Subtle but variable conformational rearrangements in the replication cycle of *Sulfolobus solfataricus* P2 DNA polymerase IV (Dpo4) may accommodate lesion bypass

YANLI WANG, KARUNESH ARORA, AND TAMAR SCHLICK

Department of Chemistry and Courant Institute of Mathematical Sciences, New York University, New York, New York 10012-2005, USA

(RECEIVED July 25, 2005; FINAL REVISION October 10, 2005; ACCEPTED October 11, 2005)

Abstract

The possible conformational changes of DNA polymerase IV (Dpo4) before and after the nucleotidyl-transfer reaction are investigated at the atomic level by dynamics simulations to gain insight into the mechanism of low-fidelity polymerases and identify slow and possibly critical steps. The absence of significant conformational changes in Dpo4 before chemistry when the incoming nucleotide is removed supports the notion that the "induced-fit" mechanism employed to interpret fidelity in some replicative and repair DNA polymerases does not exist in Dpo4. However, significant correlated movements in the little finger and finger domains, as well as DNA sliding and subtle catalytic-residue rearrangements, occur after the chemical reaction when both active-site metal ions are released. Subsequently, Dpo4's little finger grips the DNA through two arginine residues and pushes it forward. These metal ion correlated movements may define subtle, and possibly characteristic, conformational adjustments that operate in some Y-family polymerase members in lieu of the prominent subdomain motions required for catalytic cycling in other DNA polymerases like polymerase β. Such subtle changes do not easily provide a tight fit for correct incoming substrates as in higher-fidelity polymerases, but introduce in low-fidelity polymerases different fidelity checks as well as the variable conformational-mobility potential required to bypass different lesions.

Keywords: Dpo4; dynamics simulations; replication mechanism; conformational changes; catalytic metal ion; nucleotide-binding metal ion

Supplemental material: see www.proteinscience.org

Reprint requests to: Tamar Schlick, Department of Chemistry and Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, NY 10012-2005, USA; e-mail: schlick@nyu.edu; fax: (212) 995-4152.

Abbreviations: Dpo4, *Sulfolobus solfataricus* P2 DNA polymerase IV; 8-oxoG, 8-oxoguanine; pol β, DNA polymerase β; RMSD, root-mean-square deviation; dCTP, 2′-deoxyribocytidine 5′-triphosphate; dNTP, 2′-deoxyribonucleoside 5′-triphosphate; ddATP, 2′,3′-dideoxyriboadenosine 5′-triphosphate; LF, little finger; PP_i, pyrophosphate; PCA, principal component analysis; bp, base pair.

Article published online ahead of print. Article and publication date are at http://www.proteinscience.org/cgi/doi/10.1110/ps.051726906.

High-fidelity DNA synthesis is primarily performed by replicative DNA polymerases from the A, B, C, D, and X families. However, when these processive polymerases stall at damaged DNA sites, special Y-family polymerases that lack proofreading activity and exhibit low fidelity and processivity can bypass the lesions. Y-family polymerases exist in all life forms and play a vital role in an organism's survival (Friedberg et al. 2002; Yeiser et al. 2002).

One of the well-characterized Y-family polymerases is the thermophilic archaeal *Sulfolobus solfataricus* P2

DNA polymerase IV (Dpo4). Dpo4 has an error rate of 10^{-3} to 10^{-4} when synthesizing undamaged DNA and can traverse various lesions, such as abasic sites, bulky DNA adducts, and *cis-syn* thymine–thymine dimmers (Boudsocq et al. 2001).

Our understanding of low-fidelity DNA synthesis and lesion-bypassing mechanisms of Dpo4 has been advanced with various kinetic and in vitro experimental studies as well as the determination of a series of atomic structures. The pioneering Yang group has provided invaluable crystallographic ternary structures of the Dpo4/DNA complex with an incoming nucleotide paired with lesioned and nonlesioned templating bases (Ling et al. 2001, 2003, 2004a,b). Kinetic and in vitro experimental works have measured Dpo4 fidelity and proposed possible mechanisms for Dpo4's enzymatic function (Boudsocq et al. 2004; Fiala and Suo 2004a,b).

Recently, the *ternary* crystal structure of the Dpo4/DNA complex with 8-oxoguanine (8-oxoG) paired with dCTP in a matched *anti*:*anti* conformation at the active site, and the corresponding *binary* structures before and after the chemical reaction of primer extension have been solved in the Patel laboratory (O. Rechkoblit, L. Malinina, Y. Cheng, V. Kuryavyi, S. Broyde, N.E. Geacintov, and D.J. Patel, in prep.; Fig. 1A,B). Specifically, the binary crystal structure before chemistry is substrate-free and only comprised of the Dpo4 protein and a DNA duplex; the ternary crystal structure consists of the Dpo4/DNA complex and an incoming nucleotide dCTP pairing with 8-oxoG at the active site; and the binary crystal structure after chemistry is the product complex with the incoming nucleotide dCTP incorporated to the DNA duplex primer terminus opposite 8-oxoG. 8-OxoG (Fig. 1D) is an abundant lesion in DNA caused by the oxidation of guanine by reactive oxygen species and is known to contribute to carcinogenesis and premature aging in higher organisms (Ames and Gold 1991). It is mutagenic because of the potential of 8-oxoG to mispair with adenine during replication and generate G to T transversions (Cheng et al. 1992; Moriya 1993; Shibutani and Grollman 1994).

The superimposition of the ternary and binary crystal structures of the Dpo4/DNA complex with 8-oxoG according to their palm domains (Fig. 1B) indicates no large conformational changes in the active site before and after the nucleotidyl-transfer reaction. However, the little finger (LF) domain of Dpo4 rotates away from and toward the finger domain by 12° before and after the reaction, respectively, and the α-helices H and K of the thumb domain contact the next nucleotide backbone of the DNA duplex after chemistry. The DNA appears to slide by a half base pair's distance in both the pre- and post-chemistry processes so that the DNA moves forward by a base pair after one complete reaction cycle. A comparison of the Dpo4 protein residues between the binary crystal structure before and after chemistry shows that they are nearly superimposable to one another, with the root-mean-square deviation (RMSD) of all the backbone C^{α} atoms of only 0.8 Å; the difference mainly stems from the little finger domains. Thus, the major difference between the binary substrate-free crystal structure before chemistry and the binary product complex after chemistry is in the sliding of the DNA duplex of the former by 1 bp.

The catalytic core of Dpo4 is shaped like a right hand comprised of thumb, palm, and finger domains as in high-fidelity DNA polymerases, but it has an additional "little-finger" (LF) domain located at the C terminus and tethered to the thumb by a positively-charged loop. The LF domain enhances the attachment between Dpo4 and DNA by gripping on the DNA major groove. The active site of the Dpo4 ternary structure containing 8-oxoG is depicted in Figure 1C, and some key residues are listed in Table 1 and compared to their analogs in human DNA polymerase β (pol β). Significantly, the residues from the finger domain of Dpo4 (Tyr12, Val32, Ala42, Ala44, and Gly58) that interact with the template and incoming nucleotide bases are much smaller than their analogs in pol β (Table 1). Tyr10, Thr45, Tyr48, Arg51, and Lys159 form hydrogen bonds with the oxygen atoms on the β- and γ-phosphates of the incoming dCTP. Moreover, Arg51 and Pro160 stack with Tyr48 to stabilize the active site. Experimental and computational studies have shown that the spacious active site of Dpo4 can accommodate bulky lesions (Ling et al. 2003, 2004b; Perlow-Poehnelt et al. 2004; Wang et al. 2005), and this is believed to contribute to Dpo4's low fidelity (Yang 2003). In particular, it was shown that bulky carcinogen–DNA adducts, 10S-(+)-*trans-anti*-[BP]-N2-dG and 1S-(−)-Benzo[c]phenanthrenyl-N^6-dA, can be accommodated in the active site and bypassed by Dpo4 in different conformations (Perlow-Poehnelt et al. 2004; Wang et al. 2005).

Three conserved catalytic carboxylate residues (Asp7, Asp105, and Glu106) in the Dpo4 palm domain are coordinated with two metal ions for the phosphoryl-transfer reaction. The coordination details of the metal ions in the ternary crystal structure of the Dpo4/DNA complex with 8-oxoG are shown in Figure 1C. The catalytic metal ion A coordinates Asp7, Glu106, a phosphate oxygen atom (O1P) of primer dG, and three water oxygen atoms, while the nucleotide-binding ion B coordinates with Asp7, Asp105, main-chain carbonyl (C=O) group of Phe8, and three phosphate oxygen atoms on dCTP. In each of the binary crystal structures, the nucleotide-binding ion and the PP_i group have been released from the active site, and only the catalytic ion (Ca^{2+}) is present. This ion coordinates with Asp7, Asp105, Asp106, a phosphate oxygen atom of the primer

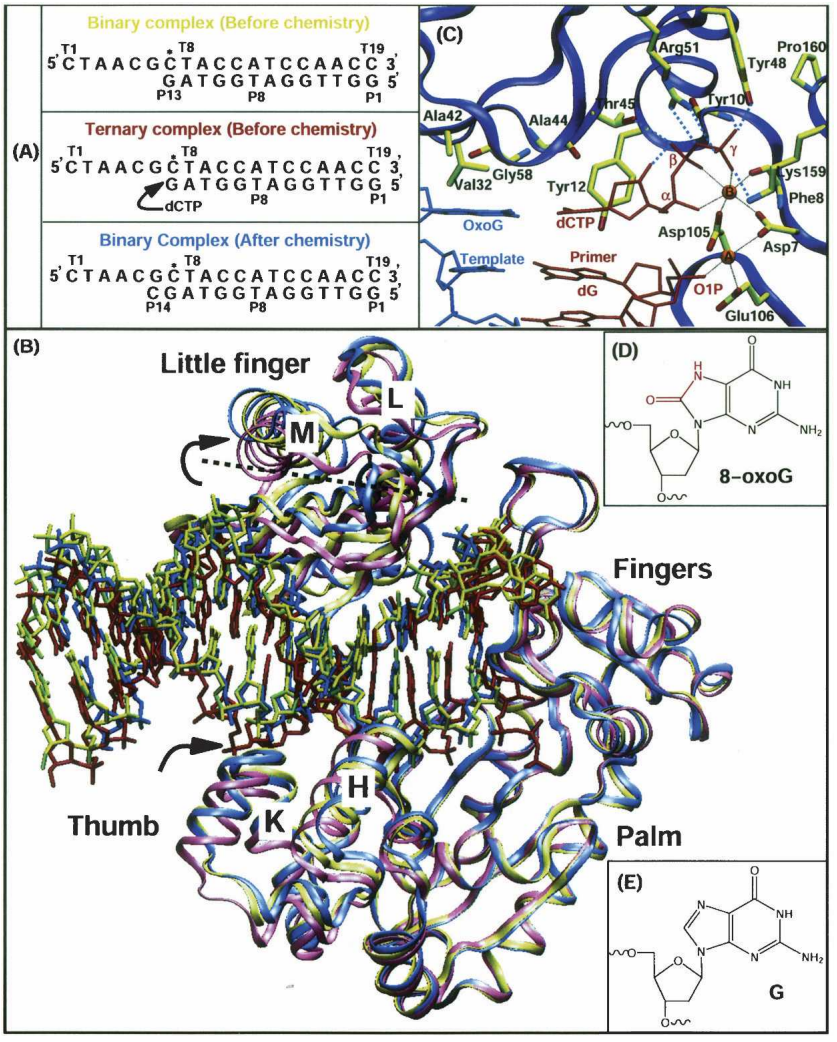

Figure 1. (*A*) DNA sequences of the binary and ternary crystal structures of the Dpo4/DNA complexes with 8-oxoG before chemistry, and the binary complex after chemistry as solved by the Patel group (O. Rechkoblit, L. Malinina, Y. Cheng, V. Kuryavyi, S. Broyde, N.E. Geacintov, and D.J. Patel, in prep.). 8-OxoG on the template strand is marked with an asterisk on *top* of G. (*B*) Structural comparison between the ternary before chemistry (orange) and the binary before (light blue) and after (light green) chemistry crystal structures. Their corresponding bound DNA duplexes are in red, blue, and green, respectively. The α-helices H, K, L, and M are labeled on the structures. Black arrows in *B* represent the domain motion direction after chemistry. A rotation axis of the LF domain is shown by a dashed line (discussed later). (*C*) Active site of Dpo4 in the ternary crystal structure of Dpo4/DNA/8-oxoG:dCTP. The residues interacting with the nascent base pair, the triphosphate moiety, and the metal ions are labeled. Hydrogen bonds and coordinations of the two metal ions (*A*, catalytic ion; *B*, nucleotide-binding ion) are shown by blue and black dashed lines, respectively. The metal ions are identified as Ca^{2+} in the crystal structure. *D* and *E* display the molecular formulas of 8-oxoG and guanine, respectively.

dG, and two water oxygen atoms in the binary substrate-free structure before chemistry, whereas it binds Asp7, Asp105, a phosphate oxygen atom of the newly inserted dC, and three water molecules in the binary product structure after chemistry.

The two-metal ion mechanism (Steitz 1993) has long been considered a general strategy used by various DNA polymerases to incorporate nucleotides to the DNA primer strand. The ions also help assemble the catalytic carboxylate groups before the phosphoryl-transfer reaction (Yang et al. 2004) and stabilize the structure of the pentacovalent transition state (Abashkin et al. 2001). Furthermore, the nucleotide-binding ion assists the departure of the pyrophosphate byproduct from the active site. Dynamics simulations performed to investigate the effects of the two metal ions (Mg^{2+}) on pol β's conformational changes before and after chemistry (Yang et al. 2004) indicate that the metal ions are not only crucial for the chemical reaction itself, but also play an important role in determining the outcome of the conformational changes

Table 1. *Important residues in the active site of Dpo4 and their analogs in pol β*

Residues in Dpo4	Domain	Analogs in Pol β	Subdomain	Function
Y12, V32, A42, A44, G58	Finger	Y271, F272, T273, G274, S275, D276, R283	Thumb	Interact with the template and incoming nucleotide bases
Y10, T45, Y48, R51, K159	Finger and Palm	S180, R183, G189	Palm	Interact with the triphosphate of the incoming dNTP
D7, D105, E106	Palm	D190, D192, D256	Palm	Conserved catalytic triad; D7, D105, and E106 reorient after the metal ions are released (discussed later)
Y108	Palm	R258	Palm	Y108 stabilizes E106 when the catalytic ion is released (discussed later)

that affect the DNA synthesis fidelity. The ions are also associated with slow substeps in the conformational change. Thus, the metal ions are a crucial part of the machinery employed by pol β to regulate fidelity, and associated correlated movements of the active-site residues and the metal ions define slow steps in the enzyme's cycle (Radhakrishnan and Schlick 2004, 2005).

DNA pol β is part of a group of moderate and high-fidelity polymerases, which also include T7 DNA polymerase and HIV-1 RT, that undergo large-scale subdomain conformational transition from an "open" to a "closed" state upon binding the correct dNTP. This "induced-fit" mechanism has been deduced on the basis of several structural and gained supports from computational studies (Sawaya et al. 1997; Arora and Schlick 2004; Krahn et al. 2004). In contrast, clear evidence of such conformational changes is lacking for Y-family polymerases including Dpo4, except that the LF domain exhibits displacements upon comparison of the binary and ternary crystal structures of the Dpo4/DNA complexes with 8-oxoG, as well as from the ternary crystal structure of the Dpo4/DNA complex with an abasic site (Ling et al. 2004a).

While recent pre-steady-state kinetic data (Fiala and Suo 2004a) suggest that the "induced-fit" mechanism (involving yet undefined structural changes) is applicable to Dpo4, other studies predict that no large conformational changes at the active site of Y-family polymerases exist when incorporating nucleotides (Ling et al. 2001; Silvian et al. 2001). Recently, Fleck and Schär proposed instead an "induced-grip" mechanism by the LF domain for Dpo4's catalysis (Fleck and Schaär 2004) based on the crystal structure of Dpo4 with an abasic lesion (Ling et al. 2004a). The particular role of the LF domain was further investigated in a clever experimental work (Boudsocq et al. 2004), in which the LF domains of Dpo4 and Dbh (both archeal Y-family polymerases) were swapped; this experiment showed that the newly built chimeric proteins adopted the functions associated with the original proteins that contain the corresponding LF domains. Thus, Dpo4's LF domain clearly plays a central role in determining Dpo4's fidelity, processivity, and lesion bypass functions. However, an atomic-level structure/function understanding of how exactly the LF domain helps accommodate lesions and affects Dpo4 fidelity is still lacking. Modeling and simulation, subject to the usual limitations of approximate force fields and limited sampling (Schlick 1999), can also pinpoint the sequential steps in the enzyme's cycle that may regulate fidelity (Radhakrishnan and Schlick 2004, 2005).

To investigate possible conformational changes implied by the three crystal complexes of Dpo4 with 8-oxoG and interpret the function of the LF domain and the DNA synthesis mechanism of Dpo4, we explore by dynamics simulations both the processes of prior to, and following, chemistry at atomic level as done for DNA pol β (Yang et al. 2002a; Arora and Schlick 2004). We use the available ternary crystal structure of the Dpo4/DNA complex with 8-oxoG and dCTP at the template/primer junction (O. Rechkoblit, L. Malinina, Y. Cheng, V. Kuryavyi, S. Broyde, N.E. Geacintov, and D.J. Patel, in prep.) to construct initial models; only for RMSD analysis, we use the analogous binary crystal complexes (binary substrate-free complex before chemistry for the simulations before chemistry, and the binary product complex after chemistry for the simulations after chemistry). Specifically, we performed two dynamics simulations before the nucleotidyl-transfer reaction and two simulations after the reaction.

Results

Simulations before chemistry

Domain motions

A simulation of the ternary crystal complex of Dpo4/DNA with 8-oxoG and the incoming nucleotide dCTP was performed as a benchmark, and, as expected, no large conformational changes were observed during the 15-nsec simulation. The simulation was very stable throughout.

The second pre-chemistry simulation was performed with dCTP and metal ions removed from the active site of the ternary complex. We expected that if the "induced-fit" mechanism existed, Dpo4 would undergo a transition from the ternary (closed) to the binary (open) state as pol β (Arora and Schlick 2004). However, no clear transition to the "open" state in Dpo4 was captured, although the finger and LF domains of Dpo4 displayed more fluctuations than in the first trajectory with the incoming nucleotide. This can be seen from the RMSD plots of the finger and LF domains with respect to the binary and ternary crystal complexes for both trajectories in Figure 2, A to D.

These simulations lead us to suggest that either there are no large conformational changes involved in the pre-chemistry phase of DNA replication by Dpo4 or that

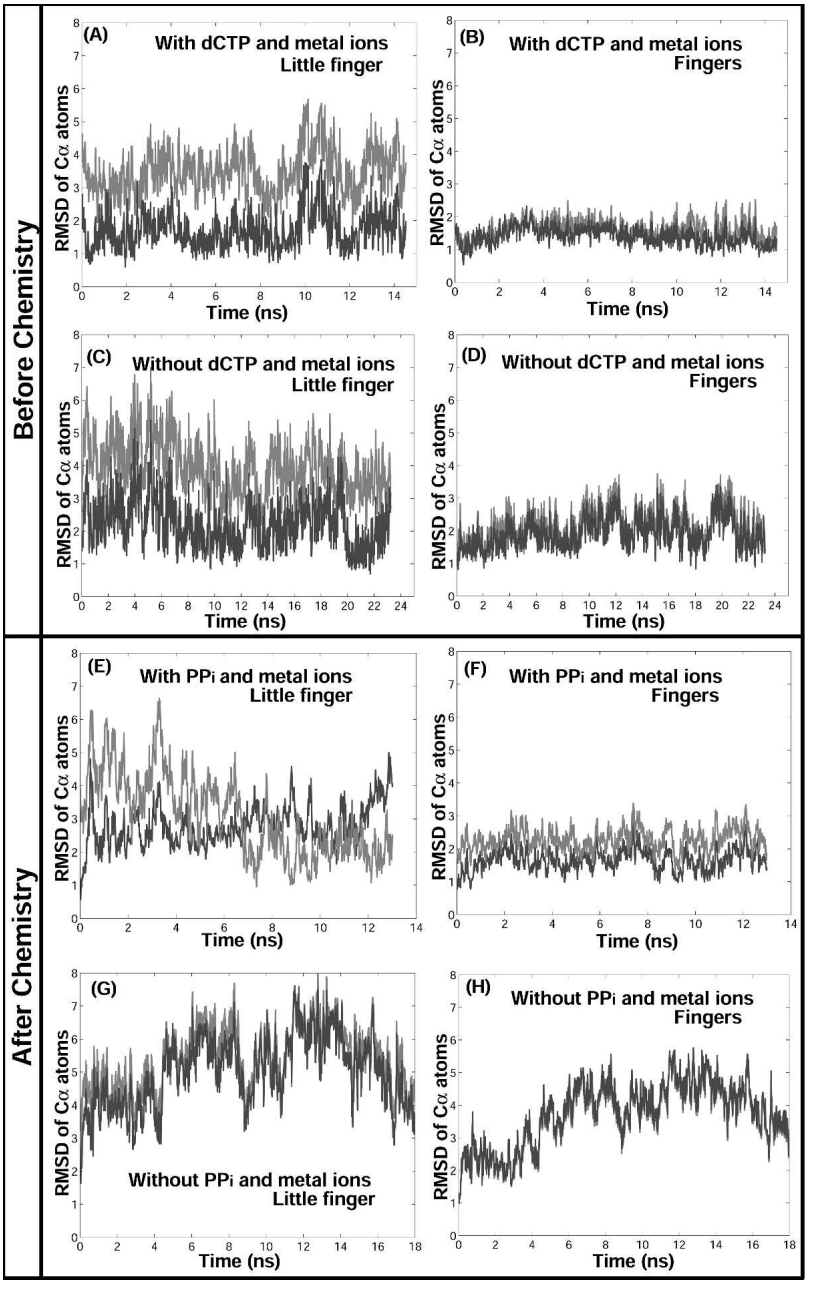

Figure 2. Evolution of the RMSD for the C^{α} atoms in the LF and finger domains of Dpo4 in the simulations before and after chemistry relative to the crystal binary (gray) and ternary complexes (black). (*A–D*) The RMSD plots of the LF and finger domains in the simulations before chemistry with and without the incoming nucleotide in the active site (see labels). (*E–H*) Those in the simulations after chemistry with and without metal ions in the active site. Superimpositioning is performed according to the palm domains.

notable deformations indicated by the binary substrate-free and ternary crystal structures before chemistry, including the ~12° rotation in the LF domain as well as the DNA translocation of a half base pair's rise, occur over longer timescales than 23 nsec simulated here.

Metal ions coordination and nucleotidyl transfer geometry

In all the available ternary crystal structures of Dpo4/DNA complexes with incoming nucleotide (Ling et al. 2001, 2003, 2004a,b), the primer terminus lacks the 3′-OH group. Thus, the effect of the missing 3′-OH on the magnesium ion coordination and the polymerase active-site geometry cannot be inferred from the crystal data. In our simulations, the missing 3′-OH group at the primer terminus was modeled in the starting structure. Figure 3A shows the magnesium coordination in the final simulated structure of the Dpo4/DNA complex with 8-oxoG and the incoming dCTP (in the control simulation). As seen in Figure 3A, both metal ions are hexa-coordinated. The catalytic ion is coordinated with Asp7, Glu106, the phosphate oxygen (O1P) of the primer terminus, and three water molecules. The nucleotide-binding ion is coordinated with three oxygen atoms on the triphosphate, Asp7, Asp105, and Phe8. The average distances between the metal ions and the coordinating ligands in the trajectory are listed in Table 2. Our simulated structures mimic the metal ion coordination well as in ternary crystal complex, but, as expected, the ligand/ion coordination distances are slightly shorter on average than in the crystal complex.

Significantly, the arrangement of the ligands bound to the catalytic metal ion is different in comparison with (moderate-fidelity) DNA pol β (Sawaya et al. 1997). Specifically, the catalytic ion in Dpo4 coordinates the phosphate oxygen (O1P) of the primer terminus rather than the α-phosphate oxygen (O1A) of the incoming nucleotide, and thus the distance between the ion and the O3′ of the incoming nucleotide is more than 6 Å throughout the simulation. We also observe that the crucial distance between P_α of the incoming nucleotide and the O3′ of the primer terminus is ~4.6 Å on average, larger than the ideal P_α–O3′ distance (3.3 Å) for the phosphoryl transfer reaction to proceed via a dissociative mechanism (Mildvan 1997). Note that the conformations of the active-site residues and especially the coordination of the metal ions captured in the crystal and simulated ternary structures do not necessarily represent the final reaction-competent state for the chemical reaction, because these structures may only be a local minimum around the reaction path. For the chemical reaction to occur, additional barriers must be surpassed, which might include the rearrangements of the catalytic residues as well as the metal ions to achieve the ideal geometry of reactants for the reaction. Indeed, for pol β, we have proposed a "pre-chemistry avenue" that lies between the closed state and the reaction state of pol β (Arora et al. 2005; Radhakrishnan and Schlick 2005).

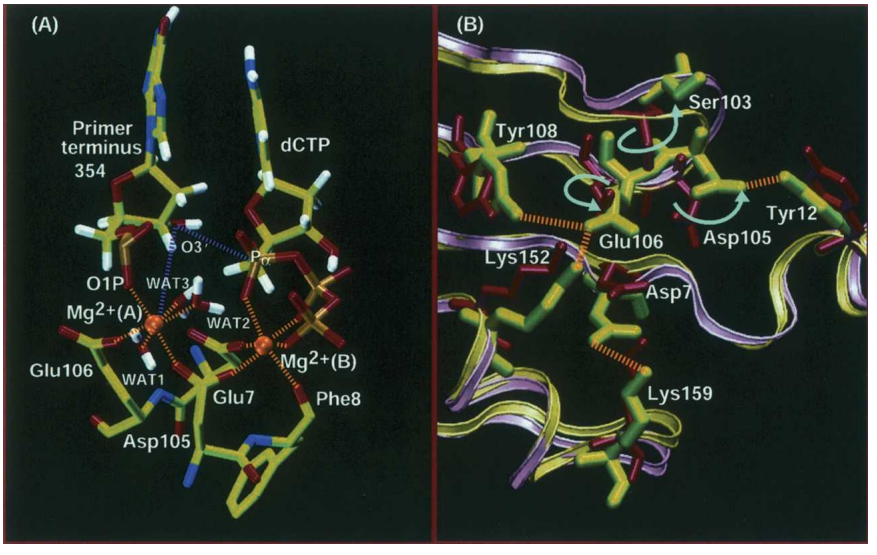

Figure 3. (*A*) Coordination of metal ions in the Dpo4/DNA ternary complex at the simulation's end. Interactions between the metal ions and ligands are shown in yellow. Corresponding average coordination distances are listed in Table 2. (*B*) The conformations of the catalytic triad (conserved carboxylates Asp7, Asp105, and Glu106) before and after the simulation before chemistry with metal ions and dCTP removed. Protein backbones in the simulated and ternary crystal structures are represented by light green and light red traces, respectively. The corresponding amino acids in the complexes are shown by green and red stick models, respectively.

Table 2. *Magnesium ions coordination distances (in Å) in the simulated Dpo4 ternary complex and its crystal form*

Distance	Dpo4/DNA/dCTP	Dpo4[a]
Mg^{2+}(A)–Asp7:$O^{\delta1}$	1.83	2.47
Mg^{2+}(A)–Glu106:$O^{\delta2}$	1.83	2.31
Mg^{2+}(A)–354:O1P	1.89	2.32
Mg^{2+}(A)–WAT1	1.99	N.A.
Mg^{2+}(A)–WAT2	2.01	N.A.
Mg^{2+}(A)–WAT3	1.98	N.A.
Mg^{2+}(B)–Asp7:$O^{\delta2}$	1.89	2.30
Mg^{2+}(B)–Phe8:O	1.89	2.30
Mg^{2+}(B)–Asp105:$O^{\delta2}$	1.84	2.30
Mg^{2+}(B)–dCTP:$O2^{\alpha}$	1.92	2.48
Mg^{2+}(B)–dCTP:$O2^{\beta}$	1.90	2.28
Mg^{2+}(B)–dCTP:$O2^{\gamma}$	1.82	2.31

Mg^{2+}(A), catalytic magnesium; Mg^{2+}(B), nucleotide-binding magnesium; dCTP, 2′-deoxyadenosine 5′-triphosphate; 354, primer terminus; N.A., not available.
[a] In the crystal complex metal ions were resolved to be calcium.

The proposed scenario may also exist in Dpo4 before the reaction, and we speculate based on the greater deviations in Dpo4's active site from the ideal geometry for the reaction (Mildvan 1997; Lahiri et al. 2003) that the energy barrier for the "pre-chemistry avenue" of Dpo4 might be larger than pol β.

The active-site residues of the Dpo4/DNA complex undergo only subtle conformational changes in response to the removal of the incoming nucleotide and metal ions. In the trajectory with dCTP and metal ions removed from the ternary complex, the catalytic triad Asp7, Asp105, and Glu106 flip or rotate from their original positions (Fig. 3B). Notably, Glu106 rotates toward Tyr108 and is stabilized by forming a hydrogen bond with Tyr108. Asp7 forms a hydrogen bond with Lys159, and Asp105 has a tendency to interact with Tyr12 in the enzyme's active site. Comparatively, the conformational changes of the acidic aspartate residues in pol β in response to the incoming nucleotide and metal ions are correlated with other interacting residues' side chains (Arg258, Phe272, and Asp192) as well as the thumb subdomain motion (Yang et al. 2002a; Arora and Schlick 2004; Radhakrishnan and Schlick 2004).

Simulations after chemistry

Domain and DNA motions

In our post-chemistry simulation where the bound ions and PP_i are retained in the active site of the Dpo4/DNA complex, we capture the LF domain rotation back to the binary conformation within 13 nsec. The simulated structure (averaged over the last 600 psec) is superimposed with the ternary and binary crystal structures (according to the palm domains) in Figure 4. The LF domain of the starting structure rotates toward the finger domain by ~12°, and the final conformation of the LF domain overlaps with that in the after-chemistry binary crystal complex. The RMSD plot of the LF domain relative to the ternary and binary structures (Fig. 2E) indicates that the conformational transition starts at 6.8 nsec and has a magnitude of rotation of ~3 Å. The finger and the thumb domains, however, show no significant transitions, and their final conformations are similar to those in the ternary crystal structure (Figs. 2F, 4, S1A). The motions of the LF, finger, and palm domains observed in this simulation are consistent with the crystal data that the LF domain rotates by ~12° and that there are no large movements in the finger and the thumb domains when superimposing the palm domains of the ternary before chemistry and binary after chemistry crystal structures (Fig. 1).

Motivated by the critical effects of the metal ions (Mg^{2+}) on pol β's thumb conformational change (Yang et al. 2004), we performed the second post-chemistry simulation where both metal ions and PP_i were removed from the active site. In contrast to the above results, in this trajectory we found that the finger and the LF domains exhibit large rigid-body movements. The RMSD evolution of the LF domain in Figure 2G shows that the LF domain starts altering conformation at the beginning of the simulation and attains the maximum deviation from the ternary crystal structure by 12 nsec. The thumb domain does not change conformation

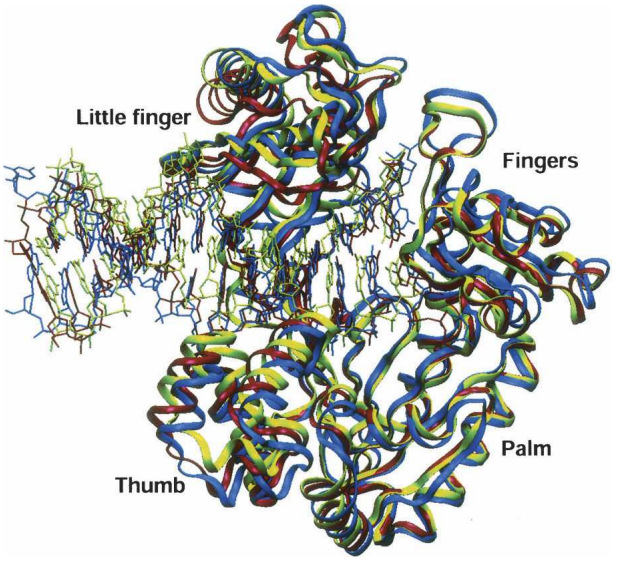

Figure 4. Superimposition of the simulated average structure (blue) in the trajectory after chemistry with metal ions and PP_i in the active site to the binary (green) and ternary (red) crystal structures according to the palm domains. Their DNA double strands are represented by blue, green, and red sticks, respectively.

as dramatically as the LF domain, although it displays motion (Fig. S1C). The shape similarity in the RMSD plots of the LF and the finger domains (Fig. 2G,H) suggests a correlation of these conformational changes for the two domains.

During the period from 4.3 nsec to 14 nsec in the second post-chemistry simulation, the finger and the LF domains deviate most significantly from the ternary structure. The averaged structure of the trajectory from 12 nsec to 12.5 nsec indicates that the LF domain moves forward (direction shown by a black arrow in Fig. 5A) by as much as 3.5 Å and rotates around the DNA major groove (see the simulated [light green] and initial [light red] structures in Fig. 5C) by ~12°, as measured by the rotation of α-helices L and M on the LF domain (labeled in Fig. 1B). Similarly, the finger domain also translates in the same direction as the LF domain by ~3.5 Å. The LF and the finger domains interact with each other and with the DNA; thus, the translation and rotation motions of the LF domain might be transmitted to the finger domain. Because the DNA template strand is in greater contact with the LF domain, it slides forward much more than the primer strand. Overall, in this period, the DNA translocates by almost a base pair's rise (Fig. 5B), such that the templating and primer strands slide by ~3.4 Å and 2.0 Å (closest atom-to-atom distance of the two base planes) from the ternary conformation, respectively.

By superimposing in Figure 5C the ternary crystal structure and simulated structure of Dpo4 in the second post-chemistry trajectory with the apo-structure of Dbh (Silvian et al. 2001), we find that Dpo4's LF domain rotates toward its conformation in Dbh, although the magnitude of the rotation is small (~12°) compared to the apo-structure of Dbh (which would require a 77° rotation and a 2 Å translation) (Ling et al. 2004a). This suggests that the LF domain of Dpo4 might loosen its grip of DNA completely after it is translocated and then return to the position in the binary crystal structure, so that the next reaction cycle could occur.

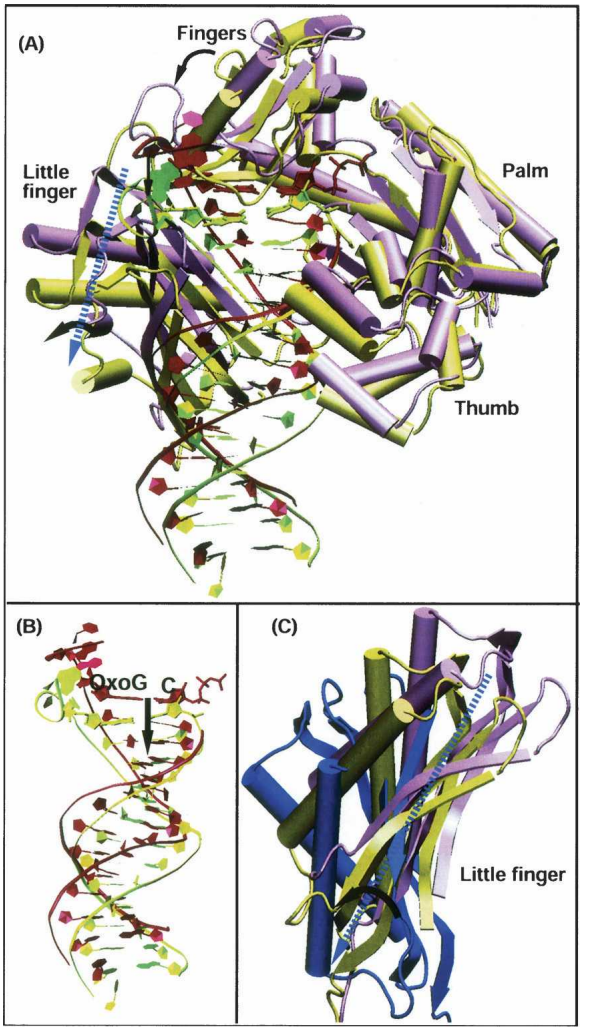

Figure 5. Structural comparison of the starting and final Dpo4/DNA complexes in the second simulation after chemistry. (*A*) Superimposition of the simulated structure (light green) in the trajectory after chemistry with metal ions and PP_i removed to the ternary crystal structure (light red) according to the palm domains. (*B*) Enlarged view of the DNA duplexes before (red) and after (green) the simulation. 8-OxoG and dCTP are labeled as OxoG and C, respectively. Black arrow indicates the direction of their movements. (*C*) Comparison of the LF domains before (light red) and after (light green) simulation to that of the Dbh apo-structure (Silvian et al. 2001) (blue) by superimposing the palm domains.

PCA analysis of the motions

Principal component analysis (PCA) was applied to extract the most significant modes of domain motions in the two simulations after chemistry. The modes with the two largest eigenvalue–eigenvector pairs for each trajectory are shown in Figure 6. The eigenvalues of the first 50 eigenvectors for each trajectory are plotted in Supplementary Figure S2.

For the first simulation trajectory after chemistry, the LF moves toward the finger domain in the largest principal mode (41% of the total motions) (Fig. 6A). Both β strands 2 and 3 (numbered according to Ling et al. [2001]) in the finger domain as well as the loop connecting them also display similar motions as the LF domain. At the same time, the thumb moves away from the LF and the DNA in a perpendicular orientation.

In the second principal mode (15% of the total motions) (Fig. 6B) of the first trajectory, the LF rotates around the axis as in Figure 1, but the finger domain does not display much motion; the thumb also moves away from the LF and DNA. In both of the principal modes, the palm and DNA primer strand that reside in

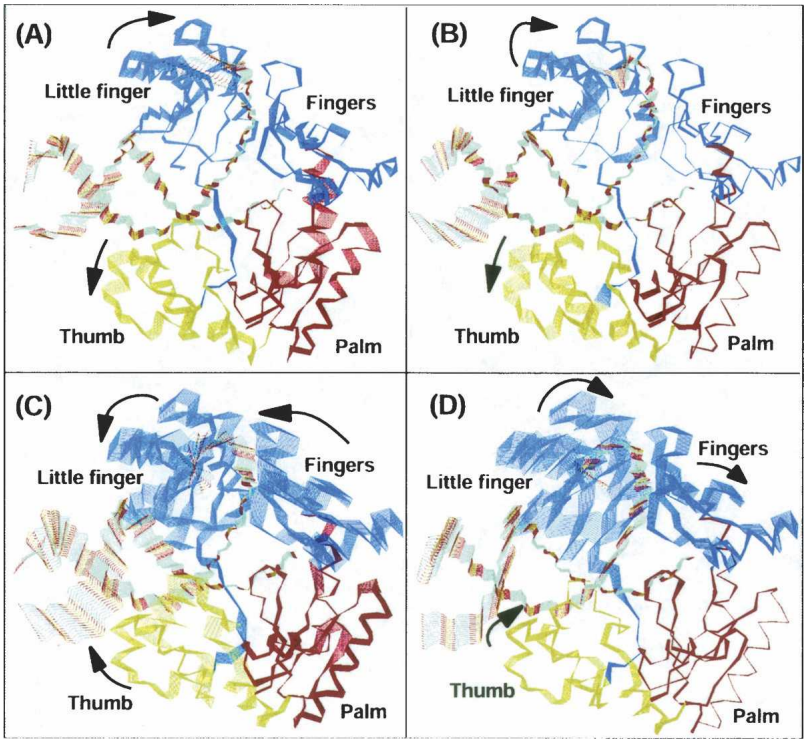

Figure 6. Dpo4 domain motions revealed by PCA from the two trajectories after chemistry. (*A*,*B*) The two largest principal modes extracted from the simulation with the metal ions and PP_i present in the active site. (*C*,*D*) The two largest principal modes from the simulation without the binding metal ions and PP_i present in the active site. Black arrows indicate the direction of domain movements. In each mode, the palm is red, thumb is green, finger is blue, and LF is light blue. The DNA backbone is traced, with P atoms in brown, O atoms in red, and C atoms in cyan.

the active site are rigid compared to the other parts of the system, while the DNA template strand attached to the LF exhibits more mobility.

With the ions released from the active site, the two largest eigenvectors represent 61% of motions (Fig. S2B). In the largest eigenvalue–eigenvector pair (Fig. 6C), both the finger and the LF domains move along the DNA duplex away from the active site at the same time as the LF rotates around the axis as in Figure 1. Since the LF domain binds DNA in the major groove, the translation of LF pushes both the DNA template and primer strands forward, and its rotation also pulls the primer strand away from the active site. Consequently, the template and the primer strands display large magnitude motions in both the DNA duplex axis direction and the LF rotation direction away from the palm domain. The thumb domain also moves slightly along the DNA and toward the minor groove: The thumb binds the DNA minor groove via hydrogen bonding and van der Waals interactions. Thus, the DNA and thumb motions are interdependent. Overall, the rigid body movements of the finger and LF domains appear to push the nascent base pair as well as the DNA duplex away from the active site.

In the second mode (Fig. 6D), the LF domain moves together with the finger domain along the DNA toward the active site. The thumb does not exhibit as much motion as in the first mode, and the α-helix H is particularly stable.

Conformational rearrangements in the active site

In the first simulation of the Dpo4/DNA complex after chemistry, with the metal ions and PP_i present in the active site, the catalytic residues (Asp7, Asp105, and Glu106) are bound to the metal ions so no conformational changes occur in these and nearby residues.

However, in the second simulation after chemistry with metal ions and PP_i released, the catalytic triad rearranges with Asp105 and Glu106 rotated away from Asp7. These residues are stabilized in new conformations by forming hydrogen bonds with other residues (Fig. 7A). Asp7 interacts with Lys159 after the simulation, while Asp105 rotates toward Ser103 and forms a weak hydrogen bond with Ser103. Glu106 is anchored by Lys152 and by the metal ions in the ternary crystal structure. By the end of the simulation, Glu106's side chain flips 120° so that it forms an additional hydrogen bond with Tyr108.

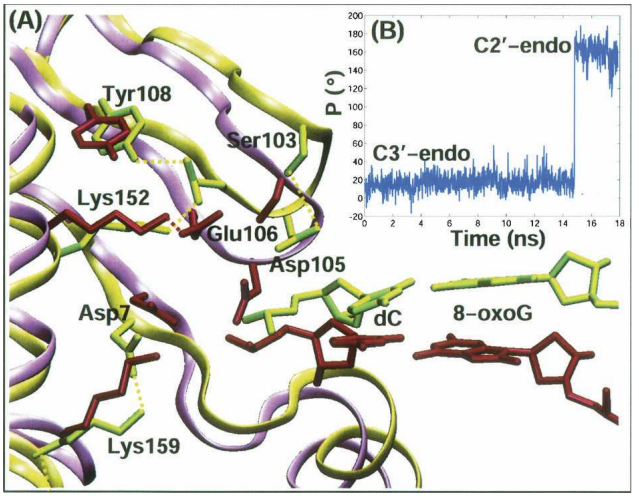

Figure 7. (*A*) The conformations of the catalytic triad (Asp7, Asp105, and Glu106) in the initial and final structures in the simulation after chemistry with metal ions and PP_i removed. Protein backbones in the simulated and ternary crystal structures are represented by light green and light red traces, respectively. The corresponding amino acids in the complexes are shown by green and red sticks, respectively. (*B*) Time evolution of the sugar pseudorotation angle P (Altona and Sundaraligam 1972; Arora and Schlick 2003) of the incoming nucleotide (dC) in this trajectory.

These new hydrogen bonds between the carboxylate residues and neighboring groups after chemistry indicate that the reaction-competent state is unstable after the two metal ions are released and that each of the catalytic residues must revert to another stable state.

Interestingly, we find that with the functional metal ions released from the simulation after chemistry, a sodium ion diffuses into the active site during the simulation and occupies a similar position as the catalytic metal ion in the binary crystal structure after chemistry (O. Rechkoblit, L. Malinina, Y. Cheng, V. Kuryavyi, S. Broyde, N.E. Geacintov, and D.J. Patel, in prep.). In our pol β simulations, we captured a sodium and magnesium ion diffusion into the active site in the absence of both the nucleotide-binding and catalytic Mg^{2+} ions (Yang et al. 2004). The sodium ion in our current simulation binds to the active site within 14 nsec and interacts with residues Asp7 and Phe8 as well as O2P atom of the nascent primer nucleotide (Fig. S3). Since the diffusion occurs on the nanosecond scale, the association of the catalytic metal ion may be a fast step in the overall reaction profile of Dpo4.

Although the sugar pseudorotation angle (Altona and Sundaraligam 1972; Arora and Schlick 2003) of dCTP remains at C3′-endo before chemistry, this angle changes from C3′-endo (~18°) to C2′-endo (~160°) ~14 nsec (Fig. 7B) after chemistry following release of the metal ions and PP_i. The pucker angle of dCTP in the ternary before chemistry and binary after chemistry crystal structures is C3′-endo (~14.4°) and C2′-endo (~137.7°), respectively. For pol β, we also captured sugar repuckering in the incoming dCTP before chemistry and found this motion to be a metastable state (Radhakrishnan and Schlick 2004) in the cascade of events that lead to thumb closing. In the ternary crystal structure of Dpo4, the 3′-OH group on the sugar ring of dCTP forms a hydrogen bond with one nonbridging oxygen atom on β-phosphate (Fig. 1C) to stabilize the incoming nucleotide. After the PP_i release, the hydrogen bond is broken and the sugar ring may evolve to a more stable C2′-endo conformation. This captured sugar repuckering in Dpo4 suggests that the C3′-endo (A-DNA-like) before and during the chemical reaction of phosphoryl transfer may help stabilize the transition state; thus, following the reaction, the sugar reverts to the C2′-endo (B-DNA-like) form to facilitate translocation of the new base pair.

Note that with the metal ions present at the active site, our simulations did not capture any conformational changes in the active-site residues both before and after chemistry. This is consistent with a comparison between the binary and the ternary crystal structures of Dpo4 (which include the catalytic metal ion binding the DNA primer strand terminus). Our simulations with both metal ions removed from the active site, however, indicate that the active site undergoes rearrangements and thus suggest possible conformational changes involved.

LF's Arg336 and Arg298 may facilitate DNA translocation

An ingenious recent study (Boudsocq et al. 2004) established that the LF domain plays a critical role in determining Dpo4's enzymatic properties, such as processivity, fidelity, and lesion bypassing. The investigators suggest that different surface curvature and electrostatic potential of the LF domains in Y-family polymerases may explain such different properties.

To explore specific residue changes during the movements of the LF domain shown in the simulation after chemistry with the metal ions and PP_i released, we analyze the side-chain profile of each charged or aromatic residue using dials and windows (see Figs. S4.1–4.5), for these residues are expected to have greater contributions to the LF's surface electrostatic potential than other noncharged ones. We find that the Arg336 side chain (C^{α}–C^{β}–C^{γ}–C^{δ}) rotates by 120° (Fig. 8A,E) at 8.1 nsec, shortening its length by ~1.5 Å. Before the rotation, Arg336 interacts with the O2P atom of T9 (Adenine) on the DNA template strand via two hydrogen bonds (see nucleic acid residue labels in Fig. 1A). One of these hydrogen bonds remains and strengthens after Arg336

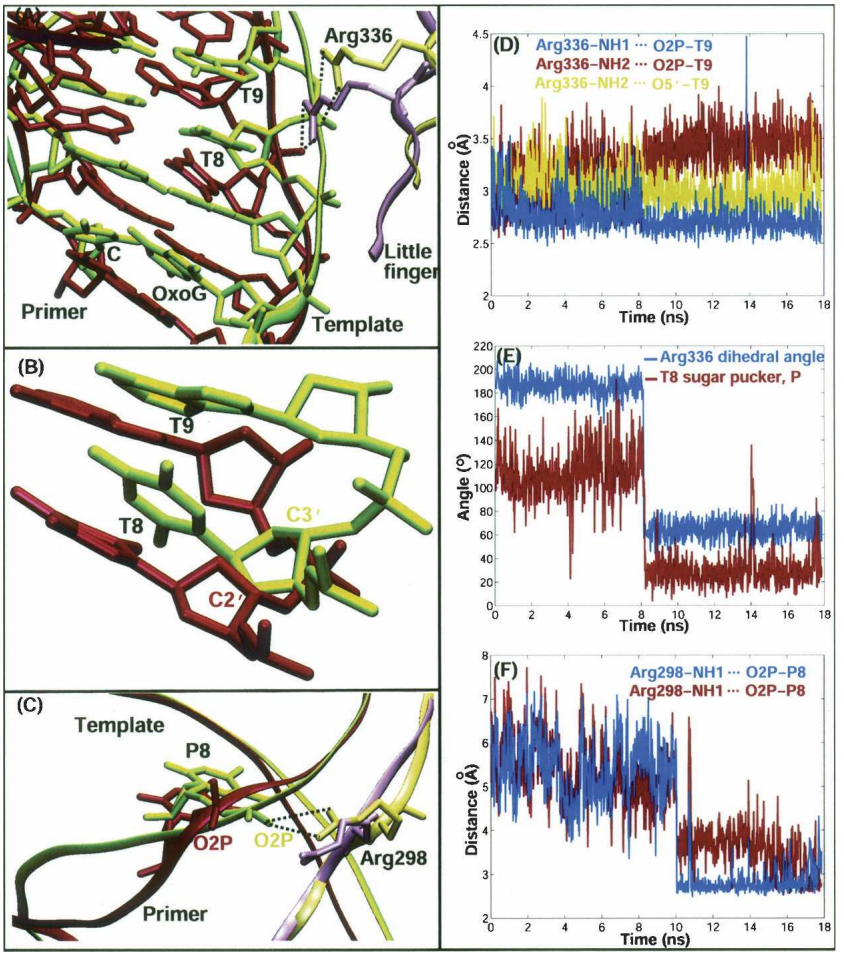

Figure 8. Conformational changes of Arg336 and Arg298 in the LF domain in the simulation after chemistry with ions and PP_i removed. (*A*) Conformations of Arg336 before (light red) and after (light green) side-chain rotation. (*B*) Sugar pucker transition of T8 from C2′-endo (red) to C3′-endo (green). (*C*) Arg298 conformations before (light red) and after (light green) hydrogen bond formation. The DNA duplexes in *A* and *C* are red and green in the structures before and after simulation, respectively. (*D*) Time evolution of atom–atom distances between Arg336 and O2P and O5′ atoms of T9. (*E*) Time evolution of the Arg336 dihedral angle (C^{α}–C^{β}–C^{γ}–C^{δ}) and T8 sugar pucker pseudorotation angle P. (*F*) Time evolution of atom–atom distances between Arg298 and O2P atom of P8.

rearranges, while the other weakens and is replaced by a hydrogen bond between Arg336 and O5′ atom of T9 (Fig. 8D). The shortened side chain of Arg336 consequently pulls the template strand toward the LF domain and triggers the sugar repuckering of T8 (Thymine) from C2′-endo to C3′-endo (Fig. 8B,E). Arg336 and T8 remain at the new conformations for another 10 nsec until the end of the simulation (Fig. 8E), indicating their stability; our simulation thus captures a transition between two stable states.

In addition to Arg336, another positively charged residue on the LF domain, Arg298, not interacting with DNA originally, becomes bonded to the O2P atom of P8 (Thymine) on the primer strand via two hydrogen bonds after the LF domain changes conformation (Fig. 8C,F). The first hydrogen bond forms when the distance between the Arg298 and P8 decreases significantly at 10 nsec. The second hydrogen bond starts to form at 10 nsec and forms completely by 16.7 nsec.

Thus, our studies reveal that, among all the charged residues on the LF domain, mainly Arg336 and Arg298 interact with the DNA differently after the conformational change of the LF domain. The sugar pucker transition of T8 triggered by the side-chain rotation of Arg336, and the formation of two hydrogen bonds between Arg298 and DNA primer strand point to possible key roles for LF's Arg336 and Arg298 to facilitate DNA translocation.

Since no such large domain motions and side-chain conformational changes of active-site residues were noted with both binding metal ions and PP_i present, the release of the metal ions and PP_i may trigger these

conformational changes and the correlated LF and finger motions facilitate DNA translocation.

Discussion

Our simulations, based on new crystal complexes of Dpo4 determined in the Patel lab (O. Rechkoblit, L. Malinina, Y. Cheng, V. Kuryavyi, S. Broyde, N.E. Geacintov, and D.J. Patel, in prep.), were aimed at exploring possible transitions and mechanisms associated with Dpo4's catalytic cycle. The fact that the ternary structure is highly stable before chemistry, both with and without the incoming nucleotide at the active site, and captures no large conformational changes, confirms the notion that the "induced-fit" mechanism might not exist in Dpo4. The finger and LF domains of Dpo4 show fluctuations around their ternary crystal structure conformation in the absence of the incoming nucleotide but no definite transitions between the ternary and the binary states. The absence of conformational changes in Dpo4 is markedly different from observations in pol β simulations, where the thumb subdomain closes or opens depending on the presence or the absence of incoming nucleotide, respectively (Arora and Schlick 2004).

Our simulations for Dpo4 after chemistry suggest, however, two scenarios depending on the presence of ions in the active site. First, only slight LF rotation (~12°) occurs when all the binding ions and PP_i group are in the active site, so that the final conformation is superimposable with the binary crystal structure (Fig. 4). This low-magnitude and rapid LF rotation suggests a low energy barrier for the transition in the interchange between ternary and binary conformations. Compared to the ternary complex of Dpo4 with DNA and incoming nucleotide ddATP opposite an abasic site (Ling et al. 2004a), where the LF domain rotates by 56° toward the finger domain, the LF rotation observed here when Dpo4 bypasses 8-oxoG is small, likely because of the modest distortion introduced to the active site by the 8-oxo group.

In contrast, the second post-chemistry simulation without specific binding ions captures a series of conformational changes in the active site (including rearrangements of the catalytic residues and sugar puckering transition of the nascent primer nucleotide), in the LF and the finger domains, and in the DNA. This suggests that the release of the metal ions may trigger more systematic and correlated conformational changes in the protein and the DNA (translocation). Since the catalytic metal ion remains in the active site of both binary crystal structures of Dpo4 before and after the chemical reaction (O. Rechkoblit, L. Malinina, Y. Cheng, V. Kuryavyi, S. Broyde, N.E. Geacintov, and D.J. Patel, in prep.), binding Asp7, Asp105, and Glu106, as well as the O1P atom of the primer terminus nucleotide, these possible conformational changes may occur after the catalytic metal ion is released from the binary crystal structure following the chemical reaction. This involves moving the finger and LF domains of Dpo4 along the DNA duplex axis, thereby pushing DNA away from the palm domain. During the largest deviation from the ternary crystal structure, the LF rotates 12° and translates by ~3.5 Å (Fig. 5A,C). Since the LF binds the major groove of DNA tightly by interacting with both template and primer strands, the large movements of the LF consequently affect the DNA: The DNA base pairs in the active site of Dpo4 slide as much as 1 bp's distance relative to the palm domain, with the template strand translocating by ~3.4 Å (Fig. 5B) and the primer strand by ~2 Å.

Thus, the ability of Dpo4 to bypass lesions may not only be due to its flexible active site but also to the motions of the LF domain that help adjust the active site environment as needed. Depending on the lesion type and size, the LF domain rotation magnitude around the DNA major groove may be adjusted to accommodate a particular lesion. This requirement for variable distortions to bypass various lesions is further made possible by a correlation between this LF rotation and a more subtle finger rearrangement.

Indeed, studies have indicated that the LF domain may play a critical role in DNA translocation (Ling et al. 2001, 2004a; Boudsocq et al. 2004). In our post-chemistry simulation without the metal ions, we found that Arg298 and Arg336 on the LF domain interact differently with the DNA after the LF rearrangement and thus may facilitate DNA translocation. The shortened Arg336 side chain induces sugar repuckering of T8 from C2′-endo to C3′-endo, and Arg298 approaches the primer strand to form two hydrogen bonds with the P8 after the conformational change of the LF domain.

Interestingly, the two arginine residues, both located on the β-sheet plane of Dpo4, interact with different DNA strands; that is, Arg298 binds the primer strand while Arg336 binds the template strand. Hence, when the LF domain translates or rotates, the DNA template and primer strands move along by the two arginine residues as well as other charged residues on the LF domain such as Arg242, Arg247, Lys275, Arg331, and Arg332. The conformational changes of Arg298 and Arg336 thereby allow the LF domain to anchor the DNA major groove, and the stronger interaction between the LF domain and the DNA makes the translocation of DNA possible.

Besides suggesting the importance of the LF region, crystal data of Dpo4 complexed with DNA and 8-oxoG indicate that α-helices H and K of the thumb domain contact the next nucleotide on DNA in the binary complex after chemistry compared to the ternary state (O. Rechkoblit, L. Malinina, Y. Cheng, V. Kuryavyi, S. Broyde, N.E. Geacintov, and D.J. Patel, in prep.; Fig. 1B); this might suggest significant thumb motion.

However, such thumb domain movements were not captured in our simulations. Perhaps these thumb motions and DNA translocation require milliseconds or longer in vivo. Nevertheless, the motions of the LF observed here demonstrate the dynamic process of how DNA might be translocated after chemistry; specifically, we propose that the LF domain grips the DNA tightly via Arg298 and Arg336 as well as other charged residues after both metal ions are released, and that the movements of the LF domain as in Figure 5C help translocate both DNA strands by 1 bp; the α-helices H and K of the thumb domain function as stopping points and interact with the next nucleotide on the double strand when the LF domain returns to the starting conformation.

Other details of crucial residue motions have also emerged from the present studies. In the simulations without the metal ions present in the active site both before and after chemistry, the rotation of Glu106 by forming a hydrogen bond with Tyr108 (Figs. 3A, 7A) resembles the flip of Asp192 in pol β after chemistry (Sawaya et al. 1997; Yang et al. 2002b; Radhakrishnan and Schlick 2004; Table 1). In addition, Asp7 and Asp105 are stabilized by forming hydrogen bonds with surrounding residues after simulation. These side-chain rearrangements may define alternative stable states of the catalytic residues when both metal ions are released from the active site. To perform the next cycle of nucleotide incorporation, the carboxylate residues must rotate back to bind the metal ions for catalysis.

Furthermore, the sugar repuckering in the newly incorporated nucleotide from C3′-endo (A-DNA-like) to C2′-endo (B-DNA-like) (Fig. 7B) after chemistry following ion release suggests that the A-DNA characteristics facilitate the chemical reaction and that this motion may direct the enzyme to the next round in the replication cycle. Indeed, a recent experimental work by Marquez et al. (2004) suggests that cellular polymerases prefer exclusively nucleotide substrates with sugar puckers in the C3′-endo conformation before the chemical reaction. The sugar puckering of the incoming nucleotide was also identified as the first substep of pol β's conformational closing for a GC pair (Radhakrishnan and Schlick 2004).

Since this manuscript was submitted, the Yang group (Vaisman et al. 2005) published complexes of Dpo4 with a matched or mismatched incoming nucleotide, superseding previous ternary complexes with degraded incoming units (ddADP) (Ling et al. 2001). Their results suggest that proper metal ion coordination associated with subtle conformational changes may be a key step in catalysis. Our simulations are thus in excellent agreement with this notion and provide further details on these motions. Note that this theme of involvement of metal ions in slow, crucial steps in the catalytic cycle of the DNA polymerase was the focus of other studies for pol β (Radhakrishnan and Schlick 2004, 2005; Yang et al. 2004; Arora et al. 2005). Furthermore, more recently, we have added the hypothesis that slow steps, in pol β cycle occur *after* closing but *prior to* the chemical reaction in separate path we call "pre-chemistry avenue." Recent experiments with a fluorescence resonance energy transfer (FRET) system (Klentaq1 DNA polymerase) (Rothwell et al. 2005) have also suggested that the rate-limiting step crucial for fidelity involves the rearrangements in the active site prior to the chemical reaction but after the closing of the fingers subdomain of the DNA polymerase.

The fact that Dpo4 and other low-fidelity polymerases may not have a significant conformational closing before the chemical reaction but have significant metal ion-dependent rearrangements further propels the idea that such a "pre-chemistry" avenue with crucial active-site rearrangements and high barriers operate more widely. Because Dpo4's active site is more distorted than pol β's after closing, a higher energy barrier for the "pre-chemistry" step may be involved (see the schematic drawing of the possible free energy profiles for Dpo4 and pol β in Supplementary Fig. S5). This may be related to the low catalysis efficiency for Dpo4.

Finally, another implication of lack of finger domain motion in Dpo4 (away from the nascent base pair) is that Dpo4 may not have a more "open" state than the ternary structure. However, a sequence of subtle conformational changes in and beyond the active site might occur in Dpo4 and DNA before and after chemistry. Our analyses support a subtle coordinated "induced-grip" fidelity mechanism (Fleck and Schär 2004) for Dpo4 rather than pronounced substrate-induced subdomain motion as in other polymerases like pol β. This lack of strict geometric fit for selecting the correct incoming unit and the open active site hinder Dpo4's ability to discriminate correct from incorrect substrates, resulting in lower fidelity of Dpo4 compared to pol β, and the correlated but subtle subdomain rearrangements introduce a variability potential to Dpo4 to handle various lesions.

This idea of poor geometric fit is sketched in Figure 9 and Supplementary Figure S5: Stricter geometric/energetic/dynamic selection rules operate in pol β to tightly orchestrate the assembly of the active site prior to, and during, nucleotide insertion, while Dpo4 discriminates correct from incorrect units more poorly due to a more flexible active site. However, in Dpo4, the more subtle yet coordinated motions of the LF with finger domains and commensurate DNA translocation, as revealed here, echo the theme of tightly orchestrated events in polymerase mechanisms (Radhakrishnan and Schlick 2004) and provide an opportunity for variable conformational rearrangements as needed to handle different lesions during Dpo4's bypass cycle.

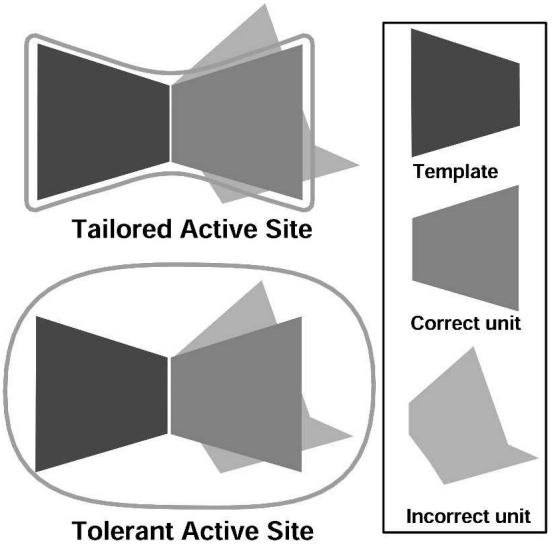

Figure 9. DNA polymerases (*top*) like pol β can tailor-fit the correct incoming substrate much more strongly than low-fidelity polymerases like Dpo4 (*bottom*) that have a more permissive active site.

Materials and methods

Parameterization of 8-oxoguanine

The force field parameters for 8-oxoG were obtained by using the parameterization protocol for nucleic acids in CHARMM based on small-molecule and condensed-phase macromolecular data (Foloppe and MacKerell 2000). The parameters for the sugar and phosphate part of 8-oxoG were adopted from guanine in the c26a2 version of CHARMM force field (Brooks et al. 1983; MacKerell et al. 1998); with this force field, the initial bond lengths and angles for 8-oxoG were optimized. The minimized bond lengths and angles were then used to calculate the vibrational spectra using the MOLVIB subroutine implemented in CHARMM. GAUSSIAN98 (Gaussian Inc.) was used to calculate the vibrational frequencies of the optimized 8-oxoG, and the bond lengths and angles were adjusted to match the results from CHARMM and GAUSSIAN98. The electrostatic potential of 8-oxoG was computed by GAUSSIAN98 at the 6-31G* level and the charge for each atom on 8-oxoG was fitted using the RESP package from AMBER (University of California, San Francisco).

System preparation and dynamics simulations before chemistry

Two initial models based on the ternary crystal structure of the Dpo4/DNA complex with 8-oxoG and an incoming nucleotide (O. Rechkoblit, L. Malinina, Y. Cheng, V. Kuryavyi, S. Broyde, N.E. Geacintov, and D.J. Patel, in prep.) before the chemical reaction were constructed. In one of the models, the incoming nucleotide (dCTP) along with the binding metal ions in the active site was kept intact as in the ternary crystal structure; the incoming nucleotide and associated binding ions were removed in the second model. In both models the hydroxyl group was added to the 3′ terminus of the primer DNA strand. CHARMM's subroutine HBUILD (Brünger and Karplus 1988) was employed to add all hydrogen atoms to the crystallographic heavy atoms. Each system was solvated in the shell of water 10 Å thick around the solute and neutralized with ions (Na^+ and Cl^-) at the physiological concentration obeying Debye-Hückel distribution using SOLVATE (H. Grubmüller). Resulting system was immersed in an equilibrated box of water, and the overlapping water molecules were deleted to obtain a system in a cubic box of water with cell dimensions 90 Å × 90 Å × 90 Å. The total system size is ~70,000 atoms (341 protein residues, 32 DNA base pairs, 69 ions, and 23,000 water molecules).

Energy minimizations and molecular dynamics simulations were carried out using the program NAMD (Kale et al. 1999) with version C26a2 of the CHARMM force field. First, each system was energy minimized with fixed position of all protein and DNA heavy atoms using the Powell algorithm. The system was then equilibrated for 150 psec at constant pressure and temperature. Pressure was maintained at 1 atm using the Langevin piston method (Feller et al. 1995), with a piston period of 100 fsec, a damping time constant of 50 fsec, and piston temperature of 300 K. Temperature coupling was enforced by velocity reassignment every 2 psec. Then, the production dynamics was performed at constant temperature and volume. Constant temperature was maintained at 300 K using weakly coupled Langevin dynamics of nonhydrogen atoms with a damping coefficient γ of 5 $psec^{-1}$. Bonds to all hydrogen atoms were kept rigid using SHAKE (Ryckaert et al. 1977), permitting a time step of 2 fsec. The system was simulated in periodic boundary conditions, with full electrostatics computed using the Particle-Mesh-Ewald method (Darden et al. 1993) with a grid spacing on the order of 1 Å or less. Short-range nonbonded terms were evaluated every step using a 12 Å cutoff for van der Waals interactions and a smooth switching function. The total simulation length for the first (control) and the second model is 14.5 nsec and 23.2 nsec, respectively.

System preparation and dynamics simulations after chemistry

To build the structure of the Dpo4/DNA complex with 8-oxoG after chemistry, the nucleotidyl-transfer reaction was performed manually by connecting the α-phosphate of the dCTP with the 3′-OH of the DNA primer strand in the ternary crystal structure (O. Rechkoblit, L. Malinina, Y. Cheng, V. Kuryavyi, S. Broyde, N.E. Geacintov, and D.J. Patel, in prep.) and breaking the covalent bond between P_α and the oxygen atom that binds with both P_α and P_β. The resulting pyrophosphate (PP_i) group from the bond breaking and the two original specific binding ions from the ternary crystal structure were kept in the first after-chemistry simulation of the complex but they were removed to accelerate the conformational changes in the second after-chemistry simulation. MODELLER (Sali et al. 1995) was used to construct the three dimensional structure of the 11 missing residues in the C terminus of Dpo4 using the crystal structure of polymerase IV from *Escherichia coli* complexed with β-clamp processivity factor (Bunting et al. 2003; PDB code 1UNN) as template. The resulting C terminus was then connected to the overall Dpo4 structure. Hydrogen atoms were added to the final model by CHARMM (Brünger and Karplus 1988).

The system was solvated in a face-centered cube using the programs Simulaid (Mezei 1997) and PBCAID (Qian et al.

2001). The smallest image distance was chosen as 16 Å. Water molecules within 1.8 Å of the heavy atoms of the protein, DNA, crystallographic water, and ions were removed. Sodium and chloride ions were added to neutralize the system and produce an ionic strength of 150 mM. Oxygen atoms in water molecules having the most negative electrostatic potentials were replaced by sodium ions, while those with the most positive electrostatic potentials were replaced by chloride ions. The electrostatic potentials were calculated using DELPHI (Klapper et al. 1986; Gilson et al. 1988). All ions added were placed at least 8 Å away from the protein or DNA or from each other. The systems contain 48,798 and 48,804 atoms in the presence and the absence of the specific metal ions and PP_i, respectively.

Periodic boundary conditions and the CHARMM all-atom force field for nucleic acids and lipids were used for all energy minimization and molecular dynamics simulations in CHARMM. The nonbonded interactions are truncated at 14 Å, with van der Waals interactions treated by the switch cutoff method and electrostatic interactions treated by the force-shift method.

The solvated system was minimized and equilibrated as follows: The water molecules and hydrogen atoms were minimized with all other heavy atoms in the system fixed using steepest decent (SD) for 10,000 steps, followed by an adopted-basis Newton-Raphson (ABNR) minimization for 20,000 steps; an equilibration of 10 psec at 300 K with the Langevin method ensured that all the sodium and chloride ions were located at the potential energy minima or maxima around the protein/DNA complexes; the entire system was again minimized by SD for 10,000 steps and ABNR for 20,000 steps with all the protein and DNA heavy atoms fixed; and finally, the system was equilibrated with the stochastic LN approach (Schlick et al. 1997; Barth and Schlick 1998a,b; Schlick 2001) for 150 psec with all the atoms released.

For the production runs, the multiple time step Langevin integrator, LN was used with time steps $\Delta\tau/\Delta t_m/\Delta t$ of 1/2/150 fsec for fast/medium/slow force components and a medium range cutoff of 7 Å and healing and buffer lengths of 4 Å each (the stability and the reliability of LN for biomolecular systems in terms of thermodynamic, structural, and dynamic properties are examined thoroughly in Yang et al. [2002a] compared to single time step Langevin as well as Newtonian [Velocity Verlet] propagators). The damping constant γ was set as 10 $psec^{-1}$ to couple the system to a 300°C heat bath. The first trajectory was simulated for 13 nsec with both divalent metal ions present in the active site, and the second was simulated for 18 nsec without the presence of the metal ions and PP_i. SHAKE was applied to all bonds containing hydrogen atoms. Coordinates were saved every 3 psec. The 13 and 18 nsec trajectories took ~86 and 129 d on eight 600 MHz R14000 processors of NCI SGI Origin 3800 computers, respectively.

Note that our simulations before and after chemistry were performed using different molecular dynamics software packages but the same force field, CHARMM version c26a2. Calculations with NAMD employed the Particle-Mesh-Ewald approach for better electrostatic treatment as well as parallelization. Namely, we found the computational performance using NAMD on highly parallel architecture to be very efficient (Kale et al. 1999), and thus used it in our pre-chemistry simulations. Approximately 15 CPU h are required to complete a 1-nsec run for our Dpo4 system using 64 CPUs in parallel on NCSA TUNGSTEN LINUX cluster, compared to 158 CPU h/nsec on eight processors using the LN package in CHARMM.

Principal component analysis (PCA)

PCA helps analyze complex motions of biomolecules in terms of independent modes. The first few modes normally capture most of the motion characteristics revealed in the trajectory. We applied PCA here to each trajectory (frames collected every 15 psec) of the Dpo4 system after chemistry, as follows:

A covariance matrix **C** is constructed using the average structure from the merged configurational ensemble as the following sum of outer products

$$C = \frac{1}{M}\sum_{k=1,M}(X_k - \langle X\rangle)(X_k - \langle X\rangle)^T$$

where $\mathbf{X}_k$ is the coordinate vector at the kth snapshot, and $\langle X\rangle$ is the average structure from the dynamics simulation:

$$\langle X\rangle = \frac{1}{M}\sum_{k=1,M} X_k.$$

The average structure is the reference for the covariance matrices **C**. Diagonalization of **C** produces the eigenvalues and eigenvectors as entries of Λ from the decomposition

$$V^T C V = \Lambda \tag{1}$$

or

$$CV_n = \lambda_n V_n, \qquad n = 1, 2, \cdots, 3N, \tag{2}$$

where Λ is the diagonal matrix with the eigenvalues

$$\lambda_i = diag(\lambda_1, \lambda_2, \cdots, \lambda_{3N}). \tag{3}$$

Each eigenvector $\mathbf{V}_n$ defines the direction of motion of N atoms as an oscillation about the average structure $\langle X\rangle$. The normalized magnitude of the corresponding eigenvalue ($\lambda_n/\Sigma_{n=1,3N}\lambda_n$) indicates the relative percentage of the trajectory motions along eigenvector $\mathbf{V}_n$. Thus, an arbitrary structure **Y** can be generated from the average structure by a displacement **D** along the linear combination of all eigenvectors $\mathbf{V}_n$ with *3N* scalars α_n, where

$$Y = \langle X\rangle + D = \langle X\rangle + \sum_{n=1,3N}\alpha_n V_n, \qquad \alpha_n = V_n^T D. \tag{4}$$

Electronic supplemental material

RMSD plots of Dpo4's palm and thumb domains and the eigenvalues from PCA in both simulations after chemistry; coordination of the sodium ion diffused into the active site; and "dials and windows" plots of the side-chain dihedral angles of the charged and aromatic residues in the LF domain of Dpo4 in the simulation after chemistry with metal ions removed from the active site.

Acknowledgments

We are grateful to Dr. Dinshaw J. Patel for providing us the coordinates of the Dpo4/DNA binary and ternary structures

with 8-oxoG prior to publication. We thank Dr. Wei Yang for critical comments on our simulations and manuscript. We are indebted to Dr. Olga Rechkoblit, Dr. Yuan Cheng, Dr. Vitaly Kuryavyi, Dr. Linjing Yang, and Dr. Ravi Radhakrishnan for discussions. This work was supported by NSF Grant ASC-9318159, NIH Grant R01 GM55164, and the donors of the American Chemical Society Petroleum Research Fund to T.S. Research described in this article was also supported in part by Philip Morris USA Inc. and by Philip Morris International. Computations were performed using granted time from Advanced Biomedical Computing Center at NCI-Frederick and the NCSA supercomputing facility. Support from these computing centers is highly appreciated. The molecular graphics software VMD (Humphrey et al. 1996) was used to generate figures in this article.

References

Abashkin, Y.G., Erickson, J.W., and Burt, S.K. 2001. Quantum chemical investigation of enzymatic activity in DNA polymerase β. A mechanistic study. *J. Phys. Chem. B* **105**: 287–292.

Altona, C. and Sundaraligam, M. 1972. Conformational analysis of the sugar ring in nucleosides and nucleotides. A new description using the concept of pseudorotation. *J. Am. Chem. Soc.* **94**: 8205–8212.

Ames, B.N. and Gold, L.S. 1991. Endogenous mutagens and the causes of aging and cancer. *Mutat. Res.* **250**: 3–16.

Arora, K. and Schlick, T. 2003. Deoxyadenosine sugar puckering pathway simulated by the stochastic difference equation algorithm. *Chem. Phys. Lett.* **378**: 1–8.

———. 2004. In silico evidence for DNA polymerase-beta's substrate-induced conformational change. *Biophys. J.* **87**: 3088–3099.

Arora, K., Beard, W.A., Wilson, S.H., and Schlick, T. 2005. Mismatch-induced conformational distortions in polymerase β/DNA complex support an induced-fit mechanism for fidelity. *Biochemistry* **44**: 13328–13341.

Barth, E. and Schlick, T. 1998a. Extrapolation versus impulse in multiple-timestepping schemes. II. Linear analysis and applications to Newtonian and Langevin dynamics. *J. Chem. Phys.* **109**: 1633–1642.

———. 1998b. Overcoming stability limitations in biomolecular dynamics. I. Combining force splitting via extrapolation with Langevin dynamics in LN. *J. Chem. Phys.* **109**: 1617–1632.

Boudsocq, F., Iwai, S., Hanaoka, F., and Woodgate, R. 2001. *Sulfolobus solfataricus* P2 DNA polymerase IV (Dpo4): An archaeal DinB-like DNA polymerase with lesion-bypass properties akin to eukaryotic pol eta. *Nucleic Acids Res.* **29**: 4607–4616.

Boudsocq, F., Kokoska, R.J., Plosky, B.S., Vaisman, A., Ling, H., Kunkel, T.A., Yang, W., and Woodgate, R. 2004. Investigating the role of the little finger domain of Y-family DNA polymerases in low fidelity synthesis and translesion replication. *J. Biol. Chem.* **279**: 32932–32940.

Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S., and Karplus, M. 1983. Charmm—A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **4**: 187–217.

Brünger, A.T. and Karplus, M. 1988. Polar hydrogen positions in proteins—Empirical energy placement and neutron-diffraction comparison. *Proteins* **4**: 148–156.

Bunting, K.A., Roe, S.M., and Pearl, L.H. 2003. Structural basis for recruitment of translesion DNA polymerase Pol IV/DinB to the β-clamp. *EMBO J.* **22**: 5883–5892.

Cheng, K.C., Cahill, D.S., Kasai, H., Nishimura, S., and Loeb, L.A. 1992. 8-Hydroxyguanine, an abundant form of oxidative DNA damage, causes G-T and A-C substitutions. *J. Biol. Chem.* **267**: 166–172.

Darden, T., York, D., and Pedersen, L. 1993. Particle mesh ewald—An N.Log(N) nethod for Ewald sums in large systems. *J. Chem. Phys.* **98**: 10089–10092.

Feller, S.E., Zhang, Y.H., Pastor, R.W., and Brooks, B.R. 1995. Constant-pressure molecular-dynamics simulation—The Langevin piston method. *J. Chem. Phys.* **103**: 4613–4621.

Fiala, K.A. and Suo, Z. 2004a. Mechanism of DNA polymerization catalyzed by *Sulfolobus solfataricus* P2 DNA polymerase IV. *Biochemistry* **43**: 2116–2125.

———. 2004b. Pre-steady-state kinetic studies of the fidelity of *Sulfolobus solfataricus* P2 DNA polymerase IV. *Biochemistry* **43**: 2106–2115.

Fleck, O. and Schär, P. 2004. Translesion DNA synthesis: Little fingers teach tolerance. *Curr. Biol.* **14**: R389–R391.

Foloppe, N. and MacKerell, A.D. 2000. All-atom empirical force field for nucleic acids. I. Parameter optimization based on small molecule and condensed phase macromolecular target data. *J. Comput. Chem.* **21**: 86–104.

Friedberg, E.C., Wagner, R., and Radman, M. 2002. Molecular biology—Specialized DNA polymerases, cellular survival, and the genesis of mutations. *Science* **296**: 1627–1630.

Gilson, M.K., Sharp, K.A., and Honig, B.H. 1988. Calculating the electrostatic potential of molecules in solution—Method and error assessment. *J. Comput. Chem.* **9**: 327–335.

Humphrey, W., Dalke, A., and Schulten, K. 1996. VMD: Visual molecular dynamics. *J. Mol. Graph.* **14**: 33–38.

Kale, L., Skeel, R., Bhandarkar, M., Brunner, R., Gursoy, A., Krawetz, N., Phillips, J., Shinozaki, A., Varadarajan, K., and Schulten, K. 1999. NAMD2: Greater scalability for parallel molecular dynamics. *J. Comput. Phys.* **151**: 283–312.

Klapper, I., Hagstrom, R., Fine, R., Sharp, K., and Honig, B. 1986. Focusing of electric fields in the active site of Cu-Zn superoxide dismutase: Effects of ionic strength and amino-acid modification. *Proteins* **1**: 47–59.

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

Lahiri, S.D., Zhang, G.F., Dunaway-Mariano, D., and Allen, K.N. 2003. The pentacovalent phosphorus intermediate of a phosphoryl transfer reaction. *Science* **299**: 2067–2071.

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

Ling, H., Boudsocq, F., Plosky, B.S., Woodgate, R., and Yang, W. 2003. Replication of a *cis-syn* thymine dimer at atomic resolution. *Nature* **424**: 1083–1087.

Ling, H., Boudsocq, F., Woodgate, R., and Yang, W. 2004a. Snapshots of replication through an abasic lesion: Structural basis for base substitutions and frameshifts. *Mol. Cell* **13**: 751–762.

Ling, H., Sayer, J.M., Plosky, B.S., Yagi, H., Boudsocq, F., Woodgate, R., Jerina, M., and Yang, W. 2004b. Crystal structure of a benzo[a]pyrene diol epoxide adduct in a ternary complex with a DNA polymerase. *Proc. Natl. Acad. Sci.* **101**: 2265–2269.

MacKerell, A.D., Bashford, D., Bellott, M., Dunbrack Jr., R.L., Evanseck, J.D., Field, M.J., Fischer, S., Gao, J., Ha, S., Joseph-McCarthy, D., et al. 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102**: 3586–3616.

Marquez, V.E., Ben-Kasus, T., Barchi Jr., J.J., Green, K.M., Nicklaus, M.C., and Riald, A. 2004. Experimental and structural evidence that herpes 1 kinase and cellular DNA polymerase(s) discriminate on the basis of sugar pucker. *J. Am. Chem. Soc.* **126**: 543–549.

Mezei, M. 1997. Optimal position of solute for simulations. *J. Comput. Chem.* **18**: 812–815.

Mildvan, A.S. 1997. Mechanisms of signaling and related enzymes. *Proteins* **29**: 401–416.

Moriya, M. 1993. Single-stranded shuttle phagemid for mutagenesis studies in mammalian-cells—8-oxoguanine in DNA induces targeted G·C→T·A transversions in simian kidney-cells. *Proc. Natl. Acad. Sci.* **90**: 1122–1126.

Perlow-Poehnelt, R.A., Likhterov, I., Scicchitano, D.A., Geacintov, N.E., and Broyde, S. 2004. The spacious active site of a Y-family DNA polymerase facilitates promiscuous nucleotide incorporation opposite a bulky carcinogen-DNA adduct—Elucidating the structure–function relationship through experimental and computational approaches. *J. Biol. Chem.* **279**: 36951–36961.

Qian, X.L., Strahs, D., and Schlick, T. 2001. A new program for optimizing periodic boundary models of solvated biomolecules (PBCAID). *J. Comput. Chem.* **22**: 1843–1850.

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

———. 2005. Fidelity discrimination in DNA polymerase β: Differing closing profiles for a mismatched G:A versus matched G:C basepair. *J. Am. Chem. Soc.* **127**: 13245–13252.

Rothwell, P.J., Mitaksov, V., and Waksman, G. 2005. Motions of the fingers subdomain of klentaq1 are fast and not rate limiting:

Implications for the molecular basis of fidelity in DNA polymerases. *Mol. Cell* **19**: 345–355.

Ryckaert, J.P., Ciccotti, G., and Berendsen, H.J.C. 1977. Numerical-integration of cartesian equations of motion of a system with constraints—Molecular-dynamics of *n*-alkanes. *J. Comput. Phys.* **23**: 327–341.

Sali, A., Potterton, L., Yuan, F., Vanvlijmen, H., and Karplus, M. 1995. Evaluation of comparative protein modeling by modeler. *Proteins* **23**: 318–326.

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

Schlick, T. 1999. Computational molecular biophysics today: A confluence of methodological advances and complex biomolecular applications. *J. Comput. Phys.* **151**: 1–8.

———. 2001. Time-trimming tricks for dynamic simulations: Splitting force updates to reduce computational work. *Structure* **9**: R45–R53.

Schlick, T., Barth, E., and Mandziuk, M. 1997. Biomolecular dynamics at long timesteps: Bridging the timescale gap between simulation and experimentation. *Annu. Rev. Biophys. Biomol. Struct.* **26**: 181–222.

Shibutani, S. and Grollman, A.P. 1994. Miscoding during DNA-synthesis on damaged DNA templates catalyzed by mammalian-cell extracts. *Cancer Lett.* **83**: 315–322.

Silvian, L.F., Toth, E.A., Pham, P., Goodman, M.F., and Ellenberger, T. 2001. Crystal structure of a DinB family error-prone DNA polymerase from *Sulfolobus solfataricus*. *Nat. Struct. Biol.* **8**: 984–989.

Steitz, T.A. 1993. DNA-dependent and RNA-dependent DNA-polymerases. *Curr. Opin. Struc. Biol.* **3**: 31–38.

Vaisman, A., Ling, H., Woodgate, R., and Yang, W. 2005. Fidelity of Dpo4: Effect of metal ions, nucleotide selection and pyrophosphorolysis. *EMBO J.* **24**: 2957–2967.

Wang, L.H., Wu, M., Yan, S.F., Patel, D.J., Geacintov, N.E., and Broyde, S. 2005. Accommodation of a 1S-(–)-Benzo[c]phenanthrenyl-N-6-dA adduct in the Y-family dpo4 DNA polymerase active site: Structural insights through molecular dynamics simulations. *Chem. Res. Toxicol.* **18**: 441–456.

Yang, W. 2003. Damage repair DNA polymerases Y. *Curr. Opin. Struct. Biol.* **13**: 23–30.

Yang, L.J., Beard, W.A., Wilson, S.H., Broyde, S., and Schlick, T. 2002a. Polymerase β simulations suggest that Arg258 rotation is a slow step rather than large subdomain motions per se. *J. Mol. Biol.* **317**: 651–671.

Yang, L.J., Beard, W.A., Wilson, S.H., Roux, B., Broyde, S., and Schlick, T., 2002b. Local deformations revealed by dynamics simulations of DNA Polymerase β with DNA mismatches at the primer terminus. *J. Mol. Biol.* **321**: 459–478.

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

Yeiser, B., Pepper, E.D., Goodman, M.F., and Finkel, S.E. 2002. SOS-induced DNA polymerases enhance long-term survival and evolutionary fitness. *Proc. Natl Acad. Sci.* **99**: 8737–8741.