Structure, Vol. 13, 1225–1233, August, 2005, Published by Elsevier Ltd DOI 10.1016/j.str.2005.05.010

Nucleotide-Induced DNA Polymerase Active Site Motions Accommodating a Mutagenic DNA Intermediate

Vinod K. Batra, William A. Beard, David D. Shock, Lars C. Pedersen, and Samuel H. Wilson*
Laboratory of Structural Biology
National Institute of Environmental Health Sciences
National Institutes of Health
P.O. Box 12233
Research Triangle Park, North Carolina 27709

Summary

DNA polymerases occasionally insert the wrong nucleotide. For this error to become a mutation, the mispair must be extended. We report a structure of DNA polymerase β (pol β) with a DNA mismatch at the boundary of the polymerase active site. The structure of this complex indicates that the templating adenine of the mispair stacks with the primer terminus adenine while the templating (coding) cytosine is flipped out of the DNA helix. Soaking the crystals of the binary complex with dGTP resulted in crystals of a ternary substrate complex. In this case, the templating cytosine is observed within the DNA helix and forms Watson-Crick hydrogen bonds with the incoming dGTP. The adenine at the primer terminus has rotated into a *syn*-conformation to interact with the opposite adenine in a planar configuration. Yet, the 3′-hydroxyl on the primer terminus is out of position for efficient nucleotide insertion.

Introduction

Although DNA replication and repair synthesis attempts to preserve Watson-Crick base-pairing rules, an incorrect nucleoside triphosphate (dNTP) is occasionally inserted. The efficiency at which this occurs is only weakly dependent on the identity of the polymerase (Beard et al., 2002a). A DNA mismatch is an obstacle to further DNA synthesis. For DNA polymerases that have an intrinsic proofreading exonuclease, this provides an opportunity to excise the misincorporated nucleotide. For polymerases that lack this proofreading activity, inefficient mismatch extension increases the probability that the polymerase will dissociate from the mismatched primer terminus. This provides an extrinsic proofreading enzyme access to the mismatch. In some instances, however, the DNA mismatch is extended, resulting in a mutagenic event (base substitution error). Structures of DNA mismatches in the confines of the DNA polymerase (pol) β active site indicate that the mispaired bases are in a staggered conformation and do not form hydrogen bonds with one another (Krahn et al., 2003, 2004). In contrast, the structure of mismatched base pairs positioned at the boundary of the nascent base pair binding pocket of the *Bacillus* DNA polymerase I fragment indicates that the mispaired bases form hydrogen bonds with one another in a planar conformation (Johnson and Beese, 2004). In most instances, the hydrogen bonding pattern is similar to that observed for crystal structures of mismatches in duplex DNA in the absence of protein.

Crystallographic structures of DNA polymerases from several polymerase families indicate that they have a modular domain organization. The polymerase domain is typically composed of functionally distinct subdomains. The catalytic subdomain coordinates two divalent metal cations that assist the nucleotidyl transferase reaction. Two other subdomains have principal roles in duplex DNA and nascent base pair (dNTP and templating nucleotide) binding and are spatially situated on opposite sides of the catalytic subdomain. These subdomains are referred to as C (catalytic), D (duplex DNA binding), and N subdomains (nascent base pair binding) to highlight their intrinsic functions (Beard et al., 2002b). These would correspond to the palm, thumb, and fingers subdomains, respectively, according to the nomenclature that utilizes the architectural analogy to a right hand (Ollis et al., 1985). Comparison of DNA polymerase structures bound to DNA and an incoming complementary dNTP reveals that the nascent base pair is "sandwiched" between the N subdomain and the growing DNA terminus (Doublié et al., 1998; Franklin et al., 2001; Garcia-Diaz et al., 2005; Huang et al., 1998; Johnson et al., 2003; Li et al., 1998; Ling et al., 2001; Nair et al., 2004; Pelletier, 1994).

Mammalian DNA polymerase β is a multifunctional enzyme that contributes two key activities to the base excision repair (BER) pathway: DNA synthesis and deoxyribose phosphate (dRP) lyase (Beard and Wilson, 2000). As part of its DNA synthesis capacity during BER, pol β usually inserts a nucleotide into a single-nucleotide gapped DNA repair intermediate (single-nucleotide BER). In addition, the polymerase activity of pol β is necessary for alternate repair pathways that require longer gap-filling DNA synthesis (long-patch BER) (Dianov et al., 1999; Horton et al., 2000). These alternate pathways are needed to remove modified 5′-dRP residues that may arise during DNA damage and that are refractory to the dRP lyase activity of pol β. Long-patch BER can be accomplished through the sequential activities of flap endonuclease 1 (FEN1) cleavage, which generates a one-nucleotide gap, and pol β gap-filling DNA synthesis (Liu et al., 2005). These alternating activities can result in a DNA repair patch of 2–11 nucleotides. Consequently, long-patch BER occurs through a series of one-nucleotide gap-filling reactions. The efficiency by which pol β inserts an incorrect nucleotide, and extends that mispair, is highly dependent on the identity of the mispair and the local DNA sequence context (Beard et al., 2004; Kunkel, 1985; Osheroff et al., 1999). Generally, pol β creates and extends purine-pyrimidine mispairs (transition intermediates) more easily than pyrimidine-pyrimidine or purine-purine mispairs (transversion intermediates) (Beard et al., 2002a, 2004). A mispaired primer terminus lowers the catalytic efficiency for correct insertion. This is usu-

*Correspondence: wilson5@niehs.nih.gov

ally due to a decreased insertion rate and binding affinity for the incoming correct dNTP.

The structural attributes of a mismatch situated at the boundary of the dNTP binding pocket of a polymerase are expected to be dependent on the mismatch and the polymerase since the propensity to extend mismatches is dependent on the identity of the mismatch and the polymerase. More importantly, insight into the structural attributes that promote or deter incorporation of a mismatch is fundamental to a molecular description of mutagenesis and will require binary (polymerase/DNA) and ternary substrate complex structures of these mutagenic intermediates. Here, we report two such crystallographic structures: 1) an open binary DNA complex with an A-A mismatch situated at the boundary of the pol β active site; and 2) the resulting closed ternary complex produced after soaking the crystals with the next correct dNTP. These structures indicate that the nucleotides of the primer terminus, and its templating partner (both adenines), as well as the coding templating base (cytosine) undergo significant conformational changes to accommodate the incoming dGTP at the active site. In doing so, the 3′-hydroxyl at the primer terminus is positioned too far from the αP of the incoming nucleotide, thereby discouraging extension of this mutagenic intermediate. These results highlight the dynamic nature of the polymerase/DNA complex.

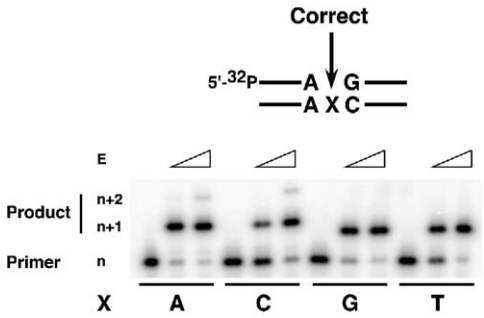

Figure 1. Gap-Filling DNA Synthesis on an A-A Mismatch

Correct nucleotide (1 mM) synthesis from an A-A mismatch was assayed on a series of single-nucleotide gapped DNA substrates (200 nM) where the templating base (X) was altered. Product formation (n+1 and n+2) was assessed after 30 min with 0 (lane 1), 100 (lane 2), or 1000 nM (lane 3) pol β (E). Products were analyzed as described in Experimental Procedures. In contrast to what was observed with the G-G primer terminus (Beard et al., 2004), the results indicate that the unpaired adenosine at the primer terminus does not block gap-filling DNA synthesis by hydrogen bonding with a templating base (e.g., template T). The minor n+2 product observed with the templating A suggests that multiple thymidines can be inserted into a gap with a dinucleotide run of adenines. The n+2 product observed with the templating C is consistent with strand displacement synthesis on the next cytosine. The position of the unextended primer (n) is indicated.

Results and Discussion

Nucleoside Triphosphate Binding

DNA polymerases pause after inserting an incorrect nucleotide since formation of a correct base pair on a mismatched primer terminus is kinetically challenging. This is partly due to the diminished binding affinity for the correct incoming dNTP utilizing the aberrant primer terminus. In contrast, the DNA binding affinity is not generally affected by a terminal mismatch (Beard et al., 2004; Goodman et al., 1993). The low binding affinity of the incoming dNTP has hampered structural characterization of relevant mutagenic intermediates (Krahn et al., 2004). A kinetic survey of the extension of the 12 possible mismatched primer termini by pol β had indicated that homopurine mispairs are extended very poorly, but they permitted binding of the next nucleotide with high affinity (Beard et al., 2004). Consistent with this observation, the apparent DNA binding affinity is increased in the presence of low concentrations of the correct nucleotide ($K_{d,binary}$ = 14 nM, $K_{d,ternary}$ = 0.3 nM), indicating the formation of a tight ternary substrate complex with single-nucleotide gapped DNA with an A-A mismatch at the primer terminus. Such a shift in DNA binding affinity is not observed for an incoming nucleotide that binds weakly (Beard et al., 2004).

Characterization of the extension of a G-G mismatched primer terminus indicated that the primer terminus and incoming nucleotide attempted to satisfy the hydrogen bonding character of their Watson-Crick edge through misalignment (Beard et al., 2004). Accordingly, a templating C in a one-nucleotide gap did not permit dGTP insertion on the G-G mispair since the primer terminus apparently hydrogen bonded with the templating C, effectively occluding the dNTP binding pocket. In contrast, an A or T in the gap permitted correct nucleotide insertion, albeit at low levels, on the G-G terminus. In view of that possibility, we analyzed the ability of pol β to extend the A-A mismatch with alternate templating bases in the one-nucleotide gap (Figure 1). In contrast to the G-G terminus, pol β is able to slowly extend the A-A mispair independent of the identity of the templating base.

Structure Determination

Since single-nucleotide gapped DNA substrates with an A-A mismatch situated at the primer terminus did not occlude the dNTP binding pocket and permitted the incoming nucleotide to bind to the binary complex with high affinity, we sought to structurally characterize two sequential intermediate structures produced during mutagenesis: the structure of the A-A mismatch positioned at the boundary of the dNTP binding pocket prior to (binary complex) and after nucleotide binding (ternary complex). Crystallization conditions were designed to be similar to those used in previous crystal structures of pol β with nicked, gapped, and mismatched DNA substrates (Krahn et al., 2003, 2004; Sawaya et al., 1997). These conditions produced similar monoclinic crystals that are compatible with a range of enzyme conformations. Although the A-A mismatched terminus permitted high-affinity dGTP binding, dGTP insertion was severely reduced (Beard et al., 2004). The dramatic decrease in insertion rate provided the opportunity to solve the ternary substrate complex structure without employing a dideoxy-terminated primer. The ternary complex was formed by soaking crystals of the binary complex with dGTP (Figure 2).

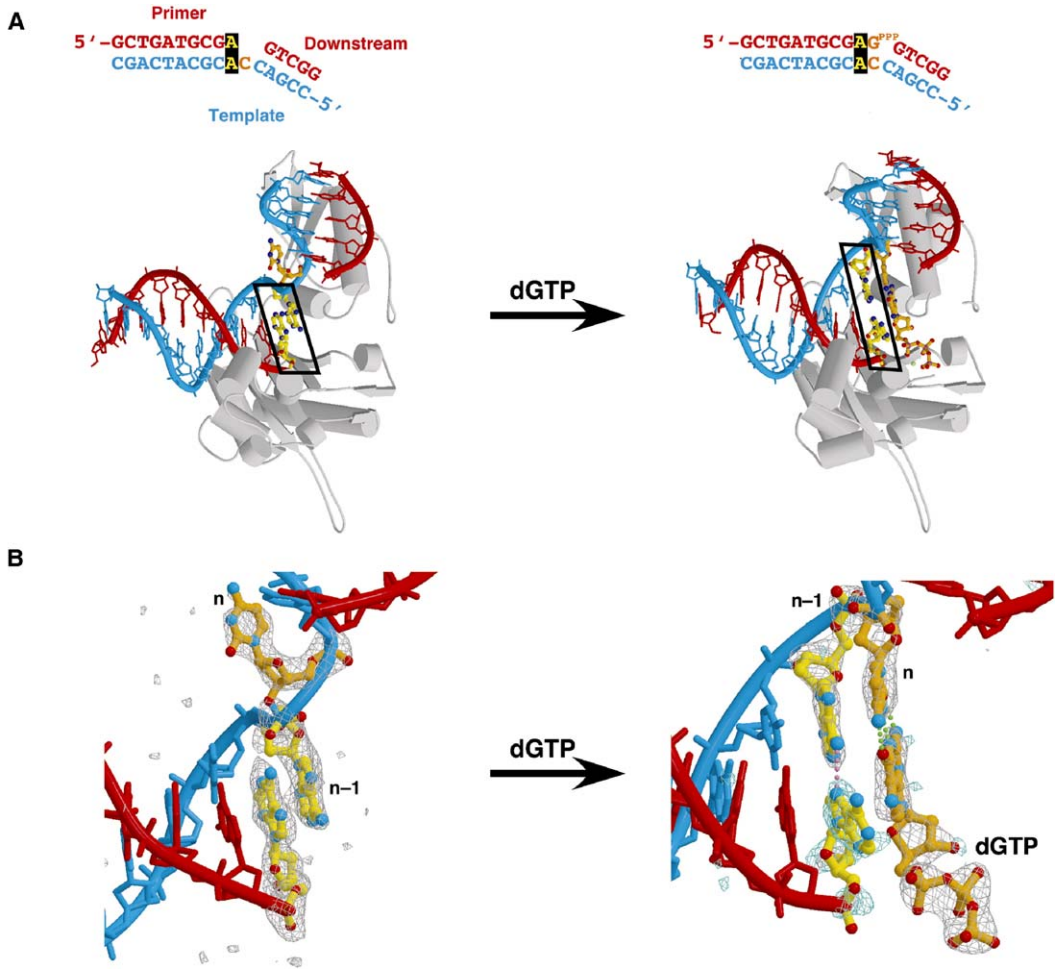

Figure 2. Binary DNA and Ternary Substrate Complex Structures of DNA Polymerase β with a Mismatch at the Boundary of the Active Site

(A) The global conformation of the polymerase domain (gray) of pol β bound to DNA with an A-A mismatch (boxed yellow base pair) at the boundary of the nascent base pair binding pocket. The 8 kDa amino-terminal lyase domain is omitted for clarity. Primer and downstream oligonucleotides (red) were annealed to template DNA (blue) to generate a single-nucleotide gapped DNA substrate with a templating cytosine (orange). An adenosine residue is positioned at the 3′ primer terminus "opposite" an adenosine in the template strand, resulting in an A-A mismatch. The structure of the binary DNA complex is illustrated in the left panel. Soaking dGTP into crystals of the binary complex generated the ternary substrate complex (right panel).

(B) F_o-F_c-simulated annealing electron density omit maps (gray) contoured at 3.2 or 3.3σ showing electron density corresponding to the A-A mismatch and the templating cytosine in the structure of the binary (left panel) or ternary (right panel) substrate complexes, respectively. Electron density for the incoming dGTP is also illustrated in the ternary complex structure. The density for the adenosine in the *syn*-conformation at the primer terminus of the ternary complex is contoured at 4σ (cyan). The densities are superimposed on the refined models of the DNA. The templating base (n) and upstream adenosine (n–1) are indicated. The adenines in the mismatch of the binary complex are stacked with one another. In doing so, the templating cytosine is flipped out of the DNA helix, with the templating adenine occupying the position vacated by the cytosine. In contrast, the adenines in the mismatch of the ternary complex are in a planar conformation where N7 of the adenine of the primer terminus forms a single weak hydrogen bond with the templating adenine. Figures were created in Molscript (Kraulis, 1991) and rendered with Raster3D (Merritt and Bacon, 1997).

Binary Complex Structure

Comparison of the binary DNA complexes of DNA polymerases from several families with those that include an incoming dNTP (i.e., ternary complex) has indicated that several important conformational changes must occur to produce an active complex. One such conformational change that is often noted is the movement of the N subdomain to close upon the nascent base pair binding pocket after binding the correct incoming nucleotide. The polymerase/DNA binary complex is referred to as "open," whereas the ternary substrate complex is "closed." As observed with other binary complexes, the pol β complex with the mismatch at the boundary of the nascent base pair binding pocket is also open (Figures 3A and 4). This is most easily observed by comparing the position of α helix N from the mismatch structure to that observed in either the binary or ternary complex structures with a matched primer terminus (Figure 3A, inset).

In contrast to what has been observed previously with mismatched base pairs at the primer terminus for binary DNA complexes (Johnson and Beese, 2004), the adenines of the A-A mismatch are in a staggered conformation; thus, their purine rings stack with one an-

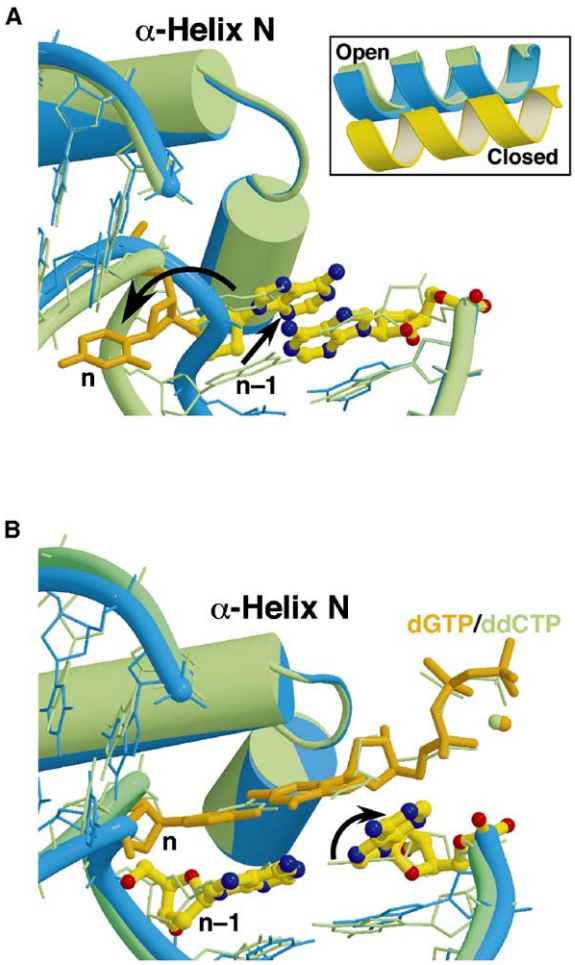

Figure 3. Structural Comparison of DNA Polymerase β with Matched and Mismatched Base Pairs at the Boundary of the Active Site

(A) A comparison of the polymerase and DNA conformations near the polymerase active site in the crystallographic structures of the binary DNA complexes with a matched (1BPX, green) and A-A mismatched (blue) base pair at the nascent base pair binding pocket. The subdomain conformation, as defined by α helix N, of both structures is the inactive open conformation (inset). The adenines in the mismatch (yellow) do not hydrogen bond with one another, but are in a staggered conformation. The templating adenosine has slid into the helix axis to stack with the 3′ primer-terminal base and occupies the position vacated by the templating cytosine (orange) that is observed to be outside of the DNA helix.

(B) A comparison of the polymerase and DNA conformations near the polymerase active site in the crystallographic structures of the ternary substrate complexes with a matched (1BPY, green) and A-A mismatched (blue) base pair at the nascent base pair binding pocket. The ternary complex with the matched primer terminus (G-ddC; template-primer terminus) has an incoming ddCTP positioned opposite the templating G. With the A-A mismatched primer terminus, there is an incoming dGTP forming Watson-Crick hydrogen bonds with the templating cytosine that has positioned itself in the DNA helix. The subdomain conformation of both structures is the closed active conformation. The adenines in the mismatch are no longer in a staggered conformation but now form a single weak hydrogen bond with one another. To accomplish this, the templating adenosine has repositioned itself from the templating position to the n–1 template position normally occupied by the base pair immediately upstream of the nascent base pair binding pocket. Concomitantly, the adenine of the primer terminus has altered its glycosidic preference from *anti* to *syn* to accommodate the two large purine rings in the DNA helix.

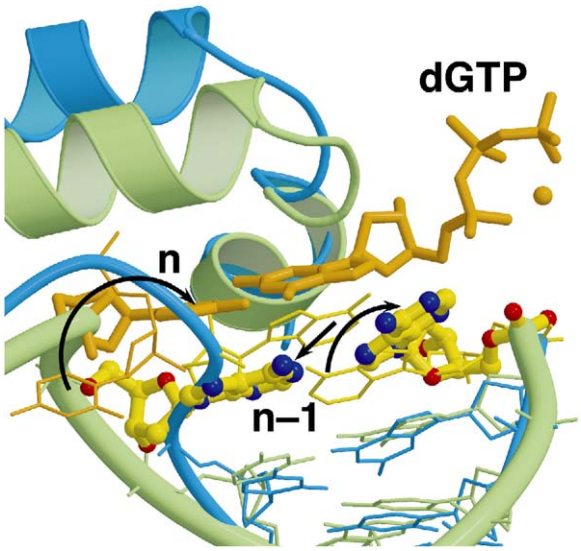

Figure 4. Structural Comparison of the Binary DNA and Ternary Substrate Complexes of DNA Polymerase β with an A-A Mismatch at the Boundary of the Active Site

A comparison of the polymerase and DNA conformations near the polymerase active site in the crystallographic structures of the binary (blue) and ternary (green) complexes with an A-A mismatch (yellow) at the primer terminus. As with matched termini, the N subdomain is observed to be open in the binary complex and closed in the ternary complex. Three striking conformational changes in the DNA are observed in forming the closed ternary substrate complex: (1) the adenine at the primer terminus rotates into a *syn*-conformation; (2) this permits the templating adenine of the mismatch to reposition itself opposite the primer terminus (n–1); (3) the coding template cytosine (orange) flips into the DNA helix, thereby forming Watson-Crick hydrogen bonds with the incoming dGTP (orange).

other (Figures 2 and 3A). The templating adenine "opposite" the adenine of the primer terminus has slid into the binding pocket for the base of the incoming nucleotide so that it occupies a position approximating that of the coding templating base (i.e., position n). However, this misplaced adenine does not support dTTP insertion (data not shown). The authentic coding templating nucleotide (deoxycytidine) is extra helical, and the template strand is shifted toward the dNTP binding site.

Ternary Complex Structure

Soaking crystals of the binary DNA complex with a mismatched primer terminus with the nucleoside triphosphate (dGTP) complementary to the authentic templating base resulted in a ternary substrate complex (Figures 2–4). Comparison of the polymerase domain with that observed with matched DNA indicates that subdomain motions produced a closed nascent base pair binding pocket (Figure 3B). Specifically, the N sub-

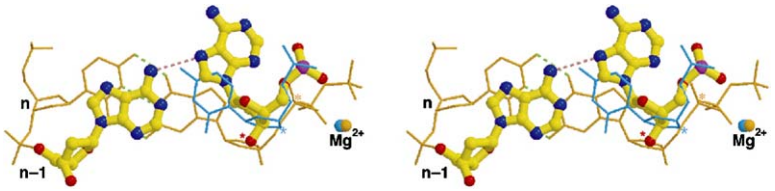

Figure 5. Glycosidic Preferences for the A-A Mismatch in the Ternary Substrate Complex

A stereoview down the DNA helix illustrating the *syn*-conformation of the adenosine at the primer terminus paired with the templating adenosine in an *anti*-conformation (n–1). The nascent base pair (dC-dGTP) is orange and forms Watson-Crick hydrogen bonds (green). A single weak (d = 3.5 Å) hydrogen bond is observed between N7 (primer) and N6 (template) of the mispaired adenines. Although the αP of the incoming dGTP is observed to be in the same position as that observed for ddCTP situated opposite a templating guanine (Figure 3B), the *syn*-conformation of the adenine at the primer terminus has repositioned the 3′-hydroxyl (red asterisk) so that it is observed to be 6.5 Å from the αP (orange asterisk). The primer terminus of a matched nucleotide (ddC; 1BPY) and the dNTP-coordinating Mg^{2+} are superimposed on the mismatch structure (blue). An asterisk identifies the position of the C3′ of this dideoxy-sugar.

domain has repositioned itself to form one face of the nascent base pair binding pocket. In contrast with earlier ternary complex structures with a matched primer terminus, the deoxyribose of the terminus with the mismatch has a 3′-OH. Kinetic analysis of dGTP insertion on this particular mismatch indicates that the rate of insertion is reduced approximately 10^4-fold (Beard et al., 2004). The structure of the mismatch indicates that this 3′-OH is too far from the αP of the incoming dNTP (6.5 Å) to promote rapid catalysis.

In addition to the dramatic subdomain motions inferred to occur upon binding the correct dNTP, even more dramatic and complex DNA conformational changes occur. In contrast to the binary complex where the mismatch is in a staggered conformation, the adenine in the template strand is now in register (i.e., positioned at n–1), so that it is planar with the 3′-terminal adenine of the primer strand (Figures 2–4). To make room for the templating strand adenine at n–1, the adenine at the primer terminus rotates into a *syn*-conformation (Figure 5), forming a weak hydrogen bond between N6 of the templating strand adenine and N7 of the terminal primer adenine. This effectively positions the adenine ring toward the major groove, resulting in a loss of stacking interactions with the incoming nucleotide.

With the templating strand adenine in register (i.e., opposite the adenine of the primer terminus), the authentic templating base (cytosine) is now observed to be stacking within the DNA helix, thereby effectively coding for the incoming dGTP. The repositioning of the coding nucleotide and adenines of the mismatch permits closure of the N subdomain. Kinetic analysis of dGTP insertion on the A-A mismatch indicates that ground-state nucleotide binding is similar to that observed for a correctly matched primer (Beard et al., 2004). Additionally, the DNA binding affinity for the ternary complex is also increased significantly in the presence of the correct incoming nucleotide, suggesting that a significant conformational change has occurred to hasten substrate binding. Repositioning of the N subdomain to produce the closed complex is consistent with these characteristics.

Catalytic Activation

Structural characterization of DNA polymerases in various liganded states suggests that numerous conformational changes must occur to position important catalytic residues. As noted above, the largest conformational change occurs when the N subdomain repositions itself upon binding the correct dNTP. In addition, there are numerous subtle conformational events that occur during this subdomain repositioning. These events comprise both DNA and protein residues. For correct nucleotide insertion, molecular modeling suggests that these events are not concerted, but occur in a systematic series of events (Radhakrishnan and Schlick, 2004; Yang et al., 2002, 2004). Previously, we had identified a group of protein side chains that could detect the position of the N subdomain and transmit this signal to the metal binding site (Sawaya et al., 1997; Vande Berg et al., 2001). These residues have distinct conformations in the binary DNA and ternary substrate complexes. Except for Arg258, these residues are observed in the same conformation as those for a matched primer terminus in both the binary (inactive conformation) and ternary (active conformation) complexes. However, Arg258 appears to assume a unique position in the binary complex with the terminal A-A mismatch (Figure 6). Specifically, Arg258 is indirectly hydrogen bonded through a water molecule to N3 of the terminal primer adenine. In the binary complex with a matched primer terminus, Arg258 hydrogen bonds with a metal binding ligand (Asp192). In the corresponding ternary complex, Asp192 is free to ligand both active site metals, and Arg258 is observed to hydrogen bond with Glu295 and Tyr296. This is also observed to occur in the structure of the ternary complex with the A-A mismatch. It appears that when the terminal adenine on the primer strand assumes the *syn*-conformation, Arg258 is free to assume the active conformation. Although pol β assumes an active conformation in the ternary substrate complex with the terminal mismatch, it exhibits very low activity. Thus, the active conformations of the residues described above are necessary, but not sufficient, to achieve efficient catalytic activation.

The structure of the ternary complex with the mismatch at the primer terminus identifies a single magnesium cation bound to the nucleotide metal binding pocket. The catalytic metal binding site appears to be empty. This is in part due to the aberrant position of the deoxyribose of the primer terminus, removing one of the metal binding ligands and thereby minimizing the influence of the catalytic metal, when it does bind, on the pK_a of the 3′-hydroxyl.

Concluding Remarks and Implications

From previous kinetic analysis of mismatch extension with pol β, we identified conditions that could allow for

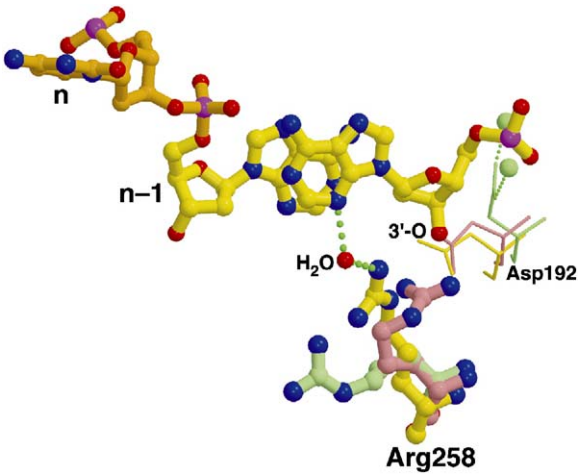

Figure 6. Alternate Arg258 Conformations

Arg258 has been implicated in catalytic activation due to transient interactions with Asp192. This aspartate coordinates the catalytic metal in the closed, but not open, complex. Arg258 (light red) interacts with Asp192 in the binary DNA complex with a matched primer terminus (1BPX) precluding metal binding necessary for catalysis. In the ternary substrate complex (1BPY), Arg258 (light green) hydrogen bonds with Glu295 and Tyr296 (not shown), allowing Asp192 to participate in metal binding (light-green spheres). The structure of the binary complex with the A-A mismatch at the primer terminus indicates that the Arg258 side chain conformation (yellow) is intermediate between that described for the open binary gapped complex and the closed ternary complex. In the mismatched gapped structure, Arg258 indirectly hydrogen bonds with N3 of the adenine at the primer terminus through a water molecule. The Arg258 conformation in the ternary substrate complex with a mismatch is identical to that observed with a matched primer terminus. Thus, the closed structure with the mismatch has induced side chain motions consistent with activation. However, this ternary substrate complex is "inactive" since the incoming nucleotide is not inserted, even though there is a 3′-hydroxyl on the sugar of the primer terminus. Accordingly, the polymerase is poised for catalysis, but induced substrate conformational changes have resulted in an inactive ternary substrate complex.

tight binding of the correct incoming nucleotide. This allowed us to capture a structure of a terminal mismatch before and after binding the correct nucleoside triphosphate. Previous attempts to produce polymerase ternary complex structures with DNA mismatches have been hampered by the weak binding affinity of the incoming nucleotide. Except for with DNA polymerase κ (Washington et al., 2002), mismatch extension is a kinetically challenging step. This is generally due to both weak binding of the next dNTP and weak insertion of the nucleotide on the mismatch. The reduced catalytic efficiency results in termination of DNA synthesis. This permits transfer to an intrinsic proofreading exonuclease site or polymerase dissociation, providing access to an extrinsic exonuclease. In rare instances, the polymerase will rebind to the mismatch and continue DNA synthesis, effectively producing a mutation.

Many of the structural differences observed between the matched and A-A mismatched primer termini of the binary DNA and ternary substrate complexes occur at the primer terminus and templating base. In the binary complex, the adenines of the mismatch attempt to maximize stacking interactions by forming a staggered conformation where the adenines overlap (Figures 2–4). A staggered conformation of mispaired nucleotides in the nascent base pair binding pocket has recently been observed (Krahn et al., 2003, Krahn et al., 2004). However, in contrast to what is observed here, in those cases, the templating base of the mispair is upstream of the base at the primer terminus (Figure 7).

The staggered conformation of the mismatch in the binary complex suggested that the closed complex might be difficult to achieve. However, addition of the complementary dGTP to crystals of the binary complex resulted in significant protein and DNA conformational changes that led to a closed complex. In the binary complex, the templating base is situated outside of the DNA helix. This is reminiscent of the position of the templating base in the binary complex of A family DNA polymerases (Kiefer et al., 1998, Li et al., 1998). In those cases, an aromatic protein side chain occupies the position of the templating base. With the A-A mismatch, the templating adenine at the n–1 position is out of register and is situated in the templating base binding pocket (Figures 2–4). Upon formation of the ternary complex, the templating coding base moves into the DNA helix to form Watson-Crick hydrogen bonds with the incoming nucleotide (Figures 2 and 4). The planarity of the nascent base pair encourages the N subdomain to position itself to effectively close the nascent base pair binding pocket. The template adenine that was out-of-register moved to its correct n–1 position, thereby opening the templating site and permitting the coding cytosine to move into the correct position. Additionally, the adenine at the primer terminus has rotated around its glycosyl bond into a *syn*-conformation to accommodate the template adenine. This 3′ adenine is less well resolved in the ternary complex than in the binary complex (B_{ave} = 84 and 38, respectively); in contrast, the template adenine of the mispair is better resolved in the ternary complex than in the binary complex (B_{ave} = 37 and 58, respectively). The average B factor for the DNA in the ternary and binary complexes is 46 and 34, respectively (Table 1). Thus, repositioning the extra-helical templating cytosine into the DNA helix to code for the incoming dGTP stabilized the template adenine of the mispair in the ternary complex. However, the adenine at the primer terminus, which has moved to a *syn*-conformation, is now in a volatile position owing to the loss of stacking interactions with the template adenine.

Kinetic analysis suggests that the incoming nucleotide binds with an affinity similar to that observed with a matched primer terminus. Notably, the stacking interactions of the adenine at the primer terminus are optimized with the upstream primer base (i.e., n–2), but are eliminated with the base of the incoming nucleotide owing to its *syn*-conformation (Figure 5). Thus, the optimal binding affinity appears to correlate with the state (open, weak dNTP binding; closed, strong dNTP binding) of the nascent base pair binding pocket as defined by the position of the N subdomain rather than the observed stacking interactions with the primer terminus. Consistent with this notion, the ternary complex structure of the R283A mutant crystallizes in an open conformation, and the binding affinity for the correct nucleotide is reduced significantly (Beard et al., 1996).

DNA polymerases, as well as pol β, produce homo-

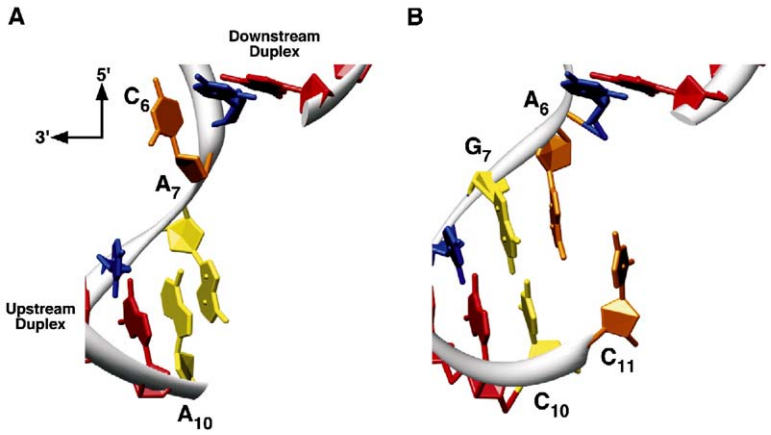

Figure 7. Structural Comparison of Binary DNA Complexes of DNA Polymerase β with Terminal Mismatches

(A) The nucleotide conformations near the polymerase active site in the structure of the binary one-nucleotide gapped DNA complex with the A-A mismatch (yellow) at the primer terminus. The templating (coding) cytosine (orange) is flipped outside of the DNA helix, permitting the 3′ adenine to occupy a position that stacks downstream of the primer terminus adenine. The polarity of the template strand is indicated.

(B) The nucleotide conformations near the polymerase active site of the binary nicked DNA complex with an A-C mismatch (orange) in the nascent base pair binding pocket (i.e., equivalent to a misinserted dCTP opposite A) (Krahn et al., 2004). Since the DNA is nicked, rather than a single-nucleotide gap, the mispair resides in the nascent base pair binding pocket. Again, the mispair was observed to be in a staggered conformation, as observed with other mispairs (T-C and 8-oxoguanine-A) in the pol β nascent base pair binding pocket (Krahn et al., 2003, Krahn et al., 2004). However, in these cases, the template base is upstream of the nucleotide situated at the incoming nucleotide binding pocket. The DNA sequence near the active site is given; the numbers refer to the position from the 5′ terminus of that DNA strand. The sequence of the remaining downstream and upstream duplex DNA is identical in these structures (see Figure 2A). This figure was created by using the USCF Chimera package (Pettersen et al., 2004).

Table 1. Crystallographic Statistics

Complex[a]	Binary	Ternary
Data Collection		
a (Å)	54.1	49.1
b (Å)	79.1	79.5
c (Å)	54.9	55.1
β (°)	105.9	106.2
d_{min} (Å)	2.1	2.6
R_{merge} (%)[b]	8.1 (43.6)[c]	8.8 (42.6)
Completeness (%)	99.2 (92.3)	99.0 (99.1)
I/σ_I	12.2 (2.7)	14.3 (2.7)
Number of observed reflections	128,734	46,243
Number of unique reflections	25,752	12,287
Refinement		
Rms deviations		
Bond lengths (Å)	0.005	0.013
Bond angles (°)	1.043	1.435
B factors (Å^2)		
Main chain	1.3	1.3
Side chain	2.0	1.9
R_{work} (%)[d]	19.9	22.6
R_{free} (%)[e]	23.7	29.9
Average B factor (Å^2)		
Protein	31.38	40.14
DNA	33.97	46.07
Ramachandran analysis (%)[f]		
Favored	97.5	94.4
Allowed	100.0	99.7

[a] The DNA binary complex has an A-A mismatch at the primer terminus adjacent to the nascent base pair binding pocket. The ternary complex includes a complementary dGTP (i.e., templating cytosine).

[b] $R_{merge} = 100 \times \Sigma_h\Sigma_i|I_{h,i} - I_h|/\Sigma_h\Sigma_i I_{h,i}$, where I_h is the mean intensity of symmetry-related reflections, $I_{h,i}$.

[c] Numbers in parentheses refer to the highest resolution shell of data (10%).

[d] $R_{work} = 100 \times \Sigma||F_{obs}| - |F_{calc}||/\Sigma|F_{obs}|$.

[e] R_{free} for a 5% subset of reflections.

[f] As determined by MolProbity (Lovell et al., 2003).

purine mismatches very poorly (Beard et al., 2002a, 2004). Mismatches pose a kinetic obstacle to further DNA synthesis. If a DNA polymerase binds the next nucleoside triphosphate tightly, but does not insert it, its DNA synthesis capacity is severely inhibited. This is the situation for the extension of homopurine mismatches by pol β (Beard et al., 2004). Although this facilitates structural characterization of the ternary complex, it severely limits its biological activity. For most mispairs, the polymerase could dissociate from the aberrant primer terminus to allow a proofreading exonuclease access to the mutagenic intermediate. Since the next correct dNTP typically binds weakly with a mispaired primer terminus, DNA dissociation is not affected. However, if the next dNTP binds, the dissociation of the polymerase from the mutagenic intermediate is reduced (i.e., $k_{off,\ DNA}$ decreased) and a "dead-end" complex is formed. This appears to be structurally equivalent to the closed form of the complex. The tight binding of the next correct nucleotide observed for pol β could serve to protect mispaired adenines from an extrinsic proofreading exonuclease, thereby increasing the probability that it may be extended during long-patch BER. The inability to discriminate between nucleotides that result in active or inactive complexes is also pertinent to the biological activity of DNA polymerase I from *Escherichia coli* (A family), DPO4 (Y family), and DNA polymerase λ (X family). In these instances, the DNA polymerase discriminates between correct and incorrect nucleotides poorly at binding (i.e., $K_{d,correct} \sim K_{d,incorrect}$), but only slowly inserts the incorrect nucleotide (Fiala et al., 2004, Fiala and Suo, 2004, Kuchta et al., 1988). Accordingly, a similar dead-end complex is formed, effectively inactivating the enzyme under biological conditions (i.e., in the presence of all four dNTPs).

The solution structure of an A-A mismatch in duplex DNA indicates that the mismatch is well accommodated: the bases are planar and in an *anti*-conformation (Gervais et al., 1995). Several structures of DNA mis-

matches at the primer terminus (i.e., postinsertion site) in the confines of an A family DNA polymerase, *Bacillus* DNA polymerase I fragment, have recently been reported (Johnson and Beese, 2004). In contrast to the binary pol β A-A mismatch structure, the binary *Bacillus* fragment DNA mismatches are planar, rather than staggered, and in many instances, they are structurally similar to those reported previously in duplex DNA in the absence of protein. Although an A-A mismatch at the primer terminus was not examined, the G-G mismatch indicated that the 3′ primer terminal guanine was in a *syn*-conformation. This is observed in the pol β ternary complex structure, but not the binary complex. This difference could be due to the different nature of the N subdomain motions with these distinct polymerases (Beard and Wilson, 2003). Additionally, the binary *Bacillus* DNA complex with an A-G mismatch at the postinsertion site indicates that both bases prefer the *anti*-conformation, but the guanine will assume a *syn*-conformation if it oxidized at C8 (Hsu et al., 2004). The structure of a ternary complex of T7 DNA polymerase (A family) with 8-oxodG-A at the postinsertion site exhibits the same glycosidic preference as observed for the *Bacillus* fragment binary complex (Brieba et al., 2004; Hsu et al., 2004). A structure of a ternary complex with a G-T mismatch at the primer terminus has been reported for a Y family DNA polymerase from *Sulfolobus solfataricus* DPO4 (Trincao et al., 2004). In this situation, the mispair is in a planar reverse wobble conformation. A binary complex with the mismatch was not determined, so the conformational accommodations after binding the correct nucleotide cannot be accessed. Since the structure of an A-A mismatch at the postinsertion site in complex with other DNA polymerases is lacking, the structure of other mispairs at the primer terminus of pol β will be revealing.

Experimental Procedures

Sample Preparation

Wild-type human DNA polymerase β was expressed and purified as described (Beard and Wilson, 1995). The protein was washed four times with several volumes of 20 mM Bis-Tris (pH 7.0) and concentrated to 15 mg ml^{-1}.

DNA substrates consisted of a 16-mer template, a complementary 9-mer primer strand with an additional noncomplementary 3′ nucleotide (i.e., 10-mer total), and a 5-mer downstream oligonucleotide (Oligos Etc., Redding Center, CT). Proper annealing produces a one-nucleotide gapped DNA substrate with a cytosine serving as the templating base. The DNA sequence was nearly identical to that used in previous studies (Sawaya et al., 1997) in order to minimize sequence-dependent structural differences. The sequence of the downstream oligonucleotide was 5′-GTCGG-3′, and the 5′ terminus was phosphorylated. The template sequence was 5′-CCGA CCACGCATCAGC-3′, and the primer sequence was 5′-GCTGAT GCGA-3′ (the underlined base identifies the mismatched base). Proper annealing results in an A-A mismatch at the 3′ primer terminus (Figure 2A). Oligonucleotides were dissolved in 20 mM $MgCl_2$, 100 mM Tris/HCl (pH 7.5). Each set of template, primer, and downstream oligonucleotides was mixed in a 1:1:1 ratio and annealed with a PCR thermocycler by heating for 10 min at 90°C and cooling to 4°C (1°C min^{-1}), resulting in a 1 mM mixture of gapped duplex DNA. This solution was then mixed with an equal volume of pol β at 4°C. The mixture was warmed to 35°C and gradually cooled to 4°C.

Crystallization and Data Collection

Pol β-DNA complexes were crystallized by sitting drop vapor diffusion versus a reservoir solution of 14%–20% PEG-3350, 350 mM sodium acetate, and 50 mM imidazole (pH 7.5). Drops were incubated at 18°C and streak seeded after 1 day. Crystals grew in ~2–4 days after seeding, were transferred to cryoprotectant solutions containing artificial mother liquor with 12% ethylene glycol, and were then flash-frozen at 100 K in a nitrogen stream. Data were collected at 100 K on an R-Axis IV area detector system mounted on a RU3H rotating anode generator. Data were integrated and reduced with HKL2000 software (Otwinowski and Minor, 1997).

Structure Determination

The structure of pol β containing gapped DNA with an A-A mismatch was determined by molecular replacement from the structure of pol β complexed with fully complementary gapped DNA (PDB accession 1BPX). The two crystal structures have similar lattices and are sufficiently isomorphous to determine the molecular replacement model position by rigid body refinement. The 1BPX model was modified by substituting oligonucleotide bases with the appropriate sequence included in the A-A mismatch crystallization. The working R factor for the final binary complex model is 19.9% for all data to 2.1 Å, with an R_{free} of 23.7%. The ternary substrate complex structure was determined by a similar procedure by using the 1BPY structure as the starting model. The working R factor for the final ternary complex model is 22.6% for all data to 2.6 Å, with an R_{free} of 29.9%. Crystallographic data collection and refinement statistics are summarized in Table 1.

Gap-Filling DNA Synthesis

DNA synthesis was assayed on a series of single-nucleotide gapped DNA substrates where the templating base was altered. The oligonucleotide sequences were as described previously (Beard et al., 2004). The sequences of the oligonucleotides used for crystallization are identical to the core sequences of those used for kinetic analyses. Enzyme activities were determined by using a reaction mixture (50 μl) containing 50 mM Tris-HCl (pH 7.4), 100 mM KCl, 5 mM $MgCl_2$, 1 mM complementary dNTP, and 200 nM single-nucleotide gapped DNA. Further details are given in the legend to Figure 1. Reactions were stopped with 20 μl 0.5 M EDTA and were mixed with an equal volume of formamide dye, and the products were separated on 15% denaturing polyacrylamide gels. The substrate and products in the dried gels were visualized with a PhosphorImager (Amersham Biosciences).

Acknowledgments

We thank J. Krahn and R. London for critical reading of the manuscript.

Received: March 8, 2005
Revised: May 4, 2005
Accepted: May 5, 2005
Published: August 9, 2005

References

Beard, W.A., and Wilson, S.H. (1995). Purification and domain-mapping of mammalian DNA polymerase β. Methods Enzymol. *262*, 98–107.

Beard, W.A., and Wilson, S.H. (2000). Structural design of a eukaryotic DNA repair polymerase: DNA polymerase β. Mutat. Res. *460*, 231–244.

Beard, W.A., and Wilson, S.H. (2003). Structural insights into the origins of DNA polymerase fidelity. Structure (Camb) *11*, 489–496.

Beard, W.A., Osheroff, W.P., Prasad, R., Sawaya, M.R., Jaju, M., Wood, T.G., Kraut, J., Kunkel, T.A., and Wilson, S.H. (1996). Enzyme-DNA interactions required for efficient nucleotide incorporation and discrimination in human DNA polymerase β. J. Biol. Chem. *271*, 12141–12144.

Beard, W.A., Shock, D.D., Vande Berg, B.J., and Wilson, S.H. (2002a). Efficiency of correct nucleotide insertion governs DNA polymerase fidelity. J. Biol. Chem. *277*, 47393–47398.

Beard, W.A., Shock, D.D., Yang, X.-P., DeLauder, S.F., and Wilson,

S.H. (2002b). Loss of DNA polymerase β stacking interactions with templating purines, but not pyrimidines, alters catalytic efficiency and fidelity. J. Biol. Chem. *277*, 8235–8242.

Beard, W.A., Shock, D.D., and Wilson, S.H. (2004). Influence of DNA structure on DNA polymerase β active site function: extension of mutagenic DNA intermediates. J. Biol. Chem. *279*, 31921–31929.

Brieba, L.G., Eichman, B.F., Kokoska, R.J., Doublie, S., Kunkel, T.A., and Ellenberger, T. (2004). Structural basis for the dual coding potential of 8-oxoguanosine by a high-fidelity DNA polymerase. EMBO J. *23*, 3452–3461.

Dianov, G.L., Prasad, R., Wilson, S.H., and Bohr, V.A. (1999). Role of DNA polymerase β in the excision step of long patch mammalian base excision repair. J. Biol. Chem. *274*, 13741–13743.

Doublié, 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.

Fiala, K.A., and Suo, Z. (2004). Pre-steady-state kinetic studies of the fidelity of *Sulfolobus solfataricus* P2 DNA polymerase IV. Biochemistry *43*, 2106–2115.

Fiala, K.A., Abdel-Gawad, W., and Suo, Z. (2004). Pre-steady-state kinetic studies of the fidelity and mechanism of polymerization catalyzed by truncated human DNA polymerase λ. Biochemistry *43*, 6751–6762.

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

Garcia-Diaz, M., Bebenek, K., Krahn, J.M., Kunkel, T.A., and Pedersen, L.C. (2005). A closed conformation for the Pol λ catalytic cycle. Nat. Struct. Mol. Biol. *12*, 97–98.

Gervais, V., Cognet, J.A.H., Le Bret, M., Sowers, L.C., and Fazakerley, G.V. (1995). Solution structure of two mismatches A•A and T•T in the *K*-ras gene context by nuclear magnetic resonance and molecular dynamics. Eur. J. Biochem. *228*, 279–290.

Goodman, M.F., Creighton, S., Bloom, L.B., and Petruska, J. (1993). Biochemical basis of DNA replication fidelity. Crit. Rev. Biochem. Mol. Biol. *28*, 83–126.

Horton, J.K., Prasad, R., Hou, E., and Wilson, S.H. (2000). Protection against methylation-induced cytotoxicity by DNA polymerase β-dependent long patch base excision repair. J. Biol. Chem. *275*, 2211–2218.

Hsu, G.W., Ober, M., Carell, T., and Beese, L.S. (2004). Error-prone replication of oxidatively damaged DNA by a high-fidelity DNA polymerase. Nature *431*, 217–221.

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.

Johnson, S.J., and Beese, L.S. (2004). Structures of mismatch replication errors observed in a DNA polymerase. Cell *116*, 803–816.

Johnson, S.J., Taylor, J.S., and Beese, L.S. (2003). Processive DNA synthesis observed in a polymerase crystal suggests a mechanism for the prevention of frameshift mutations. Proc. Natl. Acad. Sci. USA *100*, 3895–3900.

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.

Krahn, J.M., Beard, W.A., Miller, H., Grollman, A.P., and Wilson, S.H. (2003). Structure of DNA polymerase β with the mutagenic DNA lesion 8-oxodeoxyguanine reveals structural insights into its coding potential. Structure (Camb) *11*, 121–127.

Krahn, J.M., Beard, W.A., and Wilson, S.H. (2004). Structural insights into DNA polymerase deterrents for misincorportion support an induced-fit mechanism for fidelity. Structure (Camb) *12*, 1823–1832.

Kraulis, P. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of proteins. J. Appl. Crystallogr. *24*, 946–950.

Kuchta, R.D., Benkovic, P., and Benkovic, S.J. (1988). Kinetic mechanism whereby DNA polymerase I (Klenow) replicates DNA with high fidelity. Biochemistry *27*, 6716–6725.

Kunkel, T.A. (1985). The mutational specificity of DNA polymerase-β during *in vitro* DNA synthesis: production of frameshift, base substitution, and deletion mutations. J. Biol. Chem. *260*, 5787–5796.

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.

Ling, H., Boudsocq, F., Woodgate, R., and Yang, W. (2001). Crystal structure of a Y-family DNA polymerase in action: a mechanism for error-prone and lesion-bypass replication. Cell *107*, 91–102.

Liu, Y., Beard, W.A., Shock, D.D., Prasad, R., Hou, E.W., and Wilson, S.H. (2005). DNA polymerase β and flap endonuclease 1 enzymatic specificities sustain DNA synthesis for long patch base excision repair. J. Biol. Chem. *280*, 3665–3674.

Lovell, S.C., Davis, I.W., Arendall, W.B., III, de Bakker, P.I.W., Word, J.M., Prisant, M.G., Richardson, J.S., and Richardson, D.C. (2003). Structure validation by Cα geometry: φ,ψ and Cß deviation. Proteins *50*, 437–450.

Merritt, E.A., and Bacon, D.J. (1997). Raster3D: photorealistic molecular graphics. Methods Enzymol. *277*, 505–524.

Nair, D.T., Johnson, R.E., Prakash, S., Prakash, L., and Aggarwal, A.K. (2004). Replication by human DNA polymerase-ι occurs by Hoogsteen base-pairing. Nature *430*, 377–380.

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.

Osheroff, W.P., Jung, H.K., Beard, W.A., Wilson, S.H., and Kunkel, T.A. (1999). The fidelity of DNA polymerase β during distributive and processive DNA synthesis. J. Biol. Chem. *274*, 3642–3650.

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

Pelletier, H. (1994). Polymerase structures and mechanism. Science *266*, 2025–2026.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 1605–1612.

Radhakrishnan, R., and Schlick, T. (2004). Orchestration of cooperative events in DNA synthesis and repair mechanism unraveled by transition path sampling of DNA polymerase β's closing. Proc. Natl. Acad. Sci. USA *101*, 5970–5975.

Sawaya, M.R., Prasad, P., 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 mechanism. Biochemistry *36*, 11205–11215.

Trincao, J., Johnson, R.E., Wolfle, W.T., Escalante, C.R., Prakash, S., Prakash, L., and Aggarwal, A.K. (2004). Dpo4 is hindered in extending a G•T mismatch by a reverse wobble. Nat. Struct. Mol. Biol. *11*, 457–462.

Vande Berg, B.J., Beard, W.A., and Wilson, S.H. (2001). DNA structure and aspartate 276 influence nucleotide binding to human DNA polymerase β: implication for the identity of the rate-limiting conformational change. J. Biol. Chem. *276*, 3408–3416.

Washington, M.T., Johnson, R.E., Prakash, L., and Prakash, S. (2002). Human *DINB1*-encoded DNA polymerase κ is a promiscuous extender of mispaired primer termini. Proc. Natl. Acad. Sci. USA *99*, 1910–1914.

Yang, L., Beard, W.A., Wilson, S.H., Broyde, S., and Schlick, T. (2002). Polymerase β simulations suggest that Arg258 rotation is a slow step rather than large subdomain motions *per se*. J. Mol. Biol. *317*, 679–699.

Yang, L., Arora, K., Beard, W.A., Wilson, S.H., and Schlick, T. (2004). Critical role of magnesium ions in DNA polymerase β's closing and active site assembly. J. Am. Chem. Soc. *126*, 8441–8453.

Accession Numbers

Coordinates and structure factors for the binary and ternary complex structures have been deposited in the Protein Data Bank under accession codes 1ZJM and 1ZJN, respectively.