Crystal Structures of Human DNA Polymerase β Complexed with DNA: Implications for Catalytic Mechanism, Processivity, and Fidelity[†,‡]

Huguette Pelletier,*[,§] Michael R. Sawaya,[||] William Wolfle,[⊥] Samuel H. Wilson,[⊥] and Joseph Kraut[||]

Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, California 92093-0506, Sealy Center for Molecular Science, University of Texas Medical Branch, Galveston, Texas 77555-1051, and Verna and Marrs McLean Department of Biochemistry, Baylor College of Medicine, One Baylor Plaza, Houston, Texas 77030

Received December 14, 1995; Revised Manuscript Received June 18, 1996[⊗]

ABSTRACT: Mammalian DNA polymerase β (pol β) is a small (39 kDa) DNA gap-filling enzyme that comprises an amino-terminal 8-kDa domain and a carboxy-terminal 31-kDa domain. In the work reported here, crystal structures of human pol β complexed with blunt-ended segments of DNA show that, although the crystals belong to a different space group, the DNA is nevertheless bound in the pol β binding channel in the same way as the DNA in previously reported structures of rat pol β complexed with a template−primer and ddCTP [Pelletier, H., Sawaya, M. R., Kumar, A., Wilson, S. H., & Kraut, J. (1994) *Science 264,* 1891−1903]. The 8-kDa domain is in one of three previously observed positions relative to the 31-kDa domain, suggesting that the 8-kDa domain may assume only a small number of stable conformations. The thumb subdomain is in a more open position in the human pol β−DNA binary complex than it is in the rat pol β−DNA−ddCTP ternary complex, and a closing thumb upon nucleotide binding could represent the rate-limiting conformational change that has been observed in pre-steady-state kinetic studies. Intermolecular contacts between the DNA and the 8-kDa domain of a symmetry-related pol β molecule reveal a plausible binding site on the 8-kDa domain for the downstream oligonucleotide of a gapped-DNA substrate; in addition to a lysine-rich binding pocket that accommodates a 5′-PO_4 end group, the 8-kDa domain also contains a newly discovered helix−hairpin−helix (HhH) motif that binds to DNA in the same way as does a structurally and sequentially homologous HhH motif in the 31-kDa domain. DNA binding by both HhH motifs is facilitated by a metal ion. In that HhH motifs have been identified in other DNA repair enzymes and DNA polymerases, the HhH−DNA interactions observed in pol β may be applicable to a broad range of DNA binding proteins. The sequence similarity between the HhH motif of endonuclease III from *Escherichia coli* and the HhH motif of the 8-kDa domain of pol β is particularly striking in that all of the conserved residues are clustered in one short sequence segment, LPGVGXK, where LPGV corresponds to a type II β-turn (the hairpin turn), and GXK corresponds to a part of the HhH motif that is proposed to be critical for DNA binding and catalysis for both enzymes. These results suggest that endonuclease III and the 8-kDa domain of pol β may employ a similar mode of DNA binding and may have similar catalytic mechanisms for their respective DNA lyase activities. A model for productive binding of pol β to a gapped-DNA substrate requires a 90° bend in the single-stranded template, which could enhance nucleotide selectivity during DNA repair or replication.

DNA repair is an ongoing biological process that is vital to all organisms for the protection and preservation of genomic DNA (Friedberg et al., 1995). The importance of DNA repair in human biological function was demonstrated as early as 1968 when it was reported that the autosomal recessive genetic disorder xeroderma pigmentosum (XP)[1] was caused by a breakdown in the DNA repair mechanism of the cell (Cleaver 1968). More recently, DNA repair has received much attention because of evidence linking yet another genetic disease, hereditary nonpolyposis colorectal cancer (HNPCC), with mutations in genes encoding DNA repair proteins (Peltomaki et al., 1993; Aaltonen et al., 1993; Fishel et al., 1993; Leach et al., 1993; Parsons et al., 1993). Although models of many different DNA repair mechanisms have now emerged [for instance, XP is associated with a breakdown in nucleotide excision repair (NER) (Cleaver & Kramer, 1994), while HNPCC is associated with a breakdown in DNA mismatch repair (Modrich, 1994)], most DNA repair mechanisms have at least one thing in common: once the damaged or mutated DNA has been located and excised by DNA glycosylases and/or endonucleases, the resulting

[†] Supported by NIH Grants GM52860 and ES06839 and by R. A. Welch Foundation Grant H-1265.

[‡] Coordinates and reflection data are available from the Brookhaven Protein Data Bank and may also be obtained by sending an e-mail request to hug@bcm.tmc.edu.

* To whom correspondence should be addressed.

[§] Baylor College of Medicine.

[||] University of California, San Diego.

[⊥] Sealy Center for Molecular Science.

[⊗] Abstract published in *Advance ACS Abstracts,* September 15, 1996.

[1] Abbreviations: AP, apurinic−apyrimidinic (abasic); BER, base excision repair; ddCTP, dideoxycytidine triphosphate; ddNTP, dideoxynucleoside triphosphate; ddNMP, dideoxynucleoside monophosphate; dNMP, deoxynucleoside monophosphate; dRP, deoxyribophosphate; HEPES, *N*-(2-hydroxyethyl)piperazine-*N*′-2-ethanesulfonic acid, HNPCC, hereditary nonpolyposis colorectal cancer; HhH, helix−hairpin−helix; MES, 2-(*N*-morpholino)ethanesulfonic acid; NER, nucleotide excision repair; nt, nucleotide; PEG, poly(ethylene glycol); pdT4, a 5′-phosphorylated oligonucleotide with the sequence TTTT; HIV-1 RT, human immunodeficiency virus type 1 reverse transcriptase; XP, xeroderma pigmentosum; 6bp, structure of pol β complexed with 6 base pairs of DNA; 7bp, structure of pol β complexed with 7 base pairs of DNA.

S0006-2960(95)02955-2 CCC: $12.00

gap that is left behind in the DNA must be filled in by one or more DNA polymerases and then sealed with a DNA ligase. Polymerases therefore play a critical role in most DNA repair processes, and as a last line of defense against DNA errors, it is important that repair polymerases fill gaps accurately and completely so as not to make the entire repair process futile.

Although it cannot be ruled out that every one of the four known nuclear polymerases (α, β, δ, and ϵ) may participate in DNA repair to some extent, only pol β and pol ϵ have repeatedly been singled out primarily as repair polymerases (Kornberg & Baker, 1992). Pol ϵ's participation in DNA repair was demonstrated by studies with permeabilized human fibroblasts that had been damaged with UV irradiation, in which DNA repair activity could be restored by addition of pol ϵ (Nishida et al., 1988). In addition, studies of temperature-sensitive mutants in yeast genes encoding pol α, pol δ, and pol ϵ showed that pol ϵ participates in base excision repair (BER) processes in yeast (Wang et al., 1993). In neither study could a direct correlation be made between DNA repair and either pol α or pol δ. Much of the early evidence linking pol β with DNA repair came from the observation that pol β's characteristics seem to be compatible with a role in DNA repair. For instance, (i) the gene encoding pol β is constitutively expressed (Zmudzka et al., 1988) (that is, pol β is present in the cell at all times, which is expected for a repair polymerase that must work continuously), (ii) expression of the pol β gene can be enhanced by DNA-damaging agents (Fornace et al., 1989), and (iii) the preferred substrate of pol β, at least in vitro, has been shown to be DNA gaps that are 6 nucleotides in length or shorter (Singhal & Wilson, 1993; Prasad et al., 1994). In addition, numerous studies with mammalian cell extracts have shown that pol β is more efficient than other polymerases at filling in single-nucleotide DNA gaps during BER processes (Matsumoto & Bogenhagen, 1989; Wiebauer & Jiricny, 1990; Dianov et al., 1992; Singhal et al., 1995). Finally, the recent discovery that the 8-kDa domain of pol β possesses lyase activity (Matsumoto & Kim, 1995) suggests that, in addition to filling in DNA gaps, pol β may also function to remove the deoxyribophosphate (dRP) moiety of processed apurinic−apyrimidinic (AP) sites just prior to the gap-filling step in BER (Matsumoto & Kim, 1995). Although all of the aforementioned in vitro studies have thus far offered only indirect evidence of pol β's role in DNA repair, they have since been confirmed by more direct studies with mouse cell lines that are homozygous for a deletion mutation in the pol β gene (Gu et al., 1994), where recent results suggest that pol β participates specifically in BER processes in vivo[2] (Sobol et al., 1995).

In addition to its importance in DNA repair studies, pol β, at 39 kDa, is conveniently small for a polymerase and can serve as a model for study of the mechanism of nucleotidyl transfer (the polymerase reaction), as well as other enzymological properties of polymerases, such as processivity (the ability to slide along the DNA or RNA template and incorporate more than one nucleotide on a growing primer strand before dissociation) and fidelity (the ability to copy a DNA or RNA template accurately). Before accepting pol β as a model system, however, it is critical to establish two things. First, it must be shown that what is observed in the crystal structures is relevant to pol β's function in vivo; and second, it is important to distinguish the structural and mechanistic features of pol β that are general and applicable to all polymerases from those that may be peculiar only to pol β. As an example, The X-ray crystal structure determinations of pol β ternary complexes (two substrates bound, a DNA template−primer and ddCTP) allowed a detailed nucleotidyl transfer mechanism for pol β to be presented (Pelletier et al., 1994), and because structural alignments with other polymerases revealed a common active-site geometry comprising highly conserved, catalytically important carboxylate residues (Sawaya et al., 1994), it was proposed that this mechanism is common to all polymerases (Pelletier et al., 1994; Sawaya et al., 1994). In contrast, pol β has other structural features, such as an 8-kDa domain or a *cis* peptide bond located near the active site, which have not been observed in other polymerase crystal structures and which may have arisen specifically to enhance pol β's short gap-filling activities in vivo. There has been some debate, however, concerning interpretations of the crystallographic data (Steitz et al., 1994; Pelletier, 1994; Wohrl et al., 1995; Arnold et al., 1995; Patel et al., 1995) with two reservations being (i) that the pol β−DNA−ddCTP ternary complex crystal structures might not represent catalytically productive binding of pol β to the DNA or (ii) because of large structural and functional differences between pol β and other polymerases, it is not clear how information about pol β can be extended to other polymerases. As an example of one of the anomalies that has arisen, the orientation of the DNA template−primer in the pol β−DNA−ddCTP ternary complex structures (Pelletier et al., 1994) differs significantly from a proposed pol β 31-kDa domain−DNA model (Davies et al., 1994), which was based on analogies with the HIV-1 RT−DNA crystal structure (Jacobo-Molina et al., 1993). Anomalies such as this have brought into question the orientation of DNA binding for all polymerase−DNA crystal structures and models.

The crystal structures presented here of human pol β complexed with DNA have shed some light on the debate. First, the DNA in the new cocrystals is bound in the pol β binding channel in the same way as the DNA in two previously observed pol β−DNA complex structures (Pelletier et al., 1994), offering a third example of the same mode of DNA binding and further supporting the idea that irrelevant crystal-packing phenomena or crystallization conditions are not the dominant forces determining protein−DNA interactions in any of the pol β−DNA cocrystals. In addition, and perhaps the strongest evidence in support of biological relevance (not just for pol β but for any enzyme crystal structure) is that catalysis occurs directly in the crystals; X-ray structural analysis has shown that the nucleotidyl transfer reaction takes place, at least for one turnover, when the human pol β−DNA binary complex crystals are soaked in the presence of nucleotide substrates (Pelletier et al., 1996). Finally, although it has been observed that pol

[2] Unlike BER, where only a single-nucleotide gap is left to be filled in by pol β, other DNA repair mechanisms, particularly those invoked for NER, have been shown to leave gaps in the DNA that can be as long as 30−50 nucleotides. Taking into consideration the observation that pol ϵ, unlike pol β, may not be capable of filling in DNA gaps to completion at physiological ionic strengths (Randahl et al., 1988), a working hypothesis has emerged for these other repair mechanisms whereby a relatively large DNA gap may be partially filled in by pol ϵ, followed by pol β taking over to fill the smaller gap to completion. Although this is an attractive hypothesis, it is as yet not clear whether pol β is directly involved in NER (Horton et al., 1995; Sobol et al., 1996), nor is it clear that pol ϵ is incapable of filling in DNA gaps to completion in vivo (Mozzherin & Fisher, 1996).

β's thumb subdomain displays more flexibility than the fingers subdomain (Pelletier et al., 1994), as is the case with other polymerase crystal structures (Beese et al., 1993; Sousa et al., 1994; Rodgers et al., 1995), it is now evident that the position of the thumb in a pol β−DNA complex depends on whether or not a nucleotide is bound in the active site. These observations not only allow the common polymerase mechanism to be expanded to include a mobile thumb subdomain but also raise the possibility that a change from an open thumb, as in the pol β−DNA binary complex, to a closed thumb, as in the pol β−DNA−ddCTP ternary complex, represents the rate-limiting conformational change that is commonly observed in pre-steady-state kinetic studies performed on DNA and RNA polymerases (Kuchta et al., 1987; Patel et al., 1991; Wong et al., 1991; Erie et al., 1993; Spence et al., 1995; Frey et al., 1995; Werneburg et al., 1996).

In addition to helping to extend our knowledge about the nucleotidyl transfer mechanism, the work presented here attempts to address questions about polymerase processivity and fidelity as well. With regard to processivity, the new observation that there are two metal ion binding sites in pol β, away from the active site, which interact directly with the DNA backbone suggests that metal ions may act as processivity factors at the pol β−DNA interface—that is, metal ions may increase the affinity of pol β for the DNA substrate, while at the same time allowing pol β to move smoothly along the DNA. With regard to fidelity, close inspection of the protein scaffolding in the immediate vicinity of the metal ion binding sites in pol β has revealed a pair of previously overlooked, structurally homologous DNA binding motifs. One motif is located in the fingers subdomain of pol β and binds to the DNA primer near the 3′-OH terminus, whereas the other motif, which is located in the 8-kDa domain, is adjacent to a lysine-rich 5′-PO_4 binding pocket and most likely forms the binding site for the 5′ terminus of a gapped-DNA substrate. The newly discovered DNA binding motifs allow a reasonable model for the pol β−gapped DNA complex to be presented, and the most salient feature of this model is that it requires a 90° bend in the single-stranded DNA template just above pol β's active site. Such a sharp bend in the template could increase pol β's sensitivity to DNA mismatches prior to catalysis and may be a general mechanistic feature utilized by polymerases to improve nucleotide selectivity (increase fidelity) during DNA replication or transcription.

EXPERIMENTAL PROCEDURES

Crystallizations. Recombinant human DNA pol β (Sen-Gupta et al., 1986) was overexpressed in *Escherichia coli* and purified as described (Abbotts et al., 1988). After purification, the protein was washed at least three times in a microconcentrator (Centricon-10, Amicon) with a buffer solution (10 mM ammonium sulfate and 0.1 M MES, pH 7.0), then concentrated to 20 mg/mL and stored at −80 °C in sealed Eppendorf tubes (120-μL portions). All protein−DNA sample preparations and crystallizations were performed at room temperature.

In a previous report it was suggested that the DNA template of the template−primer utilized in crystallization experiments was probably too short (only 11 nucleotides) to interact with the 8-kDa domain of pol β, causing the template overhang to be disordered and unobservable in the crystal structures (Pelletier et al., 1994). In addition, recent results suggest that the best substrate for pol β is not a template−primer, as it is with most other polymerases, but is instead DNA with a gap of about 6 nt or less that possesses a 5′-PO_4 group on the downstream oligonucleotide (Prasad et al., 1994). Finally, although crystal structures of pol β apoenzyme (Sawaya et al., 1994) and of ternary complexes with template−primer and ddCTP substrates (Pelletier et al., 1994) have been determined, still missing from our structural studies of pol β was a binary complex, in which pol β was bound to a DNA template−primer alone, with no nucleotide substrate present in the active site. We therefore set out to try to crystallize various binary and ternary complexes of human DNA pol β with the following substrates: (i) a DNA template−primer in which the template strand was either 16, 18, or 20 nt long and the primer was 6 nt; (ii) the template−primers in (i) but with the addition of excess dideoxynucleoside triphosphate (ddNTP), so that a ternary complex such as those previously reported (Pelletier et al., 1994) could result; (iii) gapped DNA substrates in which the template was 13−20 nt long, the primer was 6 nt, and the downstream oligonucleotide not only varied in length from 4 to 9 nt but also varied in that some oligonucleotides possessed a 5′-PO_4 group and others did not (the anticipated result was a broad range of gapped DNA substrates that varied from 1 to 7 nt in gap size); and (iv) many of the gapped substrates in (iii) in the presence of excess ddNTP.

The most common crystal that resulted from numerous crystallization experiments was diamond-shaped and ranged in size from as large as 1.0 mm in length to microcrystalline showers throughout a given crystallization drop. Preliminary X-ray diffraction studies showed that these crystals diffracted only to a resolution of about 5−6 Å and were very sensitive to X-rays, thus making space group determination difficult. In that rat DNA pol β apoenzyme crystals also diffracted relatively poorly (Sawaya et al., 1994), it may be that the diamond-shaped crystals that showed up in so many of our crystallization drops were actually just crystals of human DNA pol β apoenzyme—after failure to form a strong complex with DNA.

A second crystal form that was more amenable to crystallographic studies did occur, however, but only in the presence of gapped-DNA substrates (a possible explanation for why all of the nongapped template−primer samples, which possessed relatively long template overhangs, repeatedly failed to cocrystallize with pol β became apparent later, as discussed below). The useful human DNA pol β crystals are orthorhombic and belong to space group $P2_12_12_1$ (a = 179.0, b = 57.5, c = 48.5 Å) with one pol β molecule in an asymmetric unit. They can be grown in the presence of the following gapped DNA substrates: (i) a 14-nt template, a 6-nt primer, and pdT4 as the downstream oligonucleotide; (ii) a 15-nt template, a 6-nt primer, and pdT4; (iii) a 18-nt template, a 6-nt primer, and a 6-nt downstream oligonucleotide; and (iv) a 20-nt template, a 6-nt primer, and a 7-nt downstream oligonucleotide. Although a few crystals grown in the presence of pdT4 [(i) and (ii) above] grew large enough for space group determinations and unit cell parameter measurements with an X-ray precession camera, obtaining crystals suitable for data collection and structure determination proved difficult, and consequently only detailed crystallization conditions for the two solved and refined human pol β−DNA structures, those from crystals grown in the presence of (iii) and (iv), are described here. In that

Table 1: Data Collection and Refinement Statistics

crystal[a]	PDB[b] code	d_{min} (Å)	$\langle I/\sigma_I \rangle$[c]	total obs./unique refs.	completeness (%)	R_{sym}[d] (%)	number of atoms: protein	DNA	water	final R[e] (%)	rms deviation[f]: bond (Å)	angle (deg)
6bp	9ICW	2.6	1.8	42717/15219	94	6.3	2624	247	156	17.5	0.021	2.9
7bp	9ICJ	3.1	1.8	18480/8473	88	5.6	2624	288	136	14.9[g]	0.020	2.9
6bp, no 5′-PO_4	9ICM	2.9	2.0	20690/10458	89	4.8	2624	242	153	15.7[g]	0.018	2.8
7bp, no 5′-PO_4	8ICC	2.8	1.9	20810/11845	91	4.2	2624	283	150	17.0[g]	0.019	3.0
6bp, no gap	9ICX	2.6	2.0	24645/14354	89	5.0	2625	248	149	17.8	0.021	3.0
7bp, no gap	9ICY	3.0	1.9	19438/9958	94	6.3	2623	294	139	18.1	0.020	3.0

[a] See Experimental Procedures for the definitions of 6bp, 7bp, no 5′-PO_4, and no gap crystals. Data from two crystals were combined for the first data set listed in the table and all other data sets were obtained from single crystals. [b] PDB = Brookhaven Protein Data Bank. Reflection data and coordinate files are available under the PDB code name listed for each structure. [c] $\langle I/\sigma_I \rangle$ is the average ratio of observed intensity to background in the highest resolution shell of reflections. [d] $R_{sym} = \sum|I_{obs} - I_{avg}|/\sum I_{avg}$. [e] Final $R = \sum|F_{obs} - F_{calc}|/\sum F_{obs}$, including all data between 20 Å and the maximum resolution. As a part of the TNT program package (Tronrud et al., 1987), a scaling function was applied during refinements to compensate for a lack of solvent continuum in the model for the low-order reflections (Moews & Kretsinger, 1975). [f] The rms bond lengths and angles are the root-mean-square deviations from ideal values. [g] These *R*-factors are relatively low because the refined, high-resolution 6bp structure (the first structure listed in the table) had been utilized as a starting structure for the lower resolution data sets. Refinement of a high-resolution structure against a closely related low-resolution data set results in a lower *R*-factor than would be expected if phases for the low-resolution structure had been obtained by other means, such as multiple isomorphous replacement (MIR) phasing. See the PDB files for additional statistics and information about refinements.

the two types of gapped-DNA substrates utilized for crystallizations differed in the number of base pairs (from 6 to 7) at the downstream end of the gap, the two different pol β−DNA crystals are henceforth referred to as the 6bp and 7bp crystals, respectively (Table 1).

DNA oligonucleotides were purchased from Integrated DNA Technologies, Inc. (Coralville, IA). DNA sequences corresponding to the template, the primer, and downstream oligonucleotide for the 6bp samples are 5′-CATCTGTTCTAAGCGCCG-3′, 5′-CGGCGC-3′, and 5′-PO_4-CAGATG-3′, respectively, and for the 7bp samples, 5′-CATTAGAAAGGGAAGCGCCG-3′, 5′-CGGCGC-3′, and 5′-PO_4-TCTAATG-3′, respectively, where 5′-PO_4 indicates that the 5′ termini of these oligonucleotides were phosphorylated. Approximately 1.0−1.2 mg of the template, 0.4−0.6 mg of the primer, and 0.4−0.6 mg of the downstream oligonucleotide were mixed together in 240 μL of buffer (20 mM $MgCl_2$ and 0.1 M MES, pH 6.5) in a 1-mL Eppendorf tube. To facilitate annealing of the gapped-DNA sample, the sealed Eppendorf tube was placed in a water bath, heated to 100 °C for 3 min, and allowed to cool slowly, over the course of an hour, to room temperature. Two portions of human DNA pol β (240 μL of a solution containing 20 mg/mL) were then thawed and added, resulting in a protein−DNA mixture containing pol β at approximately 10 mg/mL, 10 mM $MgCl_2$, and an excess of template:primer:downstream oligonucleotide in a molar ratios of 1.5:2.0:2.0 relative to the amount of protein. It was common practice during our initial crystallization experiments to divide the protein−DNA sample at this stage into two equal portions—one to be left as is and the other to have added to it a 30-fold molar excess, relative to protein, of a corresponding ddNTP, so that both samples (a potential pol β−DNA binary complex and a potential pol β−DNA−ddNTP ternary complex) could be then subjected to similar crystallization conditions. However, subsequent results showed that, in the case of 6bp, only those samples that lacked ddNTP crystallized, whereas in the case of 7bp, only those samples that contained ddNTP crystallized, and in particular, either ddCTP or ddTTP served equally well for crystallization, despite the fact that only ddTTP was capable of forming a correct base pair with the template. Due to the nature of our results (Figure 1), we do not know whether a ddNTP was added onto the originally designated primer terminus in the case of the 7bp samples, and in fact, the ddNTP requirement for obtaining the 7bp crystals could possibly reflect nothing more than a need for a nonspecific triphosphate moiety in the crystallization medium.

Crystals that were about 0.5 by 0.4 by 0.3 mm in size grew within 2 weeks by the vapor diffusion method in MVD-24 sitting drop trays (Charles Supper Co.) which had been sealed with clear packaging tape (Manco, purchased from Sears) after 10 μL of the reservoir solution had been mixed with 10 μL of either the 6bp or the 7bp sample. The two most successful reservoir solutions for the 6bp crystals consisted of (i) 10−12% (w/v) PEG 3350 (Sigma), 0.1 M MES, pH 6.5, and 0.1 M NaCl or (ii) 12−14% PEG 3350, 0.1 M HEPES, pH 7.5, and 50 mM Li_2SO_4. The most successful reservoir solutions for the 7bp crystals consisted of 15−17% PEG 3350, 0.1 M citrate, pH 6.0−6.5, and either 50 mM NaCl or 50 mM Li_2SO_4. Because citrate is a known metal ion chelator, attempts were made to replace citrate with MES, pH 5.5−7.0, but no crystals formed under these conditions, suggesting that citrate may be a requirement for obtaining the 7bp crystals.

Preparation of Nonphosphorylated and Nongapped DNA Complexes. Later, it was discovered that isomorphous 6bp and 7bp crystals could form under the same conditions as described above even when the downstream oligonucleotides were not 5′-phosphorylated (designated as no 5′-PO_4 in Table 1). Under these conditions, the presence of a ddNTP was not a requirement for obtaining the 7bp crystals. In addition, it was later discovered that isomorphous 6bp and 7bp crystals could be grown from 6 base-paired and 7 base-paired DNA substrates alone (designated as no gap in Table 1 because the substrate is nongapped DNA). The sequences of the oligonucleotides for the DNA in these crystals corresponded to the same sequences described above for the intended downstream oligonucleotides of the 6bp and 7bp crystals (5′-PO_4-CAGATG-3′ and 5′-PO_4-TCTAATG-3′, respectively) with their corresponding complementary oligonucleotides (5′-CATCTG-3′ and 5′-CATTAGA-3′) (Figure 1). In that the metal ion binding behavior of pol β was also being explored at the time that these crystals were grown (Pelletier et al., 1996), X-ray diffraction data had only been collected on no gap crystals that were grown in the presence of either Mn^{2+} or Zn^{2+}. The crystals designated as 6bp, no gap and 7bp, no gap in Table 1 were therefore grown in the presence of 10 mM $MnCl_2$ and 10 mM $ZnCl_2$, respectively. All other

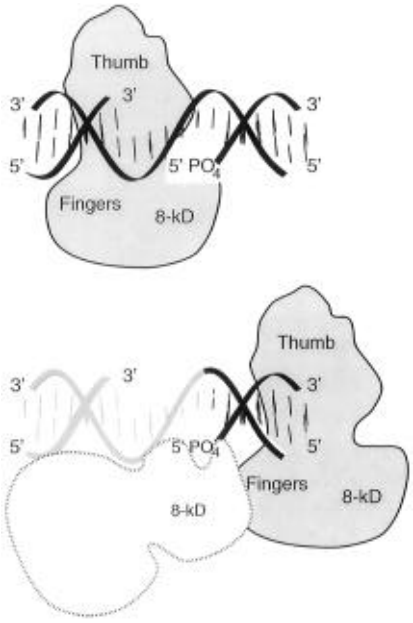

FIGURE 1: Schematic diagrams of pol β−DNA complexes depicting (top) the structure that was expected to result from cocrystallization of pol β with a gapped DNA substrate and (bottom) what actually cocrystallized. We have adopted a common nomenclature that is now utilized for all polymerase structures (Ollis et al., 1985; Kohlstaedt et al., 1992) according to which, because of the resemblance to a hand, the subdomains of the pol β 31-kDa domain are termed fingers, palm, and thumb (Sawaya et al., 1994). The view looks down into the DNA binding channel so that the palm subdomain, which is not labeled, is beneath the DNA. The DNA segment that is lightly shaded in the bottom panel is too disordered to be observable in our crystal structures. We have not analyzed our crystals to check for possible DNA cleavage products, but some of our structural work, particularly with crystals soaked in the presence of Zn^{2+} (Pelletier et al., 1996), occasionally revealed uninterpretable electron density in the crystal solvent channels that most likely corresponds to the remaining missing DNA in our structures, that is, 6 base pairs of template−primer tethered to a flexible, single-stranded DNA gap of 6 or 7 nt. We have subsequently shown that isomorphous pol β−DNA cocrystals can grow in the presence of DNA substrates that lack the disordered DNA portions shown in the bottom panel (designated as "no gap" in Table 1). As outlined in the bottom panel, the 8-kDa domain of a symmetry-related pol β molecule in the crystal binds to the protruding 5′-PO_4 terminus of the DNA and may represent a physiologically relevant protein−DNA interaction, as described in the text.

conditions were as described above, with the exception that salt concentrations had to be lowered slightly to improve crystal quality. The primary reason why we do not have X-ray data for no gap crystals that were grown in the absence of transition metal ions is that, although the no gap crystals possessed diffraction qualities that were similar to those of the original 6bp and 7bp crystals, they deteriorated relatively quickly in the crystallization drops, regardless of whether or not they were grown in the presence of transition metal ions. This is in contrast to original 6bp and 7bp crystals, which have remained stable in some crystallization drops for over a year (Pelletier & Sawaya, 1996a).

Structure Determinations and Refinements. X-ray diffraction data (Table 1) were collected on a multiwire area detector (Hamlin, 1985) (San Diego Multiwire Systems) with monochromatized CuKα radiation (Rigaku RU200 rotating-anode X-ray generator), and intensity observations for each data set were processed with a local UCSD Data Collection Facility software package (Howard et al., 1985). We utilized a slightly modified version of the high-resolution crystal structure of the 31-kDa domain of rat DNA pol β (Sawaya et al., 1994), in which all residues corresponding to sequence differences between rat and human pol β (Ito & Braithwaite, 1991)[3] were changed to alanines, as a search model to obtain a partial solution of the 6bp structure with the molecular replacement and Patterson correlation programs of X-PLOR (Brunger, 1992). Rigid-body refinement in X-PLOR was followed by least-squares positional refinement cycles with the TNT program package (Tronrud et al., 1987), resulting in an *R*-factor of about 45%. Difference Fourier maps ($F_o - F_{c(partial)}$, $\alpha_{c(partial)}$) revealed clear electron density for the 8-kDa domain, as well as some electron density for the phosphate backbone of a double-stranded DNA molecule in the pol β binding channel. In addition, electron density for the 14 residues corresponding to sequence differences between rat and human pol β (Ito & Braithwaite, 1991) indicated side chains that were in agreement with the published amino acid sequence for human pol β. The entire 8-kDa domain was taken from the $P6_1$ rat pol β ternary complex structure (Pelletier et al., 1994) and was positioned in the electron density as a rigid body, and the full 39-kDa structure was once again subjected to rigid-body refinement in X-PLOR (Brunger, 1992), followed by least-squares positional refinement with TNT (Tronrud et al., 1987). Individual isotropic temperature factor refinements were also carried out at this stage, with restrained correlation for bonded atoms.

Although difference Fourier maps revealed electron density for six base pairs of double-stranded DNA in the pol β binding channel that could have been interpreted as the template and the 6-nt primer of the gapped-DNA sample, strong electron density corresponding to a 5′-PO_4 group on one of the DNA strands told us that what had actually bound in the DNA binding channel was instead the double-stranded downstream end of the gapped DNA substrate (Figure 1, bottom). This was further confirmed by the structure determination of isomorphous 7bp crystals by the difference Fourier method, where $F_{o(7bp)} - F_{o(6bp)}$, $\alpha_{c(6bp)}$ difference Fourier maps indicated that the 7bp structure differed from the 6bp structure only by one extra base pair added to the double-stranded DNA in the pol β binding channel. Once again, strong electron density was observed at the 5′ end of one of the DNA strands. As a final check on our modeling efforts, subsequent structure determinations of isomorphous 6bp and 7bp crystals that lacked 5′-phosphorylated downstream oligonucleotides (Table 1) gave very large negative peaks at the 5′-PO_4 site in $F_o - F_o$, α_c difference Fourier maps. Once again least-squares positional and individual isotropic temperature factor refinements were carried out with TNT (Tronrud et al., 1987), resulting in the structures reported in Table 1. No restraints were placed on hydrogen bonding between DNA base pairs at any time during refinement.

RESULTS AND DISCUSSION

Description of the Structures. The 14 amino acid residue differences (out of 335) between rat and human pol β (Ito

[3] There are 14 (out of 335) residue differences in the amino acid sequences of human and rat pol β, and they are as follows (residue numbers are preceded by the one-letter code for the human residue and are followed by the one-letter code for the rat residue): T20V, S104T, F114L, G144E, V169L, S171P, S199N, T205S, C213R, H222R, K244E, K248N, K317Q, and K326R. For the complete sequences of human and rat pol β, see Ito and Braithwaite (1991).

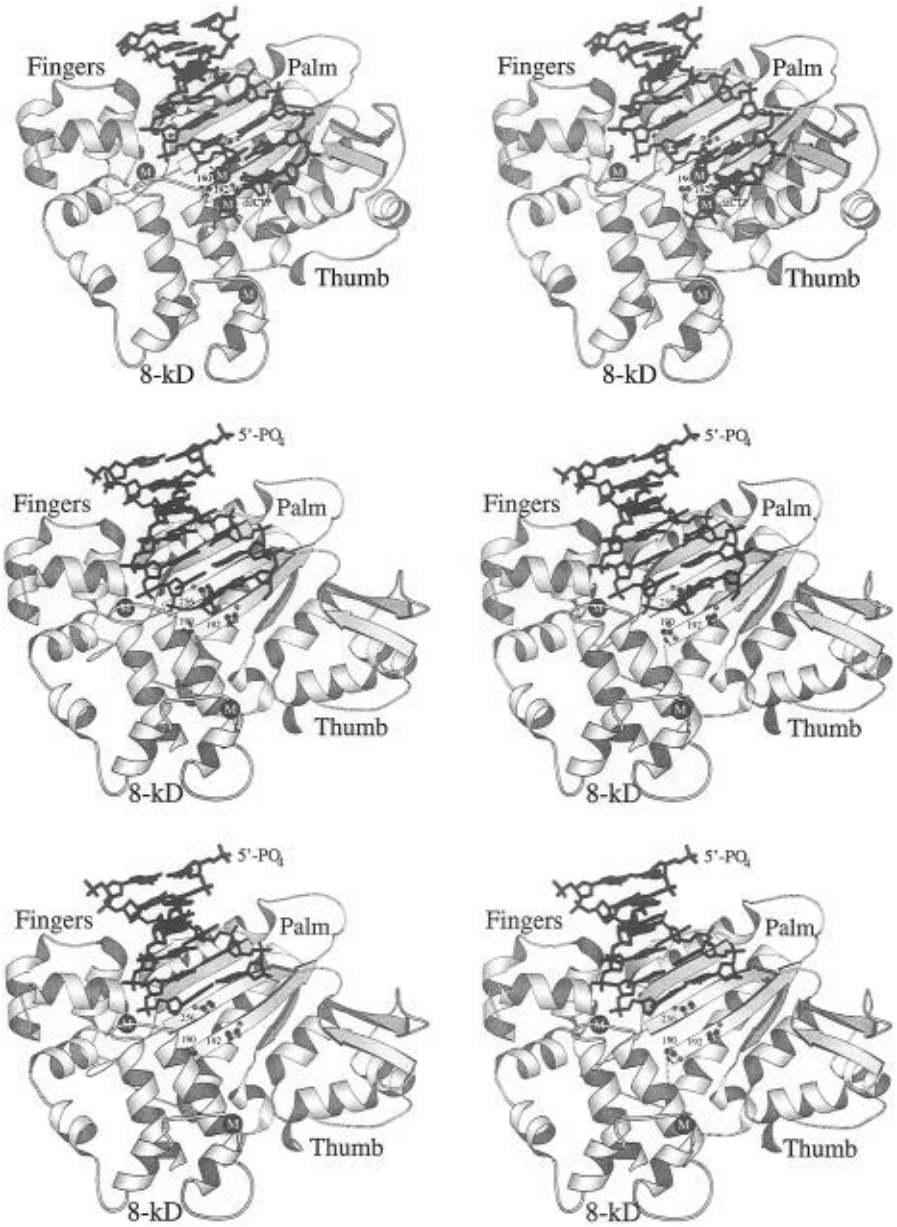

FIGURE 2: SETOR stereo diagrams (Evans, 1993) showing DNA binding for three different pol β−DNA complexes: (top) the crystal structure of rat pol β complexed with DNA and ddCTP (the ternary complex) (Pelletier et al., 1994), (middle) the crystal structure of human pol β complexed with 7 base pairs of DNA (the 7bp binary complex) (Table 1), and (bottom) the crystal structure of human pol β complexed with 6 base pairs of DNA (the 6bp binary complex) (Table 1). The view looks down into the DNA binding channel and is similar to the view in Figure 1 with the exception that the structures have been rotated clockwise by about 90° about an axis perpendicular to the plane of the page. Black spheres with a white M represent metal sites. Conserved active-site carboxylate residues, aspartic acids 190, 192, and 256, are labeled as 190, 192, and 256, respectively, and in the middle and bottom panels the protruding 5′-phosphorylated end of the primer strand is depicted as 5′-PO_4.

& Braithwaite, 1991) are all located on the surface of the molecule and do not appear to make any critical inter-or intramolecular contacts. For this reason—and because these two enzymes are also indistinguishable kinetically (Abbotts et al., 1988)—rat and human pol β are assumed to be the same protein, unless otherwise noted, in the following discussions. Although the pol β−DNA cocrystals reported here have resulted in an unexpected complex (Figure 1), it was nevertheless reassuring to discover that the DNA substrate is bound in the pol β binding channel the same way as in the previously reported ternary complex structures (Figure 2) (Pelletier et al., 1994). In fact, the bound DNA in the 7bp binary complex, in particular, is nearly identical with the bound DNA in the earlier ternary complexes with only the following exceptions: (i) the DNA sequence is different; (ii) the DNA strand that we now designate as the primer[4] in the 7bp structure has a 3′-OH positioned in the active site, whereas the dideoxy method for crystallizing the ternary complexes necessarily resulted in the terminal nucleotide of the primer being a ddNMP, which lacks a 3′-OH group (Pelletier et al., 1994); (iii) although unintentional, the 5′ terminus of the primer in the 7bp structure is phosphorylated (5′-PO_4) (Figure 2, middle), whereas 5′-phosphorylated oligonucleotides were never utilized in the ternary complex work (Pelletier et al., 1994); and (iv) in the

[4] Although the primer strands in the human pol β−DNA structures presented here were originally intended to be the downstream oligonucleotides of gapped-DNA substrates (including the 5′-PO_4 group) (Figure 1), they are nevertheless referred to as primers, and not downstream oligonucleotides, in all discussions of the human pol β−DNA structures in the text—by analogy with the template-primers in the ternary complex structures (Pelletier et al., 1994).

7bp structure, the DNA is blunt-ended at the active site so there is no template overhang available for base pairing with an incoming nucleotide (also unintentional), whereas in the ternary complex there is a template overhang of 4 nt—at least one of the 4 nt can be observed in the crystal structure, and it is base-paired with the incoming ddCTP in the active site (Pelletier et al., 1994) (Figure 2).

The 6bp structure is very similar to the 7bp structure with the most significant difference being that the DNA in the 6bp structure is one base pair shorter at the blunt end so that the 3′-OH group on the primer terminus is not positioned in the active site in the 6bp complex (Figure 2, bottom). That the shorter DNA substrate in the 6bp structure resulted in a missing base pair at the active site rather than at the other end of the DNA suggests that an intermolecular contact between the protruding end of the DNA and the 8-kDa domain of a symmetry-related pol β molecule in the unit cell (Figure 1, bottom) may have been a dominant force determining the packing arrangement in this crystal. The nature of this 8-kDa domain−DNA interaction and its implications for the biological relevance of our crystal structure are addressed below only after other aspects of the pol β−DNA structure are described.

Although it was not the structure initially being sought, the combination of two features in the 7bp structure—a primer 3′-OH group available in the pol β active site for reaction and the absence of a template overhang for possible base-pairing with incoming nucleotides—has nevertheless created a potent experimental system for studying indiscriminate, non-template-directed nucleotidyl transfer directly in these crystals. Using this system, we have demonstrated that nucleotidyl transfer can occur directly in the 7bp crystals when they are soaked in the presence of of dATP and Mn^{2+} but that under similar soaking conditions, no reaction is observed in the 6bp crystals (Pelletier et al., 1996). This observation is in agreement with the idea that, in order for catalysis to occur, the primer 3′-OH group must be positioned near the trio of carboxylate residues that have been identified as the active site in pol β (Pelletier et al., 1994). In addition to confirming the previously proposed mechanism for the nucleotidyl transfer reaction, our characterization and subsequent exploitation of this fortuitous experimental system has also allowed a structural basis for metal ion mutagenicity and nucleotide selectivity to be presented for pol β (Pelletier et al., 1996).

In that the 8-kDa domain had been observed in three different positions, relative to the 31-kDa domain, in previous studies with rat pol β, it was proposed that the position of the seemingly mobile 8-kDa domain was probably determined predominantly by crystal packing forces (Pelletier et al., 1994). It was therefore unexpected when we discovered early in our structure determinations of the human pol β−DNA binary complexes that the position of the 8-kDa domain does not differ significantly from that observed in the $P2_1$ ternary complex structure of rat pol β (Figure 3). This result encouraged us to finish some uncompleted structural studies of rat pol β apoenzyme crystals that had resulted after failure to form ternary complexes with 3′-azido-2′,3′-dideoxythymidine triphosphate (AZT-TP) (Pelletier et al., 1994). Once again this additional, newly determined apoenzyme structure revealed a position of the 8-kDa domain corresponding to that observed in the previously reported apoenzyme structure (Sawaya et al., 1994), despite the differences in space groups and crystal packings. We therefore have now observed only three different positions of the 8-kDa domain in five different crystal packing arrangements of mammalian pol β, suggesting that the 8-kDa domain is not as freely mobile as previously thought and that it may be capable of assuming only a small number of stable positions that depend on the presence or absence of a DNA substrate.

A Moveable Thumb Subdomain. The only significant difference between the pol β-DNA binary complex and the pol β−DNA−ddCTP ternary complex is observed in the thumb subdomain; a large change in the position of the thumb subdomain, with residues at the tip of the thumb moving as much as 12 Å, results in a more open active site than in the binary complex (Figure 3, bottom). In that the thumb subdomain is also in a more open conformation in the apoenzyme structure (Sawaya et al., 1994), perhaps this is the most stable conformation of pol β in solution. However, there appear to be very few electrostatic and hydrophobic interactions, all of which are relatively weak, between the thumb and the DNA or nucleotide substrates in the closed ternary complex (Pelletier et al., 1994), suggesting that the energy difference between the open and closed conformations may be low. Most of the thumb movement can be described by a change in four torsion angles in a loop connecting the palm and thumb subdomains [ψ 263 ($\Delta = 37°$), ϕ 264 ($\Delta = -44°$), ψ 264 ($\Delta = -18°$), and ψ 265 ($\Delta = 25°$)], and movement is constrained by extensive contacts between β strands 1, 2, and 5 of the palm and helix M of the thumb. The thumb pivots about the axis of helix M, which is approximately parallel with the strands of the palm (Figure 3, bottom).

The palm−thumb interface is lined with numerous hydrophobic groups on both sides of the hinge so that there appear to be an equal number of favorable hydrophobic interactions when the thumb is either in an open or in a closed conformation. Hydrophobic residues lining the inside of the hinge (nearest the active site) are leucine 194, isoleucine 260, phenylalanine 272, and tyrosine 296, and hydrophobic residues lining the outside of the hinge are isoleucine 174, threonine 196, tyrosine 265, and valine 269. In addition to hydrophobic interactions, two salt bridges, one between arginine 258 and glutamic acid 295 on the inside of the hinge and the other between arginine 182 and glutamic acid 316 on the outside of the hinge, participate in stabilizing the closed and the opened thumb positions, respectively. Finally, two additional hydrogen bonds between backbone carbonyl oxygens and amide nitrogens occur on the inside of the palm−thumb hinge (between glycine 179 and phenylalanine 272 and between proline 261 and glutamine 264), but only when the thumb is in the closed position. The backbone hydrogen bond between proline 261 and glutamine 264 is particularly interesting because glutamine 264 is a pivotal residue in the loop that connects the palm and thumb subdomains, as described above. It should be noted that none of the aforementioned pol β residues differs in sequence between the rat and human enzymes.

An apparently mobile thumb subdomain has been observed in other polymerase crystal structures (Beese et al., 1993; Sousa et al., 1994; Rodgers et al., 1995), suggesting that thumb movement may be a common mechanistic feature of polymerases. In addition, pre-steady-state kinetic experiments performed with the Klenow fragment of *E. coli* DNA pol I (Kuchta et al., 1987; Frey et al., 1995), with bacteriophage T7 DNA polymerase (Patel et al., 1991; Wong et

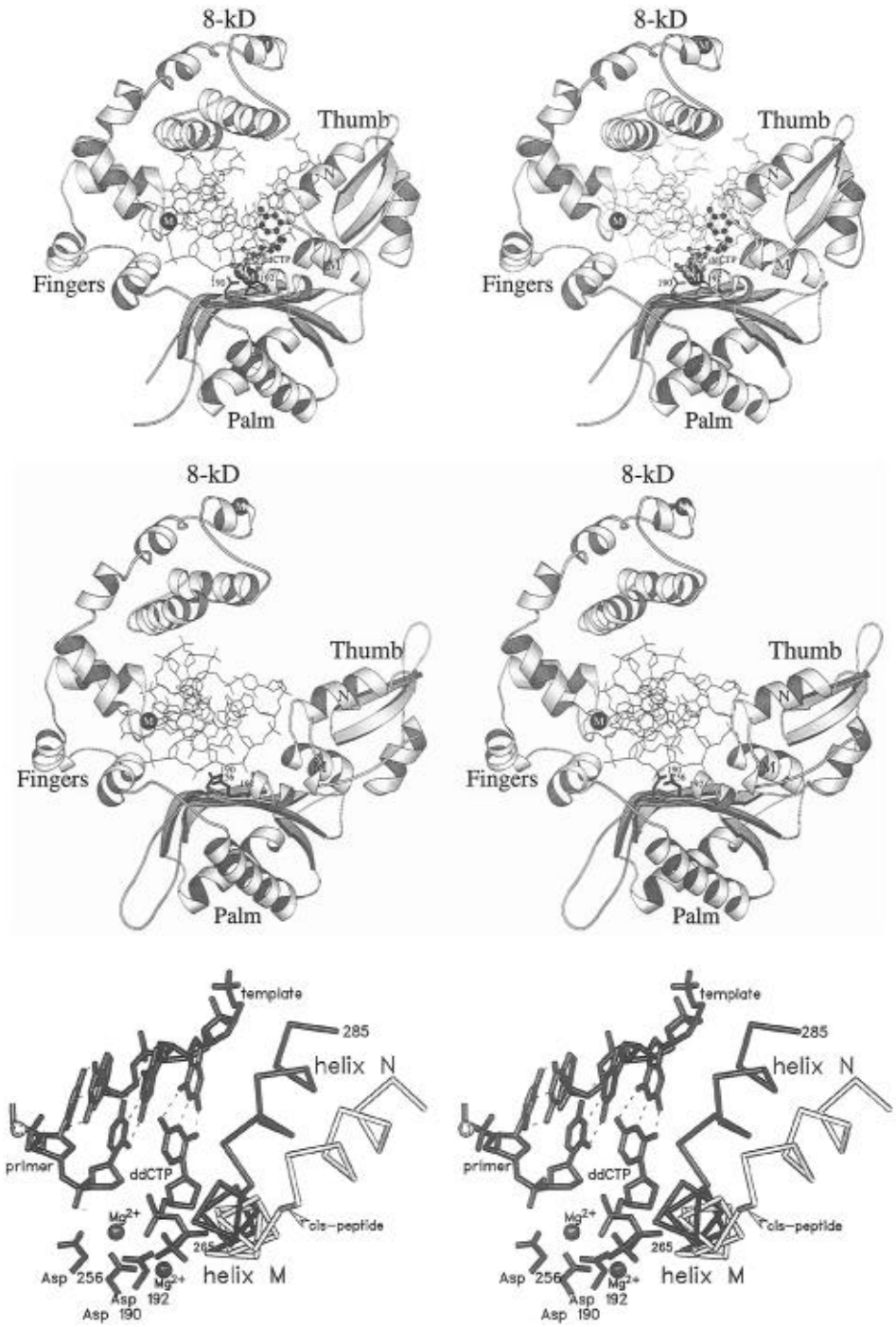

FIGURE 3: MOLSCRIPT stereo diagrams (Kraulis, 1991) showing (top) the structure of rat DNA pol β complexed with DNA and ddCTP (the ternary complex), crystallized in space group $P2_1$ (Pelletier et al., 1994) and (middle) the structure of human DNA pol β complexed with 7 base pairs of DNA (the 7bp binary complex), crystallized in space group $P2_12_12$. These figures were generated by rotating the corresponding structures in Figure 2 by 90° about a horizontal axis in the plane of the page. Helices M and N of the thumb subdomain are labeled. It is evident from this view that the position of the 8-kDa domain is similar for both structures, despite different crystal packing arrangements. A close-up view in the bottom panel highlights the most significant difference between the two structures, namely, a change in the position of the thumb subdomain, relative to the palm. To generate the figure in the bottom panel, the ternary and binary complex structures (top and middle panels, respectively) were aligned by superposing the Cα atoms of their palm subdomains. For clarity, the only part of the human pol β−DNA binary complex structure shown (lighter shade) are helices M and N of the thumb subdomain. The *cis* peptide shown in the bottom panel occurs between glycine 274 and serine 275. We have not observed a *trans* conformation for this peptide bond in any pol β structures studied thus far.

al., 1991), with *E. coli* RNA polymerase (Erie et al., 1993), with bacteriophage T4 DNA polymerase (Frey et al., 1995), with HIV-1 reverse transcriptase (RT) (Spence et al., 1995), and with pol β (Werneburg et al., 1996) have all given evidence for a rate-limiting conformational change that precedes the catalytic step in these polymerases. The change that we observe from the open conformational in the pol β−DNA binary complex to a closed conformation in the pol β−DNA−ddCTP ternary complex might therefore represent the rate-limiting conformational change that has been observed in pre-steady-state kinetic experiments. It has been proposed that a rate-limiting conformational change may be a common mechanistic feature utilized by polymerases to increase replication fidelity (Johnson, 1993). That thumb movement, in particular, may be the conformational change in question is supported by the observation that residues from

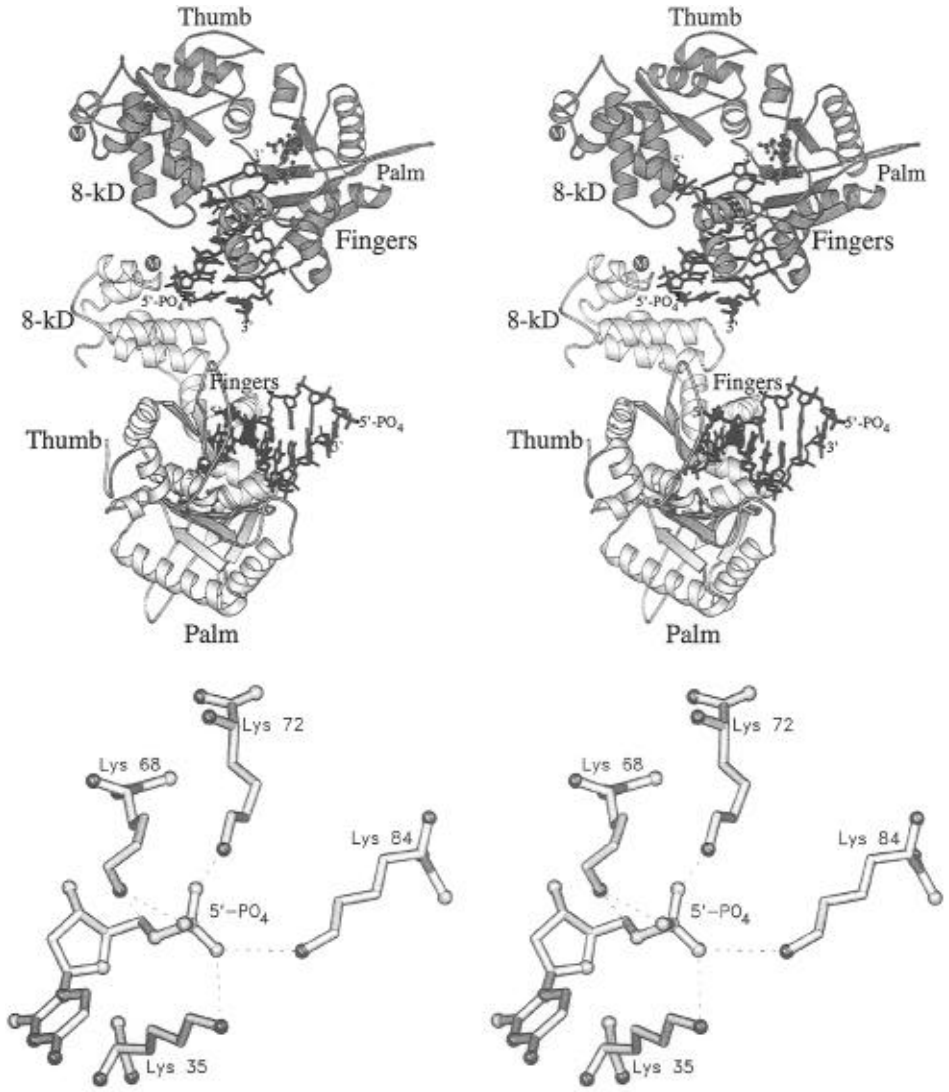

FIGURE 4: (Top) SETOR stereo diagram (Evans, 1993) showing two symmetry-related human pol β−DNA complexes in the unit cell. As outlined in Figure 1 (bottom), the 8-kDa domain of one pol β−DNA complex (lighter shade) interacts with the protruding DNA of a symmetry-related pol β−DNA complex (darker shade). (Bottom) Close-up view of the 8-kDa−DNA interface which reveals a lysine-rich 5′-PO_4 binding pocket on the 8-kDa domain. All hydrogen-bond distances are 3.3 Å or less, and the shortest hydrogen bond is with lysine 72 which, because of its high affinity for phosphate groups, has already been proposed as being a part of the 5′-PO_4 binding site (Sawaya et al., 1994).

helices M and N of the thumb subdomain in pol β protrude into the active site and interact with the nucleotide in the closed ternary complex structure (Pelletier et al., 1994). Similar structural features can be found in other polymerase thumb subdomains and are proposed to participate in nucleotide selectivity as well, by analogy with pol β (Pelletier et al., 1994; Pelletier, 1994). It is interesting to note that there is as yet no strong evidence supporting a rate-limiting conformational change for RT during reverse transcription when an RNA template is utilized (Reardon, 1993), in agreement with the idea that RT might employ an entirely different mode of binding for A-form RNA substrates than for B-form DNA substrates (Pelletier et al., 1994; Pelletier, 1994).

Although a mobile thumb appears to be a common feature of most polymerases, it should be kept in mind that other conformational changes, such as an alteration in the sugar pucker of the incoming nucleotide, might account for the kinetic results observed for other polymerases. In addition, there is some evidence that it might not be necessary for all polymerases to return to an open conformation, whatever that may be, for every subsequent catalytic event during processive polymerization (Erie et al., 1993). In the case of pol β, however, an open thumb appears to be a requirement for free movement of nucleotides in and out of the active site because all of our attempts to back-soak the ddCTP substrate out of the closed $P6_1$ ternary complex crystals (Pelletier et al., 1994) have thus far failed and have only resulted in slight movements of the ddCTP molecule within the binding pocket (Pelletier and Sawaya, unpublished data).

A 5′-PO_4 Binding Site in the 8-kDa Domain. Only three nucleotides of the primer strand are bound in the DNA binding channel of pol β in the catalytically productive complex (Pelletier et al., 1994), and in the case of the 6bp and 7bp structures presented here, at least three excess nucleotides of the primer, as well as the 5′-PO_4 end group, all protrude from the binding channel and interact with the 8-kDa domain of a symmetry-related pol β molecule [Figures 1 (bottom) and 4 (top)]. Although intermolecular contacts are common in crystals, what makes this interaction especially noteworthy is that the 5′-PO_4 group on the end of the primer, which mimics the 5′-PO_4 terminus of a gapped DNA substrate, binds to the 8-kDa domain in a lysine-rich pocket (Figure 4, bottom). There is much evidence that the 8-kDa−

DNA interaction depicted in Figure 4 (bottom) is physiologically relevant, and in fact, this particular lysine-rich section of the 8-kDa domain was previously proposed to be the binding site for the 5′-PO_4 terminus of gapped-DNA substrates (Sawaya et al., 1994). This proposal was based, in part, on results from chemical modification studies with pyridoxal 5′-phosphate indicating that lysine 72 has a high affinity for phosphate groups (Basu et al., 1989) (Figure 4, bottom). Also, solution studies have shown that the 8-kDa domain of pol β preferentially binds gapped-DNA substrates that are 5′-phosphorylated over gapped-DNA substrates that are not phosphorylated (Prasad et al., 1994).

In addition to increasing pol β's affinity for gapped-DNA substrates (Prasad et al., 1994), and in addition to increasing pol β's processivity during the gap-filling process (Singhal & Wilson, 1993), the 8-kDa domain has recently been assigned yet another possible role in pol β's DNA repair function in vivo—the excision and removal of a dRP intermediate that arises just prior to the DNA gap-filling step in the BER pathway (Matsumoto & Kim, 1995). These findings suggest that the 5′-PO_4 binding pocket in our crystal structure (Figure 4, bottom) not only represents a binding site on the 8-kDa domain for the 5′ terminus of a gapped-DNA substrate but also may form a part of the catalytic site for the newly discovered enzymic activity as well. That the 5′-PO_4 binding pocket is made up of a cluster of lysine residues lends some support to this proposal because the cleavage reaction carried out by the 8-kDa domain occurs via a β-elimination mechanism (Matsumoto & Kim, 1995), and it has been shown that a similar β-elimination reaction can be catalyzed by polyamines in solution (Bailly & Verly, 1988).

Although more structural and kinetic work will be required to fully elucidate the mechanism of the excision reaction catalyzed by the 8-kDa domain, two possible scenarios would be consistent with our current structural data. One possibility is that the 8-kDa domain binds a substrate bearing a 5′-dRP group the same way that it binds a 5′-dNMP (for example, imagine the DNA structure in Figure 4 (bottom), except without a base moiety), and the 5′-dRP is then removed via a β-elimination mechanism that is most likely facilitated by lysine 68 (because of its proximity to the scissile sugar−phosphate bond). Then again, although not shown in Figure 4 (bottom), base-pairing occurs in our crystal structure between the 5′-dNMP that is bound in the lysine-rich pocket and a complementary dNMP in the template strand (Figure 2), and it is possible that the mode of DNA binding depicted in Figure 4 (bottom) is highly selective against a base-free residue at this location. Under these circumstances, a different scenario emerges: the structure in Figure 4 (bottom) may represent the end product of an excision reaction whereby the 5′-dRP substrate either was bound in an adjacent binding pocket or perhaps was not bound to the protein at all but was simply tethered to an anchored dNMP—prior to being cleaved away. Although the latter scenario seems more attractive because it allows the cleavage reaction and the following gap-filling process to proceed with little change in 8-kDa−DNA interactions, the former scenario is more plausible because of the close proximity of lysine 68 to the proposed scissile sugar−phosphate bond. There is much evidence in support of the idea that the DNA lyase reaction catalyzed by the 8-kDa domain of pol β proceeds with a catalytic mechanism that is very similar to that of the bacterial DNA repair enzyme endonuclease III, and a more detailed description of the DNA lyase reaction for both enzymes is presented below.

Two Structurally Homologous DNA Binding Motifs in pol β. As discussed above, that the DNA in the 6bp structure is one base pair short at the active site rather than at the protruding end (Figure 2, bottom) seems to imply that the interaction between the lysine-rich pocket in the 8-kDa domain and the 5′-PO_4 DNA end group is relatively strong and could be a dominant force in determining the packing arrangement in our crystals (Figure 4). However, in that we have also been able to obtain isomorphous 6bp crystals with nonphosphorylated DNA substrates (Table 1), there is some evidence that the 5′-PO_4 alone is not entirely responsible for the strong 8-kDa interactions observed in our crystals. Inspection of the protein scaffolding around the 5′-PO_4 binding site has, indeed, uncovered additional protein−DNA interactions, revealing a previously overlooked helix−hairpin−helix (HhH) DNA-binding motif in the 8-kDa domain that is structurally homologous to yet another HhH motif that had been overlooked in the fingers subdomain (Figure 5, top). It is intriguing that both motifs not only bind a metal ion in the hairpin region but also interact with backbone phosphates of DNA strands (namely, the DNA primer and the proposed downstream oligonucleotide for the motifs in the fingers subdomain and the 8-kDa domain, respectively) in a highly similar fashion (Figure 5, middle). In fact, the only significant structural difference between the two HhH motifs is that the HhH motif in the fingers subdomain, unlike the HhH motif in the 8-kDa domain, does not have a lysine-rich pocket (Figure 4, bottom) adjacent to the motif for binding a 5′-PO_4 end group. Crystal soaking experiments with various metal ions have since confirmed that the metal ion sites of the HhH motifs are most likely occupied by Na^+ ions for the crystal structures reported here (Figure 5, bottom) and are most likely occupied by K^+ ions in vivo (Figure 6) (Pelletier & Sawaya, 1996).

The HhH motifs are made up of helices C and D of the 8-kDa domain (residues serine 55−threonine 79) and helices F and G of the fingers subdomain (residues aspartic acid 92−glycine 118), and there is evidence of sequence similarity between the two, as shown in boldface type below, where the region marked hairpin designates the four residues of a type II β-turn.

```
8-kD domain:       55-SGAEAKK LPGV GTKIAEK1-DEFLAT-79
Fingers:       92-DTSSSINFLTR VSGI GPSAARKFVDEG-118

                      helix   hairpin   helix
```

DNA Lyase Mechanism. It has been brought to our attention (Seeberg et al., 1995) that a motif that is sequentially homologous to the HhH motifs of pol β is present in the bacterial DNA repair enzyme endonuclease III which, in addition to possessing DNA glycosylase activity (Boorstein et al., 1989), catalyzes the cleavage of the DNA phosphodiester backbone at AP sites via a β-elimination mechanism (Bailly & Verly, 1987; Kow & Wallace, 1987; Kim & Linn, 1988; Mazumder et al., 1991). The crystal structure of endonuclease III has been determined (Kuo et al., 1992; Thayer et al., 1995), and structural comparisons reveal that the sequentially homologous motifs of pol β and endonuclease III are also structurally homologous (Figure 7). In fact, the term helix−hairpin−helix or HhH was originally coined for the motif in endonuclease III (Thayer et al., 1995), and we have since adopted the HhH nomen-

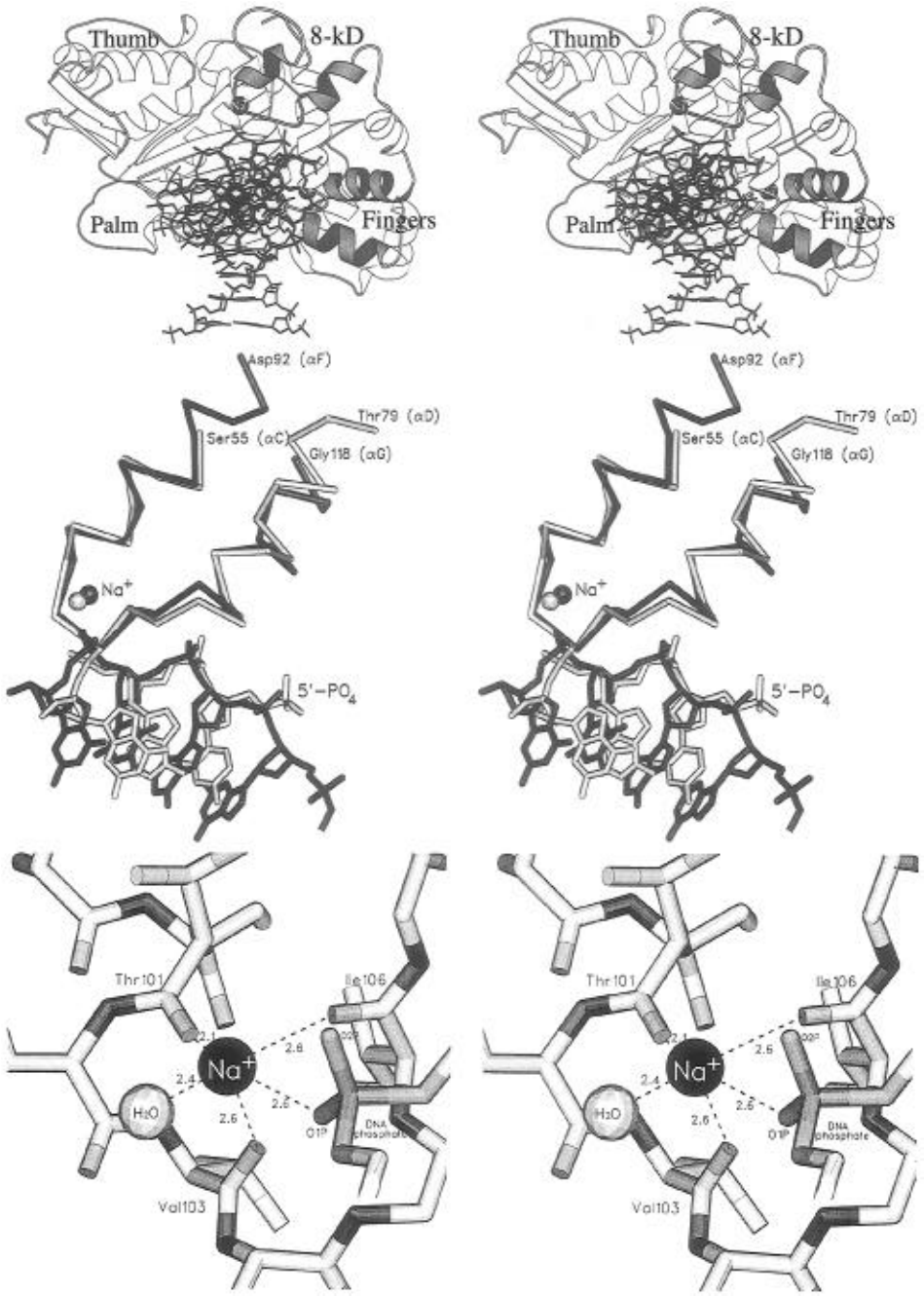

FIGURE 5: SETOR stereo diagrams (Evans, 1993) highlighting the HhH motifs of the 8-kDa domain and the fingers subdomain of pol β. (top) The HhH motifs are depicted in dark grey and are made up of α-helices C and D of the 8-kDa domain (residues serine 55−threonine 79) and helices F and G of the fingers subdomain (residues aspartic acid 92−glycine 118). Although not clear from this view, the axis of the DNA segment bound to the HhH motif of the 8-kDa domain is perpendicular to the plane of the page, while the axis of the DNA segment bound to the HhH motif of the fingers subdomain is in the plane of the page, thus creating a 90° angle between the two DNA segments. A model of pol β bound to a gapped DNA substrate has been proposed on the basis of these HhH−DNA interactions (as described in the text) and has since been confirmed by crystal structures of pol β bound to gapped and nicked DNA substrates (M. R. Sawaya, T. Rawson, S. H. Wilson, J. Kraut, and H. Pelletier, in preparation). (middle) Superposition of the HhH DNA binding motifs in the fingers subdomain (lighter shade) and the 8-kDa domain (darker shade). The hairpin regions consist of type II β-turns. Both motifs bind two DNA backbone phosphates each. The DNA segment associated with the 8-kDa domain is akin to the downstream oligonucleotide of a gapped DNA substrate and terminates right at the lysine rich pocket on the 8-kDa domain which, although not shown (middle panel), readily accommodates the 5′-PO_4 group (Figure 4, bottom panel). The DNA segment associated with the fingers subdomain is a part of the primer strand that is bound in the binding channel of pol β, and the metal ion coordinates a phosphate oxygen of the penultimate dNMP from the primer terminus, as can be seen in Figure 3 (bottom panel). An adjacent DNA phosphate (to the right of the metal site as shown in the middle panel) is hydrogen bonded to a cluster of pol β backbone amide nitrogens located at the amino-terminal end of helix G, as previously reported (Pelletier et al., 1994). The same amide nitrogen cluster is located at the amino-terminal end of helix D in the 8-kDa domain. (bottom) Closeup of the metal ion site in the fingers subdomain which, in the structures presented here, is thought to be occupied by a Na^+ ion (Pelletier & Sawaya, 1996a). The coordination geometry of the Na^+ ion is square pyramidal, and its five ligands (and their coordinate bond lengths) are as follows: one DNA phosphate oxygen (2.6 Å), one water molecule (2.4 Å), and three backbone carbonyl oxygens from threonine 101 (2.1 Å), valine 103 (2.6 Å), and isoleucine 106 (2.6 Å), respectively. It is possible that the Na^+ ion has other water ligands that are not detected at this resolution. The three protein carbonyl oxygens that act as ligands for the HhH metal ion site in the 8-kDa domain are from lysine 60, leucine 62, and valine 65, respectively. See Pelletier and Sawaya (1996) for a more detailed analysis of the HhH metal site.

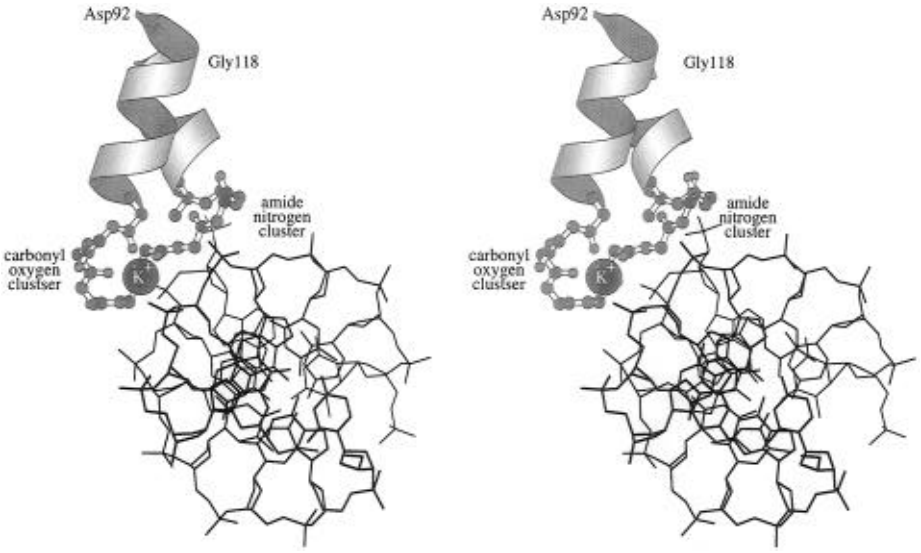

FIGURE 6: SETOR stereo diagram (Evans, 1993) depicting the non-sequence specific nature of the HhH−DNA interactions in pol β. This figure was prepared with coordinates from the crystal structure of a pol β−DNA cocrystal that had been soaked in an artificial mother liquor containing 150 mM KCl (PDB code 1ZQI) (Pelletier & Sawaya, 1996). The carbonyl oxygen cluster is made up of backbone carbonyl groups from threonine 101, valine 103, and isoleucine 106, and the amide nitrogen cluster is made up of backbone amide nitrogens from glycine 105, glycine 107, serine 109, and alanine 110. Glycine 105 and glycine 107 are invariant among the four known HhH motifs thus far identified (Table 2). In that the carbonyl oxygen cluster is located at the carboxy-terminal end of an α-helix, it possesses a negative dipole and correspondingly stabilizes binding of the positively charged metal ion. Likewise, the amide nitrogen cluster, which is located at the amino-terminal end of an α-helix and conversely possesses a positive helix dipole, stabilizes binding of the negatively charged phosphate backbone of the DNA substrate.

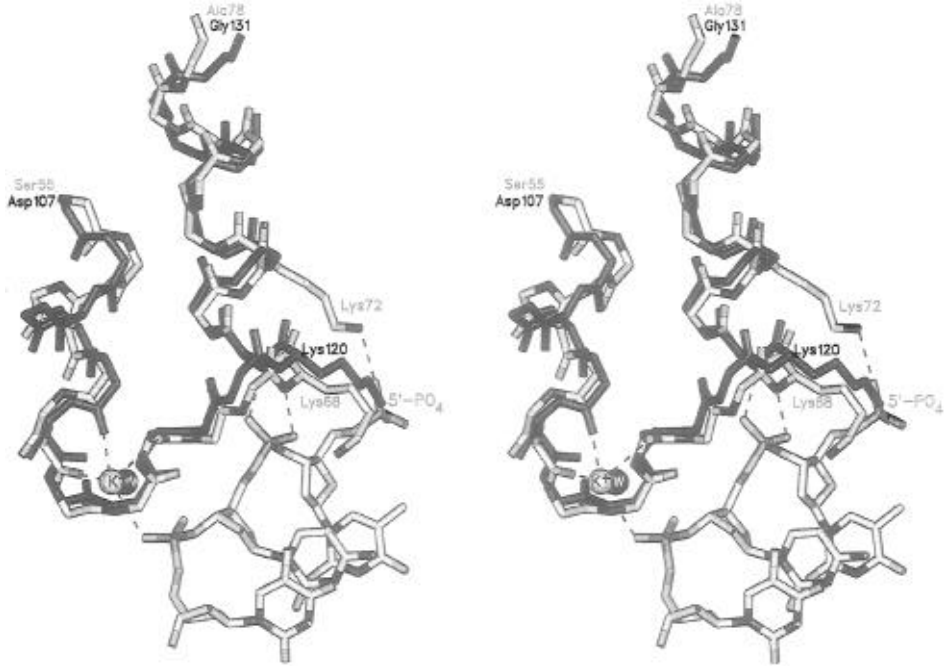

FIGURE 7: SETOR stereo diagram (Evans, 1993) depicting the HhH motif of endonuclease III (dark shade) (PDB code 2ABK) superposed on the HhH motif of the 8-kDa domain of pol β (lighter shade) (PDB code 1ZQI). A water molecule in the HhH metal site of endonuclease III is depicted as a black sphere with a white W. Lysine 72 of the HhH motif from the 8-kDa domain, which helps to stabilize binding to the negatively charged 5′-PO_4 end group (Figure 3), is not conserved in endonuclease III (Table 2), which has slightly different substrate requirements, as shown in Figure 8.

clature for the motifs of pol β as well. Although there is as yet no reported crystal structure of endonuclease III in complex with an ordered DNA substrate, the HhH motif has nevertheless been implicated in DNA binding (Kuo et al., 1992; Thayer et al., 1995), and mutational analysis of lysine 120 of endonuclease III, which corresponds to lysine 68 of the 8-kDa domain of pol β (Figures 4, bottom, and 7), shows that this residue is critical for catalysis (Thayer et al., 1995). These results suggests that endonuclease III and the 8-kDa domain of pol β may employ a similar mode of DNA binding and may have similar catalytic mechanisms for their respective β-elimination reactions (Figure 8).

HhH Motifs in Other Proteins. With the HhH motif of endonuclease III as a model, the BLAST sequence-searching program (Altschul et al., 1990; Henikoff & Henikoff, 1992) had identified potential HhH motifs in other DNA binding proteins (Thayer et al., 1995), many of which were thought to function specifically in DNA repair. However, as described above, it was only until a similar HhH sequence search was carried out with the program SpeedGenie

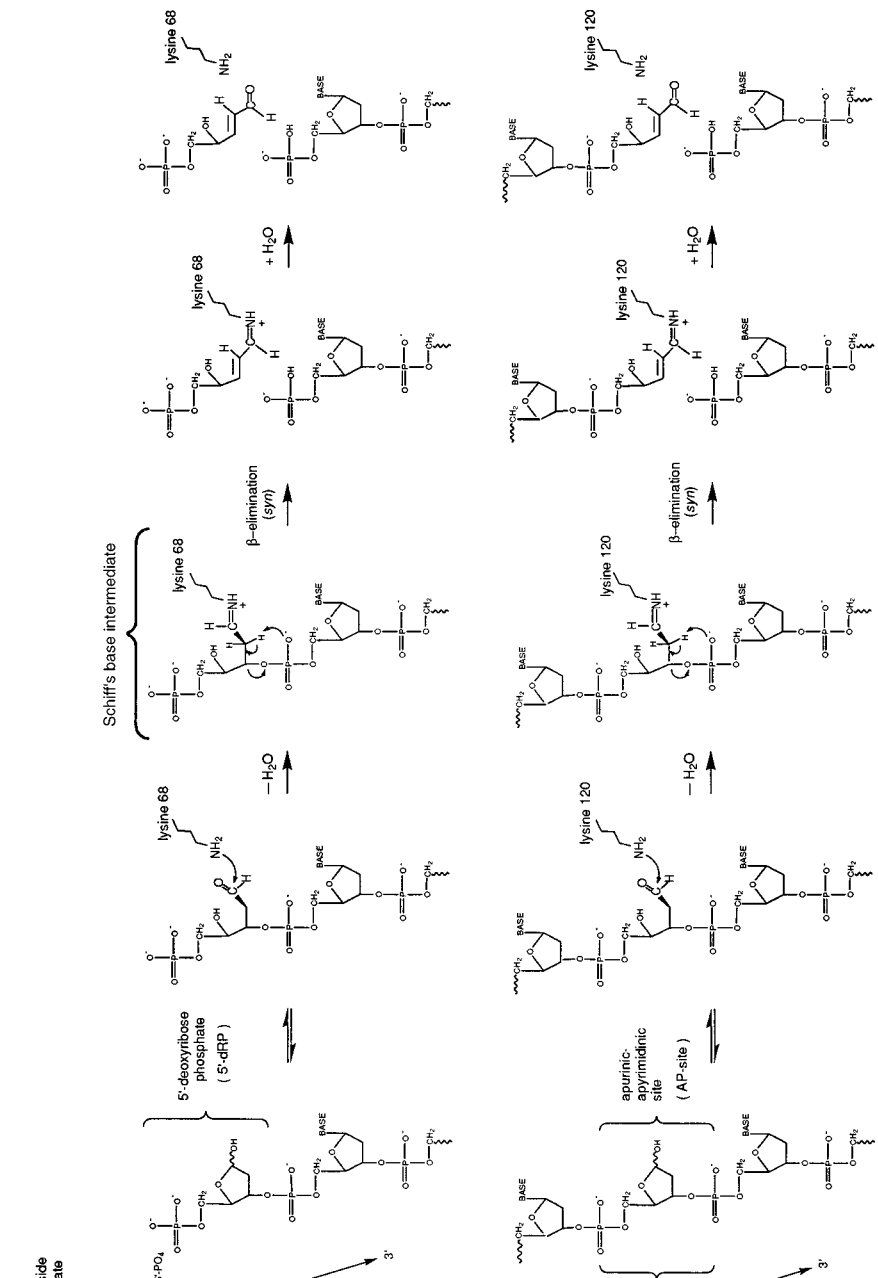

FIGURE 8: (top left) Schematic of what is bound to the 8-kDa domain in the crystal structure and (bottom right) actual substrates and proposed catalytic mechanisms for the 8-kDa domain of pol β (top) and endonuclease III (bottom). A squiggle at either end of the DNA substrate indicates that the DNA backbone continues. The mechanism for the 8-kDa domain is assumed to be similar to that of endonuclease III, which is the more studied reaction. What is known about the 5′-deoxyribose (dRP) cleavage reaction of the 8-kDa domain is that it proceeds via a β-elimination mechanism (Matsumoto & Kim, 1995). In the case of endonuclease III, however, not only has the apurinic−apyrimidinic site (AP site) cleavage reaction been shown to proceed via a β-elimination mechanism (Bailly & Verly, 1987; Kim & Linn, 1988), but also much is known about the intermediate steps of the reaction pathway, such as the possible formation of a Schiff's base intermediate (Kow & Wallace, 1987) (bottom right panel, middle figure) and the fact that the β-elimination proceeds with *syn* stereochemistry (Mazumder et al., 1991). Although the predominant form of the AP site (and the dRP site by analogy) is the hemiacetal shown to the far left in the bottom right panel (Manoharan et al., 1988), it is estimated that about 1% of the AP sites are in the aldehyde form (Wilde et al., 1989), which is the required form for reaction with endonuclease III (Mazumder et al., 1991). As noted by Mazumder and co-workers (1991), the idea that the leaving phosphate group acts as a base and catalyzes its own elimination has been proposed for other catalytic mechanisms as well, namely that for the enzyme dehydroquinate synthase (Widlanski et al., 1989; Bender et al., 1989) and that for the self-cleaving reaction of 3-hydroxypriopionaldehyde 3-phosphate (HPAP) (Gallopo & Cleland, 1979). The HPAP reaction is particularly interesting because formation of a Schiff's base with the buffer morpholine results in a 15-fold increase in the rate of phosphate elimination (Gallopo & Cleland, 1979), in support of the view that formation of a Schiff's base with a protein lysine residue may stabilize the elimination reactions carried out by pol β and endonuclease III.

Table 2: HhH Sequence Comparisons

species[a]	protein name and function	code[b]		sequence[c] helix	hairpin	helix		motif[d]
H. sapiens	DNA polymerase β (8-kDa domain), lyase	9ICJ	51	HKIKSGAEAKK	**LPGV**	**GTKIA**EKIDEF	NT	HhH
H. sapiens	DNA polymerase β (fingers subdomain)	9ICJ	92	DTSSSINFLTR	**VSGI**	**GPSAA**RKFVDE		HhH
E. coli	endonuclease III, *N*-glycosylase/AP lyase	2ABK	103	EVPEDRAALEA	**LPGV**	**GRKTA**NVVLNT		HhH
T. aquaticus	DNA polymerase I, 5′ nuclease domain	P19821[e,f]	185	LTGDESDNLPG	**VKGI**	**GEKTA**RKLLEE		HhH
E. coli	DNA polymerase I, 5′ nuclease domain	P00582[e,f]	182	LMGDSSDNIPG	**VPGV**	**GEKTA**QALLQG		HhH?
	consensus				V **GI**	**G** K **A**		
		Other Sequences of Interest						
Mu bacteriophage	transposase, N-terminal DNA binding domain	1TNS[e]	2	**E**LWV**S**PK**E**CAN	**LPG**L	PKTS**A**G**VI**YVA	NT	HtH
M. musculus	Brn3, POU domain protein 3	P17208[1]	4	ADVG**S**AL**A**NLK	**IPGV**	**G**SLSQST**I**CR**F**	NT	HtH?
E. coli	TreR, trehalose operon repressor, Lac1 family	P36673[e]	2	QNRLT**I**KDIA**R**	**LSGV**	**GKST**VSR**VLN**N	NT	HtH?
R. leguminosarum	nodulation protein nodE	P04684[f]	1	MDRRVV	**ITGI**	**G**GLCGLGTNAA	NT	HtH?
H. sapiens	p21 H-ras, transforming protein, GTP-binding	4Q21[f]	1	**M**T**E**YKLVVV**G**	AG**GV**	**GKSA**LTIQ**L**IQ	NT	StH
H. sapiens	RFC, replication factor C, large subunit	P35251[e,p2]	641	GS**S**FKA**AL**L**SG**	**PPGV**	**GKTTTAS**L**V**CQ		StH?
M. voltae	Adk, adenylate kinase	P43411[1,p1,p2]	1	**M**KNKVVVV**TG**	**VPGV**	**G**STT**S**XQ**L**AMD	NT	StH?
D. differens	ADH, alcohol dehydrogenase	P22245[f]	4	IAN**S**NII**F**VA**G**	**LGGI**	**GLDTSREIV**KS	NT	StH?
M. musculus	PTP-2, protein tyrosine phosphatase 2	S52655[1]	206	TPDHGP**AVI**HC	SA**GI**	**GRS**GTFS**LVDT**		StH?
S. scrofa	DLDH, dihydrolipoamide dehydrogenase	P09623[f]	237	ADVTAVEL**L**GH	**VGGI**	**G**IDMEVSKNFQ		StH?
B. taurus	elastin (structural protein of ligaments)	P04985[p2]	461	V**VP**GAPG**AIPG**	**LPGV**	**G**GVPGVG**IPAA**		HhH?
H. sapiens	collagen, α-2(V) chain	A25874[p2]	368	GIAGTPGPKGD	RG**GI**	**GEKGAE**GTAGN		HhH?
S. pneumoniae	PBP, penicillin-binding protein 2b, strain R6	P10524[f]	646	VAVVFPH**N**TNL	TN**GV**	**GPSIAR**DII**N**L	CT	HhH?
S. pneumoniae	PBP, penicillin-binding protein 2b (resistant)	X16022	646	VAVVFPH**N**TNL	TND**V**	**GPSIAR**DII**N**L	CT	HhH?
S. pneumoniae	PBP, penicillin-binding protein 2b (resistant)	U20073	646	VAVVFPH**N**TNL	TKN**V**	**GPTIAR**DII**N**L	CT	HhH?

[a] *H. sapiens* = *Homo sapiens*, *M. musculus* = *Mus musculus*, *R. leguminosarum* = *Rhizobium leguminosarum*, *M. voltae* = *Methanococcus voltae*, *D. differens* = *Drosophila differens*, *S. scrofa* = *Sus scrofa*, *B. taurus* = *Bos taurus*, and *S. pneumoniae* = *Streptococcus pneumoniae*. [b] Sequence searches were carried out with BLAST-P (Altschul et al., 1990; Henikoff & Henikoff, 1992), which was accessed on the World Wide Web through the National Center for Biotechnology Information (NCBI) homepage. Parameters were set to default values for all searches. Four-digit codes that begin with a number represent structural coordinates from the Brookhaven Protein Data Bank, while six-digit codes that begin with a letter are protein sequence accession numbers, where the first letter represents the following data banks: P = SwissProt, A or S = pir (Protein Identification Resource), U = GenPept, and X = EMBL. Superscripts denote the HhH query sequence that identified the sequence in question, according to the following codes: [1]8-kDa domain (lyase domain) of DNA pol β, [f]fingers subdomain of the 31-kDa domain of DNA pol β, [e]endonuclease III from *E. coli*, [p1]5′ nuclease domain of DNA pol I from *T. aquaticus*, and [p2]5′ nuclease domain of DNA pol I from *E. coli*. [c] The residue number for the first amino acid of each sequence listed precedes the sequence, and a NT or a CT proceeding the sequence indicates that the segment is near the amino terminus or the carboxy terminus of the protein, respectively. Modeling efforts show that the fourth residue of the hairpin turn of the HhH motif can only accommodate a valine or an isoleucine side chain (H. Pelletier and M. R. Sawaya, unpublished data). In addition, the first residue following the hairpin turn must be a glycine in order to prevent steric hindrance with the DNA substrate. Identified sequences were therefore screened mostly for those that possessed a GXG consensus sequence, where X is valine or isoleucine. The first set of sequences listed are from the proteins that are known to have HhH motifs by their crystal structures. For clarity, the four residues of the type II β-turn have been set apart from the rest of the sequence, and this format was maintained throughout the table, despite the fact that, according to the subsequent analyses with those sequences for which related structural data are available, many of the sequences identified in our search results probably do not have a type II β-turn in this region (H. Pelletier and M. R. Sawaya, unpublished data). Among the first five sequences listed, residues that are invariant or conserved are outlined in boldface type, where conserved is defined as varying by only two residues among the five sequences. However, only the invariant residues (two glycines and an alanine) are highlighted in boldface type in the consensus sequence listed below these five sequences. A different criteria for boldface type was applied to the remaining sequences. An amino acid residue in any of these sequences is represented in boldface type if the same amino acid residue appears in an equivalent position in any one of the first five sequences. This table represents only a select few of the sequence segments identified in our searches. The full list of proteins identified, as well as a more detailed analysis of the results, can be obtained by sending an E-mail inquiry to hug@bcm.tmc.edu. [d] Sequence segments for which corresponding structural data are not available are designated with a ?. HhH represents helix−hairpin−helix. HtH represents helix−turn−helix. StH represents strand−turn−helix, which is more commonly known as the phosphate-binding, or P-loop, motif (Saraste et al., 1993).

(Medprobe AS, Norway) that one of the two HhH motifs in pol β was identified in a similar list of repair enzymes (Seeberg et al., 1995). It is not yet known why SpeedGenie, unlike BLAST, was successful at identifying at least one of the HhH motifs of pol β. Ironically, SpeedGenie identified the HhH motif from the fingers subdomain of pol β (Seeberg et al., 1995) even though the HhH motif from the 8-kDa domain appears to show greater sequence similarity with the HhH motif of endonuclease III (Table 2).

Although the limited sequence similarities between the HhH motif of endonuclease III and the HhH motifs of pol β may have eluded some sequence-searching programs, both of the previous searches (Thayer et al., 1995; Seeberg et al., 1995) have nevertheless identified yet another protein that contains a HhH motif, the 5′ nuclease of DNA polymerase I (pol I) from bacteria. This result has been confirmed by the recent crystal structure determination of DNA pol I from *Thermus aquaticus* (Kim et al., 1995). Inspection of the pol I structure as presented in the paper (the coordinates for this structure have not been made available) shows that, with the exception of the hairpin turn, which is not visible in the electron density maps, the sequence that has been identified as a potential HhH motif in pol I corresponds to two α-helices that appear to be very similar to the two α-helices of HhH motifs found in pol β and endonuclease III. Therefore, structurally homologous HhH motifs have apparently been correctly identified in at least three different DNA repair enzymes by amino acid sequence searches (Table 2).

We have expanded on the previous sequence searches, which were carried out only with the sequence of the HhH motif from endonuclease III (Thayer et al., 1995; Seeberget al., 1995), by including sequences from the HhH motifs of pol β, as well as the HhH motif from pol I, as a part of the search protocol. Results from our sequence searches with the program BLAST (Altschul et al., 1990; Henikoff & Henikoff, 1992) did not reveal any sequence similarities between the two HhH motifs of pol β or between the HhH motifs of pol β and endonuclease III, but the HhH sequences of both enzymes did identify the HhH motif of DNA pol I (Table 2). Also, in addition to the identification of many of the same proteins as those found in previous searches (Thayer et al., 1995; Seeberg et al., 1995), a large array of new proteins have been added to the list (Table 2) (H. Pelletier and M. R. Sawaya, unpublished data). Analysis of the few proteins for which structural data are available indicates that, although many of the sequences identified by our search protocol probably do not correspond to HhH motifs in other proteins, all nevertheless correspond to critical surface loops in the protein structure. In some cases, particularly for those proteins involved in DNA recombination processes, the identified sequence corresponded to an area of the protein structure that is thought to be flexible and has been implicated in the catalytic mechanism. In most cases, however, identified sequences correspond to well-known phosphate-binding structural motifs, including DNA-binding helix−turn−helix (HtH) motifs, mononucleotide-binding strand−turn−helix (StH) motifs (more commonly known as P-loops), and even dinucleotide-binding protein folds (more commonly known as Rossmann folds), suggesting a possible evolutionary relationship between the HhH motif and other phosphate-binding motifs (Table 2).

In that we did not screen our results for only those sequences implicated in phosphate binding, many non-phosphate binding proteins have been identified in our sequence searches as well. Of this category of proteins, only one, elastin, has some structural data available (Cook et al., 1980; Urry, 1992). Similar to results with the phosphate binding proteins, a short pentapeptide segment (LPGVG) of the amino acid sequence identified in elastin corresponds to a well-known structural motif, a type II β-turn (Cook et al., 1980), and is identical to the sequence of the type II β-turn found in the HhH motifs of endonuclease III and the 8-kDa domain of pol β (Table 2), once again suggesting possible evolutionary relationships.

Among the non-phosphate binding proteins for which no structural data are available is perhaps one of our more peculiar findings. A protein segment that has strong sequence similarities with the HhH motif from the 31-kDa domain of pol β has been identified near the carboxy terminus of penicillin binding protein 2b (PBP-2b) (Table 2), a membrane-bound bacterial enzyme that is a target for penicillin inhibition. Although not thought to contain catalytically critical residues, mutations concentrated in this segment of PBP-2b have nevertheless been implicated in helping to confer resistance to penicillin in certain strains of bacteria (Hakenbeck et al., 1994; Smith & Klugman, 1995). These results further confirm the idea that, regardless of structure or function, searches with the HhH sequence seem to identify critical surface segments of other proteins, many of which are implicated in substrate binding and specificity.

Implications for Template−Primer Binding. The somewhat odd crystal-packing phenomenon that we observe in the human pol β−DNA structures presented here [Figures 1 and 4 (top)] prompted us to examine more closely the intermolecular crystal packing contacts in the ternary complex structures (Pelletier et al., 1994). Interestingly, we found the same intermolecular contacts between a segment of protruding DNA and the HhH motif in the 8-kDa domain of a symmetry-related pol β molecule—in both ternary complex structures. Perhaps the absence of a 5′-PO_4 group on the protruding end of the primer had previously obscured these critical intermolecular contacts with the 8-kDa domain (compare the top panel with the middle and bottom panels in Figure 2). In addition, the metal ions utilized in all pol β−DNA crystallizations to date have been Mg^{2+} and Li^+ or Na^+ which, because they all contain only a relatively small number of electrons, are difficult to detect crystallographically. As a result, ordered water molecules had been modeled into electron density that most likely corresponded to Na^+ ions in the HhH motifs of previous pol β−DNA structures (Figure 5) (Pelletier & Sawaya, 1996).

Putting aside possible reasons why the special HhH−DNA interactions in pol β had been overlooked, the final analysis is that, of the six HhH motifs from three different pol β−DNA crystal structures determined to date, all six HhH motifs interact with a DNA backbone segment in the same manner. Specifically, the 5′ to 3′ orientation of the bound DNA strand, with respect to the HhH motif, is always from right to left as viewed in Figure 5 (middle panel). This characteristic not only facilitates modeling efforts for a pol β−gapped-DNA complex, as described below, but also may explain why we were never able to obtain pol β−DNA cocrystals with template−primer samples that possessed relatively long template overhangs (see Experimental Procedures). Modeling experiments indicate that a single-stranded template overhang would have to loop out considerably from the active site in order to bind to the HhH motif in the 8-kDa domain with the correct DNA orientation. It is therefore likely that the 8-kDa domain did not bind to the template, at least not in a specific way, in any of our template−primer samples, and the long, flexible template overhangs probably only served to inhibit crystal formation. These results are in agreement with the observation that pol β is nonprocessive on template−primer DNA substrates (Wang & Korn, 1982) but is highly processive on gapped-DNA substrates (Singhal & Wilson, 1993; Prasad et al., 1994) where the 8-kDa domain most likely binds to the downstream oligonucleotide during polymerization.

Possible Effects of Metal Ions on DNA Binding and Processivity. As described above, in a crystallographic experiment it is difficult to distinguish low atomic weight metal ions such as Mg^{2+}, Li^+, and Na^+ (all of which are present in our crystallization drops) from water molecules, particularly at the resolution of data for the crystal structures presented here (Table 1). However, the coordination geometry and short bond lengths of the entity bound in the hairpin turn of the HhH motifs in pol β (Figure 5, bottom) suggest that this site is occupied by a metal ion in our structures and not a water molecule. In an attempt to further establish whether metal ions are bound in the HhH motifs in our structures, crystal soaking experiments were carried out with metal ions that are more easily detected in a crystallographic experiment, that is, metal ions like K^+, Ca^{2+}, Mn^{2+}, Ni^{2+}, Cs^+, and Ba^{2+}. Results from these experiments

show that the hairpin turns of both HhH motifs do, indeed, bind metal ions (Pelletier & Sawaya, 1996). Furthermore, the HhH metal site is selective for larger, more polarizable metal ions like K^+ or Ba^{2+}, with Na^+ showing the next strongest binding affinity (Pelletier & Sawaya, 1996). These results suggest that all of the crystal structures presented here most likely have Na^+ ions bound in the hairpin turn of the HhH motifs (Figure 5, bottom).[5]

Although there are as yet no reports of similar metal ion binding behavior at the protein−DNA interface in other polymerase crystal structures, there are nevertheless many attractive features to the idea that Na^+ and K^+ ions, which are present at relatively high concentrations in the cell, may facilitate DNA binding and processivity for polymerases in general. For instance, one such feature is that a metal ion can dramatically increase the strength of polymerase−DNA interactions in a non-sequence-dependent manner−a desirable characteristic if high processivity on arbitrary DNA substrates is the goal. Then again, speed is also an important factor for any enzyme, and the ability of the polymerase to form a tight, non-sequence-specific complex with DNA has to be counterbalanced by the ability of a polymerase to slide along the DNA quickly as nucleotides are added to a growing primer strand. A metal ion at a polymerase−DNA interface may therefore serve a dual function, one that may be analogous to a "magnetic ball bearing"−the metal ion could help to increase affinity for the DNA substrate (act as a magnet) while at the same time facilitate movement along the DNA (act as a ball bearing). In the case of pol β, only the metal ion in the fingers subdomain (Figure 6) would be required to act in this way because the HhH motif in the 8-kDa domain, which binds the 5′ terminus of gapped DNA substrates (Figure 4, bottom) (Prasad et al., 1994), probably does not have to facilitate movement along the DNA substrate, unless translocation is required for the mechanism of dRP cleavage, as discussed above.

Some caution should be exercised, however, with the idea of a magnetic ball bearing at the polymerase−DNA interface because this analogy implies that the same metal ion remains tightly bound in the HhH metal site as the polymerase slides along the DNA. Although there is evidence for strong metal ion binding in the presence of DNA substrate, we have yet to observe metal ions bound to the HhH motifs in the absence of DNA (Pelletier & Sawaya, 1996), so another possibility suggests itself: the polymerase may slide along the DNA on a "magnetic railing" instead. In this case, metal ions like K^+ and/or Na^+ associate more strongly with the negatively charged DNA phosphate backbone than with the polymerase and form a metal ion coating around the DNA. During processive polymerization the polymerase travels along the DNA backbone guided by this metal ion coating, or magnetic railing, with a new metal ion binding to the polymerase with each new tranlocation event. As with the pol β HhH metal sites in the absence of DNA, however, we also have not observed strong metal ion binding to free-standing DNA phosphates in any of our crystal structures, so it is possible that neither model is relevant. The validity of either model depends on the binding constants and the rates of dissociation of the metal ions (measured for both the free-standing HhH metal site and the free-standing DNA phosphate group) and on the rate of the translocation event itself, which is thought to be extremely rapid (Johnson, 1993). In that none of these numbers is known with great accuracy, the detailed dynamics of metal ion-mediated processivity for pol β remains to be elucidated.

A Model of Pol β Bound to a Gapped-DNA Substrate. As demonstrated above, cocrystallizing pol β with its natural, gapped-DNA substrate has proven to be difficult. In vivo, a typical gapped-DNA substrate is flanked by long stretches of double-stranded DNA on each side of the gap, with no blunt ends of DNA nearby to interfere with binding. Although desirable, this type of substrate is impractical for crystallization. Thus the intermolecular pol β−DNA contacts observed in any one of our cocrystals might be thought to represent nothing more than crystal-packing phenomena. However, one strong argument against this idea, as discussed above, is that crystal structures of pol β−DNA complexes have now been determined in three different space groups [the $P6_1$ and $P2_1$ rat pol β ternary complexes (Pelletier et al., 1994) and the $P2_12_12$ human pol β binary complexes reported here], and all reveal nearly identical protein−DNA interactions in the 8-kDa domain as well as in the fingers subdomain (Figure 5, middle), despite the varying crystal packing arrangements. In addition, the observed pol β−DNA interactions, particularly those involving the 8-kDa domain, are in accord with solution studies (Singhal & Wilson, 1993; Prasad et al., 1994; Matsumoto & Kim, 1995), which suggests that the pol β−DNA crystal structures correctly represent at least some aspects of productive DNA binding. Finally, we have demonstrated that catalysis occurs directly in our 7bp crystals (Pelletier et al., 1996) which is perhaps the most definitive argument in support of catalytically productive binding of a substrate for any enzyme crystal structure. Therefore, specific DNA binding sites in the 8-kDa domain and the fingers subdomain now appear to be well established (three examples each), and a reasonable model for productive binding of pol β to a gapped-DNA substrate can be presented (Figure 9).

The most salient feature of our model is a sharp 90° bend in the single-stranded DNA template, which appears to be stabilized by binding of the 8-kDa domain to the downstream oligonucleotide of a gapped-DNA substrate (Figure 9). It has been proposed that binding of pol β to a gapped-DNA substrate may occur via a pathway whereby the first step is the binding of the 8-kDa domain to the 5′-PO_4 group of the downstream end of the gap (Prasad et al., 1994), and because the 8-kDa domain in the apoenzyme is most likely in an extended, open conformation (Pelletier et al., 1994; Sawaya et al., 1994), perhaps the bend in the DNA template is created when the 31-kDa domain of pol β subsequently folds down and binds to the template−primer segment of the gapped-DNA substrate. In fact, the initial, intermediate, and final position of the 8-kDa domain, relative to the 31-kDa domain, might be represented by the apoenzyme, the $P6_1$ ternary complex, and the $P2_1$ pol ternary complex structures, respectively [see Figure 3 of Pelletier et al. (1994)]. As discussed above, the position of the 8-kDa domain in our model has now been observed twice, offering further support for the idea that this conformation, in particular, probably represents a very stable, final configuration for the 8-kDa domain in its binding to a gapped-DNA substrate. In addition, modeling experiments show that when the 8-kDa domain is in the position shown in Figure 9, it is possible

[5] In that our MES buffer had been titrated with NaOH, all crystallization drops, including those in which NaCl had been replaced by Li_2SO_4 as the crystallization salt, contained a relatively high concentration of Na^+ ions.

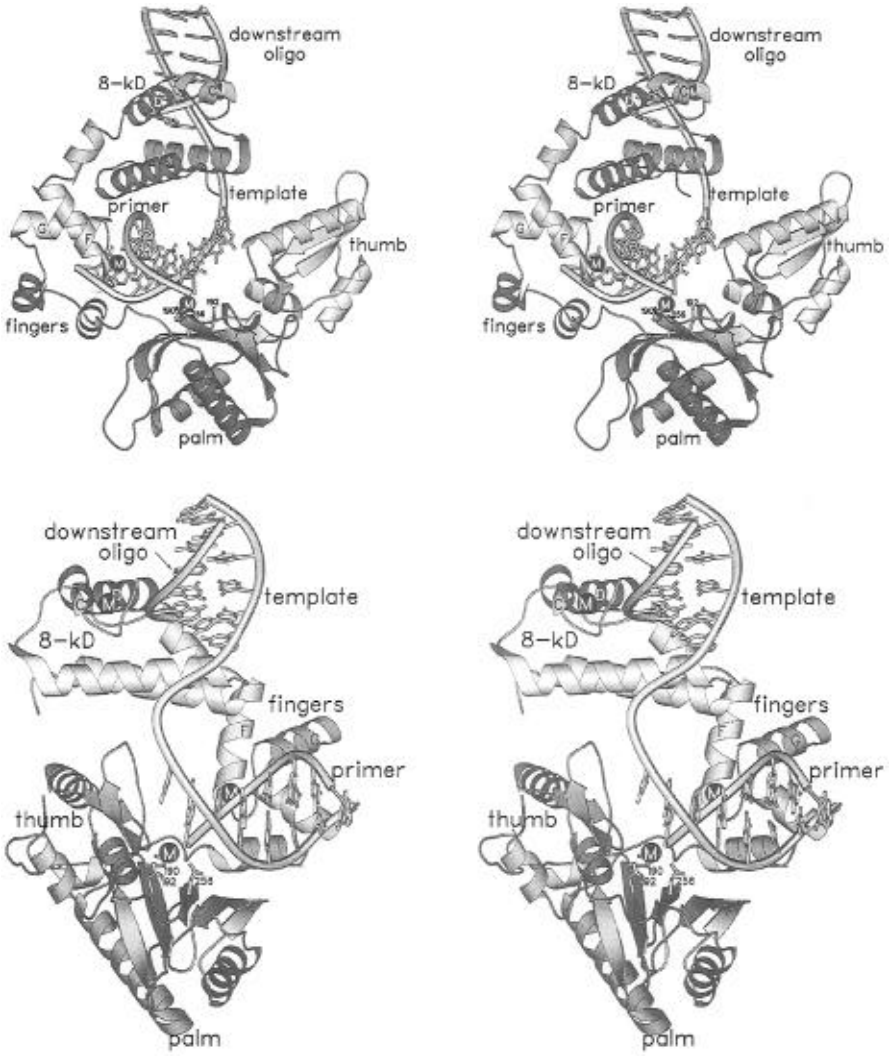

FIGURE 9: SETOR stereo diagrams (Evans, 1993) showing a model for productive binding of pol β to a gapped-DNA substrate. The view of pol β in the top panel is similar to the view in Figure 3. The view in the bottom panel was generated by rotating the top structure by 90° about a vertical axis in the plane of the page and is similar to the view of the lower pol β−DNA complex depicted in Figure 4, top panel. Helices C and D of the HhH motif in the 8-kDa domain and helices F and G of the HhH motif in the fingers subdomain are labeled. Three metal sites (one in the active site, one in the HhH motif of the fingers subdomain, and one in the HhH motif of the 8-kDa domain) are shown as black spheres. A second active-site metal ion is observed only when accompanied by a nucleotide or pyrophosphate (Pelletier et al., 1994, 1996; Sawaya et al., 1994; Pelletier, 1994; Pelletier & Sawaya, 1996) and consequently is not depicted in the model. In that the DNA phosphate groups are not shown, the distances between the HhH metal sites and the DNA backbone may appear further than they really are. The pol β molecule and the DNA template−primer in the binding channel are from one asymmetric unit of the human pol β−DNA binary complex crystal structure, whereas the 5′-phosphorylated downstream oligonucleotide (as well as its complementary template) that is bound to the 8-kDa domain are from a DNA molecule that is associated with the DNA binding channel of a symmetry-related pol β−DNA complex in the unit cell (the symmetry-related pol β molecule is not shown—compare the bottom panel with Figure 4, top panel). Therefore, the only modification of crystallographic data which can be termed modeling is the addition of the connecting template segment for which no bases are shown. We estimate that this model represents a gap in the DNA that is about 3−4 nucleotides long, assuming that the single-stranded template does not bulge or loop out in any way. In that pol β is capable of carrying out strand-displacement synthesis under certain conditions (Mosbaugh & Linn, 1983), another model for productive binding of pol β to a gapped-DNA substrate can be presented whereby the 8-kDa domain displaces the downstream oligonucleotide from the template during processive polymerization. However, such a model would deviate even further from the crystallographic data than our primary model and therefore will not be explored any further at present. Although we have yet to observe a *trans* conformation for the peptide bond between glycine 274 and serine 275 in the thumb subdomain, our model suggests that perhaps some of the steric hindrance that inevitably occurs in the active site when the DNA gap has been reduced to 1 nt may be circumvented somewhat by a collapsible *cis* peptide bond in the thumb subdomain, that is, a *cis* peptide that can easily flip to a *trans* conformation whenever steric hindrance becomes a problem for incorporation of the last nucleotide.

for a DNA gap of about 6 nt to be completely filled in by pol β with only a very slight pivoting of the 8-kDa domain about its connecting hinge with the fingers subdomain. More specifically, we propose that this rotation involves the ϕ and ψ torsion angles of aspartic acid 92, which is located at the amino terminus of helix F, the first helix of the HhH motif in the fingers subdomain (Figures 5 and 9).

Template-Bending and Fidelity. Although structural similarities between pol β and other polymerases [with the exception of the conserved catalytic residues and the palm subdomain that supports them (Sawaya et al., 1994)], are not immediately apparent from the crystal structures, there is nevertheless some support for the idea that structural features that seem to influence nucleotide selectivity and processivity in pol β may have functionally similar counterparts in other polymerases as well. One of the best examples, as discussed above, is that many polymerases appear to have moveable thumb subdomains. However, taking into consideration the observation that a DNA template might bend into the pol β active site during

processive polymerization (Figure 9), perhaps equally thought-provoking is the observation that other polymerases also have structural entities that are similar to pol β's 8-kDa domain in that they rise high above the active site. These structural entities, like pol β's 8-kDa domain, may serve to bend the template down into the active site during processive polymerization. In fact, such a bend may be a critical mechanistic feature utilized by all polymerases to cause base-pairing between the template and an incoming nucleotide to be more sensitive to mismatches. The rationale for this notion is that base-stacking forces are generally non-sequence-specific but nevertheless account for much of the stabilization energy in helical DNA structures. Therefore, a bending away of the DNA template at the active site could function to minimize the potentially powerful, yet indiscriminate, base-stacking energy between the primer terminus and an incoming nucleotide—at least until after a check has been made against the template for proper base pairing. Along these lines, a bend in the template could also function to keep the template strand away from the thumb subdomain during the initial template-checking process (Figure 9), thus forcing all communication between an open thumb subdomain and a bent template to rely on the nature of the incoming nucleotide, which interacts with both. When polymerase structural data (Beese et al., 1993; Sousa et al., 1994; Pelletier et al., 1994, 1996; Sawaya et al., 1994; Rodgers et al., 1995) are combined with data from pre-steady-state kinetic studies (Kuchta et al., 1987; Patel et al., 1991; Wong et al., 1991; Erie et al., 1993; Spence et al., 1995; Frey et al., 1995; Werneburg et al., 1996), an induced-fit mechanism for polymerase fidelity seems plausible. According to this scenario, only a "correct" nucleotide can trigger a global, rate-determining conformational change, in which both a bent DNA template and a highly flexible thumb subdomain simultaneously close down on the active site and allow the nucleotidyl transfer reaction to occur.

Although a large bend in the DNA substrate has also been proposed in a polymerase−DNA model for the Klenow fragment of *E. coli* DNA pol I (Beese et al., 1993), that bend is different from the bend proposed in our pol β−DNA model (Figure 9). In the case of the Klenow fragment, the bend is proposed to occur in the double-stranded portion of the template−primer after nucleotide addition has occurred and after the newly synthesized DNA has exited the polymerase active site, whereas in the case of pol β, the bend is proposed to occur in the single-stranded template overhang of the template−primer before nucleotide addition has occurred and just as the template enters the active site—so that the enhanced nucleotide selectivity that results from the bend occurs right at the active site for pol β.

In addition to shedding light on a possible mechanism for nucleotide selectivity, a template-bending model offers an explanation for how a polymerase such as HIV-1 RT might employ two different modes of binding on a template yet utilize the same active site (Pelletier et al., 1994; Pelletier, 1994). It has been suggested, for instance, that two different modes of template binding would raise substantial problems at the polymerase active site because switching the direction of the template also switches the direction of the strands (Steitz et al., 1994). However, a common template binding site on RT that is located above the active site circumvents such problems because the direction of template−primer binding (that is, the direction of polymerization) then simply depends on the direction of the bend that the template makes as it enters the active site, which in turn may depend on the makeup of the template (RNA versus DNA). In view of our proposals concerning the role of thumb movement and template bending in polymerase fidelity, such a template-bending model for RT may also help to explain the various anomalies observed in studies of RT errors rates with DNA versus RNA templates (Boyer et al., 1992; Hubner et al., 1992; Ji & Loeb, 1992). With DNA templates, RT displays a much lower frequency of base-substitution errors than it does with RNA templates (Hubner et al., 1992; Ji & Loeb, 1992). This is in agreement with a dual mode of template binding because according to such a model, the enhanced nucleotide selectivity that may occur at the active site as a result of a flexible thumb subdomain can only occur when RT is bound to a DNA template (in a pol β-like mode of binding), and not an RNA template (in an anti-pol β mode of binding) (Pelletier et al., 1994). Similarly, with RNA templates, RT displays a much lower frequency of frameshift errors than it does with DNA templates (Boyer et al., 1992). This also agrees with a dual mode of template binding because according to such a model, RT binds to many more base pairs of a RNA/DNA substrate than of a DNA/DNA substrate, suggesting that polymerization on an RNA template would be more processive and accordingly, less susceptible to frameshift errors. The net result is an overall error frequency for RT that is similar for both templates, yet the types of errors observed differ.

It is evident that many more polymerase−DNA crystal structures and solution studies, including those for pol β, are needed to help confirm or to disprove the proposals presented in favor of template bending, but in light of the difficulties encountered in obtaining such cocrystals, perhaps other techniques, such as scanning force microscopy, can be utilized to help settle some of the issues concerning a bend in the DNA template. In fact, evidence in favor of a template-bending mechanism has already been presented by way of scanning force microscopy studies with RNA polymerase from *E. coli*, where 90° bends in the DNA substrate are observed whenever RNA polymerase is in a processive mode of polymerization (Rees et al., 1993). What is most reassuring about these results is that RNA polymerase at about 500 kDa, is one of the largest polymerases known, indicating that DNA bending may indeed be a common strategy employed by all polymerases, from the smallest to the largest, to increase fidelity during processive polymerization.

ACKNOWLEDGMENT

We thank the personnel of the N. H. Xuong Laboratory for aid in data collection, especially C. Nielsen, N. Nguyen, and D. Sullivan. This work was supported in part by grants from the National Institutes of Health (NIH), the R. A. Welch Foundation, and a grant of computing time from the San Diego Super Computer Center.

NOTE ADDED IN PROOF

After this paper had been accepted for publication, a report appeared in the literature describing sequential and structural similarities between the HhH motifs of pol β and potential HhH motifs of other DNA binding proteins (Doherty et al., 1996). In addition to the four known HhH motifs reported here [two from pol β, one from endonuclease III, and one from the 5′ nuclease of DNA pol I (Table 2)], a fifth HhH

motif was confirmed to be present in a recently determined crystal structure of 3-methyladenine DNA glycosylase (MAG or AlkA) (Doherty et al., 1996).

REFERENCES

Aaltonen, L. A., Peltomaki, P., Leach, F. S., Sistonen, P., Pylkkanen, L., Mecklin, J. P., Jarvinen, H., Powell, S. M., Jen, J., Hamilton, S. R., Peterson, G. M., Kinzler, K. W., Vogelstein, B., & de la Chapelle, A. (1993) *Science 260*, 812−816.

Abbotts, J., SenGupta, D. N., Zmudzka, B., Widen, S. G., Notario, V., & Wilson, S. H. (1988) *Biochemistry 27*, 901−909.

Altschal, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990) *J. Mol. Biol. 215*, 403−410.

Arnold, E., Ding, J., Hughes, S. H., & Hostomsky, Z. (1995) *Curr. Opin. Struct. Biol. 5*, 27−38.

Bailly, V., & Verly, W. G. (1987) *Biochem. J. 242*, 565−572.

Bailly, V., & Verly, W. G. (1988) *Biochem. J. 253*, 553−559.

Basu, A., Kedar, P., Wilson, S. H., & Modak, M. J. (1989) *Biochemistry 28*, 6305−6309.

Beese, L. S., Derbyshire, V., & Steitz, T. A. (1993) *Science 260*, 352−355.

Bender, S. L., Widlanski, T., & Knowles, J. R. (1989) *Biochemistry 28*, 7560−7572.

Boorstein, R. J., Hilbert, T. P., Cadet, J., Cunningham, R. P., & Teebor, G. W. (1989) *Biochemistry 28*, 6164−6170.

Boyer, J. C., Bebenek, K., & Kunkel, T. A. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 6919−6923.

Brunger, A. T. (1992) *X-PLOR Manual*, version 3.1, Yale University Press, New Haven, CT.

Cleaver, J. E. (1968) *Nature 218*, 652−656.

Cleaver, J. E., & Kraemer, K. H. (1994) *The Metabolic Basis of Inherited Disease*, 7th ed., McGraw-Hill, New York.

Davies, J. F., Almassy, R. J., Hostomska, Z., Ferre, R. A., & Hostomsky, Z. (1994) *Cell 76*, 1123−1133.

Dianov, G., Price, A., & Lindahl, T. (1992) *Mol. Cell. Biol. 12*, 1605−1612.

Doherty, A. J., Serpell, L. C., & Ponting, C. P. (1996) *Nucleic Acids Res. 24*, 2488−2497.

Erie, D. A., Hajiseyedjavadi, Q., Young, M. C., & von Hippel, P. H. (1993) *Science 262*, 867−873.

Evans, S. V. (1993) *J. Mol. Graphics 11*, 134−138.

Fishel, R., Lescoe, M. K., Rao, M. R., Copeland, N. G., Jenkins, N. A., Garber, J., Kane, M., & Kolodner, R. (1993) *Cell 75*, 1027−1038.

Fornace, A. J., Zmudzka, B., Christine Hollander, M., & Wilson, S. H. (1989) *Mol. Cell. Biol. 9*, 851−853.

Frey, M. W., Sowers, L. C., Millar, D. P., & Benkovic, S. J. (1995) *Biochemistry 34*, 9185−9192.

Friedberg, E. C., Walker, G. C., & Siede, W. (1995) *DNA Repair and Mutagenesis*, ASM Press, New York.

Gallopo, A. R., & Cleland, W. W. (1979) *Arch. Biochem. Biophys. 195*, 152−154.

Gu, H., Marth, J. D., Orban, P. C., Mossmann, H., & Rajewsky, K. (1994) *Science 265*, 103−106.

Hamlin, R. (1985) *Methods Enzymol. 114*, 416−452.

Henikoff, S., & Henikoff, J. G. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 10915−10919.

Horton, J. K., Srivastava, D. K., Zmudzka, B. Z., & Wilson, S. H. (1995) *Nucleic Acids Res. 23*, 3810−3815.

Howard, A. J., Nielsen, C., & Xuong, H. H. (1985) *Methods Enzymol. 114*, 452−472.

Hubner, A., Kruhoffer, M., Grosse, F., & Krauss, G. (1992) *J. Mol. Biol. 223*, 595−600.

Ito, J., Braithwaite, D. K. (1991) *Nucleic Acids Res. 19*, 4045−4057.

Jacobo-Molina, A., Ding, J., Nanni, R. B., Clark, A. D., Lu, X., Tantillo, C., Williams, R. L., Kamer, G., Ferris, A. L., Clark, P., Hizi, A., Hughes, S. H., & Arnold, E. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 6320−6324.

Ji, J., & Loeb, L. A. (1992) *Biochemistry 31*, 954−958.

Johnson, K. A. (1993) *Annu. Rev. Biochem. 62*, 685−713.

Kim, J., & Linn, S. (1988) *Nucleic Acids Res. 16*, 1135−1141.

Kohlstaedt, L. A., Wang, J., Friedman, J. M., Rice, P. A., & Steitz, T. A. (1992) *Science 256*, 1783−1790.

Kornberg, A., & Baker, T. A. (1992) *DNA Replication*, 2nd ed., Freeman, San Francisco, CA.

Kow, Y. W., & Wallace, S. S. (1987) *Biochemistry 26*, 8200−8206.

Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.

Kuchta, R. D., Mizrahi, V., Benkovic, P. A., Johnson, K. A., & Benkovic, S. J. (1987) *Biochemistry 26*, 8410−8417.

Kuo, C.-F., McRee, D. E., Fisher, C. L., O'Handley, S. F., Cunningham, R. P., & Tainer, J. A. (1992) *Science 258*, 434−440.

Leach, F. S., Nicolaides, N. C., Papadopoulos, N., Liu, B., Jen, J., Parsons, R., Peltomaki, P., Sistonen, P., Aaltonen, L. A., Nystrom-Lahti, M., Guan, X.-Y., Zhang, J., Chen, D. J., Cerosalettie, K. M., Fournier, R. E. K., Todd, S., Lewis, T., Leach, R. J., Naylor, S. L., Weissenbach, J., Mecklin, J.-P., Jarvinen, H., Peterson, G. M., Hamilton, S. R., Green, J., Jass, J., Watson, P., Lynch, H. T., Trent, J. M., de la Chapelle, A., Kinzler, K. W., & Vogelstein, B. (1993) *Cell 75*, 1215−1225.

Manoharan, M., Ransom, S. C., Mazumder, A., & Gerlt, J. A. (1988) *J. Am. Chem. Soc. 110*, 1620−1622.

Matsumoto, Y., & Bogenhagen, D. F. (1989) *Mol. Cell. Biol. 9*, 3750−3757.

Matsumoto, Y., & Kim, K. (1995) *Science 269*, 699−702.

Mazumder, A., Gerlt, J. A., Absalon, M. J., Stubbe, J., Cunningham, R. P., Withka, J., & Bolton, P. H. (1991) *Biochemistry 30*, 1119−1126.

Modrich, P. (1994) *Science 266*, 1959−1960.

Moews, P. C., & Kretsinger, R. H. (1975) *J. Mol. Biol. 91*, 201−225.

Mosbaugh, D. W., & Linn, S. (1983) *J. Biol. Chem. 258*, 108−118.

Mozzherin, D. J., & Fisher, P. A. (1996) *Biochemistry 35*, 3572−3577.

Nishida, C., Reinhard, P., & Linn, S. (1988) *J. Biol. Chem. 263*, 501−510.

Ollis, D. L., Brick, P., Hamlin, R., Xuong, N. G., & Steitz, T. A. (1985) *Nature 313*, 762−766.

Parsons, R., Li, G. M., Longley, M. J., Fang, W. H., Papadopoulos, N., Jen, J., de la Chapelle, A., Kinzler, K. W., Vogelstein, B., & Modrich, P. (1993) *Cell 75*, 1227−1236.

Patel, S. S., Wong, I., & Johnson, K. A. (1991) *Biochemistry 30*, 511−525.

Patel, P. H., Jacobo-Molina, A., Ding, J., Tantillo, C., Clark, A. D., Raag, R., Nanni, R. G., Hughes, S. H., & Arnold, E. (1995) *Biochemistry 34*, 5351−5363.

Pelletier, H. (1994) *Science 266*, 2025−2026.

Pelletier, H., & Sawaya, M. R. (1996) *Biochemistry 35*, 12778−12787.

Pelletier, H., Sawaya, M. R., Kumar, A., Wilson, S. H., & Kraut, J. (1994) *Science 264*, 1891−1903.

Pelletier, H., Sawaya, M. R., Wolfle, W., Wilson, S. H., & Kraut, J. (1996) *Biochemistry 35*, 12762−12777.

Peltomaki, P., Aaltonen, L. A., Sistonen, P., Pylkkanen, L., Mecklin, J. P., Jarvinen, H., Green, J. S., Jass, J. R., Weber, J. L., Leach, F. S., Peterson, G. M., Hamilton, S. R., de la Chapelle, A., & Vogelstein, B. (1993) *Science 260*, 810−812.

Prasad, R., Beard, W. A., & Wilson, S. H. (1994) *J. Biol. Chem. 269*, 18096−18101.

Randahl, H., Elliott, G. C., & Linn, S. (1988) *J. Biol. Chem. 263*, 12228−12234.

Reardon, J. E. (1993) *J. Biol. Chem. 268*, 8743−8751.

Rees, W. A., Keller, R. W., Vesenka, J. P., Yang, G., & Bustamante, C. (1993) *Science 260*, 1646−1649.

Rodgers, D. W., Gamblin, S. J., Harris, B. A., Ray, S., Culp, J. S., Hellmig, B., Woolf, D. J., Debouck, C., & Harrison, S. C. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 1222−1226.

Saraste, M., Sibbald, P. R., & Wittinghofer, A. (1993) *TIBS 15*, 430−434.

Sawaya, M. R., Pelletier, H., Kumar, A., Wilson, S. H., & Kraut, J. (1994) *Science 264*, 1930−1935.

Seeberg, E., Eide, L., & Bjoras, M. (1995) *Trends Biochem. Sci. 20*, 391−397.

SenGupta, D. N., Zmudzka, B. Z., Kumar, P., Cobianchi, F., Skowronski, J., & Wilson, S. H. (1986) *Biochem. Biophys. Res. Commun. 136*, 341−347.

Singhal, R. K., & Wilson, S. H. (1993) *J. Biol. Chem. 268*, 15906−15911.

Singhal, R. K., Prasad, R., & Wilson, S. H. (1995) *J. Biol. Chem. 270*, 949−957.
Sobol, R. W., Horton, J. K., Singhal, R. K., Prasad, R., Kuhn, R., Gu, H., Rajewsky, K., & Wilson, S. H. (1996) *Nature 379*, 183−185.
Sousa, R., Rose, J., & Wang, B. C. (1994) *J. Mol. Biol. 244*, 6−12.
Spence, R. A., Kati, W. M., Anderson, K. S., & Johnson, K. A. (1995) *Science 267*, 988−993.
Steitz, T. A., Smerdon, S. J., Jager, J., & Joyce, C. M. (1994) *Science 266*, 2022−2025.
Thayer, M. M., Ahern, H., Xing, D., Cunningham, R. P., & Tainer, J. T. (1995) *EMBO J. 14*, 4108−4120.
Tronrud, D. E., Ten Eyck, L. F., & Matthews, B. W. (1987) *Acta Crystallogr. A43*, 489−501.
Wang, T. S.-F., & Korn, D. (1982) *Biochemistry 21*, 1597−1608.
Wang, Z., Wu, X., & Friedberg, E. C. (1993) *Mol. Cell. Biol. 13*, 1051−1058.
Werneburg, B. G., Ahn, J., Zhong, X., Hondal, R. J., Kraynov, V. S., & Tsai, M.-D. (1996) *Biochemistry 35*, 7041−7050.
Widlanski, T., Bender, S. L., & Knowles, J. R. (1989) *J. Am. Chem. Soc. 111*, 2299−2300.
Wiebauer, K., & Jiricny, J. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 5842−5845.
Wilde, J. A., Bolton, P. H., Mazumder, A., Manoharan, M., & Gerlt, J. A. (1989) *J. Am. Chem. Soc. 111*, 1894−1896.
Wohrl, B. M., Georgiadis, M. M., Telesnitsky, A., Hendrickson, W. A., & Le Grice, S. F. (1995) *Science 267*, 96−99.
Wong, I., Patel, S. S., & Johnson, K. A. (1991) *Biochemistry 30*, 526−537.
Zmudzka, B. Z., Fornace, A., Collins, J., & Wilson, S. H. (1988) *Nucleic Acids Res. 16*, 9587−9596.

BI952955D