Structure and Mechanism of DNA Polymerase β

William A. Beard 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 27709-12233

Received September 28, 2005

Contents

1. Introduction

1.1. Overview of Mammalian Base Excision Repair

Cells throughout nature contain multiple and overlapping DNA repair pathways that are essential for maintaining the integrity of genomic DNA.[1,2] The DNA repair pathway known as "base excision repair" (BER) protects the genome by removing damaged nucleotides and abasic sites arising from a variety of exogenous and endogenous stressors.[3] Spontaneous and enzymatic removal of damaged bases through hydrolytic *N*-glycosidic bond cleavage leads to a mutagenic BER intermediate referred to as an abasic or apurinic/apyrimidinic (AP) site. It has been estimated that as many as 10 000 hydrolytic depurinations occur per day in a mammalian cell.[4] In addition, AP sites arise as intermediates during the BER of deaminated cytosines and adenines, base oxidation arising from reactive oxygen species (ROS), and base alkylation from metabolites (e.g., *S*-adenosylmethionine) and environmental or therapeutic alkylating agents (e.g., nitrosamines and temozolomide, respectively). Finally, natural and therapeutic perturbations of nucleotide metabolism can lead to incorporation of unnatural bases into DNA, such as uracil and 8-oxoguanine. These nucleotides are removed from DNA through BER in both prokaryotic and eukaryotic cells.

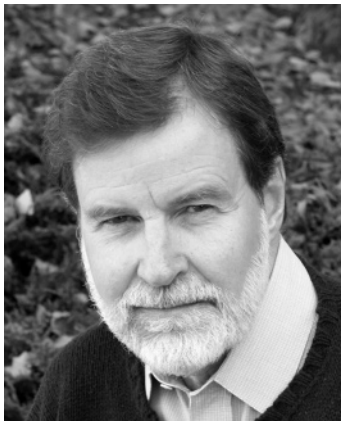
Using purified proteins, BER has been reconstituted *in vitro*, thereby revealing a biochemical pathway of sequential enzymatic steps (Figure 1a).[5,6] In operational terms, the BER pathway is distinguished from other DNA repair pathways by the short excision patch generated in double-stranded DNA after removal of the base lesion and also by the fact that DNA lesions repaired by BER are generally limited to base modifications that do not greatly alter the size and shape of the base. The five core steps of BER are (i) lesion-recognition/base-removal, (ii) strand incision, (iii) incised strand processing (enabling DNA synthesis or DNA ligation), (iv) DNA synthesis to fill the gap, and (v) DNA strand ligation. Since the DNA intermediates in BER are themselves cytotoxic or lead to genomic instability through recombination, a highly concerted handoff from one step to the next has been proposed.[7] The mechanisms of achieving this coordination are unknown. Yet, a reasonable model for explaining step-to-step coordination in BER calls for a combination of exquisite DNA substrate specificity by the respective enzymes and microenvironment effects afforded by multi-protein complexes with the necessary enzymatic machinery and accessory proteins.

Mammalian enzymes performing functions similar to the prokaryotic counterparts for each step in BER have been identified, cloned, and overexpressed as recombinant proteins. Several DNA glycosylases have been identified and characterized, including the well-known mammalian DNA glycosylases 3-methyladenine-DNA glycosylase (AAG), 8-oxoguanine DNA glycosylase (OGG), endonuclease III-like (Nth) DNA glycosylases, and the endonuclease VIII-like (NEIL) DNA glycosylases.[8] These DNA glycosylases recognize and remove altered bases in double-stranded DNA, creating an intact or cleaved AP site. There are two general classes of DNA glycosylases: the monofunctional DNA glycosylases catalyze the hydrolytic removal of a damaged base that results in an AP site, and the bifunctional glycosylases additionally catalyze strand incision at the AP site after base removal (Figure 1a). The AP site created from monofunctional glycosylase activity is subsequently 5′-incised by AP endonuclease, generating a one-nucleotide gap

* Corresponding author information: telephone, 919-541-3267; fax, 919-541-3592; e-mail, wilson5@niehs.nih.gov.

10.1021/cr0404904
Published on Web 01/14/2006

William A. Beard explored electron and proton transport coupling for ATP synthesis in the laboratory of Professor Richard Dilley at Purdue University. After receiving his Ph.D. in 1986, he did postdoctoral training with James Appleman and Raymond Blakely at St. Jude Children's Research Hospital characterizing wild-type and mutant derivatives of human dihydrofolate reductase by steady-state, pre-steady-state, and transient-state kinetic methods. In 1991, he joined Samuel H. Wilson at the National Cancer Institute, NIH, to study the enzymology of HIV-1 reverse transcriptase and DNA repair enzymes, primarily mammalian DNA polymerase β. Through a combination of site-directed mutagenesis, kinetic analysis, molecular modeling and structure determination, these systems serve as ideal models for uncovering basic catalytic mechanisms and strategies that influence genome stability. He is currently a Staff Scientist at the National Institute of Environmental Health Sciences, NIH.



Samuel H. Wilson received a B.A. degree from the University of Denver and did his graduate and postdoctoral training in medicine and biochemistry, respectively, at Harvard Medical School and the NIH, and conducted two years of graduate training in organic synthesis at the University of Denver Research Institute. He began his career in 1970 at the National Cancer Institute (NCI), National Institutes of Health (NIH), and remained there until 1991, when he moved to the University of Texas, Galveston, to found a center focused in the areas of genetic toxicology, structural biology, and functional genomics. In 1996, he moved to the National Institute of Environmental Health Sciences (NIEHS), NIH, as Deputy Director, where he has been instrumental in helping develop the Institute's programs in genetic susceptibility, functional genomics, children's health research, and community involvement, among others. As head of the DNA Repair and Nucleic Acid Enzymology Section in the Laboratory of Structural Biology, he has had a long-term research interest in mammalian DNA metabolism. Over the past several years, he and his associates have focused their efforts on the mammalian base excision DNA repair pathway and the role of DNA polymerase β. He has authored and coauthored over 300 research publications and has been editor of four reference volumes. His recent activities include membership on numerous advisory groups involved in biological and medical research. He was chair of the 2001 Mammalian DNA Repair Gordon Research Conference and co-chair of the 2002/2004 US–Japan and 2003/2005 US–EU international conferences on DNA Repair. He is an Editor of the journal *DNA Repair* and is a member of several editorial boards.

with 3′-OH and 5′-deoxyribose phosphate (dRP) termini in the gap. The primary DNA synthesis role is contributed by DNA polymerase β (pol β), which also possesses dRP lyase activity.[9,10] The lyase activity removes the 5′-sugar-phosphate residue, leaving a 5′-phosphate on the downstream DNA strand. After single-nucleotide gap-filling DNA synthesis and removal of the 5′-sugar-phosphate group, the nicked DNA will be ligated by DNA ligase I[11] or III[12] to complete BER.

Although the core BER pathway noted above is consistent with the enzymatic activities and substrate specificities of the cloned enzymes, there are many biological circumstances altering specificities and coordination of BER enzymes that can lead to alternate sub-pathways. For example, if the AP site is incised by a bifunctional DNA glycosylase, the incision event results in a one-nucleotide gap intermediate where the 3′-margin of the gap is blocked with a sugar-phosphate group (Figure 1a). Additional enzymatic activities are therefore required to tailor the 3′-margin and generate a 3′-OH for gap-filling DNA synthesis. Cells have several enzymes that can liberate a blocked 3′-end: polynucleotide kinase[13] or AP endonuclease.[14]

Finally, recent studies have shown that BER can also occur by sub-pathways that generate repair patches greater than one-nucleotide and are referred to as long-patch BER.[15,16] For example, a modified 5′-dRP group can interrupt single-nucleotide BER (Figure 1b). In this situation, the modified dRP group is not a substrate for the lyase activity of pol β. In this situation, pol β-dependent strand displacement DNA synthesis generates a longer repair patch and a single-stranded DNA-flap with the modified sugar at its 5′-end. The flap is subsequently removed by flap endonuclease-1 (FEN1).[17,18] Alternatively, FEN1 can cleave one-nucleotide downstream of the modified 5′-dRP group prior to strand displacement DNA synthesis, thereby generating a one-nucleotide gap, an ideal substrate for pol β.[19] In addition to pol β-dependent long-patch BER, pol β-independent BER has also been reported.[15,16]

Quantitative description of the BER sub-pathways within living cells is lacking but is clearly a prerequisite to understanding the contribution of each sub-pathway in the cellular responses to genotoxic stresses. This quantification of cellular BER sub-pathways is particularly important, since many of the enzymes and accessory proteins participating in the sub-pathways of BER are also integral components of other DNA repair pathways (e.g., nucleotide excision repair and mismatch repair) and metabolic processes (e.g., replication), and these processes can undoubtedly modulate BER.

1.2. DNA Polymerase β

DNA polymerases have been divided into seven families based on sequence alignment.[20,21] There are at least 14 mammalian DNA polymerases belonging to five of these families (A-, B-, RT-, X-, and Y-families). Identified in 1971,[22,23] DNA polymerase β is a member of the X-family that also includes the template-independent enzyme, terminal deoxynucleotidyl transferase (TdT), and recently identified template-dependent members DNA polymerases λ[24] and μ.[25] DNA polymerase β is found in all vertebrate species as a 39-kDa protein lacking intrinsic 3′- or 5′-exonuclease activities but containing 5′-dRP lyase and AP lyase activities.[26,27] In light of its size, the enzyme is considered the simplest naturally occurring cellular DNA polymerase, making it is an ideal model for studies of the nucleotidyl transferase and

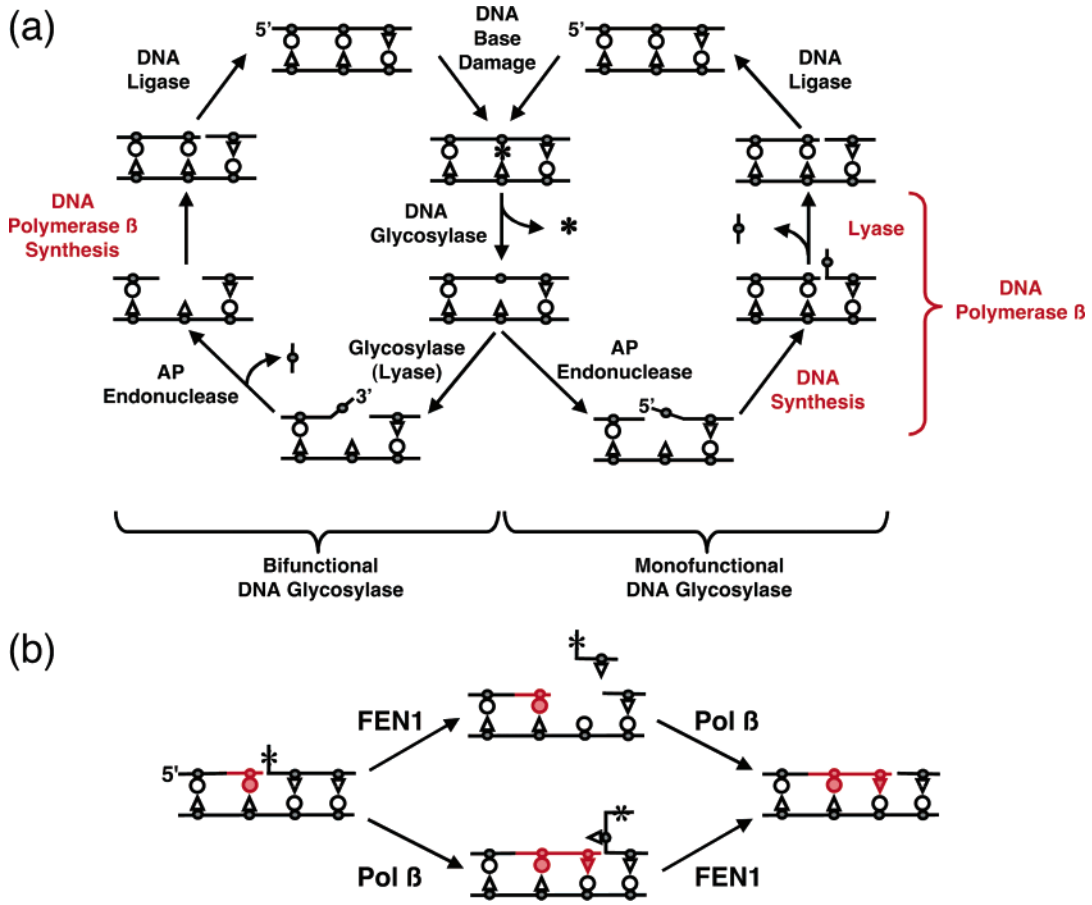


Figure 1. Alternate base excision repair pathways. (a) Single-nucleotide BER commonly occurs through either a monofunctional DNA glycosylase-initiated pathway (right loop) or a bifunctional DNA glycosylase-initiated pathway (left loop). DNA polymerase β contributes single-nucleotide gap-filling DNA synthesis in each pathway, but only its dRP lyase activity is required in the monofunctional DNA glycosylase path. (b) Two alternate routes are illustrated for pol β-dependent long-patch BER. These routes are initiated when the 5′-dRP group is modified (*) and serve as a way of removing the dRP group as part of a DNA fragment, thereby creating a 5′-phosphate required for ligation. After pol β fills a one-nucleotide gap, it can continue DNA synthesis (strand displacement, lower route), producing a flap to be removed by FEN1. In the upper route, FEN1 cleavage generates a one-nucleotide gap for another round of one-nucleotide gap filling.

lyase reaction mechanisms.[28] Upon the basis of the high level of sequence conservation of pol β among mammalian species, it seems likely that pol β conducts a role that is essential for animal survival.[29] The discovery of a pol β homologous gene in yeast through genomic sequencing and the finding that this gene encodes a protein that when overexpressed in *E. coli* has pol β-like enzymatic activity suggested that pol β homologues occur in eukaryotes from vertebrates to yeast.[30] The yeast enzyme is referred to as DNA polymerase IV and is now known to be homologous to newly identified members of the X-family of DNA polymerases.[31] Early inhibitor-based studies had implicated pol β in some types of mammalian DNA repair, such as the repair of bleomycin-induced damage.[32] Other studies also had implicated pol β in gap-filling DNA synthesis during BER of deaminated cytosine (i.e., uracil) in mammalian nuclear extracts[33,34] and in the repair of UV-damaged DNA[35] and abasic lesions[36] in *Xenopus laevis* oocyte extracts.

Although the inhibitor-based and other studies noted above implicated pol β in BER as a short gap-filling repair DNA polymerase, pol β's role in the cell had been widely considered as undocumented. A "knock-out" of the pol β gene in mice results in embryonic lethality, indicating an essential role of pol β during fetal development.[37] More importantly, the hypersensitivity of pol β null mouse embryonic fibroblasts toward monofunctional DNA-alkylating agents provides unequivocal evidence for the cellular role of pol β in BER.[38]

Mammalian pol β can be expressed in *E. coli* at high levels, hastening biophysical and kinetic characterization. The recombinant proteins from human and rat are fully active in DNA synthesis and possess substrate specificity and catalytic properties similar to those of the natural enzymes,[39,40] making them an excellent model for structure−function studies.

2. Structure

2.1. Biochemical and Genetic Domain Organization

Recombinant pol β's are monomeric in solution, like their natural analogues. DNA polymerase β was found to have an elongated structure in solution, with an axial ratio of about 5.[41] Controlled proteolytic or chemical cleavage of the native enzyme has demonstrated that pol β is folded into discrete domains and subdomains (Figure 2).[42,43] Studies of the purified domain fragments have provided important information on the relationship between particular structural elements of the enzyme and the various activities, such as single-stranded (ss) and double-stranded (ds) nucleic acid (NA) binding, nucleoside triphosphate (dNTP) binding, and the dRP lyase and nucleotidyl transferase catalytic activities.[43−45] The association of these properties with defined portions of

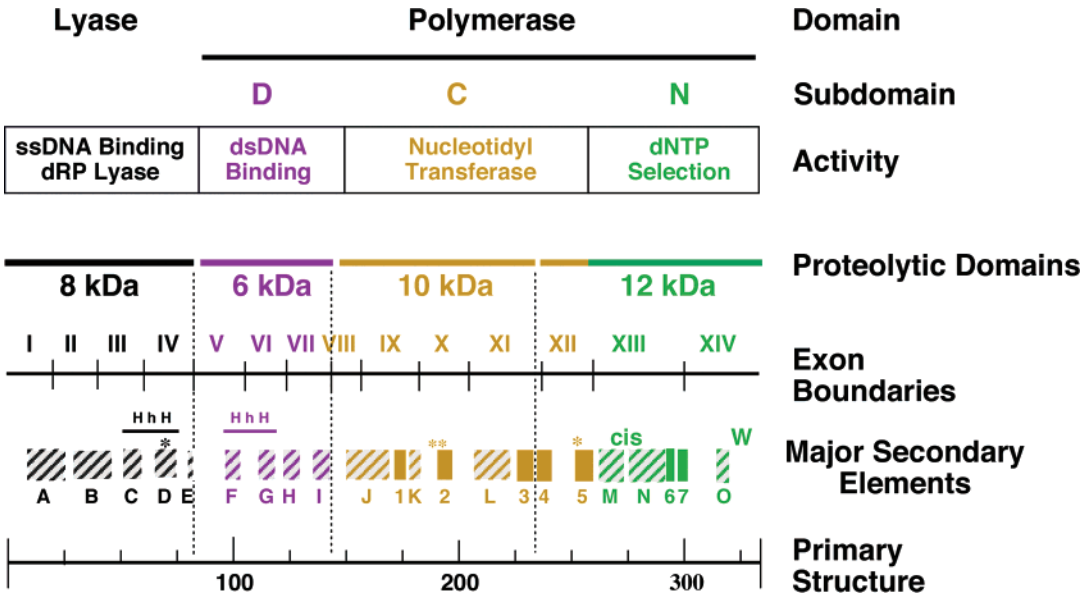


Figure 2. Relationship between the gene, primary, and protein structures of DNA polymerase β. The positions of α-helices (stripped boxes) and β-sheets (filled boxes) are indicated. HhH refers to the helix−hairpin−helix motif observed in several DNA repair proteins. The asterisks (*) represent critical active-site residues: Lys72 in the lyase domain and Asp190, Asp192, and Asp256 in the polymerase domain. The positions of the lone tryptophan residue (W) and rare *cis*-peptide (cis) bond are indicated.

the protein is consistent with the existence of tightly folded domains and subdomains linked by short solvent-exposed regions. Thus, the full-length enzyme consists of an amino-terminal lyase domain (8 kDa, ~90 residues), connected by a short protease-sensitive segment to a carboxyl-terminal polymerase domain (31 kDa, ~250 residues). Nucleic acid-binding studies have shown that the purified 8-kDa amino-terminal domain binds ssNA with an affinity almost the same as that of the intact enzyme. However, unlike the full-length protein, the 8-kDa domain has little if any dsNA-binding activity. By contrast, the 31-kDa carboxyl-terminal polymerase domain has little or no affinity for single-stranded lattices but binds dsDNA, such as the case for a template−primer. The 31-kDa domain possesses active-site residues necessary for the nucleotidyl transferase activity,[46] whereas the 8-kDa domain has dRP lyase activity (Figure 2).[9,10] The catalytic efficiencies of the isolated domains are reduced relative to that observed in the intact enzyme.[42,47]

The human gene encoding pol β spans 33 kb and contains 14 exons.[48] As expected, all the proteolytic domain boundaries occur between elements of secondary structure, with the boundary of the lyase domain and the 6-kDa proteolytic subdomain occurring at exon−intron junctions at the end of a secondary structural element (Figure 2). In contrast, the 10- and 12-kDa subdomains end and begin with exon XI, respectively. More than half of the 13 introns occur at the ends of secondary structural elements. However, a simple statistical analysis indicates that the introns do not appear to occur preferentially at the ends of structural elements.[49] This result is consistent with the idea that introns of human pol β may have been acquired in a random fashion.

2.2. X-ray Crystallography

2.2.1. Crystallographic Domain Organization

The first structure of a DNA polymerase, the *E. coli* DNA polymerase I Klenow fragment (A-family), likened its shape to a hand.[50] It was subsequently recognized that the polymerase domain of HIV-1 reverse transcriptase (RT-family) also resembled a right-hand, with fingers, thumb, and palm subdomains that could grasp DNA.[51] Crystallographic structures of DNA polymerases derived from several families and biological sources are now available and indicate that they have a modular domain organization with the polymerase domain composed of functionally distinct subdomains.[52,53] The palm subdomains of most DNA polymerases (A-, B-, Y-, and RT-families) show a similar topology and contain catalytic aspartates that coordinate the two metals involved in the nucleotidyl transferase reaction. In contrast, the fingers and thumb subdomains are structurally diverse among the different polymerase families. Crystallographic structures have been solved for pol β in a variety of liganded states and indicate that the polymerase domain is composed of three subdomains.[46,54,55] While the structure of the palm subdomain of pol β is similar to that of other DNA polymerases, its topology is unique.[54] This has resulted in a confusing subdomain nomenclature for pol β. This is because the fingers and thumb subdomains can be defined by a structural alignment of the palm subdomains from pol β and other polymerases or a functional alignment of the catalytic participants (metals, dNTP, DNA).[56] A functional alignment of pol β with other DNA polymerases of known structure results in a consistent functional nomenclature of the subdomains; the fingers and thumb subdomains of different DNA polymerases are functionally equivalent. In contrast, the nomenclature based on the original structural alignment of the palms defines the subdomains opposite to that of the functional alignment. Since the topology of the palm subdomain of X-family members is unique, one alternative is to consider members of the X-family as left-handed, rather than the original right-hand analogy.[27] This approach highlights the nonhomologous nature of the palm subdomains but requires prior knowledge of the architectural origin of the nomenclature. A functional alignment is the simplest approach, but the standard handlike analogy offers no functional insight. Accordingly, we employ a functionally based nomenclature where the subdomains are referred to as C- (catalytic), D- (duplex DNA binding), and N-

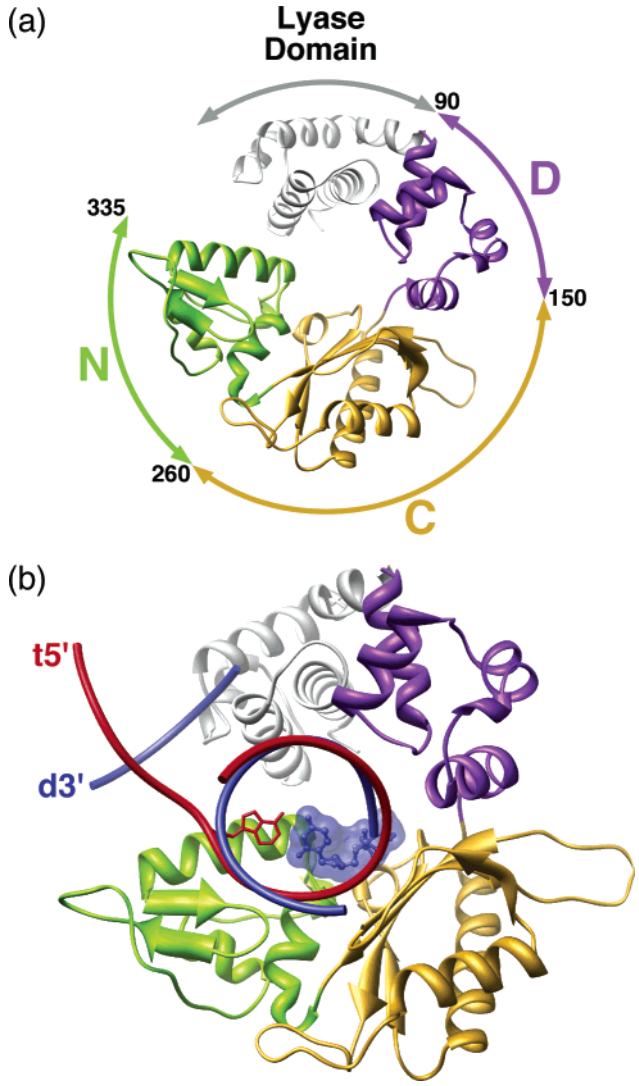


Figure 3. Domain and subdomain organization of DNA polymerase β. (a) DNA polymerase β is composed of a polymerase (colored) and an amino-terminal lyase domain (gray). The polymerase domain is composed of three subdomains: D- (purple), C- (gold), and N- (green) subdomains. These subdomains are involved in duplex DNA binding, nucleotidyl transfer, and binding of the nascent base pair, respectively. The lyase domain provides the enzymatic activity required to remove a 5′-dRP intermediate during BER. (b) A view down the upstream duplex DNA helix of pol β bound to one-nucleotide-gapped DNA. The template strand (red) is bent 90° as it exits the polymerase active site. At the 3′-end of the primer strand (blue) is an incoming dNTP (blue transparent surface). The 3′-end of the downstream strand and 5′-end of the template strand are indicated (d3′ and t5′, respectively). Molecular images were produced with Chimera[166] and MSMS.[167]

subdomains (nascent base pair binding) to highlight the intrinsic function.[57] These would correspond to the palm, thumb, and fingers subdomains, respectively, according to the nomenclature that utilizes the architectural analogy to a right-hand.[50]

While pol β appears to have evolved separately from members of other polymerase families, it shares many general structural and mechanistic features with polymerases from these other families (see below). Structures of pol β have identified molecular interactions that are involved in DNA binding and/or binding of the dNTP.[46,54,55,58−66] The structure of pol β illustrating the domain/subdomain organization is shown in Figure 3a. As summarized in Figure 2, the domains/subdomains exhibit a contiguous alignment with the D-, C-, and N-subdomains making up the carboxyl-terminal polymerase domain. The C-subdomain contributes three aspartates (190, 192, and 256) that coordinate two divalent metal cations, Mg^{2+}, that assist the nucleotidyl transferase reaction. The D- and N-subdomains are spatially situated on opposite sides of the catalytic subdomain. In addition to the polymerase domain, DNA polymerases generally have an accessory domain that contributes a complementary enzymatic activity necessary for the polymerase to fulfill its biological task. For pol β, the amino-terminal lyase domain contributes a biologically important dRP lyase activity that is required during monofunctional glycosylase-initiated single-nucleotide BER (Figure 1a).[67]

2.2.2. Liganded Complexes

All DNA polymerases require a single-stranded DNA template and divalent metal ions as cofactors while utilizing two types of substrate: a 2′-deoxynucleoside 5′-triphosphate and a primer annealed to the template (i.e., a template−primer DNA). DNA synthesis typically results in a new complementary strand obeying Watson−Crick base pairing rules of the single-stranded DNA template strand and the release of pyrophosphate. A molecular characterization of an enzyme requires structures representing different stages in the catalytic cycle. For pol β, such structures have provided structural insight into how pol β particularly, and DNA polymerases generally, achieves catalytic efficiency and fidelity. The first structures of pol β reported in 1994 were of the apoenzyme,[54] a dATP binary complex,[54] and a DNA/ddNTP ternary complex.[46] The binary dATP complex revealed that the binding pocket for the triphosphate-metal (Mn^{2+}) was near two conserved aspartates (190 and 192) in the C-subdomain (palm). Since the complex lacked DNA, the sugar and base were not expected to be in a catalytically relevant position and did not exhibit any protein interactions. Significantly, the ternary polymerase complex was the first reported structure that included both substrates (dNTP and DNA) and provided structural evidence for the DNA synthesis reaction. The novel approach used to trap this abortive ternary complex is now routinely used to trap ternary complexes of a variety of DNA polymerases for structure determination. The rationale for the method involved freezing the reaction at the insertion step by making use of dNTPs that lack the 3′-OH (i.e., ddNTP) in the crystallization solution. Polymerase-dependent insertion of a ddNMP into the primer results in a "blocked" primer terminus that permits binding of the next dNTP but impedes incorporation of the bound nucleotide. The structure of this pre-transition state is described in detail below.

The first ternary substrate complex was crystallized with a single-stranded template overhang of four nucleotides.[46] In this structure the amino-terminal lyase domain did not interact with the ssDNA template. A subsequent ternary complex structure of human pol β bound to a single-nucleotide-gapped DNA, a BER intermediate, indicated that the lyase domain bound to the downstream duplex.[55] The 5′-phosphate on the downstream strand was positioned near the putative dRP lyase active site that also displays a positive electrostatic surface potential.[27] More importantly, the amino-terminal lyase domain is now observed to interact with the N-subdomain, thereby producing a doughnut-shaped structure (Figures 3 and 4a). This conformation of the lyase domain is significant due to its proximity to the nucleotide-binding pocket and the fact that it can limit nucleotide access to the

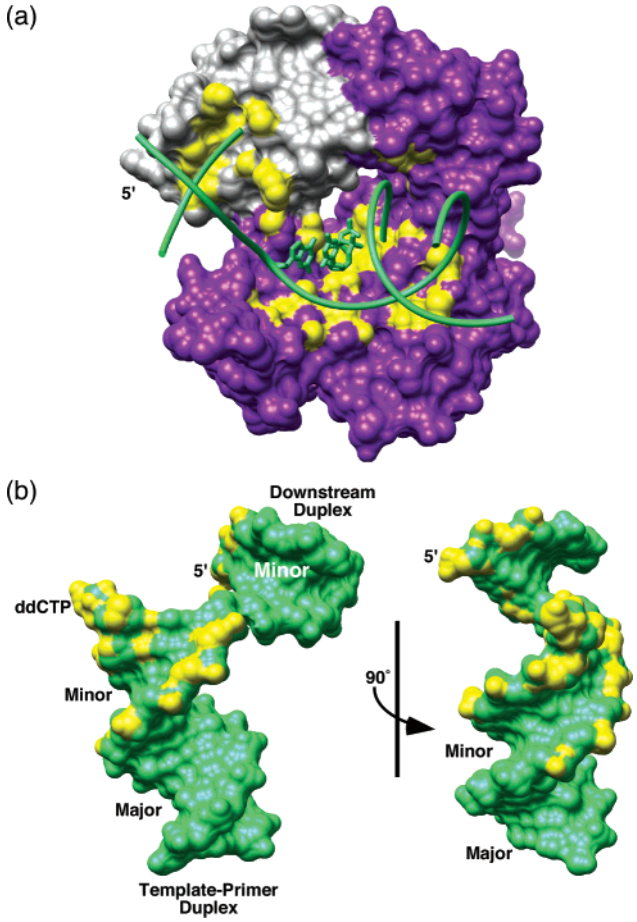


Figure 4. Molecular surfaces of DNA polymerase β and substrates.[55] (a) The solvent-excluded molecular surface of pol β is shown; lyase domain (gray) and the polymerase domain (purple). The surface of the atoms that are within 3.5 Å of the DNA or incoming ddCTP (green) is highlighted in yellow. The 5′-end of the template strand is indicated. (b) Two views of the molecular surface of the substrates from the ternary substrate complex. In addition to DNA sugar-phosphate backbone interactions (yellow), interactions are confined to the DNA minor groove. The 90° bend in the DNA exposes the terminal base pairs. His34 stacks with the first base pair of the downstream duplex, and α-helix N stacks with the nascent base pair (not shown). The 5′-end of the downstream strand is indicated.

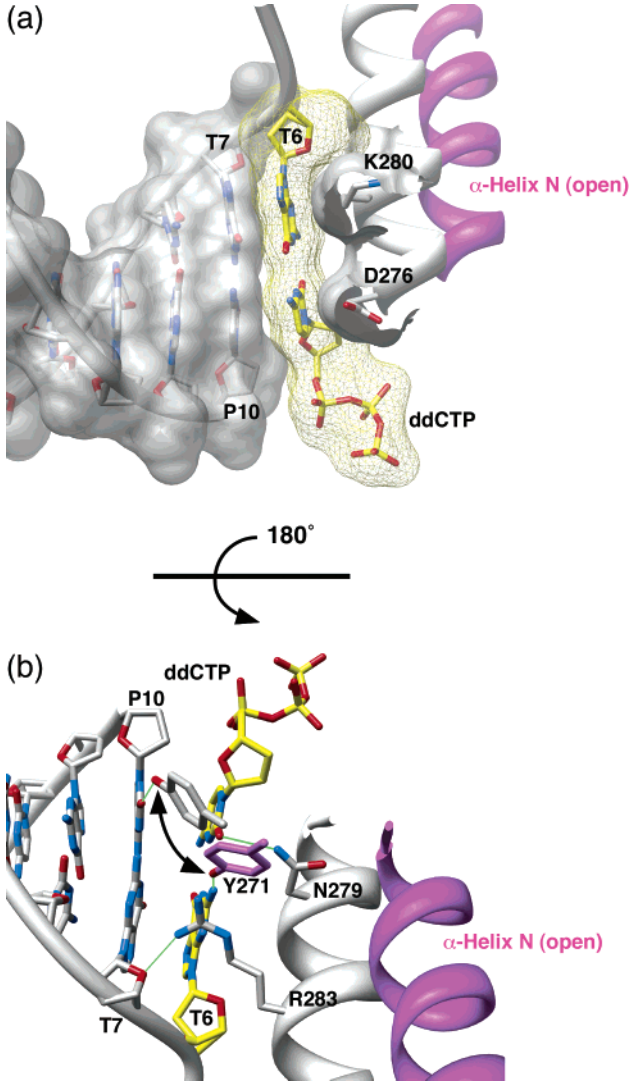


Figure 5. Open and closed conformations of α-helix N of the N-subdomain. (a) A view of the major groove edge of the nascent base pair (incoming and templating nucleotides, yellow). The semitransparent molecular surface of the upstream DNA duplex is shown (gray), and the molecular surface of the nascent base pair is also shown (mesh representation). In the closed conformation, Asp276 (D276) and Lys280 (K280) stack with the bases of the incoming and templating nucleotides, respectively. In the open conformation (magenta), these interactions are lost. (b) A view of the minor groove edge of the nascent base pair illustrating hydrogen-bonding interactions (green) with pol β. In the closed conformation, Tyr271 (Y271), Asn279 (N279), and Arg283 (R283) hydrogen bond with P10 (primer terminus), the incoming nucleotide (ddCTP), and T7 (template base for P10), respectively. In the open conformation, these interactions are lost but Tyr271 now hydrogen bonds with T6 (templating base).

polymerase active site.[49,68] Interactions between the lyase domain and the N-subdomain are also altered upon binding the correct incoming nucleotide.

Comparing the apoenzyme structure,[54] or DNA binary complex (substrate one-nucleotide-gapped DNA and product-nicked DNA),[55] with abortive ternary complex structures[46,55] indicated that significant conformational changes occur upon binding the correct dNTP and its subsequent insertion. The most notable conformational change occurs with the N-subdomain (residues 260−335). When the pol β DNA binary complex binds a correct dNTP, the N-subdomain repositions itself to "sandwich" the nascent base pair between the growing DNA terminus and the polymerase (Figure 5a).[28,53,55] The N-subdomain rotates 30° about α-helix M to achieve this conformational transition. The enzyme in the binary DNA complex is in an "open" conformation, whereas the enzyme in the ternary complex is in a "closed" form. In the latter case, there is a tight fit of the nascent base pair within a pocket formed by the polymerase and the DNA duplex terminus. The binary complex with nicked DNA suggests that the N-subdomain opens at some point after catalysis, such as PP_i dissociation. The various polymerase side chains proposed to activate the 3′-hydroxyl on the primer terminus and stabilize the transition state are not properly positioned in the open conformation. These observations suggest that the N-subdomain closes down around the incoming dNTP when the new base pair conforms to Watson−Crick geometry. Accompanying the motion of the N-subdomain are more subtle side-chains adjustments that allow the active site to monitor the state of the N-subdomain.[46,55,68] In the open inactive binary complex, one of the active-site metal ligands (Asp192) forms a salt bridge with Arg258 (Figure 6). However, closing of the N-subdomain results in repositioning of Phe272 so as to interfere with this interaction. Arg258 is

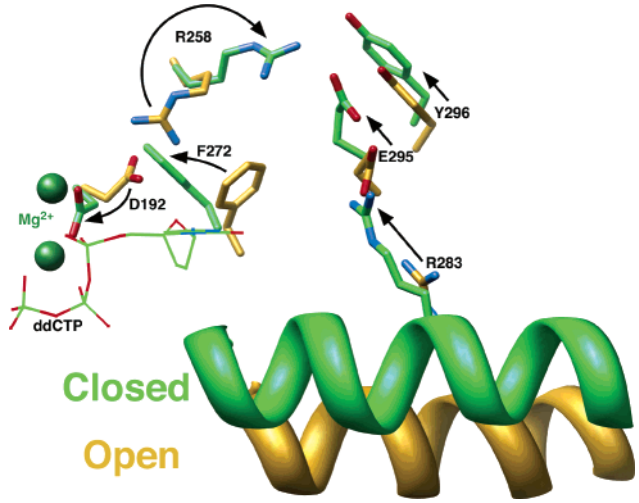


Figure 6. Active-site conformational changes associated with the open and closed conformational states of the N-subdomain. The position of the N-subdomain can be structurally transmitted to the catalytic metals (green) through altered interactions between Arg283 (R283), in α-helix N, and Asp192 (D192) that coordinates both active-site Mg^{2+} ions. This may occur through altered interaction observed between Glu295 (E295) and Arg258 (R258) in the open (gold, inactive) and closed (green, active) forms of pol β. Phe272 (F272) is postulated to transiently interfere with interactions between Asp192 and Arg258, permitting an interaction with Glu295. The incoming nucleotide (ddCTP) of the closed ternary substrate complex is illustrated with thin lines, and the templating base is omitted for clarity.

observed to interact with Glu295 and Tyr296 in the closed complex, thereby permitting Asp192 to contribute key oxygen ligands with the active-site metals.

The conformational adjustments are not limited to the protein. There are also subtle conformational adjustments in the DNA substrate. Since α-helix M is situated under the sugar of the incoming nucleotide, rotation in this helix results in significant structural changes in the vicinity of the template strand.[27,58] Specifically, the templating base and primer terminus are shifted toward the major groove in the absence of the incoming nucleotide. Additionally, the templating base 5′ (i.e., downstream) to the coding template base approaches the coding position in the binary DNA complex with a single-stranded template. This potentially represents a one-nucleotide deletion intermediate.[69] Repositioning of the N-subdomain upon correct dNTP binding is also observed for many other DNA polymerases.[53,70] In contrast, these subdomain motions originate near the template strand, resulting in significant conformational adjustments near the incoming nucleotide. In contrast, members of the Y-family of DNA polymerases[71] and pol λ[72] appear to utilize DNA conformational adjustments (e.g., template sliding), rather than subdomain motions, to open/close the polymerase active site.

The lyase domain and the D-subdomain are α-helical in nature. Each contains a structural motif that binds monovalent metals and interacts with the DNA backbone (Figure 7).[59,63,73] These helix−hairpin−helix (HhH) motifs (residues 55−79 and 92−118) have also been described in many proteins that bind either single- or double-stranded DNA in a non-sequence-specific manner.[74] The HhH motif apparently does not bind RNA, yet there are no protein groups to sterically exclude a 2′-hydroxyl. Instead, high-resolution structures of pol β suggest that binding specificity may be conferred by the ability to induce sugar puckers that are energetically unfavorable for RNA.[63] The two motifs are observed to interact with each end of the incised DNA strand, the downstream and primer strands. In the crystal structure of pol β bound to one-nucleotide-gapped DNA,[55] the DNA is bent 90° (Figures 3b and 4). Thus, the downstream duplex does not travel through the hole in the doughnut-like structure but instead interacts with the HhH motif in the lyase domain. The 90° bend is also observed in a binary complex with nicked DNA (i.e., product complex) and occurs in the 5′-phosphodiester backbone of the template (coding) base.[55] This positions the 3′-hydroxyl and the 5′-phosphate in the

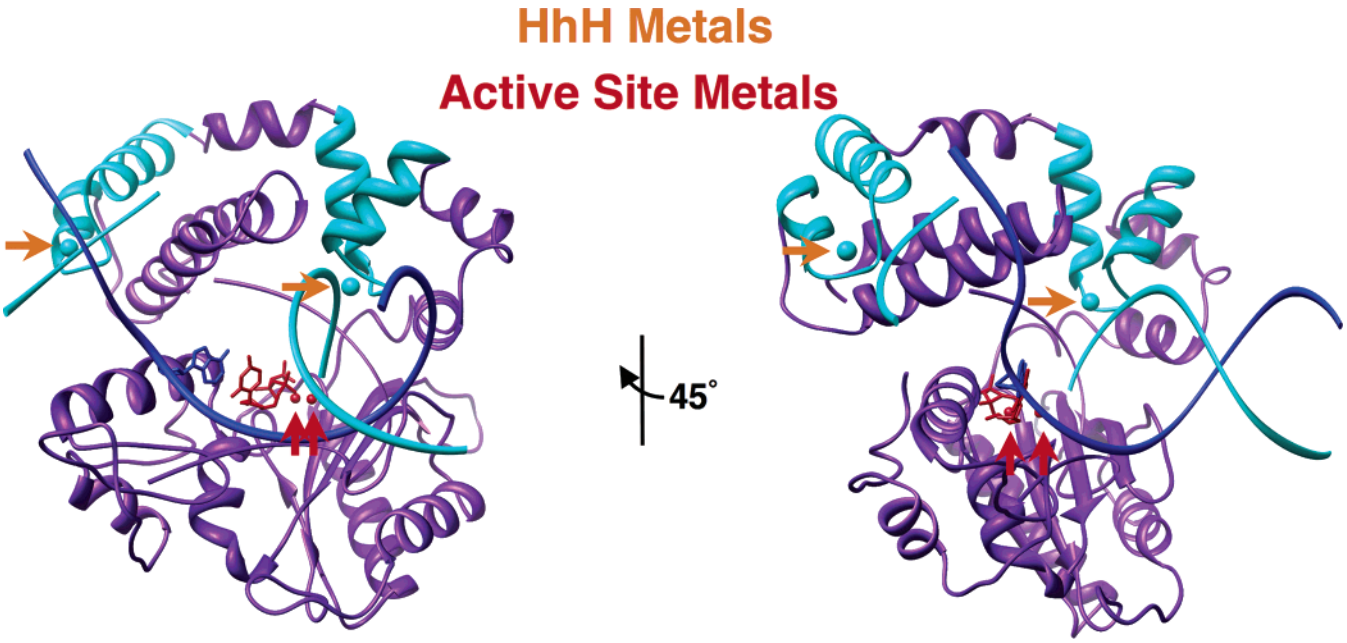


Figure 7. HhH motifs of DNA polymerase β. The two HhH motifs (light blue) bind to the DNA sugar-phosphate backbone of the incised DNA strand (light blue). The HhH of the lyase domain interacts with the downstream strand, and the HhH of the D-subdomain interacts with the upstream strand. These motifs bind monovalent ions[168] (light blue spheres highlighted with orange arrows) and stabilize the DNA substrates in their respective active sites. Also shown are the incoming dNTP (red) and the position of the polymerase active-site metals (red spheres).

nick over 27 Å apart. These ends must be ligated in the final step in BER. The function of the HhH motifs is to stabilize the bend and/or the incised DNA strand, facilitating proper positioning of the two free DNA ends. During single-nucleotide BER initiated by a monofunctional DNA glycosylase, each of these ends represents an enzymatic substrate, 3′-hydroxyl for DNA synthesis and 5′-dRP for removal to generate a 5′-phosphate.

Like other DNA polymerases, pol β primarily interacts with the DNA backbone (Figure 4b). The sharp bend in the DNA also serves to expose the nascent base pair (template base-incoming dNTP) (Figure 4b). This permits the polymerase access to probe whether a good Watson−Crick geometry is achieved and by monitoring minor groove hydrogen bonding (Figure 5). It has frequently been noted that the DNA minor groove features hydrogen bond acceptors, O2 of pyrimidines and N3 of purines, in similar minor groove positions that are independent of the base pair or sequence.[75] In the closed conformation, pol β interacts with the minor groove of the nascent base pair and the adjacent upstream base pair (primer terminus and template nucleotide) (Figure 4b). Specifically, Tyr271, Asn279, and Arg283 form hydrogen bonds in the DNA minor groove with the primer terminus, the incoming nucleotide, and the template base opposite the primer terminus, respectively (Figure 5b). In the open conformation of the binary DNA complex, these interactions are not observed, but Tyr271 is now hydrogen bonded to the orphaned templating base.

2.2.3. Pre-Transition-State Complex

DNA polymerases must select a complementary deoxynucleotide from a pool of structurally similar molecules to preserve the integrity of the genome. Failure to faithfully replicate DNA usually results in deleterious biological consequences. Differences in fidelities among polymerases are due to the divergent abilities of polymerases to insert the right nucleotide; low-fidelity enzymes insert the correct nucleotide slowly, whereas high-fidelity DNA polymerases insert the correct nucleotide rapidly.[53,76] Accordingly, an understanding of the molecular mechanisms that influence correct nucleotide insertion will provide clues to those interactions that modulate DNA replication fidelity.

An initial working model for the nucleotidyl transferase enzymatic mechanism gained structural support from pol β ternary substrate complexes.[46,55,62] These structures are consistent with a chemical mechanism that proceeds by an in-line nucleophilic attack of the Mg^{2+}-activated primer 3′-O^- on the α-phosphate of the incoming nucleotide, leading to a pentacoordinated bipyramidal α-phosphate transition state (Figure 8). The transition state is resolved by release of the pyrophosphate product from the α-phosphorus at the opposite side of the attacking 3′-O^-, resulting in stereochemical inversion about the α-phosphorus atom of the newly incorporated nucleotide. It is not clear whether the pyrophosphate separation occurs before (dissociative mechanism) or after (associative mechanism) nucleophilic attack or whether the two steps are "somewhat" concerted.[77] The C-subdomain of all DNA polymerases contributes two structurally conserved aspartate residues that coordinate two Mg^{2+} ions participating in the nucleotidyl transfer reaction.[56] The catalytic metal (metal A) lowers the pK_a of the 3′-OH of the growing primer terminus while the nucleotide binding metal (metal B) coordinates the triphosphate moiety, hastening binding of the incoming nucleotide. Additionally, metal B assists PP_i dissociation. Details on how the active-site residues actually contribute to the various steps in the chemical process are not understood.

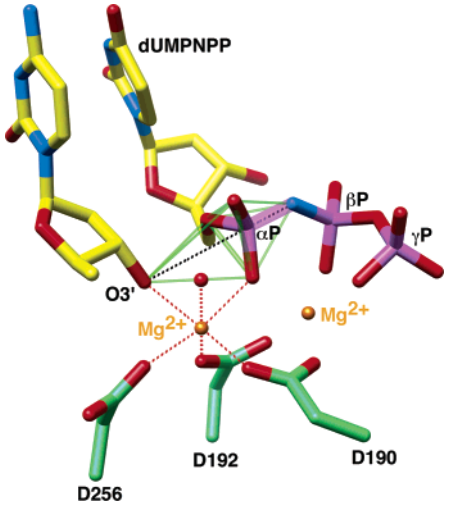


Figure 8. Pre-transition-state structure of DNA polymerase β.[66] The ternary substrate complex of pol β was trapped with an inert dUTP analogue (dUMPNPP). The developing trigonal-bipyramidal transition state is illustrated with green lines. The nucleotidyl transfer reaction involves an in-line nucleophilic substitution mechanism where the activated primer terminus (3′-O-) attacks the αP of the incoming nucleotide, resulting in dNMP transfer and pyrophosphate release. Since the primer terminus has a 3′-OH, the first coordination sphere of the catalytic Mg^{2+} is complete and exhibits good octahedral geometry (dotted red lines). The catalytic Mg^{2+} is also coordinated to all three active-site aspartates (D190, D192, D256). The active-site metals and residues interacting with the triphosphate moiety are expected to stabilize the pentacoordinated transition state and assist leaving of pyrophosphate. In this structure, O3′ of the primer terminus is 3.4 Å from αP of dUMPNPP.

As noted above, the ternary substrate complexes of pol β, as well as ternary substrate complexes of DNA polymerases from other polymerase families,[53] typically lack the 3′-hydroxyl on the primer terminus so as to abort catalysis. This results in a distorted catalytic metal site since this oxygen is believed to provide a coordinating ligand.[46] Additionally, the complete inner octahedral coordination sphere of the catalytic metal was not observed at the resolution of the ternary complex (1BPY, 2.2 Å).[55] A new high-resolution structure (1.65 Å) of pol β provides important details related to the coordination and identity of the catalytic metal.[66] A water molecule that had been suggested to participate in the octahedral coordination sphere of the catalytic Mg^{2+} is now clearly observed. The sixth coordinating ligand is proposed to be O3′ of the primer terminus that is missing in this structure. Additionally, the sugar pucker of the dideoxy-terminated primer is C3′-endo, similar to that observed in the well-defined structure of the ternary complex of the *Bacillus* DNA polymerase I fragment,[78] an A-family member, and a conformation commonly observed in A-form DNA.

The high-resolution structure also suggests that the ion occupying the catalytic metal site was Na^+ rather than Mg^{2+}. Since these ions have the same number of electrons, they cannot be distinguished by electron density. Thus, the identity of the ions is based on the average coordination distance for the coordinating ligands. In the high-resolution complex, the average coordination distance is 2.38 ± 0.06 Å. This distance is similar to that expected of Na^+ (2.42 Å)[79] rather than Mg^{2+} (2.07 Å),[80] observed in accurately determined small molecule

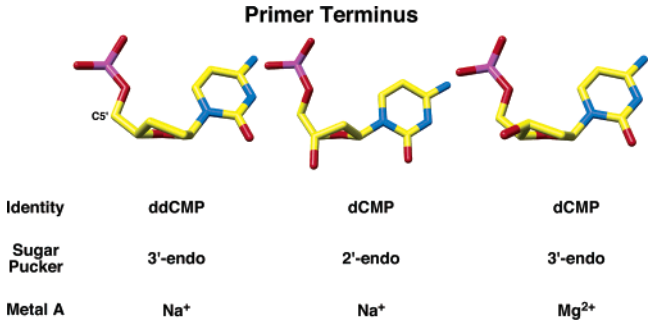


Figure 9. Factors influencing the sugar conformation of the primer terminus.[66] See text for description.

structures (Cambridge Structural Database). It should also be noted that other DNA polymerase ternary substrate complexes have a metal modeled into the catalytic metal site with average coordination distances much greater than 2.1 Å.[66]

By using a non-hydrolyzable dNTP analogue, a structure of a precatalytic complex of pol β that includes the catalytic Mg^{2+} and O3′ of the primer terminus has been determined.[66] This was achieved by using the dUTP analogue, 2′-deoxyuridine-5′-(α,β)-imido triphosphate (dUMPNPP). The bridging oxygen between the α- and β-phosphates is replaced with nitrogen, preventing insertion. DNA polymerases efficiently insert dUTP since it closely resembles dTTP and does not alter Watson−Crick hydrogen bonding. DNA polymerase β inserts dUTP with the same efficiency as dTTP and is competitively inhibited by dUMPNPP. The structure reveals that O3′ of the primer terminus completes the octahedral coordination sphere for the catalytic metal (Figure 8). Furthermore, the average coordination distance for this ion is consistent with its identity as Mg^{2+} (2.13 ± 0.03 Å). The sugar pucker of the primer terminus is also observed to respond to the identity of the ion occupying the catalytic metal site: C2′-endo with Na^+ and C3′-endo with Mg^{2+} (Figure 9). More importantly, the catalytic magnesium induces subtle conformational rearrangements to provide good octahedral geometry as well as position O3′ for an in-line nucleophilic attack on αP of the incoming nucleotide. It is now possible to confidently examine the precise sequence of events (deprotonation/protonation and bond making/breaking) and the free energy surface of the chemical reaction by computational methods.

2.2.4. Mutagenic Intermediates

Although the ability of a DNA polymerase to discriminate "right from wrong" nucleotides depends on the identity of the polymerase, all DNA polymerases insert incorrect nucleotides poorly.[76] Accordingly, strategies used by any specific DNA polymerase to discriminate against incorrect nucleotide insertion should be generally applicable to all polymerases.[53] One strategy that most DNA polymerases utilize to avert a misincorporation event is to bind incorrect dNTPs weakly. This has hampered structural characterizations of mismatched base pairs in a polymerase active site. Oligonucleotides with mispaired primer termini generally bind to a postinsertion site adjacent to the nascent base pair binding pocket.[81] In contrast, positioning a terminal mismatch in product-nicked DNA prevents translocation, thereby trapping the mispair in a conformation immediately after chemistry. This strategy has successfully been used to determine the crystallographic structure of two mismatches

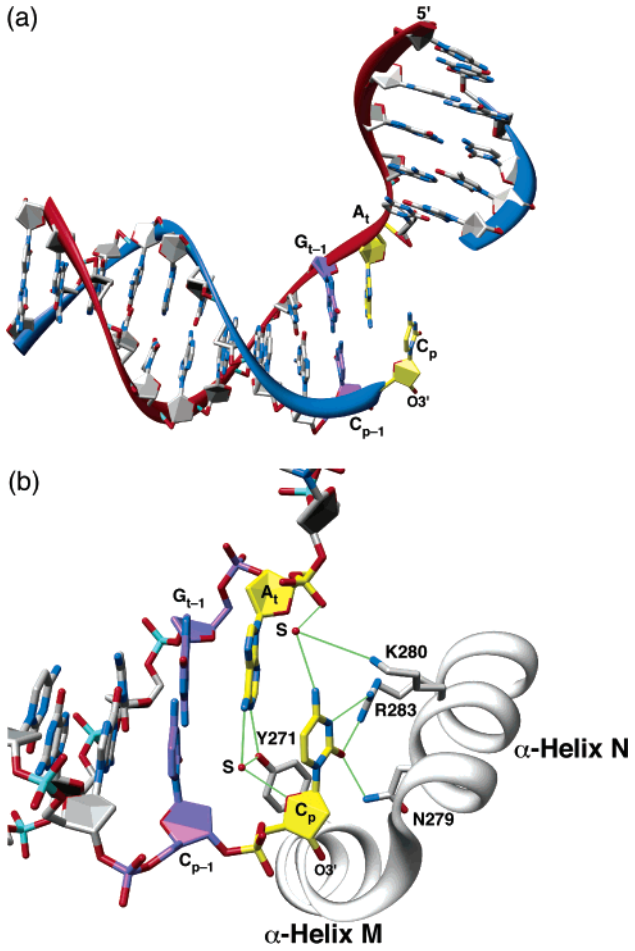


Figure 10. Structure of a mismatch in the confines of the pol β active site.[63] (a) Structure of the duplex illustrating the staggered conformation of the A−C mispair (yellow). The template strand is red, and the primer and downstream strands are blue. The adenine in the template strand (A_t) stacks with the upstream template base (G_{t-1}). The base pair that is positioned where the primer terminus would be poised for catalysis is purple. The mispaired cytosine at the primer terminus (C_t) stacks with the template adenine. The N-subdomain is in an intermediate conformation (between open and closed; not shown). (b) The hydrogen-bonding interactions (green) in the mismatch structure are significantly different than those observed in either a binary or ternary complex (see Figure 5b). For example, Arg283 (R283) that typically interacts with the template strand is observed to interact with the Watson−Crick edge of the terminal primer base that is situated in the incoming nucleotide-binding site.

(template−primer, A−C or T−C) in the confines of the pol β active site.[63]

The structure of each mispair complex indicates that the template and "incoming" nucleotide bases do not form hydrogen bonds with one another but instead form a staggered conformation where the bases of the mispair partially overlap (Figure 10a). Thus, the conformation of the mispair is dramatically different than that expected from the structures of these mispairs in duplex DNA.[82−85] The staggered conformation appears to prevent closure of the N-subdomain that is believed to be required for catalytic activation. The N-subdomain is observed in an intermediate position between a fully opened or closed state. The partially open conformation of the N-subdomain results in distinct hydrogen bonding networks that are unique for each mispair. The hydrogen-bonding pattern for the A−C mispair is illustrated in Figure 10b. The Watson−Crick edge of the

surrogate incoming adenine is hydrogen bonded to Arg283, a residue that interacts with the template strand in the closed conformation with a correct incoming nucleotide (Figure 5b). In addition, the staggered conformation of the primer terminus results in a loss of stacking interactions with the upstream ($n - 1$) base pair.

A feature of catalytic regulation, deduced from comparison of the open and closed conformations, appears to be the insertion of Phe272 between the side chains of Asp192 and Arg258, enabling Asp192 to coordinate the two active-site Mg^{2+} atoms, thus promoting both substrate binding and catalysis (Figure 6).[55,68,86] This positioning of Phe272 to activate catalysis is coupled to movement of the N-subdomain, suggesting a mechanism by which subdomain closure could be linked to catalytic activation. Although there is Phe272 backbone movement coupled with the partial closing of the N-subdomain, complete closure and catalytic activation also requires a different Phe272 side-chain rotomer. In all pol β structures, the adjacent Tyr271 is in van der Waals contact with Phe272. Tyr271 also is in contact with DNA in all three pol β conformations (open, intermediate, and closed), although specific interactions vary significantly among the three structures. Consequently, Tyr271 appears to function as a direct monitor of the state of bound DNA, with hydrogen bonds formed to the free hydrogen-bonding groups in the minor groove of the DNA substrate (Figure 5b). The Tyr271 hydroxyl group can function as either a hydrogen bond donor or acceptor, providing versatility to hydrogen-bonding capabilities with any standard template base. Given these observations, it appears that closure of the N-subdomain is necessary, but not sufficient, for catalytic activation, with the activated position specified directly by coupled movements of Tyr271 and Phe272.

As noted above, polymerases generally bind incorrect nucleotides with low affinity, indicating that the molecular interactions between the polymerase and a wrong incoming nucleotide are compromised. In the A−C structure, many of the hydrogen-bonding interactions of the incoming nucleotide observed in the closed polymerase complex are retained in the mismatched structures, albeit with different partners, suggesting that the loss of binding affinity is primarily due to the loss of polymerase and duplex DNA stacking interactions with the nascent base pair. In addition, the partially closed conformation reduces interactions with the template strand and shifts the primer terminus away from the C-subdomain, protecting it from nucleotide incorporation.

Although DNA replication and repair synthesis attempts to preserve Watson−Crick base-pairing rules, an incorrect dNTP is occasionally inserted. The efficiency at which this occurs is only weakly dependent on the identity of the polymerase.[76] 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 (Figure 11). 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. During BER, AP endonuclease has been suggested to provide this activity.[87] In some instances, however, the DNA mismatch is extended, resulting in a mutagenic event (base substitution error, Figure 11).[88] As described above, DNA mismatches in the confines of the pol β active site indicate that the mispaired bases are in a

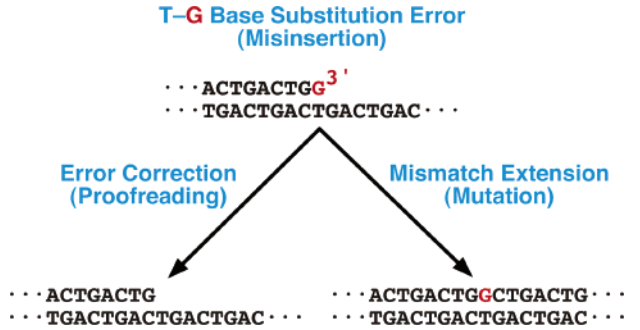


Figure 11. Fates of a DNA synthesis base substitution error. A mismatched primer terminus represents an obstacle for further DNA synthesis. This provides the opportunity to be proofread by either an intrinsic or extrinsic 3′-exonuclease. However, in some instances the mismatch will be extended, thereby generating a mutation.

staggered conformation and do not form hydrogen bonds with one another.[62,63] 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.[81] In most instances, the hydrogen bonding pattern is similar to that observed from crystal structures of mismatches in duplex DNA in the absence of protein.

The polymerase activity of pol β is necessary for alternate repair pathways that require longer gap-filling DNA synthesis (long-patch BER).[17,89] This alternate pathway is needed to remove modified 5′-dRP residues that may arise during DNA damage and not substrates for the dRP lyase activity of pol β. Long-patch BER can be accomplished through the sequential activities of FEN1 cleavage that generates a one-nucleotide gap and pol β gap-filling DNA synthesis.[19] These alternating activities can result in a DNA repair patch of 2−11 nucleotides. Consequently, long-patch BER can occur through a series of one-nucleotide gap-filling reactions or strand displacement DNA synthesis (Figure 1b). 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.[90−92] Generally, pol β creates and extends purine−pyrimidine mispairs (transition intermediates) more easily than pyrimidine−pyrimidine or purine−purine mispairs (transversion intermediates).[76,92] A mispaired primer terminus lowers the catalytic efficiency for correct insertion. This is usually due to a decreased insertion rate and binding affinity for the incoming correct dNTP. In contrast, the DNA binding affinity is not generally affected by a terminal mismatch.[92,93] The low binding affinity of the incoming dNTP has hampered structural characterization of relevant mutagenic intermediates.[63] 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 permitted binding of the following nucleotide with high affinity.[92]

The structure of a binary DNA complex with an A−A mismatch at the primer terminus indicates that the mismatch is not planar but in a staggered conformation with the adenines stacking with one another (Figure 12a).[64] 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 the coding templating base (i.e., position n). The authentic coding templating nucleotide (deoxycytidine) is

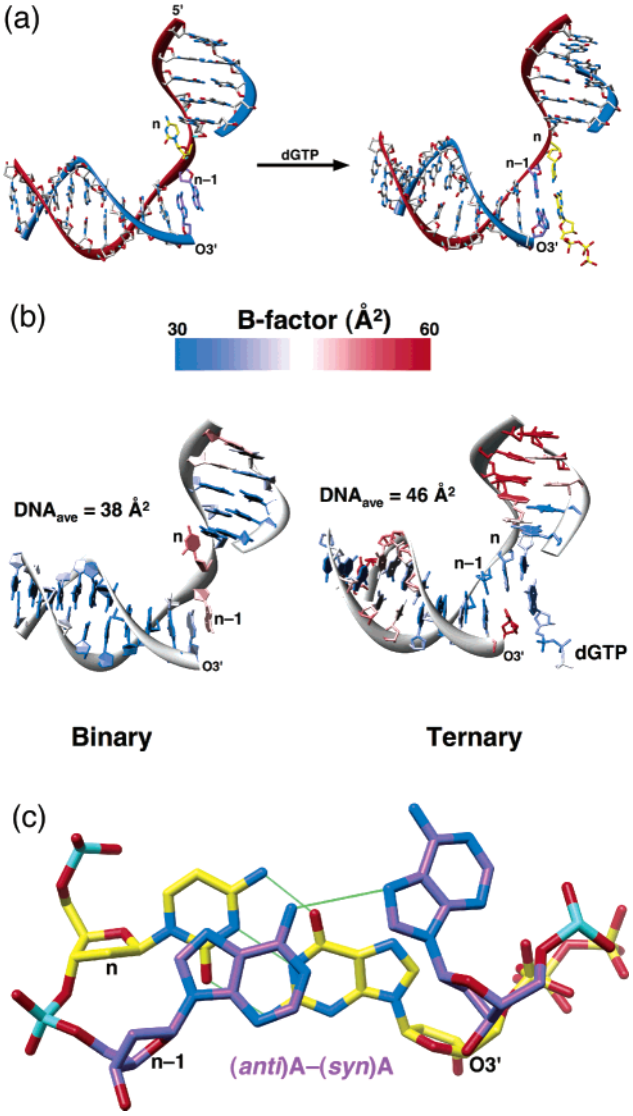


Figure 12. Structure of an A−A mismatch at the boundary of the nascent base pair binding pocket of the pol β active site.[64] (a) The polymerase is omitted for clarity. The structure of the duplex in a binary DNA complex is illustrated on the left. The A−A mismatch at the primer terminus (purple) is in a staggered conformation. The template strand is red, and the primer and downstream strands are blue. The adenine in the template strand (A_{t-1}) stacks with the primer terminal adenine (O3′) and occludes the templating base pocket. This forces the templating cytosine (n, yellow) to flip outside of the helix axis. Soaking dGTP into crystals of the binary complex resulted in a ternary complex and produced dramatic conformational changes in the DNA near the active site. The primer terminal adenine has moved to a *syn*-conformation (panel c), and the templating cytosine has reclaimed its binding pocket to pair with the incoming dGTP. (b) Mapping the respective structural B-factors of the DNA nucleotides illustrates the volatile positions of the template cytosine (n) in the binary complex and the primer terminus in the ternary complex (red, high B-factor). Note that the template base is stabilized (blue, low B-factor) in the ternary substrate complex.

extrahelical, and the template strand is shifted toward the dNTP-binding site. In contrast to the staggered conformation observed with mismatches in the active site where the templating base stacks with the upstream duplex (Figure 10), in this case, the primer terminus stacks with the upstream duplex in the binary complex (Figure 12a). As with other pol β DNA binary complexes, the N-subdomain is in the inactive open position.

Soaking crystals of the binary DNA complex with the nucleoside triphosphate (dGTP) complementary to the authentic templating base resulted in a ternary substrate complex (Figure 12a). In this situation, the N-subdomain has repositioned itself to form one face of the nascent base pair binding pocket (i.e., formed a closed conformation). In contrast with earlier ternary complex structures with a matched dideoxy-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.[92] 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 closing of the N-subdomain inferred to occur upon binding the correct dGTP, 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. To make room for the templating strand adenine at $n - 1$, the adenine at the primer terminus has rotated into a *syn*-conformation (Figure 12a and c), 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.

These structures suggest that the closed polymerase conformation is permitted only when a base pair that conforms to a good planar geometry is situated in the nascent base pair binding pocket and that this leads to tight binding of the incoming nucleotide. It is expected that proper positioning of the template base is crucial for faithful and efficient DNA synthesis.[28] An example of this principle can be seen by examining the B-factor for the respective nucleotides in the binary and ternary complexes with the A−A mismatch (Figure 12b). The templating cytosine is less well resolved (high B-factor, red) in the binary complex where it is extrahelical; in contrast, when paired with the incoming dGTP, the templating cytosine is better resolved (blue). Likewise, the template adenine that does not stack with duplex DNA in the binary complex has a higher B-factor than that when it stacks between duplex DNA and the stabilized templating base. However, the adenine at the primer terminus, which has moved to a *syn*-conformation in the ternary complex, is now in a volatile position owing to the loss of stacking interactions with the template adenine. These structures highlight the dynamic nature of DNA in the polymerase active site and the importance of proper template position. Indeed, binary DNA complexes of A-family DNA polymerases indicate that these polymerases replace the template (coding) base with an aromatic side chain, thereby forcing the unpaired base outside the helical axis.[27] As observed here with the A−A mismatch, binding of the correct nucleotide repositions the coding base inside the helix axis, thereby forming Watson−Crick hydrogen bonds with the incoming nucleotide.

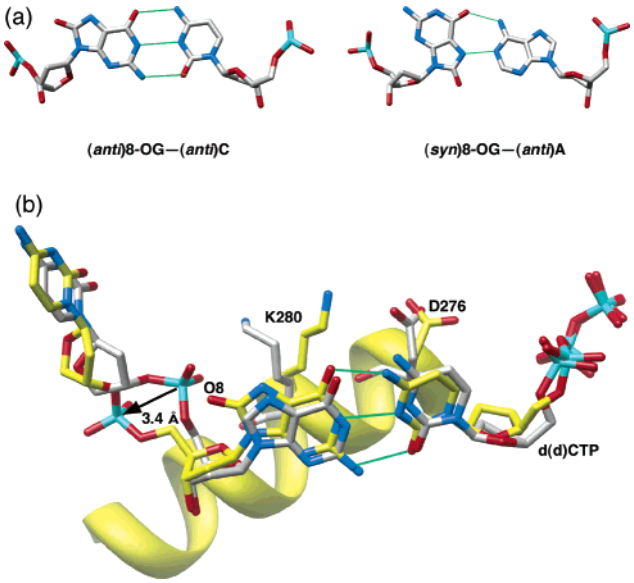


Figure 13. Ternary complex structure of pol β with a templating 8-oxodG.[62] (a) Base-pairing properties of 8-oxodG paired with cytosine or adenine. (b) Comparison of the 5′-phosphate backbone conformation of 8-oxodG relative to that observed in the structure of pol β with an unmodified deoxyguanine in the polymerase active site (gray). The presence of the sharp bend along with limited enzyme contacts with this phosphate enables flipping of the phosphate away from the carbonyl oxygen at C8 of 8-oxodG.[55] The backbone torsion angle of the templating guanine is altered ~180° when a carbonyl group is introduced at C8. In addition, the Lys280 (K280) side-chain position, but not that of Asp276 (D276), is altered in the presence of 8-oxodG.

2.2.5. Complexes with DNA Lesions

Oxidative damage to DNA generates 8-oxo-7,8-dihydro-2′-deoxyguanosine (8-oxodG). Whereas the unmodified deoxyguanine glycosidic torsion angle preference is *anti*, 8-oxodG favors a *syn*-conformation that can form a Hoogsteen base pair with adenine (Figure 13a). During DNA replication and repair synthesis, 8-oxodG can pair with cytosine or adenine, and the ability to insert the nonmutagenic C residue opposite this lesion depends on the DNA polymerase. DNA polymerase β exhibits a preference to insert dCMP, rather than dAMP, opposite 8-oxodG,[94] but this depends on the DNA sequence context.[95] The structure of pol β with a promutagenic DNA lesion, 8-oxodG, in the confines of its active site has been reported.[62] The oxidized guanine residue is in the *anti*-conformation and forms usual Watson−Crick hydrogen bonds with the incoming dCTP (Figure 13b). To accommodate the oxygen at the C8 position, the 5′-phosphate backbone of the templating nucleotide flips 180°, compared with its position in structures with normal DNA. Thus, the propensity for repositioning of the template sugar-phosphate backbone near the polymerase active site is one parameter that influences the *anti*−*syn* equilibrium of 8-oxodG and the success of the enzyme in error-free bypass of the lesion.

Although pol β preferentially inserts dCMP opposite 8-oxodG, insertion of dAMP also occurs.[94] Therefore, the structure of a ternary complex with ATP as the incoming nucleotide opposite 8-oxodG in the templating position was also assessed.[62] It was expected that the templating 8-oxodG would assume a *syn*-conformation paired with adenine. Instead, the templating 8-oxodG assumed an *anti*-conformation and was in a staggered conformation with dATP, and the N-subdomain was in the "open" conformation. The geometry of the staggered nascent base pair is similar to that observed in pol β complexes with mispairs in the active site, indicating that pol β recognizes the 8-oxodG-A combination as a mispair.

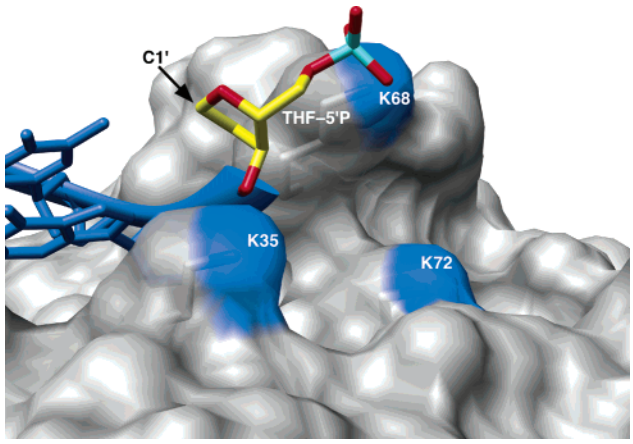


Figure 14. Closed sugar analogue bound to the lyase domain of pol β.[65] The molecular surface of the lyase domain active site indicates that the 5′-phosphorylated tetrahydrofuran (THF-5′P) is bound in a noncatalytic position since the Schiff base nucleophile (NZ of Lys72, K72)[10,96] is over 10 Å from the sugar C1′. The positions of Lys35 (K35) and Lys68 (K68) are also shown.

The chemical nature of the 5′-deoxyribose BER intermediate represents an important branch point for alternate BER pathways. A normal 5′-dRP group serves as a substrate for the lyase activity of pol β during monofunctional glycosylase-initiated single-nucleotide BER (Figure 1a). The dRP lyase active site is found in the amino-terminal 8-kDa domain.[9] This reaction proceeds via β-elimination, as evidenced by the trapping of a covalent Schiff base intermediate between the polymerase and dRP-containing DNA strand.[10] Mass spectrometric analysis of the 8-kDa domain covalently trapped with dRP-containing DNA indicated that NZ of Lys72 is the sole Schiff base nucleophile[96] consistent with site-directed mutagenesis.[97] If the dRP group is modified, however, the modified sugar is removed as part of a flap with the aid of FEN1 (Figure 1b). Tetrahydrofuran (THF) is a closed ribose ring that is not a substrate for the lyase reaction. A crystallographic structure of pol β bound to a 5′-phosphorylated THF residue in nicked DNA has been reported.[65] The global structure of pol β is similar to that described above for the binary DNA complex (i.e., open conformation). The density for the 5′-phosphorylated dRP group is less well defined than that for the surrounding structure, indicating that it is loosely bound in this position (Figure 14). The dRP group is bound in a lysine-rich pocket but, surprisingly, is distant from the Schiff base nucleophile. The distance between the reactive atoms for Schiff base formation, C1′/dRP and NZ/Lys72, is 10.1 Å, indicating that the analogue is in a noncatalytic position. Additionally, other nearby lysine residues are also too distant from C1′ to substitute as a nucleophile. Interestingly, a simple 120° rotation about the 3′-phosphate repositions the dRP group in a position near NZ of Lys72. Since this analogue is not a substrate, it cannot be concluded whether the natural substrate would favor the noncatalytic binding mode. In any case, the slow removal of the dRP group represents a key rate-limiting step for single-nucleotide BER[6] and influences the cytotoxicity of simple alkylating agents.[67] Finally, it remains to be determined whether the noncatalytic binding mode may directly influence downstream events such as the long-patch

BER sub-pathway choice: pol β-independent or pol β-dependent (strand displacement or gap-filling) pathways.

2.3. NMR Characterization

The ability to isolate domains and/or subdomains of pol β has provided the opportunity to characterize the solution structure of these smaller fragments by NMR spectroscopy. The 8-kDa lyase domain has been particularly amenable to this approach. The backbone dynamics and the refined solution structure of the amino-terminal lyase domain were characterized in order to examine potential contribution(s) of backbone motion to DNA binding and dRP lyase function of this domain.[98] NMR characterization of the binding site(s) for a ssDNA-5mer, ssDNA-8mer, ssDNA-9mer, and dsDNA-12mer revealed a "consensus surface" for the binding of these various DNA oligomers; the surface surrounds and includes the dRP lyase active-site region. The HhH motif in the lyase domain (Figures 2 and 7) displays a significant degree of picosecond motions within the peptide backbone that is probably important for phosphodiester backbone DNA binding. A second HhH motif is also found in the D-subdomain of the polymerase domain (Figures 2 and 7). These nonspecific DNA binding motifs are generally found in DNA binding proteins.[74] An Ω-loop connecting α-helices 1 and 2 and helix-2 itself display significant exchange contributions at the backbone amides due to apparent conformational type motions on a millisecond time scale. This motion is likely important in allowing the Ω-loop and helix-2 to shift toward, and productively interact with, gapped DNA substrates. The dRP lyase catalytic residues, including Lys72 that forms the Schiff's base, Tyr39 that is postulated to promote proton transfer to the C1′ aldehyde, and Lys35 that is postulated to assist in phosphate elimination, all show little or no backbone motion. Thus, these active-site residues appear to be in a conformationally rigid platform.

XRCC1 (X-ray cross-complementing group 1) is a DNA repair protein that forms complexes with pol β, DNA ligase III, and poly(ADP-ribose) polymerase in the repair of DNA single-strand breaks.[99] The domain organization of XRCC1 has been mapped, and characterization of the respective domains involved in the interaction of the XRCC1 and pol β complex has provided information on the specificity and mechanism of binding.[100] The domain structure of XRCC1, determined using limited proteolysis, was found to include an amino-terminal domain (NTD), a central BRCT-I (breast cancer susceptibility protein-1) domain, and a carboxyl-terminal BRCT-II domain. Using cross-linking experiments, XRCC1 was found to bind intact pol β and the 31-kDa polymerase domain. The XRCC1-NTD (residues 1−157) was found to bind the 22-kDa pol β fragment comprising the C- and N-subdomains (residues 140−335). Most of the backbone, C_α, and C_β assignments of the 22-kDa domain were determined.[101] Most recently, residues of pol β that interact with XRCC1 were determined by NMR chemical shift mapping (CSM) and mutagenesis.[100] A $^{15}N/^{13}C/^{2}H/^{1}H,^{13}C$-methyl[Leu,Ile,Val]-labeled pol β 22-kDa fragment was used for assignments of the ^1H, ^{15}N, and ^{13}C resonances by CSM upon forming a complex with the XRCC1-NTD. Large chemical shift changes were observed in the N-subdomain with complex formation. ^{15}N relaxation data indicate reduction in high-frequency motion for an N-subdomain loop and three C-subdomain turn/loops, which showed concomitant chemical shift changes with complex formation. Deletion of residues Val303−Val306 or a triple mutant (L301R/V303R/V306R) abolished the NTD interaction. It remains to be determined how this interaction influences the efficiency of BER or pol β-dependent activities, but progress in this direction has been made recently.[102,103]

In recent years, there has been considerable effort toward isolating and identifying natural product inhibitors of mammalian DNA polymerases. The interaction of bile acid and long-chain fatty acid pol β inhibitors with the 8-kDa lyase domain has been characterized by NMR.[104,105] The binding site for the known natural product inhibitor koetjapic acid (KJA)[106] was identified by CSM to the lyase domain.[107] Specifically, CSM suggests the binding site includes Gly66 and Ile69 that are part of the HhH motif (Figure 2) that interacts with downstream duplex DNA (Figure 7). Two structurally similar triterpenoid compounds and several other synthetic compounds containing aromatic or other hydrophobic groups in combination with two carboxylate groups were found to bind to the same or similar sites on the lyase domain. The pol β residues that interact with the inhibitor include Lys35, Gly66, Ile69, Arg83, and Leu85. These residues overlap the DNA binding and lyase active sites. In addition, Lys35 is also known to be involved in both ssDNA binding and 5′-phosphate recognition in gapped DNA.[55,97] The ability of these compounds to potentiate methyl methanesulfonate (MMS) cytotoxicity, an indicator of cellular BER capacity, in wild-type and pol β null mouse fibroblasts identified pamoic acid as the most active. *In vitro* analysis confirmed that pamoic acid could inhibit both DNA synthesis and dRP lyase activities on a BER substrate.[107] These results indicate that NMR-based mapping techniques can be used in the design of small molecule enzyme inhibitors.

3. Catalytic Mechanism

3.1. Kinetic Mechanism of DNA Synthesis

DNA polymerases utilize a similar kinetic mechanism (Figure 15a). As described below, the magnitude of each step is dependent on the specific polymerase as well as the DNA sequence. Steady-state kinetic analyses indicate that pol β follows an ordered addition of substrates.[108] After binding a DNA substrate, DNA polymerases preferably bind a nucleoside triphosphate that preserves Watson−Crick hydrogen bonding as dictated by the template base (step 1). Upon binding the correct dNTP, the ternary complex undergoes numerous conformational changes (step 2). As described above, these include subdomain motions, protein side-chain repositioning, and DNA structural alterations. In some instances, these conformational changes limit the rate of nucleotide insertion so that the chemical step (step 3) is not rate-limiting.[109,110] The identity of the rate-limiting conformational change is not known, but experimental evidence strongly suggests that subdomain motions are too fast to limit chemistry.[68,111] Following chemistry, the ternary product complex undergoes a conformational change (step 4) that facilitates PP_i release (setp 5). At this point, the extended product (DNA_{+1}) is released (single-nucleotide insertion) or serves as substrate DNA for another round of insertion (processive DNA synthesis).

3.1.1. Kinetic Characteristics of DNA Polymerase β

There are three standard kinetic approaches used to characterize the kinetic steps of pol β: (i) steady-state

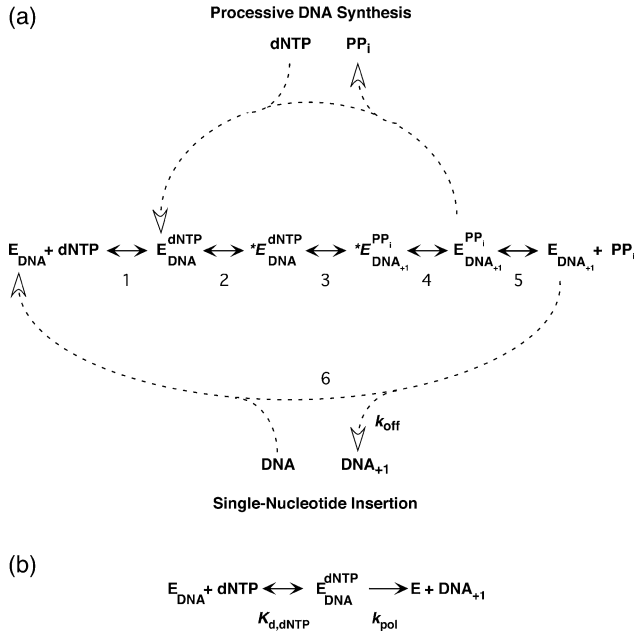


Figure 15. General kinetic pathway for DNA synthesis. (a) Complete kinetic scheme governing the kinetics of nucleotide insertion illustrating an ordered binding of substrates. Additionally, this scheme includes two central ternary complexes that have undergone a conformational change (*). For pol β, there are several conformational changes that occur upon ligand binding, but whether they are kinetically significant is unknown. The scheme also illustrates the continued synthesis of DNA without the enzyme dissociating from the DNA (processive DNA synthesis, upper loop) and catalytic cycling where the polymerase dissociates after each insertion event (single-nucleotide insertion, lower loop). (b) Simplified kinetic scheme operative for all DNA polymerases under single-turnover conditions. See text for discussion.

kinetics, (ii) pre-steady-state kinetics, and (iii) single-turnover kinetics. These approaches differ from one another by the ratio of polymerase to DNA as well as the concentration of enzyme. Traditionally, a steady-state approach employs a low concentration of enzyme relative to substrate DNA (E $\ll$ DNA). In contrast, a pre-steady approach uses a high concentration of enzyme to follow the formation and decay of intermediates (high E, E < DNA). Finally, a single-turnover analysis employs a high concentration of enzyme so as to follow catalytic events at the active site without interference from catalytic cycling (high E, E $\gg$ DNA) (Figure 15b). These later two approaches often require a rapid-mixing and quenching instrument.

Using defined template−primer oligonucleotides, it is possible to measure the kinetics of nucleotide insertion by all three techniques. For a well-defined mechanism, the results of all three techniques should be self-consistent (i.e., be predictive of the results from other approaches). The k_{cat} and K_m values ascertained from a steady-state approach must be interpreted in the context of a defined mechanism.[112] Only in this way can these values be interpreted reliably. For a polymerase that binds DNA tightly (slow k_{off}) and inserts a nucleotide rapidly (k_3 or $k_4 \gg k_{off}$), a burst of product formation occurs during the first turnover.[113] Subsequent turnovers occur at a steady-state rate that is equivalent to k_{off}. The burst phase is referred to as the pre-steady-state. The amplitude of the burst is equivalent to the active enzyme fraction, and the rate of the burst can be described by a single exponential that will saturate at high dNTP concentrations. The maximal rate of this phase is defined as k_{pol} (step 3 or 4), and the hyperbolic dNTP concentration dependence defines the $K_{d,dNTP}$ (step 2).

A pre-steady-state analysis of pol β-dependent single-nucleotide gap filling demonstrates that the pre-steady-state and steady-state phases of the reaction are not well-separated.[68] This is due to the similar magnitudes of k_{pol} and $k_{off,DNA}$, resulting in partially rate-limiting steps during the course of the reaction. An alternate approach to define k_{pol} and K_d utilizes the single-turnover experiment so that catalytic cycling does not interfere with the reaction. This is the most common approach that is used to kinetically characterize wild-type and mutant forms of pol β currently.

DNA polymerase β prefers to bind gapped DNA substrates that have a 5′-phosphate on the downstream strand in the gap.[114] In the absence of this downstream strand, pol β has very low DNA binding affinity. In contrast, tight DNA binding does not require the upstream strand that would normally be extended. This result suggests that pol β targets short DNA gaps in DNA through the 5′-phosphate on the downstream portion of the gap. Only when a 3′-OH is "within reach" will it be extended. This is consistent with the processive gap-filling DNA synthesis observed with short (<5 nucleotide) DNA gaps.[115] Surprisingly, pol β can extend template−primers that do not have a downstream strand (i.e., not gapped). This is due to the increased DNA binding affinity of pol β for these substrates in the presence of the correct nucleotide (unpublished data). Single-turnover analysis indicates that pol β prefers a single-nucleotide-gapped DNA substrate with kinetic parameters k_{pol}, $K_{d,dCTP}$, and $K_{d,DNA}$ of 10 s^{-1}, 6 μM, and 20 nM.[68] It should be noted that these parameters are dependent on DNA sequence and solution conditions. Accordingly, these should be carefully scrutinized before attempting to draw conclusions from data published from different laboratories.

The dNTP binding pocket of DNA polymerases is formed by the template base, DNA duplex terminus, and enzyme. In the ternary substrate complex of pol β, the nascent base pair (templating and incoming nucleotides) is sandwiched between the duplex DNA terminus and α-helix N (Figure 5a). It is observed that normal Watson−Crick pairing, as well as aberrant primer termini, has a strong influence on correct nucleotide insertion.[92] The insertion efficiency (k_{cat}/K_m, dGTP−dC) is highly dependent on the sequence identity of the DNA terminus: G−C ~ A−T > T−A ~ C−G (template−primer). Mismatches at the primer terminus strongly diminish nucleotide insertion but do not affect DNA binding affinity. Transition intermediates are generally extended more easily than transversions. The loss of catalytic efficiency with purine−purine mismatches is entirely due to the inability to insert the incoming nucleotide, since the $K_{d,dGTP}$ is not affected. Abasic sites and extra nucleotides in and around the duplex terminus indicate that altering the primer strand is more detrimental to the nascent base pair binding pocket than the equivalent modification in the template strand. These results define the influence of wild-type enzyme on "aberrant" DNA structures that represent mutagenic intermediates.

To determine whether a conformational change (step 3) or chemistry (step 4) limits k_{pol}, the rate constant for the incorporation of a α-thio-substituted dNTP analogue is usually examined and compared to that for the natural nucleotide. Due to the lower electronegativity of sulfur relative to oxygen, a significant decrease in rate upon sulfur substitution would suggest that chemistry is rate-limiting.

Model studies with phosphate triesters indicate a large elemental effect upon substitution of sulfur at a nonesterified position, whereas studies with phosphate diesters indicate smaller decreases in rate upon sulfur substitution.[116] The thio-elemental effect observed for pol β is only 2.1, suggesting that a step other than chemistry may be rate-limiting.[68] However, this interpretation must be tempered since the intrinsic elemental effect is not known, and there are steric considerations, in addition to the electronegativity of sulfur, that could influence the observed rate.[110]

Liu and Tsai[117] have examined the stereoselectivity of pol β for the Rp and Sp isomers of dATPαS by single-turnover experiments. Sulfur substitution at αP makes this phosphate a chiral center with the nonbridging atoms defined as Sp or Rp. As with other DNA polymerases, pol β prefers to insert the Sp isomer. However, it can also insert the Rp isomer with a 57-fold lower efficiency. Replacing Mg^{2+} with more thiophilic metals such as Mn^{2+} or Cd^{2+} did not alter the selectivity in a predictable manner, indicating again that possible steric influences are affecting insertion efficiency. The Sp isomer of the triphosphate would have the nonbridging oxygen coordinating the catalytic and nucleotide binding metals (Figure 8), and the sulfur interacting with H_2O.[55] Significantly, other A-, B-, and RT-family polymerases for which there is a ternary substrate complex indicate that the pro-Sp oxygen interacts with a basic side chain. In contrast, this oxygen coordinates a water molecule for catalytically less efficient polymerases from the X- and Y-families.[53]

Another approach commonly used to identify if a nonchemical step limits nucleotide insertion is to compare the apparent concentration of the ternary substrate complex formed when the reaction is quenched with acid (pulse-quench) with the concentration when it is chased with nucleotide (pulse-chase; use unlabeled dNTP as a chase when following radioactive dNTP incorporation). If a dNTP-bound complex exists where the dNTP is not in rapid equilibrium with free dNTP (e.g., complex between steps 2 and 3) and the chemical equilibrium (step 4) permits accumulation of the isomerized ternary substrate complex, then more product will be formed with the chase protocol than with the acid quench. This central complex will be quenched immediately with acid, but if it is not in equilibrium with free dNTP, then this complex can form product in the presence of a cold dNTP. However, since the apparent burst for pol β is partially rate limited by at least two steps, and it appears it has a strong commitment to product formation, very little isomerized ternary complex can accumulate, even if a rate-limiting conformational change limits nucleotide insertion. However, structural characterizations of pol β in different liganded states clearly identify a number of conformational changes that must occur before the transition state is achieved. Indeed, the apparent binding affinity for gapped DNA is nearly 200-fold tighter in the presence of a correct nucleotide than in its absence, indicating that a conformational change has occurred.[92]

Steady-state kinetic characterization of mouse pol β failed to demonstrate a reversal of the polymerization reaction (i.e., pyrophosphorolysis) in the presence of PP_i with activated DNA. Yet, PP_i was inhibitory for DNA synthesis.[108] Whereas pyrophosphorolysis can be measured on nicked DNA, its rate is significantly diminished on single-nucleotide-gapped DNA (unpublished data). This suggests that the primer terminus of a one-nucleotide-gapped substrate does not readily enter the polymerase active site: an event required to add PP_i to the terminal dNMP primer residue.

3.1.2. Fidelity

DNA polymerases produce several classes of errors during DNA synthesis.[88] One of the most common errors is a base substitution error where an incorrect nucleotide (does not preserve Watson−Crick base pairing) is inserted opposite the templating nucleotide. Kinetically, there are several opportunities for a polymerase to discriminate against incorrect nucleotide incorporation. It can bind the incorrect nucleotide weakly (step 1, Figure 15a) and insert it more slowly (steps 2 and/or 3). If an incorrect nucleotide is inserted, further DNA synthesis is inefficient. This provides the opportunity for an intrinsic proofreading activity (e.g., 3′-5′ exonuclease) to remove the "mistake" or allow the polymerase to dissociate to permit an extrinsic enzyme the chance to remove the mispair (e.g., AP endonuclease). The error rate for pol β single-nucleotide gap filling is about 1 error/3000 nucleotides synthesized[118] and is even greater with larger gaps.[91] This represents a moderate fidelity in terms of the wide spectrum of fidelities exhibited by natural exonuclease-deficient DNA polymerases.[76]

DNA polymerase fidelity or specificity expresses the ability of a polymerase to select a correct dNTP from structurally similar incorrect molecules. Fidelity is quantified from the ratio of specificity constants (catalytic efficiencies) for alternate substrates (i.e., correct and incorrect dNTPs). An analysis of the efficiency of dNTP (correct and incorrect) insertion from published work from exonuclease-deficient DNA polymerases from five families derived from a variety of biological sources reveals that a strong correlation exists between the ability to synthesize DNA and the probability that the polymerase will make a mistake (i.e., base substitution error).[76] Unexpectedly, this analysis indicates that the difference between low- and high-fidelity DNA polymerases is related to the efficiency of correct (varying 10^7-fold), but not incorrect, nucleotide insertion. Thus, low-fidelity DNA polymerases lack the ability to insert the correct nucleotide. Inefficient enzymes (i.e., poor correct nucleotide insertion efficiency) exhibit low fidelity, whereas efficient polymerases display high fidelity. Unexpectedly, low-fidelity naturally occurring polymerases generally insert wrong nucleotides with an efficiency that is similar to, or lower than, that of a polymerase with higher fidelity. Low-fidelity polymerases appear to be an evolutionary solution to how to replicate damaged DNA, or DNA repair intermediates, without burdening the genome with excessive polymerase-initiated errors. Enzyme efficiency is quantified by a ratio of kinetic parameters. From steady-state kinetics, efficiency is given by $k_{cat}/K_{m,dNTP}$, whereas it is defined by $k_{pol}/K_{d,dNTP}$ when determined by a pre-steady-state approach or single-turnover analysis. As predicted from the kinetic mechanism, efficiency is identical, independent of the kinetic approach (i.e., $k_{cat}/K_{m,dNTP} = k_{pol}/K_{d,dNTP}$). This has been verified for pol β,[57,68] and these kinetic parameters can be reliably used to determine enzyme specificity. Fidelity can be calculated from the expression

$$\text{fidelity} = \{(\text{eff})_{\text{correct}} + (\text{eff})_{\text{incorrect}}\}/(\text{eff})_{\text{incorrect}}$$

where (eff) refers to the catalytic efficiency for correct or incorrect insertion. Note that in most instances $(\text{eff})_{\text{incorrect}} \ll (\text{eff})_{\text{correct}}$, so that fidelity is the simple ratio of catalytic efficiencies. More importantly, it is a measure of the total number of insertions that occur before a misinsertion event. Accordingly, structure−function studies of DNA po-

lymerases must interpret site-directed mutagenesis results in terms of catalytic efficiencies. If the fidelity of a mutant polymerase is altered, then the catalytic efficiency for the alternate substrates is *differentially* affected by the protein modification, indicating that the side chain of interest provides different interactions for the correct and incorrect incoming nucleotides.

The numerous conformational changes observed when binding a correct dNTP are consistent with an induced-fit model for enzyme specificity. The induced-fit model states that, after initial correct dNTP binding, conformational changes of the ternary complex result in an activated complex where catalytic residues are "aligned" for chemistry. In contrast, the incorrect dNTP results in a poor fit such that alternate conformational changes result in a suboptimal "activated" complex. Much of the discussion in the literature relative to the induced-fit model has relied on the assumption that a rate-limiting conformational change must limit correct nucleotide insertion.[109,119] As discussed above, clear evidence for a rate-limiting conformational change is lacking for pol β. Additionally, kinetic,[68,111] structural,[61] and modeling[86,120−122] experimental approaches suggest that subdomain motions are too rapid to limit nucleotide insertion. Nevertheless, it has been pointed out that induced fit can alter enzyme specificity even when critical conformational changes are kinetically silent,[123] such as when the transition states for correct and incorrect incorporation are unique. Although an induced fit model reduces enzyme specificity relative to a situation where a unique transition state exists, the reduced specificity represents an acceptable compromise for an enzyme such as a DNA polymerase that must select a different/new substrate (DNA and dNTP) with each catalytic cycle.

The observation that low- and high-fidelity polymerases insert incorrect nucleotides with similar efficiencies suggests that the "structure" of the transition state for incorrect insertions may be similar for all DNA polymerases. In contrast, since fidelity is modulated by the efficiency of correct nucleotide insertion, the molecular interactions that contribute to efficient DNA synthesis (i.e., formation of the transition state) are dependent on the specific polymerase. More importantly, an understanding of fidelity at the molecular level requires structural insight into the polymerase-dependent attributes that contribute to correct insertion efficiency. A comparison of crystallographic ternary substrate complexes of DNA polymerases from five families exhibiting a range of nucleotide insertion rates reveals that the geometries of the nascent base pair binding pockets are similar, consistent with the similar binding affinities reported for the correct nucleotide for these DNA polymerases.[53] Thus, the loss in catalytic efficiency for correct insertion by low-fidelity polymerases is due to a decreased rate of nucleotide insertion. A comparison of the electrostatic environment of the reaction atoms indicates that low-fidelity DNA polymerases have unique aspects that influence the transition state for correct insertion.

3.1.3. Site-Directed Mutagenesis

An important approach to deduce the role of individual amino acids is through structure-based site-directed mutagenesis. DNA polymerase β mutant enzymes that have been examined fall into two broad categories, those that directly interact with the nascent base pair and those that influence catalysis from a distant site. This later category suggests that the conformational adjustments that occur during catalytic cycling can transmit a structural signal or perturbation to the active site.

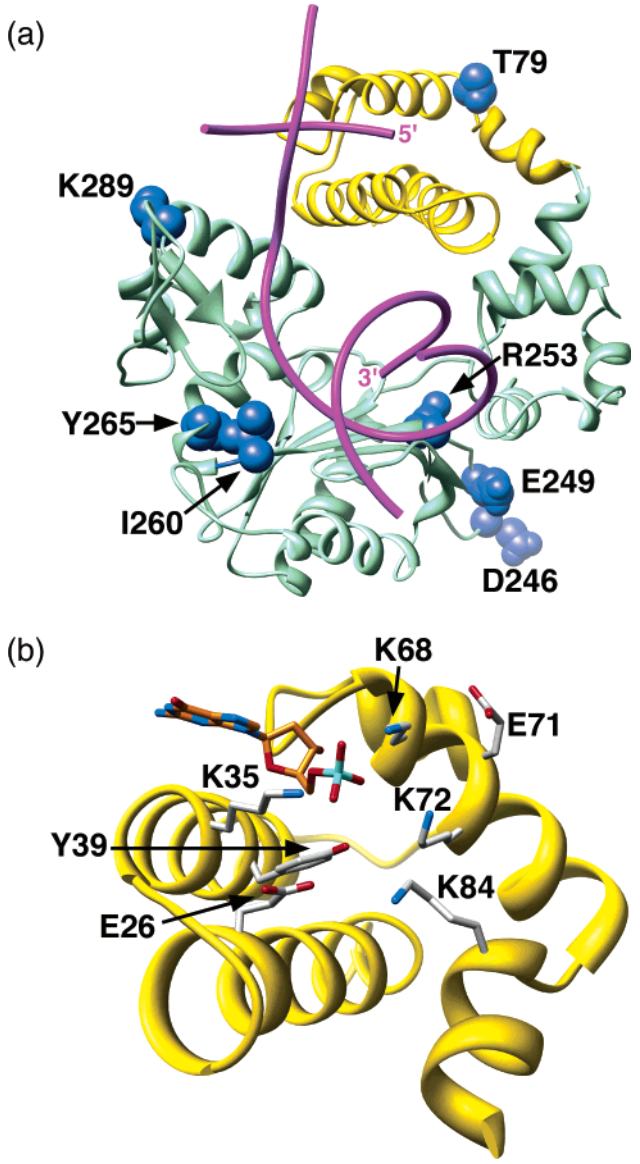


Figure 16. Position of key residues in pol β probed by site-directed mutagenesis. (a) Ribbon representation pol β where the polymerase domain is colored turquoise and the lyase domain is gold. The DNA backbone is magenta, and the 3′- and 5′-ends of the primer and downstream strands are indicated, respectively. Site-directed mutagenesis has identified several residues (blue) that do not interact with substrates but influence catalytic activity and/or fidelity when altered. See text for discussion. (b) Key dRP lyase active-site residues are identified in this product complex (i.e., after removal of the dRP group; the 5′-nucleotide is orange and is part of the first base pair in the downstream duplex). See text for discussion of the proposed role of these side chains in the lyase mechanism.

Upon binding a correct nucleotide, it is clear that the N-subdomain closes around the nascent base pair. This is accomplished by a simple rotation within α-helix M. An aromatic residue at the base of this helix, Tyr265 (Figure 16a), has been examined in detail by site-directed mutagenesis.[124−128] The results show that altered interactions at this site can alter fidelity (base substitution and frame shift), catalytic efficiency, and the ability to extend certain mispairs. In regard to catalytic efficiency, the Y265H is extraordinary in that the catalytic efficiency for insertion of dGTP opposite A is elevated 10-fold.[127] Coupled with the 10-fold loss in catalytic efficiency for correct insertion, this translates into

a 100-fold loss of fidelity. More importantly, the increase in efficiency for incorrect insertion makes this polymerase a mutagenic threat. Another hydrophobic residue near Tyr265 is Ile260. Ile260 is situated at the boundary between the C- and N-subdomains (Figure 16a). Alterations of this side chain can also influence catalytic efficiency, and His and Gln variants exhibit lower fidelity.[129] Kinetic characterization of I260Q indicates that the loss in fidelity is specifically due to a loss in dNTP binding discrimination ($K_{d,correct} \geq K_{d,incorrect}$) whereas k_{pol} is not affected.[130]

A large loop under the C-subdomain connects two adjacent antiparallel β-strands. Mutation of three residues in this region (Asp246, Glu249, and Arg253; Figure 16a) has been shown to modulate polymerase fidelity by three different apparent mechanisms. The D246V mutant exhibits decreased discrimination by enhancing the efficiency of incorrect nucleotide insertion.[131] In contrast, a R253M mutant is resistant to AZTTP, as demonstrated by a loss of efficiency for AZTTP insertion.[132] It was further shown that the R253M variant discriminates against other dNTP analogues better than wild-type enzyme.[133] Last, a E249K mutant has been shown to exhibit a mutator phenotype, but the fidelity for base substitution errors is not altered as determined by kinetic analysis.[134] It was shown that this mutant could extend mispairs more efficiently than wild-type enzyme.

Thr79 sits at the bottom of one of the α-helixes of the HhH motif in the lyase domain (Figure 16a). It is near the protease-sensitive peptide that connects the two enzymatic domains of pol β. Accordingly, this region undergoes significant conformational adjustments upon interactions of the two domains to an incised DNA strand. In the absence of downstream DNA, however, the lyase domain only makes limited, if any, contact with the ssDNA template.[28] Interestingly, a serine mutant at this position does not influence the fidelity of single-nucleotide gap filling but decreases the fidelity of DNA synthesis on nongapped substrates.[135] This suggests that the lyase domain may interact with the ssDNA template or that the mutant protein transmits an altered conformation of the lyase domain into the adjacent D-subdomain that is responsible for orienting the primer terminus in the polymerase active site.

Lys289 sits at the end of α-helix N of the N-subdomain (Figure 16a). A variant identified from a colorectal carcinoma, K289M, exhibits a mutator phenotype *in vivo*.[136] Characterization of the purified mutant polymerase suggested that it discriminated against incorrect nucleotides less readily than wild type. It was postulated that the modified protein altered DNA positioning, leading to a higher proportion of misincorporations. More importantly, it was suggested that the variant might have been responsible for tumorigenesis.

A number of substrate-interacting and active-site residues have also been examined by site-directed mutagenesis, and in most instances, they confirm the role of the individual side chain in binding and catalysis. In this way, the critical role of the active-site aspartates (Asp190, Asp192, Asp256, Figure 8) has been confirmed.[67,137,138] Additionally, Ser180 and Arg183 contribute significantly to catalysis.[139,140] These side chains hydrogen bond to nonbridging oxygens on the γ- and β-phosphates, respectively.

Structures of DNA polymerases bound with DNA reveal that the 5′-trajectory of the template strand is dramatically altered as it exits the polymerase active site (Figure 3b). This distortion provides the polymerase access to the nascent base pair to interrogate proper Watson−Crick geometry. Upon binding a correct dNTP, α-helix N of pol β is observed to form one face of the binding pocket for the new base pair. Asp276 and Lys280 stack with the bases of the incoming nucleotide and template, respectively (Figure 5a). Variants at residue 276, including glutamate, increase the incoming nucleotide binding affinity during template base recognition.[68,141,142] The implication is that α-helix N and Asp276 influence ground-state binding in the closed conformation that occurs prior to the rate-limiting step so that that subdomain closure must be very rapid and not rate-limiting.[68]

To determine the role of Lys280, seven site-directed mutants were characterized.[57] The catalytic efficiency for single-nucleotide gap filling with the glycine mutant (K280G) was strongly diminished relative to the case of wild type for templating purines due to a decreased binding affinity for the incoming nucleotide. In contrast, catalytic efficiency was hardly affected by glycine substitution for templating pyrimidines. The fidelity of the glycine mutant was identical to that of wild-type enzyme for misinsertion opposite a template thymidine, whereas the fidelity of misinsertion opposite a template guanine was modestly altered. The nature of the Lys280 side-chain substitution for thymidine triphosphate insertion (templating adenine) indicates that Lys280 "stabilizes" templating purines through van der Waals interactions. The observation that glycine substitution for Lys280 results in a decrease in catalytic efficiency that is strongly dependent on the identity of the templating base indicates that interactions with the nascent base pair may be energetically unique for formation of each Watson−Crick base pair. Thus, residue 280 interactions with templating purines are more important than they are for templating pyrimidines, suggesting that template positioning and stabilization is unique for each base pair.

Closing of the N-subdomain upon nucleotide binding positions three residues (Tyr271, Asn279, Arg283) within hydrogen bonding distances to groups in the DNA minor groove around the nascent base pair binding pocket (Figure 5b). Tyr271 hydrogen bonds to the minor groove edge of the primer terminus in the closed ternary complex but hydrogen bonds to the minor groove edge of the templating base in the open binary complex (Figure 5b). In the nicked DNA product structure, Tyr271 hydrogen bonds with the base at the primer terminus.[27] In this latter situation, the primer terminus is in the dNTP site. Surprisingly, mutants at this position have little or no effect on catalytic efficiency.[58,143] In addition, alanine substitution for Tyr271 has little or no effect on insertion efficiency with mismatched termini or when an abasic site is positioned opposite the primer terminus.[92] Unexpectedly, these results suggest that Tyr271 does not play a critically important role during correct nucleotide insertion independent of whether the primer terminus is properly base paired. Alternatively, the hydrogen bond provided by Tyr271 may be very important to catalytic cycling, but the polymerase has compensated for the loss of this hydrogen bond by altering active-site hydrogen-bonding patterns as suggested by molecular dynamics simulations.[144]

Side-chain substitution at Asn279 is also observed to have no or little effect on pol β fidelity or catalytic efficiency for Watson−Crick base pairs.[58,143] However, the N279A mutant was observed to alter the specificity of 8-oxodGTP insertion relative to the case of wild-type enzyme but not the specificity when 8-oxodG was the templating base.[94] This suggests that interactions of 8-oxodGTP with Asn279 in the polymerase active site may influence the conformation of

8-oxodGTP and therefore alter its misincorporation. A *syn*-conformation of 8-oxodGTP positions the oxygen at C8 in the minor groove that would normally be occupied by O2 or N3 or pyrimidines or purines, respectively, in a Watson−Crick base pair.[62]

Arg283 hydrogen bonds with the sugar immediately upstream of the templating base and forms a van der Waals contact with the templating base (Figure 5b). Rationale or random mutagenesis of the polymerase active site generally results in mutant enzymes that have moderately reduced or improved specificity. Alanine substitution for the Arg283 residue of pol β (R283A) results in a dramatic loss of catalytic efficiency for correct insertion and fidelity, implying that catalytic efficiency for correct nucleotide insertion and discrimination were coupled.[58] Subsequently, it was found that the nature of the side-chain substitution at this residue influences the base-substitution specificity[118] and frame-shift fidelity.[69] A kinetic analysis of the R283A mutant of human pol β indicates that there is a 33 000-fold loss in catalytic efficiency for insertion of dCTP opposite a templating guanine within a single-nucleotide-gapped DNA substrate relative to the case of the wild-type enzyme. In contrast, the mutant enzyme inserts dTTP opposite guanine 15-fold less efficiently than wild-type enzyme. Since fidelity is the ratio of catalytic efficiencies, the mutant enzyme has a 2400-fold lower ability to discriminate against dTTP insertion relative to dCTP insertion. This loss in discrimination is entirely due to the loss of the ability to insert the correct nucleotide, dCTP, since dTTP insertion efficiency is reduced in the mutant enzyme. Thus, the fidelity of this mutant is decreased simply because it can no longer synthesize DNA. This is the strategy used by natural polymerases to modulate their fidelity.[76] This is in contrast to the Y265H mutant described above that exhibited an increase in efficiency for incorrect nucleotide insertion.[127] A ternary substrate complex structure of the R283A mutant indicates that it is in the open conformation.[58] Since correct, but not incorrect, nucleotide insertion is specifically affected, this observation suggests that correct insertion occurs from a closed conformation whereas incorrect insertion occurs from an open or partially open conformation. Consistent with this idea is the observation that structures of Y-family DNA polymerases have an "open" templating site[145] that specifically lowers the efficiency of correct insertion.[53]

A rare *cis*-peptide bond occurs between Gly274 and Ser275 that creates a sharp turn between two α-helices (M and N) that contribute significant interactions with the nascent base pair (Figure 10). Proline substitution for Gly274 would be expected to clash with the incoming nucleotide and alter helix interactions with the nascent base pair. The mutant enzyme (G274P) exhibited a 10^4-fold decrease in catalytic efficiency.[76] The loss of efficiency, however, did not have a dramatic effect on fidelity, indicating that this mutation altered correct and incorrect insertion to about the same extent. Thus, although all natural low-fidelity DNA polymerases are inefficient, not every inefficient DNA polymerase has low fidelity.

As noted previously, there are several residues that reposition themselves in response to closing/opening of the N-subdomain (Figure 6). Some of these have been altered and characterized. As described above, Asp192 and Arg283 play significant roles in catalysis. Likewise, alanine and lysine substitution for Glu295 severely decreases catalysis.[140,146] Interestingly, the lysine variant has been identified in a human gastric carcinoma, and it has been found that it could compete with wild-type enzyme in an *in vitro* BER assay.[146] Phe272 positions itself between Arg258 and Asp192 in the closed conformation. Leucine substitution at this residue results in a mutant enzyme with a decreased fidelity, but it has only a modest effect on catalytic efficiency. Since leucine is a relatively bulky hydrophobic side chain, it will be interesting to see what the effect on catalytic efficiency will be of removing a greater portion of this side chain. Since Arg258 stabilizes an inactive conformation, it will also be interesting to examine the influence of alterations at this residue on correct/incorrect nucleotide insertion. A steady-state kinetic analysis on a homopolymeric template−primer system suggests that alanine substitution has little, if any, effect.[138] It remains to be seen what the influence of such a mutation is on single-nucleotide gap filling.

3.2. Polymerase Dynamics

3.2.1. Spectroscopic Analyses

Spectroscopic monitoring of rapid conformational changes has provided critical information on the dynamic nature of the polymerase and DNA substrate. This information is a prerequisite in attempting to assign molecular events to specific kinetic steps. Tsai and co-workers have used a fluorescent base (2-aminopurine) strategically positioned in the DNA to monitor enzyme and substrate conformational changes by stopped-flow analysis.[147,148] Using an exchange-inert metal−dNTP complex, they were able to form a ternary complex poised for catalysis but lacking the catalytic metal. Formation of the "poised" ternary complex was associated with a rapid fluorescence transient phase that appears to be associated with subdomain closing. Addition of the catalytic metal resulted in a second fluorescent phase that had an identical rate of product formation. Consequently, they assigned this step to chemistry.

DNA polymerase β also has one intrinsic tryptophan residue (Trp325) located in the N-subdomain. To probe the microenvironment and dynamics of the lyase domain and α-helix N in the polymerase domain, the native tryptophan was replaced with alanine and tryptophan was substituted for Phe25 or Leu287.[149] Influences of substrate on the fluorescence anisotropy decay of these single-tryptophan forms of pol β indicated that the segmental motion of α-helix N was rapid (∼1 ns) and far more rapid than the step that limits chemistry. Binding of Mg^{2+} and/or gapped DNA did not cause a noticeable change in the rotational correlation time or angular amplitude of tryptophan in α-helix N. More important, binding of a correct nucleotide significantly limited the angular range of the nanosecond motion within α-helix N. In contrast, the segmental motion of the 8-kDa domain was "frozen" upon DNA binding alone, and this restriction did not increase further upon nucleotide binding.

NMR has been used to study the dynamic aspects of ligand binding to pol β.[150,151] The six methionine residues of pol β can be labeled with [methyl-^{13}C]methionine without altering the activity or normal folding properties of the polymerase. This selective labeling of the methionine methyl carbons permits a means of probing structural and dynamic features of the protein in the vicinity of these residues when the polymerase binds substrates. The methionine residues are situated in the lyase domain (Met18) and the C- (Met155, Met158, Met191, and Met236) and N-subdomains (Met282). Addition of the single-stranded template selectively broad-

ened the methyl resonance of Met18 in the lyase domain. Further addition of a 5′-phosphorylated downstream oligonucleotide shifted and sharpened this resonance. Formation of an abortive ternary complex, enzymatically produced like that employed for crystallography studies, resulted in significant shifts for Met155, Met191, Met236, and Met282. The resonance signal for position 282 is of particular interest because this residue is in the middle of α-helix N that stacks with the nascent base pair in the crystal structure of the closed pol β ternary complex (Figure 5). The shifted and broadened resonance signal of Met282 suggests exchange among several conformational forms in the closed complex. It remains to be determined what the significance of this apparent induced motional freedom is on catalysis.

Watson−Crick hydrogen bonding plays a crucial role in stabilizing the structure of DNA. However, the role of these hydrogen bonds for substrate discrimination during DNA synthesis is unclear. Kool and co-workers[152] have demonstrated the utility of fluorinated nucleoside isosteres for separating the contributions of hydrogen bonding of the nascent base pair from steric and geometric constraints imposed by the nascent base pair binding pocket. Studies utilizing the thymine isostere 2,4-difluorotoluene (F) have demonstrated that the contribution of hydrogen-bonding interactions between the bases of the nascent base pair and the polymerase varies significantly and depends on the polymerase. For pol β, the presence of the F-isostere in the templating position fails to support polymerization.[153] Using [methyl-^{13}C]Met-labeled pol β, it was demonstrated that replacement of a templating thymine base with the F-isostere fails to support the large conformational change that is monitored by Met282.[151] These studies demonstrate that the block in catalysis is directly related to the absence of the set of conformational transitions that include the open to closed transition monitored by Met282.

3.2.2. Computational Characterization

Computational approaches provide an attractive method to bridge the gap between crystallographic and kinetic studies. This approach can provide information on dynamic and energetic contributions to catalysis and substrate specificity (fidelity). High-temperature and targeted molecular dynamics simulations suggest closure of the N-subdomain does not occur in a single concerted step (Figure 6).[86,154] Instead, there is a sequence of structural events that lead to the closed or open conformation observed in the ternary substrate or binary complex, respectively. Specifically, it was found that Arg258 side-chain rotation toward Asp192 during opening is slower than the Phe272 ring flip and that the global N-subdomain motion is very fast. Additionally, the roles of the residues involved in this conformational "signaling" pathway near the pol β active site (Figure 6) were examined by site-directed alterations and computer simulations. The results indicated that, in some instances, structural alterations occurred to compensate for the loss of critical interactions (i.e., Tyr271).[155]

These computational studies have been extended to examine the active-site accommodations that could occur when introducing a mismatch in the active site before[156] or after chemistry.[154] Analysis of the simulated structures after chemistry revealed that the degree of structural distortion provided insight to how extension of a mismatch may be averted. The distortions (GG > CC > CA) paralleled the experimentally deduced inability of pol β to extend these mispairs.[92] Furthermore, the simulations with an incorrect dNTP in the active site indicated that the mismatch distorted the coordination geometry of the catalytic metal,[156] and these distortions parallel the catalytic efficiency for insertion of the respective incorrect nucleotides.[92] These results are consistent with the catalytic-metal-induced conformational adjustments observed in the pretransition complex described earlier (Figure 8).[66] The new structure also provides the opportunity to apply hybrid quantum mechanical and molecular mechanics approaches to investigate the chemical mechanism.[157,158]

3.3. Lyase

The dRP lyase activity represents the slowest step during monofunctional glycosylase-initiated single-nucleotide BER.[6,67] However, it has not been studied as extensively as the DNA synthesis reaction and the mechanism of β-elimination is not well understood. An AP site in DNA is in equilibrium between three species: the α- and β-hemiacetals that are 2′-deoxy-D-erythro-pentofuranoses, and the open-chain aldehyde. While the predominant form of an AP site in solution is a mixture of the α- and β-hemiacetals, the ring-opened aldehyde form represents less than 1% of the total AP sites at equilibrium, as determined by NMR.[159] It is generally accepted that the ring-opened aldehyde species is the most reactive of the AP site species toward nucleophilic attack, but it is unclear whether this species arises as an enzyme-catalyzed ring-opening step or by the enzyme simply waiting for the ring-opened form to occur by uncatalyzed equilibration. A general framework for the lyase reaction catalyzed by the bifunctional DNA glycosylases has been proposed.[8,160] The enzyme employs an amine as a nucleophile to attack the C1′ of the sugar associated with the damaged base, thereby forming an imino intermediate (Schiff base). An experimental hallmark of the reaction is the formation of an imino intermediate between the enzyme and substrate DNA that may be isolated as a stable complex upon reduction with sodium borohydride. More recently, several crystal structures of borohydride-trapped enzyme/DNA intermediates have been reported.[160,161] The DNA strand cleavage reaction catalyzed by these enzymes appears to proceed by a *syn* β-elimination involving abstraction of the *2′-pro-S* proton and formation of a *trans*-α,β-unsaturated aldose product.[162]

As described earlier, the structure of pol β bound to nicked DNA with a 5′-THF residue indicated that C1′ and the nucleophile (NZ of Lys72) were distant from one another (Figure 14). This structure suggested a model where the sugar must rotate 120° to position C1′ near Lys72. On the basis of this model, a lyase mechanism for pol β has been proposed.[65] Several important residues (Glu26, Ser30, His34, Lys35, Tyr39, Lys68, Lys 72, and Lys84) define the lyase active-site pocket (Figure 16b). After correct positioning of the flexible sugar, an enzyme-mediated ring opening is proposed. Rotating the closed sugar ring 120° also brings the O4′ in close proximity (∼3.0 Å) with the NZ of Lys72, suggesting that protonation of O4′ by NZ of Lys72 may induce ring opening. Subsequently, the deprotonated and neutral NZ of Lys72 appears to be in an optimum position for nucleophilic attack at C1′. In this structural model, Tyr39 is hydrogen bonded to the deprotonated form of Lys72 and keeps it in position for nucleophilic attack. The reduction in dRP lyase activity of the alanine substitution for Tyr39 is consistent with this interpretation.[65] This model indicates that Lys84 could participate as an alternative nucleophile in the

absence of Lys72, albeit with a reduced efficiency. This is consistent with the incomplete loss of dRP lyase activity observed for Lys72 mutants[65,97] and with previous sodium borohydride trapping results.[10] Following Schiff base formation, the enzyme must catalyze β-elimination by abstracting a proton from the C2′ of the substrate.

Abstraction of the proton at C2′ is thought to trigger β-elimination, resulting in scission of the C3′−O phosphodiester bond. The imine at C1′ is expected to lower the pK_a of C2′ and make it a favorable target for proton abstraction. It has been proposed that abstraction of a proton from C2′ could be facilitated by a neighboring acidic residue such as Glu26 or Glu71.[163] However, these residues appear to be too far from C2′. Another candidate that may promote proton abstraction is Asp26. It also is too far from the closed-ring C2′ atom, but an intervening water molecule can be modeled in a position similar to that observed in the EndoIII active site[161] and is stabilized by hydrogen bonds to Tyr39 and an oxygen on the phosphate 3′ to the dRP residue. Interestingly, a water has been observed in other pol β crystal structures,[55] suggesting that this structural model for proton abstraction is reasonable. The final step in the lyase mechanism is the release of dRP from the Schiff base−dRP complex by hydrolysis. The space and rigidity provided by the dRP lyase active site accommodating the flexibility of the dRP group suggested by these structures makes the proposed events feasible.

4. Concluding Remarks

In summary, pol β is a multidomain protein whose domains/subdomains contribute distinct biochemical functions. Crystal structures of pol β have been solved in many different liganded forms, indicating that multiple conformations of the enzyme and substrates are possible. These conclusions are supported by solution studies, both spectroscopic and kinetic. This information has provided mechanistic insight at the molecular level to our understanding of DNA polymerase function in general.

Tremendous progress has been made in our understanding of the enzymatic characteristics exhibited by pol β over the past decade. With this information, it should be possible to determine the extent that other enzymes influence pol β functionally, directly or indirectly. Reciprocally, it will also be important to understand how pol β might influence other enzymes (e.g., enzymes of BER). Using fragments of pol β, it has already been shown that these fragments can interfere with BER.[164] The possibility that pol β has a role in the etiology of cancer[165] makes it all the more important to gain a better understanding of the biological role of this "simple" DNA repair enzyme. With a molecular understanding of pol β, there is an opportunity to develop substrates/inhibitors that could specifically enhance/inhibit pol β.

With atomic level detail, we now have a good picture of catalytic cycling by pol β. This was accomplished by a multifaceted approach that utilized structural, biophysical, biochemical, and kinetic techniques. Although there are many questions yet to be answered, this framework provides an excellent foundation to begin to confidently ask biological questions and decipher the roles of this polymerase during cellular development, stress, and uncontrolled growth.

5. Acknowledgments

We thank our colleagues in the DNA Repair and Nucleic Acid Enzymology group and collaborators for their insights and contributions to the study of DNA polymerase β. Additionally, we are grateful to the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences, for support.

6. References

(1) Lindahl, T.; Wood, R. D. *Science* **1999**, *286*, 1897.
(2) Wood, R. D.; Mitchell, M.; Sgouros, J.; Lindahl, T. *Science* **2001**, *291*, 1284.
(3) Barnes, D. E.; Lindahl, T. *Annu. Rev. Genet.* **2004**, *38*, 445.
(4) Lindahl, T. *Nature* **1993**, *362*, 709.
(5) Dianov, G.; Lindahl, T. *Curr. Biol.* **1994**, *4*, 1069.
(6) Srivastava, D. K.; Vande Berg, B. J.; Prasad, R.; Molina, J. T.; Beard, W. A.; Tomkinson, A. E.; Wilson, S. H. *J. Biol. Chem.* **1998**, *273*, 21203.
(7) Wilson, S. H.; Kunkel, T. A. *Nature Struct. Biol.* **2000**, *7*, 176.
(8) Fromme, J. C.; Banerjee, A.; Verdine, G. L. *Curr. Opin. Struct. Biol.* **2004**, *14*, 43.
(9) Matsumoto, Y.; Kim, K. *Science* **1995**, *269*, 699.
(10) Piersen, C. E.; Prasad, R.; Wilson, S. H.; Lloyd, R. S. *J. Biol. Chem.* **1996**, *271*, 17811.
(11) Prasad, R.; Singhal, R. K.; Srivastava, D. K.; Molina, J. T.; Tomkinson, A. E.; Wilson, S. H. *J. Biol. Chem.* **1996**, *271*, 16000.
(12) Cappelli, E.; Taylor, R.; Cevasco, M.; Abbondandolo, A.; Caldecott, K.; Frosina, G. *J. Biol. Chem.* **1997**, *272*, 23970.
(13) Wiederhold, L.; Leppard, J. B.; Kedar, P.; Karimi-Busheri, F.; Rasouli-Nia, A.; Weinfeld, M.; Tomkinson, A. E.; Izumi, T.; Prasad, R.; Wilson, S. H. *Mol. Cell* **2004**, *15*, 209.
(14) Pascucci, B.; Maga, G.; Hübscher, U.; Bjoras, M.; Seeberg, E.; Hickson, I. D.; Villani, G.; Giordano, C.; Cellai, L.; Dogliotti, E. *Nucleic Acids Res.* **2002**, *30*, 2124.
(15) Frosina, G.; Fortini, P.; Rossi, O.; Carrozzino, F.; Raspaglio, G.; Cox, L. S.; Lane, D. P.; Abbondandolo, A.; Dogliotti, E. *J. Biol. Chem.* **1996**, *271*, 9573.
(16) Fortini, P.; Pascucci, B.; Parlanti, E.; Sobol, R. W.; Wilson, S. H.; Dogliotti, E. *Biochemistry* **1998**, *37*, 3575.
(17) Dianov, G. L.; Prasad, R.; Wilson, S. H.; Bohr, V. A. *J. Biol. Chem.* **1999**, *274*, 13741.
(18) Prasad, R.; Dianov, G. L.; Bohr, V. A.; Wilson, S. H. *J. Biol. Chem.* **2000**, *275*, 4460.
(19) Liu, Y.; Beard, W. A.; Shock, D. D.; Prasad, R.; Hou, E. W.; Wilson, S. H. *J. Biol. Chem.* **2005**, *280*, 3665.
(20) Delarue, M.; Poch, O.; Tordo, N.; Moras, D.; Argos, P. *Protein Eng.* **1990**, *3*, 461.
(21) Ohmori, H.; Friedberg, E. C.; Fuchs, R. P. P.; Goodman, M. F.; Hanaoka, F.; Hinkle, D.; Kunkel, T. A.; Lawrence, C. W.; Livneh, Z.; Nohmi, T.; Prakash, L.; Prakash, S.; Todo, T.; Walker, G. C.; Wang, Z.; Woodgate, R. *Mol. Cell* **2001**, *8*, 7.
(22) Chang, L. M. S.; Bollum, F. J. *J. Biol. Chem.* **1971**, *246*, 5835.
(23) Baril, E. F.; Brown, O. E.; Jenkins, M. D.; Laszlo, J. *Biochemistry* **1971**, *10*, 1981.
(24) García-Díaz, M.; Domínguez, O.; López-Fernández, L. A.; de Lera, L. T.; Saníger, M. L.; Ruiz, J. F.; Párraga, M.; Garcia-Ortiz, M. J.; Kirchhoff, T.; del Mazo, J.; Bernad, A.; Blanco, L. *J. Mol. Biol.* **2000**, *301*, 851.
(25) Dominguez, O.; Ruiz, J. F.; Laín de Lera, T.; Garcia-Díaz, M.; González, M. A.; Kirchhoff, T.; Martínez, -. A., C.; Bernad, A.; Blanco, L. *EMBO J.* **2000**, *19*, 1731.
(26) Wilson, S. H. *Mutat. Res.* **1998**, *407*, 203.
(27) Beard, W. A.; Wilson, S. H. *Mutat. Res.* **2000**, *460*, 231.
(28) Beard, W. A.; Wilson, S. H. *Chem. Biol.* **1998**, *5*, R7.
(29) Wilson, S. H.; Abbotts, J.; Widen, S. *Biochim. Biophys. Acta* **1988**, *949*, 149.
(30) Prasad, R.; Widen, S. G.; Singhal, R. K.; Watkins, J.; Prakash, L.; Wilson, S. H. *Nucleic Acids Res.* **1993**, *21*, 5301.
(31) Bebenek, K.; Garcia-Diaz, M.; Patishall, S. R.; Kunkel, T. A. *J. Biol. Chem.* **2005**, *280*, 20051.
(32) Yamada, K.; Hanaoka, F.; Yamada, M.-a. *J. Biol. Chem.* **1985**, *260*, 10412.
(33) Dianov, G.; Price, A.; Lindahl, T. *Mol. Cell. Biol.* **1992**, *12*, 1605.
(34) Singhal, R. K.; Prasad, R.; Wilson, S. H. *J. Biol. Chem.* **1995**, *270*, 949.
(35) Jenkins, T. M.; Saxena, J. K.; Kumar, A.; Wilson, S. H.; Ackerman, E. J. *Science* **1992**, *258*, 475.
(36) Matsumoto, Y.; Bogenhagen, D. F. *Mol. Cell. Biol.* **1989**, *9*, 3750.
(37) Gu, H.; Marth, J. D.; Orban, P. C.; Mossmann, H.; Rajewsky, K. *Science* **1994**, *265*, 103.
(38) Sobol, R. W.; Horton, J. K.; Kühn, R.; Gu, H.; Singhal, R. K.; Prasad, R.; Rajewsky, K.; Wilson, S. H. *Nature* **1996**, *379*, 183.
(39) Abbotts, J.; SenGupta, D. N.; Zmudzka, B.; Widen, S. G.; Notario, V.; Wilson, S. H. *Biochemistry* **1988**, *27*, 901.

(40) Patterson, T. A.; Little, W.; Cheng, X.; Widen, S. G.; Kumar, A.; Beard, W. A.; Wilson, S. H. *Protein Expression Purif.* **2000**, *18*, 100.
(41) Kim, S. J.; Lewis, M. S.; Knutson, J. R.; Porter, D. K.; Kumar, A.; Wilson, S. H. *J. Mol. Biol.* **1994**, *244*, 224.
(42) Kumar, A.; Widen, S. G.; Williams, K. R.; Kedar, P.; Karpel, R. L.; Wilson, S. H. *J. Biol. Chem.* **1990**, *265*, 2124.
(43) Beard, W. A.; Wilson, S. H. *Methods Enzymol.* **1995**, *262*, 98.
(44) Casas-Finet, J. R.; Kumar, A.; Morris, G.; Wilson, S. H. *J. Biol. Chem.* **1991**, *29*, 19618.
(45) Casas-Finet, J. R.; Kumar, A.; Karpel, R. L.; Wilson, S. H. *Biochemistry* **1992**, *31*, 10272.
(46) Pelletier, H.; Sawaya, M. R.; Kumar, A.; Wilson, S. H.; Kraut, J. *Science* **1994**, *264*, 1891.
(47) Prasad, R.; Beard, W. A.; Strauss, P. R.; Wilson, S. H. *J. Biol. Chem.* **1998**, *273*, 15263.
(48) Chyan, Y. J.; Ackerman, S.; Shepherd, N. S.; McBride, O. W.; Widen, S. G.; Wilson, S. H.; Wood, T. G. *Nucleic Acids Res.* **1994**, *22*, 2719.
(49) Srivastava, D. K.; Evans, R. K.; Kumar, A.; Beard, W. A.; Wilson, S. H. *Biochemistry* **1996**, *35*, 3728.
(50) Ollis, D. L.; Brick, P.; Hamlin, R.; Xuong, N. G.; Steitz, T. A. *Nature* **1985**, *313*, 762.
(51) Kohlstaedt, L. A.; Wang, J.; Friedman, J. M.; Rice, P. A.; Steitz, T. A. *Science* **1992**, *256*, 1783.
(52) Steitz, T. A. *J. Biol. Chem.* **1999**, *274*, 17395.
(53) Beard, W. A.; Wilson, S. H. *Structure (Cambridge)* **2003**, *11*, 489.
(54) Sawaya, M. R.; Pelletier, H.; Kumar, A.; Wilson, S. H.; Kraut, J. *Science* **1994**, *264*, 1930.
(55) Sawaya, M. R.; Prasad, P.; Wilson, S. H.; Kraut, J.; Pelletier, H. *Biochemistry* **1997**, *36*, 11205.
(56) Steitz, T. A.; Smerdon, S. J.; Jager, J.; Joyce, C. M. *Science* **1994**, *266*, 2022.
(57) Beard, W. A.; Shock, D. D.; Yang, X.-P.; DeLauder, S. F.; Wilson, S. H. *J. Biol. Chem.* **2002**, *277*, 8235.
(58) Beard, W. A.; Osheroff, W. P.; Prasad, R.; Sawaya, M. R.; Jaju, M.; Wood, T. G.; Kraut, J.; Kunkel, T. A.; Wilson, S. H. *J. Biol. Chem.* **1996**, *271*, 12141.
(59) Pelletier, H.; Sawaya, M. R.; Wolfle, W.; Wilson, S. H.; Kraut, J. *Biochemistry* **1996**, *35*, 12742.
(60) Pelletier, H.; Sawaya, M. R.; Wolfle, W.; Wilson, S. J.; Kraut, J. *Biochemistry* **1996**, *35*, 12762.
(61) Arndt, J. W.; Gong, W.; Zhong, X.; Showalter, A. K.; Liu, J.; Dunlap, C. A.; Lin, Z.; Paxson, C.; Tsai, M.-D.; Chan, M. K. *Biochemistry* **2001**, *40*, 5368.
(62) Krahn, J. M.; Beard, W. A.; Miller, H.; Grollman, A. P.; Wilson, S. H. *Structure (Cambridge)* **2003**, *11*, 121.
(63) Krahn, J. M.; Beard, W. A.; Wilson, S. H. *Structure (Cambridge)* **2004**, *12*, 1823.
(64) Batra, V. K.; Beard, W. A.; Shock, D. D.; Pedersen, L. C.; Wilson, S. H. *Structure (Cambridge)* **2005**, *13*, 1225.
(65) Prasad, R.; Batra, V. K.; Yang, X.-P.; Krahn, J. M.; Pedersen, L. C.; Beard, W. A.; Wilson, S. H. *DNA Repair* **2005**, *4*, 1347.
(66) Batra, V. K.; Beard, W. A.; Shock, D. D.; Krahn, J. M.; Pedersen, L. C.; Wilson, S. H. *Structure (Cambridge)*, in press.
(67) Sobol, R. W.; Prasad, R.; Evenski, A.; Baker, A.; Yang, X. P.; Horton, J. K.; Wilson, S. H. *Nature* **2000**, *405*, 807.
(68) Vande Berg, B. J.; Beard, W. A.; Wilson, S. H. *J. Biol. Chem.* **2001**, *276*, 3408.
(69) Osheroff, W. P.; Beard, W. A.; Yin, S.; Wilson, S. H.; Kunkel, T. A. *J. Biol. Chem.* **2000**, *275*, 28033.
(70) Doublié, S.; Sawaya, M. R.; Ellenberger, T. *Structure (Cambridge)* **1999**, *7*, R31.
(71) Ling, W.; Boudsocq, F.; Woodgate, R.; Yang, W. *Mol. Cell* **2004**, *13*, 751.
(72) Garcia-Diaz, M.; Bebenek, K.; Krahn, J. M.; Kunkel, T. A.; Pedersen, L. C. *Nat. Struct. Mol. Biol.* **2005**, *12*, 97.
(73) Mullen, G. P.; Wilson, S. H. *Biochemistry* **1997**, *36*, 4713.
(74) Doherty, A. J.; Serpell, L. C.; Ponting, C. P. *Nucleic Acids Res.* **1996**, *24*, 2488.
(75) Seeman, N. C.; Rosenberg, J. M.; Rich, A. *Proc. Natl. Acad. Sci. U.S.A.* **1976**, *73*, 804.
(76) Beard, W. A.; Shock, D. D.; Vande Berg, B. J.; Wilson, S. H. *J. Biol. Chem.* **2002**, *277*, 47393.
(77) Mildvan, A. S. *Proteins* **1997**, *29*, 401.
(78) Johnson, S. J.; Taylor, J. S.; Beese, L. S. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 3895.
(79) Harding, M. M. *Acta Crystallogr., D* **2002**, *58*, 872.
(80) Harding, M. M. *Acta Crystallogr., D* **2001**, *57*, 401.
(81) Johnson, S. J.; Beese, L. S. *Cell* **2004**, *116*, 803.
(82) Hunter, W. N.; Brown, T.; Anand, N. N.; Kennard, O. *Nature* **1986**, *320*, 552.
(83) Kalnik, M. W.; Kouchakdjian, M.; Li, B. F.; Swann, P. F.; Patel, D. J. *Biochemistry* **1988**, *27*, 100.
(84) Sarma, M. H.; Gupta, G.; Sarma, R. H.; Bald, R.; Engelke, U.; Oei, S. L.; Gessner, R.; Erdmann, V. A. *Biochemistry* **1987**, *26*, 7707.
(85) Boulard, Y.; Cognet, J. A. H.; Fazakerley, G. V. *J. Mol. Biol.* **1997**, *268*, 331.
(86) Yang, L.; Beard, W. A.; Wilson, S. H.; Broyde, S.; Schlick, T. *J. Mol. Biol.* **2002**, *317*, 679.
(87) Chou, K.-M.; Chou, Y.-C. *Nature* **2002**, *415*, 655.
(88) Kunkel, T. A.; Bebenek, K. *Annu. Rev. Biochem.* **2000**, *69*, 497.
(89) Horton, J. K.; Prasad, R.; Hou, E.; Wilson, S. H. *J. Biol. Chem.* **2000**, *275*, 2211.
(90) Kunkel, T. A. *J. Biol. Chem.* **1985**, *260*, 5787.
(91) Osheroff, W. P.; Jung, H. K.; Beard, W. A.; Wilson, S. H.; Kunkel, T. A. *J. Biol. Chem.* **1999**, *274*, 3642.
(92) Beard, W. A.; Shock, D. D.; Wilson, S. H. *J. Biol. Chem.* **2004**, *279*, 31921.
(93) Goodman, M. F.; Creighton, S.; Bloom, L. B.; Petruska, J. *Crit. Rev. Biochem. Mol. Biol.* **1993**, *28*, 83.
(94) Miller, H.; Prasad, R.; Wilson, S. H.; Johnson, F.; Grollman, A. P. *Biochemistry* **2000**, *39*, 1029.
(95) Efrati, E.; Tocco, G.; Eritja, R.; Wilson, S. H.; Goodman, M. F. *J. Biol. Chem.* **1999**, *274*, 15920.
(96) Deterding, L. J.; Prasad, R.; Mullen, G. P.; Wilson, S. H.; Tomer, K. B. *J. Biol. Chem.* **2000**, *275*, 10463.
(97) Prasad, R.; Beard, W. A.; Chyan, J. Y.; Maciejewski, M. W.; Mullen, G. P.; Wilson, S. H. *J. Biol. Chem.* **1998**, *273*, 11121.
(98) Maciejewski, M. W.; Liu, D.; Prasad, R.; Wilson, S. H.; Mullen, G. P. *J. Mol. Biol.* **2000**, *296*, 229.
(99) Caldecott, K. W. *DNA Repair* **2003**, *2*, 955.
(100) Gryk, M. R.; Marintchev, A.; Maciejewski, M. W.; Robertson, A.; Wilson, S. H.; Mullen, G. P. *Structure (Cambridge)* **2002**, *10*, 1709.
(101) Gryk, M. R.; Maciejewski, M. W.; Robertson, A.; Mullen, M. A.; Wilson, S. H.; Mullen, G. P. *J. Biomol. NMR* **2002**, *22*, 197.
(102) Dianova, I. I.; Sleeth, K. M.; Allinson, S. L.; Parsons, J. L.; Breslin, C.; Caldecott, K. W.; Dianov, G. L. *Nucleic Acids Res.* **2004**, *32*, 2550.
(103) Wong, H.-K.; Wilson, D. M., III. *J. Cell. Biochem.* **2005**, *95*, 794.
(104) Mizushina, Y.; Ohkubo, T.; Date, T.; Yamaguchi, T.; Saneyoshi, M.; Sugawara, F.; Sakaguchi, K. *J. Biol. Chem.* **1999**, *274*, 25599.
(105) Mizushina, Y.; Ohkubo, T.; Sugawara, F.; Sakaguchi, K. *Biochemistry* **2000**, *39*, 12606.
(106) Sun, D.-A.; Starck, S. R.; Locke, E. P.; Hecht, S. M. *J. Nat. Prod.* **1999**, *62*, 1110.
(107) Hu, H. Y.; Horton, J. K.; Gryk, M. R.; Prasad, R.; Naron, J. M.; Sun, D.-A.; Hecht, S. M.; Wilson, S. H.; Mullen, G. P. *J. Biol. Chem.* **2004**, *279*, 39736.
(108) Tanabe, K.; Bohn, E. W.; Wilson, S. H. *Biochemistry* **1979**, *18*, 3401.
(109) Johnson, K. A. *Annu. Rev. Biochem.* **1993**, *62*, 685.
(110) Joyce, C. M.; Benkovic, S. J. *Biochemistry* **2004**, *43*, 14317.
(111) Rothwell, P. J.; Mitaksov, V.; Waksman, G. *Mol. Cell* **2005**, *19*, 345.
(112) Beard, W. A.; Wilson, S. H. In *A Practical Approach: HIV Volume 2*; Karn, J., Ed.; IRL Press: New York, 1995.
(113) Johnson, K. A. In *The Enzymes: Mechanisms of catalysis*; Sigman, D. S., Ed.; Academic Press: New York, 1992; Vol. XX.
(114) Prasad, R.; Beard, W. A.; Wilson, S. H. *J. Biol. Chem.* **1994**, *269*, 18096.
(115) Singhal, R. K.; Wilson, S. H. *J. Biol. Chem.* **1993**, *268*, 15906.
(116) Herschlag, D.; Piccirilli, J. A.; Cech, T. R. *Biochemistry* **1991**, *30*, 4844.
(117) Liu, J.; Tsai, M. D. *Biochemistry* **2001**, *40*, 9014.
(118) Osheroff, W. P.; Beard, W. A.; Wilson, S. H.; Kunkel, T. A. *J. Biol. Chem.* **1999**, *274*, 20749.
(119) Showalter, A. K.; Tsai, M.-D. *Biochemistry* **2002**, *41*, 10571.
(120) Radhakrishnan, R.; Schlick, T. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 5970.
(121) Arora, K.; Schlick, T. *Biophys. J.* **2004**, *87*, 3088.
(122) Arora, K.; Schlick, T. *J. Phys. Chem. B* **2005**, *109*, 5358.
(123) Post, C. B.; Ray, W. J., Jr. *Biochemistry* **1995**, *34*, 15881.
(124) Opresko, P. L.; Sweasy, J. B.; Eckert, K. A. *Biochemistry* **1998**, *37*, 2111.
(125) Clairmont, C. A.; Narayanan, L.; Sun, K. W.; Glazer, P. M.; Sweasy, J. B. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9580.
(126) Opresko, P. L.; Shiman, R.; Eckert, K. A. *Biochemistry* **2000**, *39*, 11399.
(127) Shah, A. M.; Li, S.-X.; Anderson, K. S.; Sweasy, J. B. *J. Biol. Chem.* **2001**, *276*, 10824.
(128) Shah, A. M.; Maitra, M.; Sweasy, J. B. *Biochemistry* **2003**, *42*, 10709.
(129) Starcevic, D.; Dalal, S.; Sweasy, J. *Biochemistry* **2005**, *44*, 3775.
(130) Starcevic, D.; Dalal, S.; Jaeger, J.; Sweasy, J. B. *J. Biol. Chem.* **2005**, *280*, 28388.
(131) Dalal, S.; Kosa, J. L.; Sweasy, J. B. *J. Biol. Chem.* **2004**, *279*, 577.
(132) Kosa, J. L.; Sweasy, J. B. *J. Biol. Chem.* **1999**, *274*, 3851.
(133) Hamid, S.; Eckert, K. A. *Biochemistry* **2005**, *44*, 10378.
(134) Kosa, J. L.; Sweasy, J. B. *J. Biol. Chem.* **1999**, *274*, 35866.

(135) Maitra, M.; Gudzelak, A., Jr.; Li, S.-X.; Matsumoto, Y.; Eckert, K. A.; Jager, J.; Sweasy, J. B. *J. Biol. Chem.* **2002**, *277*, 35550.
(136) Lang, T.; Maitra, M.; Starcevic, D.; Li, S.-X.; Sweasy, J. B. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 6074.
(137) Date, T.; Yamamoto, S.; Tanihara, K.; Nishimoto, Y.; Matsukage, A. *Biochemistry* **1991**, *30*, 5286.
(138) Menge, K. L.; Hostomsky, Z.; Nodes, B. R.; Hudson, G. O.; Rahmati, S.; Moomaw, E. W.; Almassy, R. J.; Hostomska, Z. *Biochemistry* **1995**, *34*, 15934.
(139) Date, T.; Yamamoto, S.; Tanihara, K.; Nishimoto, Y.; Liu, N.; Matsukage, A. *Biochemistry* **1990**, *29*, 5027.
(140) Kraynov, V. S.; Showalter, A. K.; Liu, J.; Zhong, X.; Tsai, M.-D. *Biochemistry* **2000**, *39*, 16008.
(141) Lavrik, O. I.; Prasad, R.; Beard, W. A.; Safronov, I. V.; Dobrikov, M. I.; Srivastava, D. K.; Shishkin, G. V.; Wood, T. G.; Wilson, S. H. *J. Biol. Chem.* **1996**, *271*, 21891.
(142) Skandalis, A.; Loeb, L. A. *Nucleic Acids Res.* **2001**, *29*, 2418.
(143) Kraynov, V. S.; Werneburg, B. G.; Zhong, X. J.; Lee, H.; Ahn, J. W.; Tsai, M. D. *Biochem. J.* **1997**, *323*, 103.
(144) Yang, L.; Beard, W. A.; Wilson, S. H.; Broyde, S.; Schlick, T. *Biophys. J.* **2004**, *86*, 3392.
(145) Beard, W. A.; Wilson, S. H. *Structure (Cambridge)* **2001**, *9*, 759.
(146) Iwanaga, A.; Ouchida, M.; Miyazaki, K.; Hori, K.; Mukai, T. *Mutat. Res.* **1999**, *435*, 121.
(147) Zhong, X. J.; Patel, S. S.; Werneburg, B. G.; Tsai, M. D. *Biochemistry* **1997**, *36*, 11891.
(148) Bakhtina, M.; Lee, S.; Wang, Y.; Dunlap, C.; Lamarche, B.; Tsai, M.-D. *Biochemistry* **2005**, *44*, 5177.
(149) Kim, S.-J.; Beard, W. A.; Harvey, J.; Shock, D. D.; Knutson, J. R.; Wilson, S. H. *J. Biol. Chem.* **2003**, *278*, 5072.
(150) Bose-Basu, B.; DeRose, E. F.; Kirby, T. W.; Mueller, G. A.; Beard, W. A.; Wilson, S. H.; London, R. E. *Biochemistry* **2004**, *43*, 8911.
(151) Kirby, T. W.; DeRose, E. F.; Beard, W. A.; Wilson, S. H.; London, R. E. *Biochemistry* **2005**, *44*, 15230.
(152) Kool, E. T. *Annu. Rev. Biochem.* **2002**, *71*, 191.
(153) Morales, J. C.; Kool, E. T. *J. Am. Chem. Soc.* **2000**, *122*, 1001.
(154) Yang, L.; Beard, W.; Wilson, S.; Roux, B.; Broyde, S.; Schlick, T. *J. Mol. Biol.* **2002**, *321*, 459.
(155) Yang, L.; Beard, W. A.; Wilson, S. H.; Broyde, S.; Schlick, T. *Biophys. J.* **2003**, *86*, 3392.
(156) Arora, K.; Beard, W. A.; Wilson, S. H.; Schlick, T. *Biochemistry* **2005**, *44*, 13328.
(157) Abashkin, Y. G.; Erickson, J. W.; Burt, S. K. *J. Phys. Chem. B* **2001**, *105*, 287.
(158) Florian, J.; Goodman, M. F.; Warshel, A. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6819.
(159) Wilde, J. A.; Bolton, P. H.; Mazumder, A.; Manoharan, M.; Gerlt, J. A. *J. Am. Chem. Soc.* **1989**, *111*, 1894.
(160) Zharkov, D. O.; Rosenquist, T. A.; Gerchman, S. E.; Grollman, A. P. *J. Biol. Chem.* **2000**, *275*, 28607.
(161) Fromme, J. C.; Verdine, G. L. *EMBO J.* **2003**, *22*, 3461.
(162) Mazumder, A.; Gerlt, J. A.; Absalon, M. J.; Stubbe, J.; Cunningham, R. P.; Withka, J.; Bolton, P. H. *Biochemistry* **1991**, *30*, 1119.
(163) Feng, J.-a.; Crasto, C. J.; Matsumoto, Y. *Biochemistry* **1998**, *37*, 9605.
(164) Husain, I.; Morton, B. S.; Beard, W. A.; Singhal, R. K.; Prasad, R.; Wilson, S. H.; Besterman, J. M. *Nucleic Acids Res.* **1995**, *23*, 1597.
(165) Starcevic, D.; Dalal, S.; Sweasy, J. B. *Cell Cycle* **2004**, *3*, 998.
(166) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. *J. Comput. Chem.* **2004**, *25*, 1605.
(167) Sanner, M. F.; Olson, A. J.; Spehner, J. C. *Biopolymers* **1996**, *38*, 305.
(168) Pelletier, H.; Sawaya, M. R. *Biochemistry* **1996**, *35*, 12778.

CR0404904