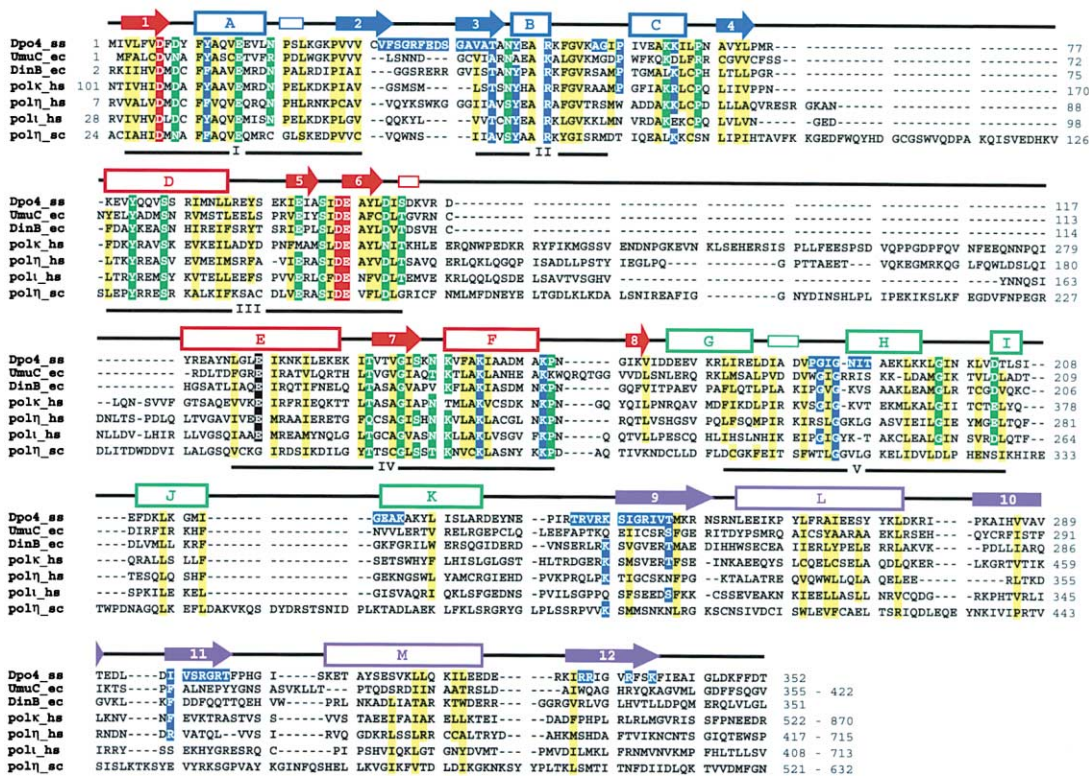
Figure 2. Structure-Based Sequence Alignment

S. solfataricus Dpo4, *E. coli* UmuC (PolV), *E. coli* DinB (PolIV), human Polη (Rad30A), ι (Rad30B), and κ (DinB1), and *S. cerevisiae* polη (Rad30) are included. The secondary structures observed in Dpo4 are indicated above the aligned sequence and colored in the same scheme as shown in Figure 1. α helices are shown as rectangular boxes, β strands as arrows, and 3₁₀ helices as narrow rectangles. Conserved residues important for the hydrophobic core formation are highlighted in yellow, for structural integrity in green, and for polymerase activity in red. Residues interacting with DNA are highlighted in blue, whether conserved or not. E127 highlighted in black appears to be conserved, but it is completely exposed to solvent. The previously identified five sequence motifs are indicated beneath the alignment.

ture with a deep cleft to bind DNA in all A-, B-, X-, and Y-polymerases (Figure 3a).

The palm domain of Dpo4, which contains the active site, is most recognizable due to a structural similarity to other polymerases. The central four β strands and two surrounding α helices are superimposable with the palm domain of T7 (Figure 3b). D7, D105, and E106 of Dpo4, which are located on the adjacent β strands 1 and 6, are in identical positions to the catalytic carboxylates found in T7, Klenow, *Bacillus*, Taq, and RB69 DNA polymerases as well as HIV RT (Doublie et al., 1998; Franklin et al., 2001; Huang et al., 1998; Kiefer et al., 1998; Li et al., 1998; Ollis et al.,1985). These three carboxylates are conserved throughout the Y-family of polymerases (Figure 2). Mutations of any one of the three conserved carboxylates in *E. coli* UmuC or DinB, *S. cerevisiae* polη, or human polι abolish or greatly reduce the polymerase activity (Kondratick et al., 2001; Tang et al., 1999; Tissier et al., 2000; Wagner et al., 1999). Similarly, replacing D105 and E106 of Dpo4 with Ala renders the polymerase inactive (Figure 4a), thus confirming the catalytic function of these three carboxylates.

The thumb and finger domains of Dpo4 are unusually small (Figure 3). The finger domain shares limited structural similarity with the DNA binding domains of the mismatch recognition protein MutS (Obmolova et al.,

2000; Lamers et al., 2000). In the fingers of T7, Taq, PB69, and β polymerase, a long α helix best known as the "O helix" in the pol A family can be induced to fit on the flat face of a perfectly matched Watson-Crick base pair between the incoming nucleotide and the template base, thus performing the fidelity check (Doublie et al., 1999; Li et al., 1998; Franklin et al., 2001). Dpo4 does not contain an equivalent helix. Instead, β strand 3 and an adjacent extended loop contact the replicating base pair, which is reminiscent of the β3 and β4 strands in the case of HIV RT (Huang et al., 1998). The thumb domain of Dpo4 is entirely helical and is structurally similar to the N-terminal lyase domain of polβ and to the finger domain of T7 polymerase, despite being much smaller (Figure 3c). The resemblance raises the possibility that these polymerases may result from domain swapping and divergent evolution.

The topology of the little finger resembles a six-stranded "jelly roll," except that the second strands from the N and the C termini are replaced by α helices. The resulting structure consists of a layer of a four-stranded antiparallel β sheet and another layer of 2 antiparallel α helices, and bears resemblance to the core of the palm domain (Figures 3b and 4c). A search of the database using DALI (Holm and Sander, 1998) found similar structures in several enzymes, including glutamine phospho-
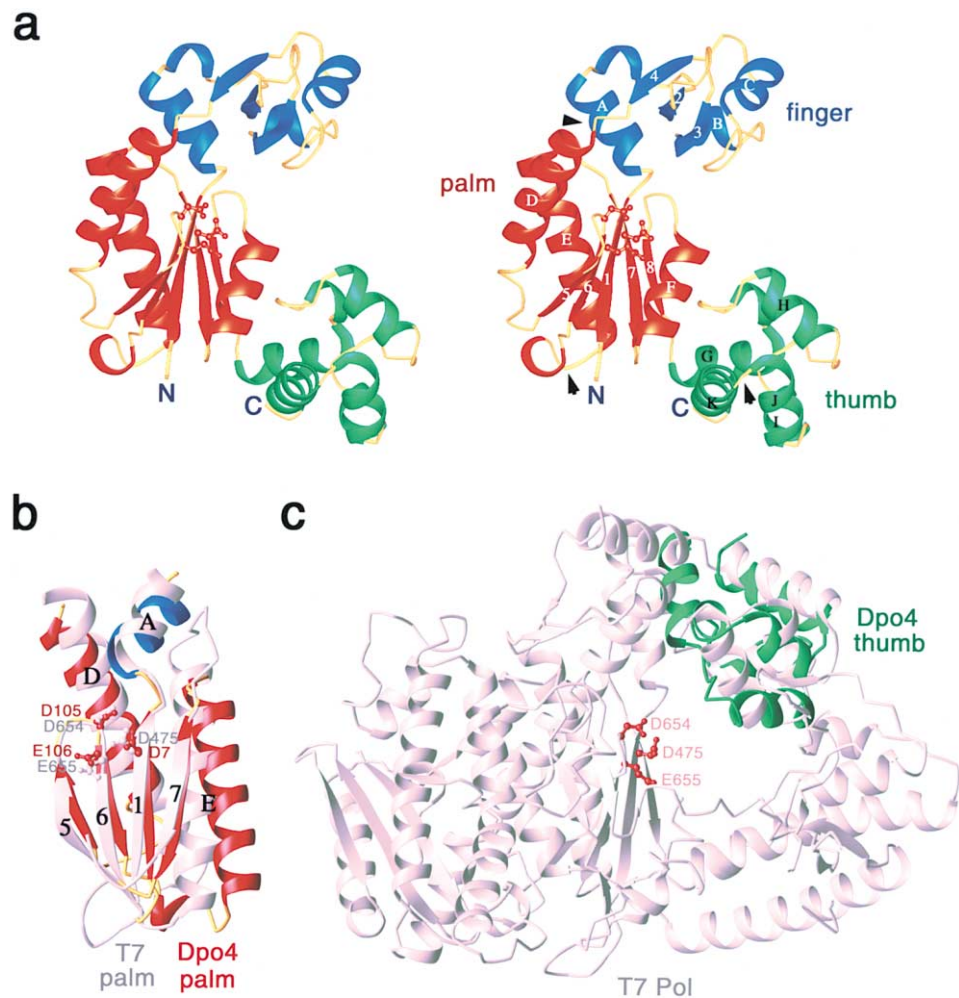
Figure 3. Structural Analysis of the Catalytic Core of Dpo4

(a) A stereo view of the palm, finger, and thumb domains shown in ribbon diagram. The little finger is removed for clarity. The conserved carboxylates D7, D105, and E106, are included. Black arrowheads indicate the locations of amino acid insertions in other Y-family polymerases.
(b) Superposition of the T7 (PDB accession: 1T7P) and Dpo4 palm domain, including the central four β strands and two α helices. The three catalytic carboxylates and an α helix of the finger domain are also shown. Dpo4 is colored as in Figure 3a and T7 Pol is in lilac and gray.
(c) The entire T7 Pol is shown as a ribbon diagram in lilac. The Dpo4 thumb is superimposed on the structurally homologous T7 finger.

ribosyl pyrophosphate amidotransferase (Muchmore et al., 1998), but it is always found as part of a larger domain and has not been observed as an independent structural entity.

DNA and DNA-Protein Interactions

The 13 base pair primer-template duplex, including the dideoxynucleotide at the 3′ end of the primer plus the 14th with the incoming ddADP, is in nearly standard B-form, instead of the A-form as previously observed near the active site of the A-family polymerase-DNA complexes (Doublie et al., 1998; Eom et al., 1996; Kiefer et al., 1998; Li et al., 1998). The average rise and twist per base pair in the Dpo4 ternary complex are 3.2 Å and 33.4°, respectively. While five of the base pairs are exposed to solvent, Dpo4 holds the remaining eight bp by both hydrogen bonds and van der Waals contacts. These interactions are mostly with the phosphodiester

moieties, a few with the deoxyriboses, and none with the bases.

The thumb and little finger grip the dsDNA across the minor groove from underneath and the major groove from above (Figure 5a). The thumb contacts the backbones of the primer and template strands on each side of the minor groove with the N termini of α helices H and K (Figures 2 and 5a). The little finger fits into the DNA major groove snugly, with the curvature of the β sheet cupping the major groove and α helices facing toward solvent. The β strands 9 and 11 at each edge of the sheet contact the template and primer strand, respectively (Figures 2 and 5). In addition, a number of Arg residues from β strands 12 and K152 of the palm domain are hydrogen bonded to the DNA backbones (Figure 5). The tether that links the thumb and little finger domain, laden with positively charged side chains, wraps around the DNA and interacts mainly with the phosphosugar backbone of the template strand. Includ-
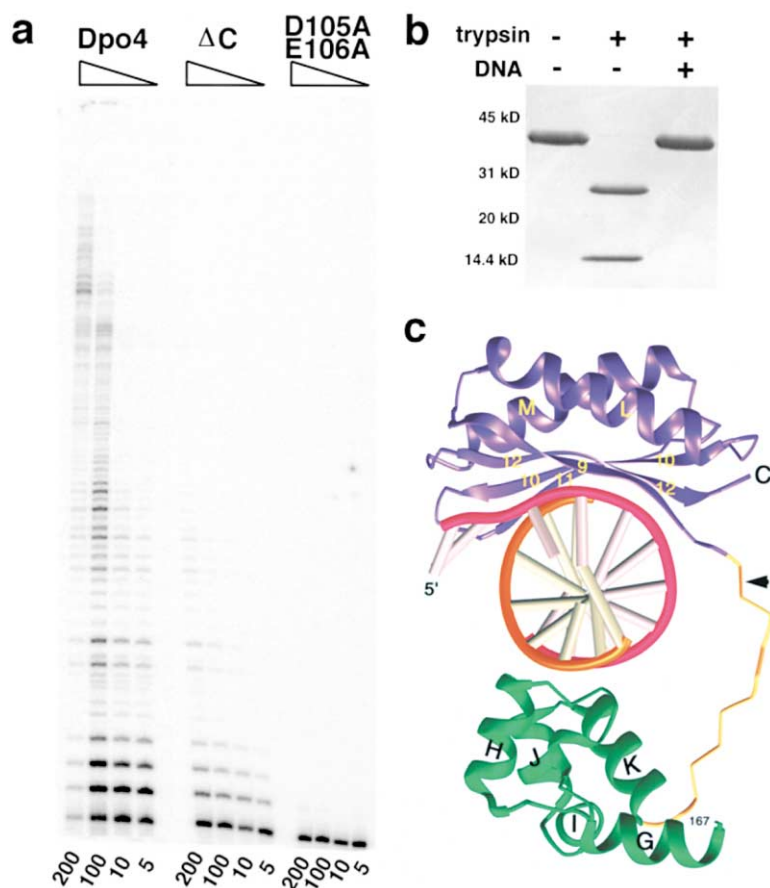
Figure 4. Activities of the Full-Length, Catalytic Core, and Mutant Dpo4 Proteins

(a) The ability to extend a primer annealed to a circular M13 DNA by the full-length Dpo4, the catalytic core (residues 1–240), and an active site mutant (D105A-E106A) was assayed over a range of enzyme concentrations. Reactions contained all four dNTPs (100 μM each), 10 nM primer/template, and 200 nM to 5 nM enzyme as indicated, and were performed for 5 min at 37°C.

(b) SDS gel showing trypsin digestion of Dpo4 in the absence and presence of DNA. The molecular weight is marked on the left of the panel.

(c) A ribbon diagram of the extended region that connects the thumb and little finger domains. The view is similar to Figure 1b, with the palm and finger removed for clarity. The black arrowhead indicates where trypsin cleaves Dpo4 in the absence of DNA.

ing the 5′ unpaired template strand, a total of 920 Å² of the DNA surface becomes buried upon forming the complex with Dpo4.

The tether that links the little finger and thumb domain is easily cleaved between R240 and V241 by trypsin in the absence of DNA (Figures 4b and 4c), and the little finger is released from the catalytic core. The interactions between the tip of the little finger and the finger observed in the ternary complex (Figure 1) are most probably absent without a bound DNA. DNA completely protects Dpo4 from proteolytic cleavage (Figure 4b). The catalytic core without the little finger is much less active and synthesizes shorter DNA stretches than does the intact Dpo4 (Figure 4a). Even though DNA synthesis by Dpo4 is essentially distributive, as evidenced by the pause sites observed at each template base (Figure 4a), the little finger apparently improves the attachment of the polymerase to DNA and increases the number of nucleotides synthesized per binding event.

The Polymerase Active Site

The Dpo4 active site is defined by the three absolutely conserved carboxylates, D7, D105, and E106, and is also marked by a metal ion. The metal ion is chelated by the carboxylates of D7 and D105, the carbonyl oxygen of Tyr8, the nonbridging oxygens of both the α and β phosphates of ddADP, and a water molecule (Figure 6a). This metal ion has been refined as Ca^{2+} because of 100 mM CaAc$_2$ and only 5 mM MgCl$_2$ in the crystallization

buffer and the high electron density. The second metal ion essential for phosphatidyl transfer (Steitz, 1999) is not observed, probably because the α-phosphate of ddADP, which is located where the γ-phosphate would normally be if it were dATP, is out of the range for the chemical reaction.

The template base and the incoming nucleotide are surrounded by the palm and finger domains (Figure 6). V32, A42, and the Cα of G58 interact with the template base from both the 5′ end and the minor groove side (Figure 6a). Residues 44–58 form a lid (including β3 and αB) on the incoming nucleotide and contact both the base and the phosphate moiety of ddADP (Figure 6a). A44 and A57 form van der Waals contacts with the adenine base; T45 and R51 are directly hydrogen bonded to the phosphates. The phosphosugar moiety of ddADP is further stabilized by the palm domain, for instance, D7, D105, K159, main chain nitrogen atom of Y10, and the Ca^{2+} ion. The sugar ring of ddADP packs against the benzene ring of Y12, which is in such close contact that the 2′ hydroxyl of a ribose would not fit. The replicating base pair between the template T and the incoming ddADP is, on the whole, less well protected in comparison with those found in complexes of high-fidelity polymerases, such as T7, Taq, and RB69 (Doublie et al., 1998; Franklin et al., 2001; Li et al., 1998) (more discussion later).

In the type II crystal, the primer is not extended when the template base is G and only ddGTP is supplied. It
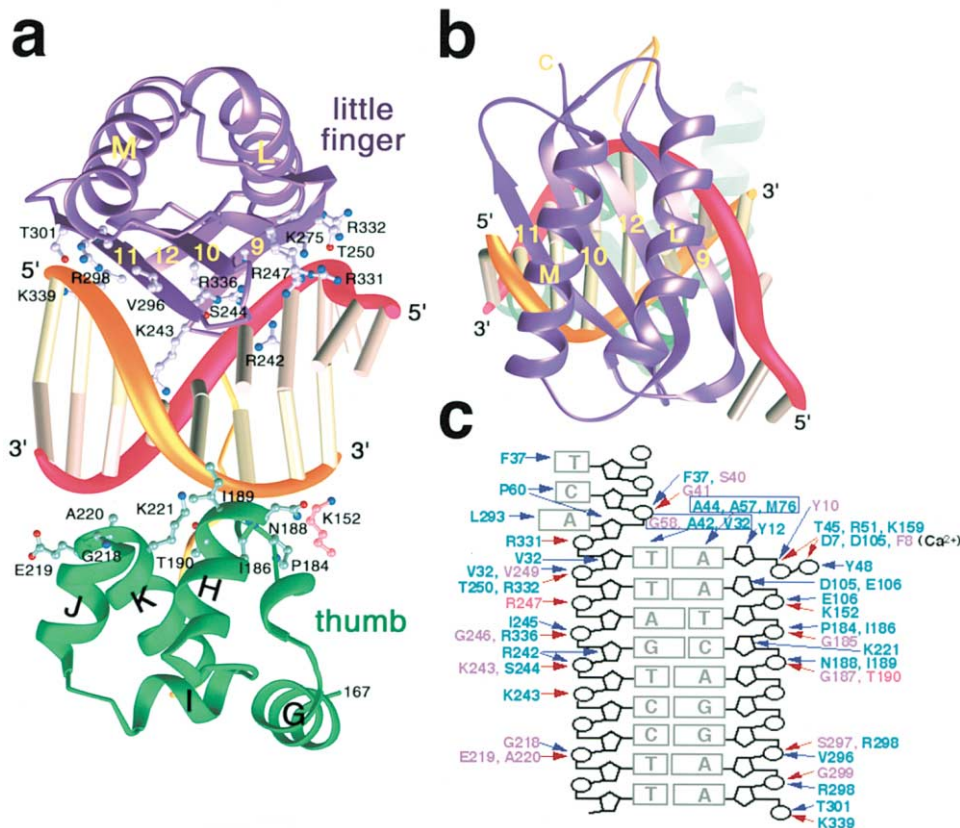
Figure 5. Interactions between Dpo4 and the Template-Primer Duplex

(a) The thumb and little finger domains represented by ribbon diagrams are shown with the contacting 8 bp. The thumb is green and the little finger is purple. The template strand is shown in red ribbon (backbone) and pink rods (bases), and the primer strand in orange (backbone) and yellow (bases). Residues hydrogen bonded to or in van der Waals contact with DNA are shown in ball-and-stick and labeled.

(b) A top view of 5a shows the perfect fit of the little finger in the major groove.

(c) A diagram of the DNA-protein interactions. Hydrogen bonds and van der Waals contacts are shown as red and blue arrows, respectively. Protein residues are labeled in purple or green for main or side chain interactions, or pink for both.

is clear that a G:G mismatch is not favored, which is no surprise since Dpo4 makes a G:G mismatch with a frequency of 3.5×10^{-4} relative to a G:C pair (Boudsocq et al., 2001). Interestingly, the template is translocated without replication of G (tG4); the next template base, tC3, is instead positioned in the active site and base paired with the incoming ddGTP (Figure 6b). The base at the 3′ end of the primer (pT12) is tilted up, and guanine of ddGTP is tilted down, skipping the missing nucleotide opposite tG4 so that the α-phosphate of ddGTP and the 3′ hydroxyl group of the primer are less than 4 Å apart (Figure 6b). The sugar-phosphate bond of ddGTP is in weak electron densities, suggesting high mobility. The phosphotidyl transfer reaction did not take place in the crystal structure, but the γ-phosphate of ddGTP is retained. Presumably, the phosphatase activity of Dpo4 is stimulated by the correctly positioned primer-template DNA in the case of type I crystals.

The conformation of the nucleotide 5′ to the template base is remarkably different between the type I and type II structures. In the type I crystal, the base of tA4 is oriented away from the protein, which is similar to that observed in other DNA polymerases, and requires a rotation of over 90° in order to be positioned in the active

site (Figure 6a). The base of tT2 in the type II crystal, however, is oriented toward the same direction as the template base (Figure 6b). It is shielded by P60 and hydrogen bonded to the carbonyl oxygen of G58, and seems ready to slide into the active site. Interestingly, only a pyrimidine can be accommodated in this unconventional position due to the tight fit.

Discussion

Structural Conservation in the Y-Family
Based on amino acid alignments of UmuC and DinB orthologs, we previously suggested that these proteins contained five conserved motifs clustered in their N termini (Kulaeva et al., 1996). According to the three-dimensional structures, we have refined the alignment (Figure 2). Motifs I and III supply the three invariant carboxylates in the active site, motifs II and IV interact with the incoming nucleotide, and motif V contacts the primer strand. All five sequence motifs are distributed within the first three structural domains, which constitute the catalytic core (Figure 3a) (Zhou et al., 2001). These three structural domains are clearly present in all members of the Y-family, because residues forming the hydrophobic
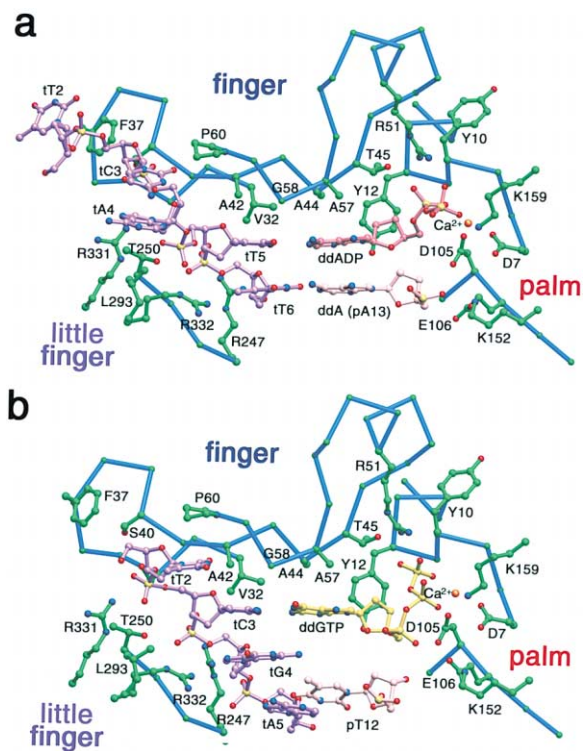
a

b

Figure 6. The Active Site

(a) Type I crystal. The protein residues surrounding the active site are shown with the Ca^{2+} ion, the 3′ end nucleotide of the primer strand (pA13), the incoming nucleotide (ddADP), the template base (tT5), and the surrounding 4 nucleotides (tT2, tC3, tA4, and tT6). Protein side chains are shown in green, the template strand in purple, the primer in pink, and ddADP in dark pink. Nitrogen and oxygen atoms are drawn in blue and red, respectively. $C\alpha$ traces of parts of Dpo4 that frame the active site (residues 7–16, 32–60, 104–108, and 247–250) are shown in blue.

(b) Type II crystal. The last base pair between the template and primer, tA5 and pT12, is shown in addition to the incoming nucleotide (ddGTP), the template base (tC3), and the unpaired tG4 and tT2. ddGTP is highlighted in yellow.

cores are either conserved or substituted by other hydrophobic residues. Visual inspection of the Dbh structure (Zhou et al., 2001) in comparison with the ternary complex structures of Dpo4 revealed some structural differences in the finger and thumb domains; however, the nature of these changes and whether an induced-fit mechanism is involved await further studies.

Eukaryotic members of the Y-family contain a unique ~30–70 residue insertion in the palm domain (Figure 2). The insertion occurs between motifs III and IV, and therefore is unambiguously placed near the N terminus far away from the central cleft (Figure 3a). It is thus unlikely to alter either the active site or the DNA binding of each polymerase. The linker between the finger and palm domains often varies in length and composition (Figure 2). This change in the primary sequence is unlikely to alter the overall structure, but potentially influences the catalytic activity (see discussion later). Finally, *S. cerevisiae* polη contains a third insertion between α helices J and K in the thumb domain. Since the N terminus of αK contacts the template strand (Figure 5a), this insertion possibly alters the protein-DNA interaction.

The sequence corresponding to the little finger is not conserved among members of the Y-family, which raises the question whether or not this domain is present in eukaryotic homologs. Secondary structure predictions indicate that all of the members can potentially form the βαββαβ structure found in the little finger of Dpo4. Residues essential for hydrophobic core formation can be roughly aligned even though sequence conservation is low (Figure 2). Considering these structural features and its function in DNA binding, we predict that the little finger domain is also present in all members of the Y-family of polymerases. The recently published structure of the catalytic core of *S. cerevisiae* polη (Trincao et al., 2001) agrees with our sequence alignment and supports this prediction.

Low-Fidelity DNA Synthesis on Undamaged DNA Template

The limited and nonspecific interactions between Dpo4 and the replicating base pair provide the structural basis for low-fidelity replication on undamaged DNA. The finger and thumb domains of Dpo4 are relatively small in comparison with T7, Taq, and RB69 polymerase. A large crevice separates the little finger of Dpo4 from the catalytic core, which has not been observed in other multidomain polymerases (Figure 7). In addition, the DNA duplex bound to Dpo4 is entirely in B-form. The minor groove, which faces the active site of Dpo4, is thus deeper and harder to reach than in A-form, as observed in the DNA complexed with the high-fidelity A-family polymerases (Doublie et al., 1998; Eom et al., 1996; Kiefer et al., 1998; Li et al., 1998). As a result, the minor groove of the primer-template DNA in the active site, which faces the gap between the finger and little finger domains, is unusually accessible to solvent (Figures 6 and 7). The major groove is also less shielded compared with the high-fidelity polymerase-DNA complexes, such as T7 polymerase (Figure 7).

The residues in direct contact with the replicating base pair, G41, A42, A44, A57, and G58 (Figure 6), are all small, unlike the high-fidelity polymerases, which often involve an aromatic ring of Tyr or Phe or the guanidinium group of Arg to form planar stacking with the incoming base (Doublie et al., 1998; Huang et al., 1998; Li et al., 1998). Except for one water-mediated hydrogen bond, the replicating base pair is unrestrained in either minor or major groove and has no specific polar interactions with Dpo4, which results in Dpo4 being able to replicate DNA with wobble base pairs and potentially to accommodate bulky adducts in template DNA.

Lesion Bypass

Dpo4 is able to bypass *cis-syn* cyclobutane thymine dimers (Boudsocq et al., 2001), as does yeast and human polη (Johnson et al., 1999; Masutani et al., 1999b). The two covalently attached thymine bases are not separable and hence cannot individually fit into the active site to serve as a template base as required by polymerases of the A-, B-, C-, and X-family. In contrast, in the type II crystals of the Dpo4 ternary complex, two adjacent template bases are admitted into the active site simultaneously, illustrating a mechanism that enables Dpo4 to possibly accommodate the thymine dimer. To
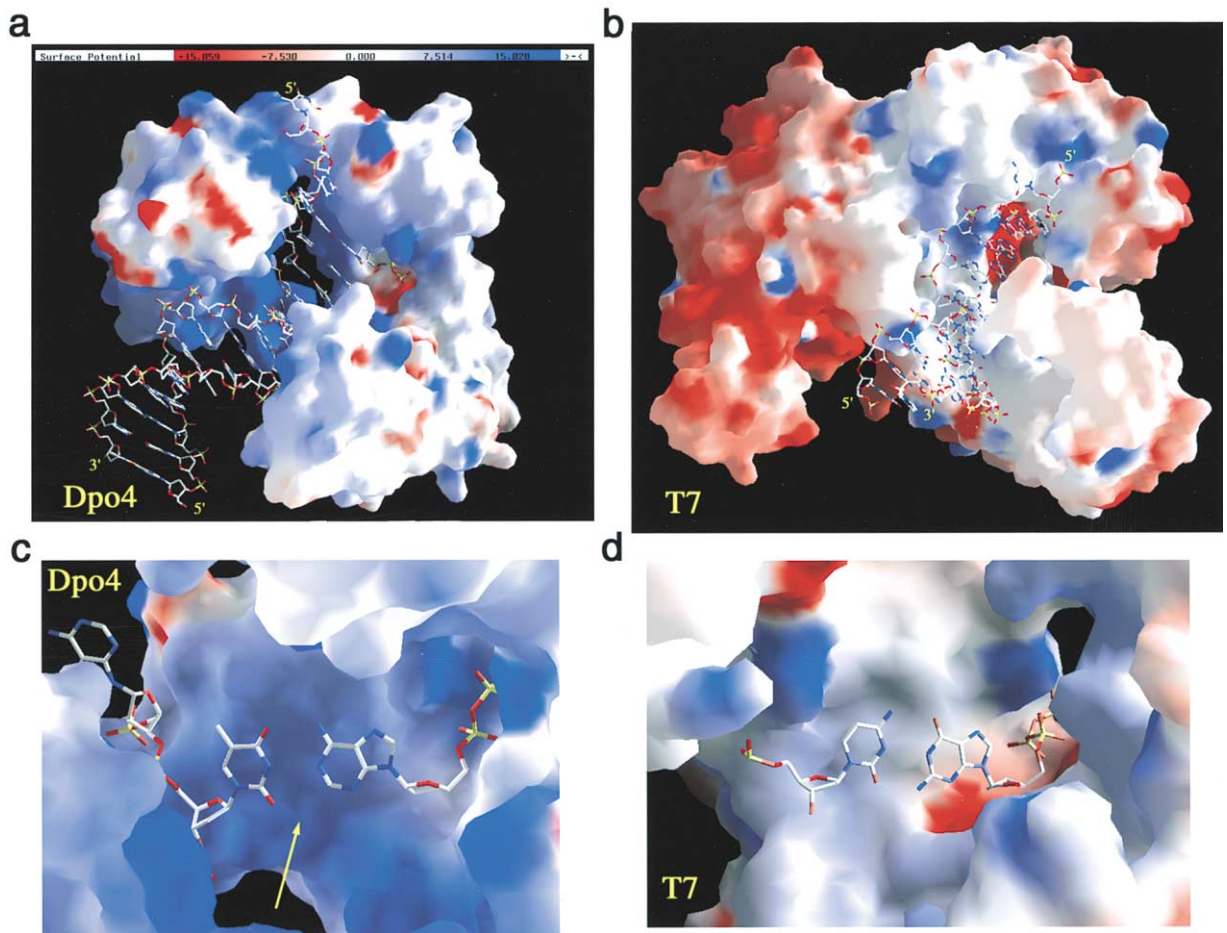
Figure 7. Comparison between Dpo4 and T7 DNA Polymerase

(a) Dpo4 and (b) T7 are shown in molecular surface representation with positive (blue) and negative (red) electrostatic potential. The DNA molecules are drawn in sticks. The minor groove adjacent to the replicating base pair is exposed to solvent in the Dpo4 complex but completely buried in T7 Pol.

(c and d) Closeup views of the replicating base pairs in Dpo4 and T7. The minor groove side of this base pair indicated by the yellow arrow is solvent accessible in Dpo4.

bypass such a lesion, Dpo4 and polη may skip the first base and replicate the second, as observed in the type II crystal. The newly incorporated adenine is then switched to pair with the first thymine, and the polymerase replicates the second again without template translocation. Alternatively, the template base and the pyrimidine 5′ to it can be accommodated in a parallel fashion, as observed in the type II structure (Figure 6b), which may enable the polymerase to faithfully replicate a covalently linked CT or TC photoproduct without slippage.

All DinB-like polymerases studied to date are able to bypass an abasic site, and the resulting daughter strand is often 1 nt shorter than the template (Ohashi et al., 2000a; Wagner et al., 1999; Zhang et al., 2000b). It is possible that a mechanism of template slippage similar to that observed in the type II crystals of Dpo4 may also be employed in replicating an abasic lesion.

Diversity of the Y-Family Polymerases
The Y-family polymerases display remarkably different abilities to traverse damaged DNA and result in different

mutation spectra when replicating either normal or damaged DNA. For instance, PolV (UmuC) tends to make base substitution; PolIV (DinB) tends to slip along the DNA template and result in deletion mutations (McKenzie et al., 2001; Ohashi et al., 2000a; Strauss et al., 2000; Tang et al., 1999; Wagner et al., 1999). While polη traverses a thymine dimer efficiently, only Dpo4 of the DinB branch is able to bypass such dimers. Interestingly, many residues surrounding the replicating base pair, that is, residues 31 to 60, are not conserved among the members of the Y family (Figure 2), which probably contributes to the different substrate specificities observed among these polymerases.

Due to the lack of close and specific contacts with the replicating base pair, the base selectivity of Dpo4 and possibly many other Y-family polymerases seem to mainly depend on the stability of a Watson-Crick base pair. In other cases, such as human polι, specific polar interactions may favor a T:G mismatch. A57 of Dpo4, which is in contact with the base of the incoming nucleotide (Figure 6), is substituted by Lys or Arg in polι and

polη (Figure 2). G58 of Dpo4, which has left-handed main chain torsion angles and is in close contact with the template base, is replaced by Leu in human polι (Figure 2). Such variation of the amino acid composition alters the shape and electrostatic potential of the protein surface and hence may influence selection of the incoming nucleotide.

The polymerase-template interaction can also influence the replication specificity, as suggested by the DNA sequence-dependent error rate of polι (Tissier et al., 2000; Vaisman et al., 2001; Zhang et al., 2000a). The protein-base interactions, aside from the replicating base pair, occur only between the 5′ single-stranded template and the β2-β3 loop of Dpo4 (Figure 6). Interestingly, residues 32–42 in the β2-β3 loop are completely nonconserved among members of the Y-family (Figure 2). Future experiments, in which site-directed mutations are generated within this region, should elucidate the role of this template-interacting loop in the substrate specificity of each polymerase.

In addition to the direct protein-DNA interactions surrounding the active site, communication between the verification of a Watson-Crick base pair by the finger domain and the chemical reaction carried out by the palm must be precise. The E16 of Dpo4, which makes two hydrogen bonds to R77 and maintains the proper orientation of the finger and palm domain, is absolutely conserved in the Y family (Figure 2). Mutation of E39 to Ala in S. cerevisisiae polη, the equivalent of Dpo4 E16, eliminates the polη polymerase activity (Kondratick et al., 2001). The number of amino acids between the finger and palm domains and R77 is conserved in DinB-like polymerases. However, human polη has 1 additional residue in this linker; UmuC has 2, human polι 7, and S. cerevisiae polη 34. The insertions are placed in the back of the polymerase (Figure 3a). They are close enough to potentially interact with the replicating base pair in the minor groove, but most certainly they will change the interactions between the finger and palm and hence the catalytic properties of these polymerases.

Finally, the little finger domain, by virtue of clamping the DNA to the polymerase, may influence the outcome of replication. The overall sequence diversity in the little finger, as well as the lack of conservation of the DNA binding residues (Figure 2), suggests that the mode of DNA binding by the little finger is different in each individual polymerase. The extent of the interaction between the little finger and the DNA major groove, and the overall orientation of the little finger domain and the substrate DNA relative to the active site, can potentially alter the processivity and the fidelity of each polymerase.

Conclusion

The two ternary complex structures of Dpo4 reveal the architecture of a full-length Y-family polymerase and its atomic interactions with substrate DNA and incoming nucleotide. Dpo4 shares a common structural configuration and a conserved active site with all high-fidelity polymerases, but it also possesses an additional DNA binding domain. The relaxed contact of the polymerase with the replicating base pair and the simultaneous translocation of two template bases to the active site observed in the crystal structures suggest a mechanism

for error-prone and lesion bypass DNA synthesis. Although Dpo4 is from the Crenarchaeon S. solfataricus, its structural features are likely to be conserved in all polymerases of the Y-family, and it therefore provides a paradigm for understanding the low-fidelity and lesion bypass properties of all members of the Y-family. It is clear that the crystal structures alone do not answer all of the outstanding questions concerning the biochemical properties of these enzymes, yet they highlight the regions critical for the polymerase function and provide a foundation to postulate possible mechanisms that are testable by future experiments.

Experimental Procedures

Purification and Crystallization

S. solfataricus Dpo4 was overproduced by cloning the dpo4 gene into pET22b (Novagen, Wisconsin) to generate p1914 (Boudsocq et al., 2001). Selenomethionine-labeled (SeMet) Dpo4 was expressed using the methionine auxotrophic E. coli strain B834 (DE3) in a defined medium (Hendrickson et al., 1990). The native and SeMet Dpo4 were purified using HiTrap Q HP and monoS columns (Pharmacia) as described (Boudsocq et al., 2001), except that the phosphate buffer was replaced by HEPES. Purified proteins were concentrated to 20 mg/ml in 20 mM HEPES (pH 7.0), 100 mM NaCl, 1 mM DTT, 0.1 mM EDTA, and 2% glycerol. Oligonucleoties were purchased from either Yale University or Oligos Etc. Inc., and gel purified.

Annealed template (5′-TTCATTAGTCCTTCCCCC-3′) and primer (5′- GGGGGAAGGACT-3′) were used to make type I complex. In the type II complex, the template was 5′-TTCGAAdU*CCTdU*CCCCC-3′ (dU* = 5-bromo-dU), and the primer was 5′-GGGGGAAGGATT-3′. Brominated oligonucleotides were intended for phase determination. The Dpo4 ternary complexes were made by mixing protein and DNA in 1:1.2 molar ratio, and 1 mM ddATP or ddGTP in 20 mM HEPES (pH 7.0), 60 mM NaCl, 5 mM MgCl₂, and 1.0 mM DTT, and incubating the mixture for 10 min at room temperature. The final protein concentration was kept at 8 mg/ml. Crystals were obtained using the hanging-drop vapor diffusion method at 20°C. The well solution contained 16% PEG 3350, 0.1 M HEPES (pH 7.0), 100 mM calcium acetate, and 2.5% glycerol. Crystals were transferred to a mother liquor containing 25% PEG 3350 and 15% ethylene glycol and flash frozen in liquid propane.

Data Collection and Structure Determination

X-ray diffraction data were collected using a Quantum4 CCD detector mounted on beamline X9B at Brookhaven National Synchrotron Light Source. For MAD phasing, four wavelength data were collected from a single type I crystal of SeMet Dpo4, which diffracted X-rays to beyond 1.7 Å at 100 K. The type II crystal diffracted to 2.0 Å. The X-ray data were processed and scaled with HKL (Table 1) (Otwinowski and Minor, 1997).

All six selenium atoms of Dpo4 were found using the four wavelength MAD data at 2.4 Å by SOLVE (Table 1) (Terwilliger and Berendzen, 1999). The resulting electron density map was further improved by RESOLVE (Terwilliger and Berendzen, 1999) and then autotraced by ARP/wARP 5.1 (Perrakis et al., 1999) at 1.7 Å resolutions. About 90% of the protein main chain and 80% of side chains were placed by ARP. The structural model was completed manually using O (Jones et al., 1991). The type II structure was solved using the type I structure as the initial model, followed by one round of rigid body refinement. All components of both ternary complexes were well defined in the electron density maps, except for the disordered C-terminal 11 residues of Dpo4 (342–352) and the 5′-end nucleotide of the template. Both were refined using CNS 1.0 (Table 1) (Brünger et al., 1998). In the type I structure, 93.2% residues of Dpo4 are in the most favored regions of Ramachandran plot and none in disallowed regions. In the type II structure, 91.3% are in the most favored regions and 8.7% in the additional allowed regions of Ramachandran plot.

S. solfataricus P2 Dpo4 Active Site Mutant

A Dpo4 active site mutant was constructed from p1914 using the Quick Change mutagenesis kit (Stratagene). Adjacent residues D105 and E106 in Dpo4 were replaced by Ala using primers P2DEADUP 5′-ATTGCAAGTATAGCTGCAGCTTATCTTGATATC-3′ and P2DEA DDW 5′-ATCAAGATAAGCTGCAGCTATACTTGCAAT-3′. For the convenience of subsequent identification, these primers also contain silent changes that result in a novel PstI restriction enzyme site. The mutant protein was expressed and purified in an identical manner to the wild-type protein described above.

Processivity of Wild-Type and Mutant Dpo4

Standard primer extension reactions were performed at 37°C for 5 min in the presence of all four dNTPs at a concentration of 100 μM. The template was circular M13mp18. The amount of radiolabeled primer/template was kept constant at 10 nM, and the concentration of Dpo4 varied from 200 nM to 5 nM. Reaction products were separated in a 20% polyacrylamide gel containing 7 M urea, and dried gels were analyzed using a PhosphorImager (Boudsocq et al., 2001).

Protease Digestion

Purified Dpo4 and Dpo4-DNA complex at 0.5 mg/ml protein concentrations were digested with 0.05 mg/ml trypsin in 50 mM Tris (pH 8.0) and 50 mM NaCl for 2 hr at room temperature and analyzed by SDS PAGE. The digested fragments were sequenced from the N terminus, which revealed that the cleavage occurred between R240 and V241. The resulting catalytic core of Dpo4 (1–240) was separated from the intact Dpo4 and the little finger domain using a monoS column.

Figures 1, 3, 4c, 5, and 6 were made with Ribbons (Carson, 1987); Figure 7 with GRASP (Nicholls et al., 1991).

References

Boudsocq, F., Iwai, S., Hanaoka, F., and Woodgate, R. (2001). Sulfolobus solfataricus P2 DNA polymerase IV (Dpo4): an archaeal DNA polymerase with lesion-bypass properties akin to eukaryotic polη. Nucleic Acids Res., in press.

Brünger, A.T., Adams, P.D., Clore, G.M., Delane, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. (1998). Crystallography and NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. D54, 905–921.

Carson, M. (1987). Ribbon models of macromolecules. J. Mol. Graphics 5, 103–106.

Doublie, S., Tabor, S., Long, A.M., Richardson, C.C., and Ellenberger, T. (1998). Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Å resolution. Nature 391, 251–258.

Doublie, S., Sawaya, M.R., and Ellenberger, T. (1999). An open and closed case for all polymerases. Structure Fold Des. 7, R31–R35.

Eom, S.H., Wang, J., and Steitz, T.A. (1996). Structure of Taq polymerase with DNA at the polymerase active site. Nature 382, 278–281.

Franklin, M.C., Wang, J., and Steitz, T.A. (2001). Structure of the replicating complex of a polα family DNA polymerase. Cell 105, 657–667.

Friedberg, E.C., and Gerlach, V.L. (1999). Novel DNA polymerases offer clues to the molecular basis of mutagenesis. Cell 98, 413–416.

Goodman, M.F., and Tippin, B. (2000). The expanding polymerase universe. Nat. Rev. Mol. Cell. Biol. 1, 101–109.

Hendrickson, W.A., Horton, J.R., and LeMaster, D.M. (1990). Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. EMBO J. 9, 1665–1672.

Holm, L., and Sander, C. (1998). Touring protein fold space with Dali/FSSP. Nucleic Acids Res. 26, 316–319.

Huang, H., Chopra, R., Verdine, G.L., and Harrison, S.C. (1998). Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. Science 282, 1669–1675.

Ishikawa, T., Uematsu, N., Mizukoshi, T., Iwai, S., Iwasaki, H., Masutani, C., Hanaoka, F., Ueda, R., Ohmori, H., and Todo, T. (2001). Mutagenic and nonmutagenic bypass of DNA lesions by Drosophila DNA polymerases dpolη and dpolι. J. Biol. Chem. 276, 15155–15163.

Johnson, K.A. (1993). Conformational coupling in DNA polymerase fidelity. Annu. Rev. Biochem. 62, 685–713.

Johnson, R.E., Prakash, S., and Prakash, L. (1999). Efficient bypass of a thymine-thymine dimer by yeast DNA polymerase, Polη. Science 283, 1001–1004.

Johnson, R.E., Washington, M.T., Prakash, S., and Prakash, L. (2000a). Fidelity of human DNA polymerase η. J. Biol. Chem. 275, 7447–7450.

Johnson, R.E., Prakash, S., and Prakash, L. (2000b). The human DINB1 gene encodes the DNA polymerase Polθ. Proc. Natl. Acad. Sci. USA 97, 3838–3843.

Johnson, R.E., Washington, M.T., Haracska, L., Prakash, S., and Prakash, L. (2000c). Eukaryotic polymerases ι and ζ act sequentially to bypass DNA lesions. Nature 406, 1015–1019.

Jones, T.A., Zou, J.-Y., and Cowan, S.W. (1991). Improved methods for building models in electron density maps and the location of errors in these models. Acta Crystallogr. A 47, 110–119.

Kiefer, J.R., Mao, C., Braman, J.C., and Beese, L.S. (1998). Visualizing DNA replication in a catalytically active Bacillus DNA polymerase crystal. Nature 391, 304–307.

Kondratick, C.M., Washington, M.T., Prakash, S., and Prakash, L. (2001). Acidic residues critical for the activity and biological function of yeast DNA polymerase η. Mol. Cell. Biol. 21, 2018–2025.

Kulaeva, O.I., Koonin, E.V., McDonald, J.P., Randall, S.K., Rabinovich, N., Connaughton, J.F., Levine, A.S., and Woodgate, R. (1996). Identification of a DinB/UmuC homolog in the archeon Sulfolobus solfataricus. Mutat. Res. 357, 245–253.

Kunkel, T.A., and Bebenek, K. (2000). DNA replication fidelity. Annu. Rev. Biochem. 69, 497–529.

Lamers, M.H., Perrakis, A., Enzlin, J.H., Winterwerp, H.H.K., de Wind, N., and Sixma, T.K. (2000). The crystal structure of DNA mismatch repair protein MutS binding to a G:T mismatch. Nature 407, 711–717.

Li, Y., Korolev, S., and Waksman, G. (1998). 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. 17, 7514–7525.

Lindahl, T., and Wood, R.D. (1999). Quality control by DNA repair. Science 286, 1897–1905.

Masutani, C., Kusumoto, R., Yamada, A., Dohmae, N., Yokoi, M., Yuasa, M., Araki, M., Iwai, S., Takio, K., and Hanaoka, F. (1999a). The XPV (xeroderma pigmentosum variant) gene encodes human DNA polymerase η. Nature 399, 700–704.

Masutani, C., Araki, M., Yamada, A., Kusumoto, R., Nogimori, T., Maekawa, T., Iwai, S., and Hanaoka, F. (1999b). Xeroderma pigmentosum variant (XP-V) correcting protein from HeLa cells has a thymine dimer bypass DNA polymerase activity. EMBO J. 18, 3491–3501.

McKenzie, G.J., Lee, P.L., Lombardo, M., Hastings, P.J., and Rosenberg, S.M. (2001). SOS mutator DNA polymerase IV functions in adaptive mutation and not adaptive amplification. Mol. Cell 7, 571–579.

Muchmore, C.R., Krahn, J.M., Kim, J.H., Zalkin, H., and Smith, J.L. (1998). Crystal structure of glutamine phosphoribosylpyrophosphate amidotransferase from Escherichia coli. Protein Sci. 7, 39–51.

Nicholls, A., Sharp, K.A., and Honig, B. (1991). Protein folding and

association: insights from the interfacial and themodynamic proper-
ties of hydrocarbons. Proteins Struct. Funct. Genet. *11*, 281–296.

Obmolova, G., Ban, C., Hsieh, P., and Yang, W. (2000). Crystal struc-
tures of mismatch repair protein MutS and its complex with a sub-
strate DNA. Nature *407*, 703–710.

Ohashi, E., Bebenek, K., Matsuda, T., Feaver, W.J., Gerlach, V.L.,
Friedberg, E.C., Ohmori, H., and Kunkel, T.A. (2000a). Fidelity and
processivity of DNA synthesis by DNA polymerase κ, the product
of the human *DINB1* gene. J. Biol. Chem. *275*, 39678–39684.

Ohashi, E., Ogi, T., Kusumoto, R., Iwai, S., Masutani, C., Hanaoka,
F., and Ohmori, H. (2000b). Error-prone bypass of certain DNA le-
sions by the human DNA polymerase κ. Genes Dev. *14*, 1589–1594.

Ohmori, H., Friedberg, E.C., Fuchs, R.P.P., Goodman, M.F., Ha-
naoka, F., Hinkle, D., Kunkel, T.A., Lawrence, C.W., Livneh, Z.,
Nohmi, T., et al. (2001). The Y-family of DNA polymerases. Mol. Cell
8, 7–8.

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

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction
data collected in oscillation mode. Methods Enzymol. *276*, 307–326.

Pelletier, H., Sawaya, M.R., Wolfle, W., Wilson, S.H., and Kraut, J.
(1996). Crystal structures of human DNA polymerase β complexed
with DNA: implications for catalytic mechanism, processivity, and
fidelity. Biochemistry *35*, 12742–12761.

Perrakis, A., Morris, R.J., and Lamzin, V.S. (1999). Automated protein
model building combined with iterative structure refinement. Nat.
Struct. Biol. *6*, 458–463.

Sawaya, M.R., Prasad, R., Wilson, S.H., Kraut, J., and Pelletier, H.
(1997). Crystal structures of human DNA polymerase β complexed
with gapped and nicked DNA: evidence for an induced fit mecha-
nism. Biochemistry *36*, 11205–11215.

Schaaper, R.M. (1993). Base selection, proofreading, and mismatch
repair during DNA replication in *Escherichia coli*. J. Biol. Chem. *268*,
23762–23765.

She, Q., Singh, R.K., Confalonieri, F., Zivanovic, Y., Allard, G.,
Awayez, M.J., Chan-Weiher, C.C., Clausen, I.G., Curtis, B.A., De
Moors, A., et al. (2001). The complete genome of the crenarchaeon
Sulfolobus solfataricus P2. Proc. Natl. Acad. Sci. USA *98*, 7835–
7840.

Steitz, T.A. (1999). DNA polymerases: structural diversity and com-
mon mechanisms. J. Biol. Chem. *274*, 17395–17398.

Strauss, B.S., Roberts, R., Francis, L., and Pouryazdanparast, P.
(2000). Role of the dinB gene product in spontaneous mutation in
Escherichia coli with an impaired replicative polymerase. J. Bacte-
riol. *182*, 6742–6750.

Tang, M., Shen, X., Frank, E.G., O'Donnell, M., Woodgate, R., and
Goodman, M.F. (1999). UmuD′₂C is an error-prone DNA polymerase,
Escherichia coli pol V. Proc. Natl. Acad. Sci. USA *96*, 8919–8924.

Tang, M., Pham, P., Shen, X., Taylor, J.S., O'Donnell, M., Woodgate,
R., and Goodman, M.F. (2000). Roles of *E. coli* DNA polymerases
IV and V in lesion-targeted and untargeted SOS mutagenesis. Nature
404, 1014–1018.

Terwilliger, T.C., and Berendzen, J. (1999). Automated structure so-
lution for MIR and MAD. Acta Crystallogr. D *55*, 849–861.

Tissier, A., McDonald, J.P., Frank, E.G., and Woodgate, R. (2000).
Polι, a remarkably error-prone human DNA polymerase. Genes Dev.
14, 1642–1650.

Trincao, J., Johnson, R.E., Escalante, C.R., Prakash, S., Prakash,
L., and Aggarwal, A.K. (2001). Structure of the catalytic core of
S. cerevisiae DNA polymerase η: implications for translesion DNA
synthesis. Mol. Cell *8*, 417–426.

Vaisman, A., Tissier, A., Frank, E.G., Goodman, M., and Woodgate,
R.F. (2001). Human DNA polymerase ι promiscuous mismatch exten-
sion. J. Biol. Chem. *276*, 30615–30622.

Woodgate, R. (1999). A plethora of lesion-replicating DNA polymer-
ases. Genes Dev. *13*, 2191–2195.

Wagner, J., Gruz, P., Kim, S.R., Yamada, M., Matsui, K., Fuchs, R.P.,
and Nohmi, T. (1999). The *dinB* gene encodes a novel *E. coli* DNA

polymerase, DNA pol IV, involved in mutagenesis. Mol. Cell *4*,
281–286.

Wood, R.D., Mitchell, M., Sgouros, J., and Lindahl, T. (2001). Human
DNA repair genes. Science *291*, 1284–1289.

Zhang, Y., Yuan, F., Wu, X., and Wang, Z. (2000a). Preferential incor-
poration of G opposite template T by the low-fidelity human DNA
polymerase ι. Mol. Cell. Biol. *20*, 7099–7108.

Zhang, Y., Yuan, F., Wu, X., Wang, M., Rechkoblit, O., Taylor, J.S.,
Geacintov, N.E., and Wang, Z. (2000b). Error-free and error-prone
lesion bypass by human DNA polymerase κ *in vitro*. Nucleic Acids
Res. *28*, 4138–4146.

Zhou, B.-L., Pata, J.D., and Steitz, T.A. (2001). Crystal structure of
a DinB lesion bypass DNA polymerase catalytic fragment reveals a
classical polymerase catalytic domain. Mol. Cell *8*, 427–437.

Accession Numbers

Coordinates reported in this paper have been deposited in the Pro-
tein Data Bank with ID codes of 1JX4 for the type I complex and
1JXL for the type II complex.