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, 251 Mercer Street, New York, New York 10012

Running title: Conformational changes in Dpo4 cycle

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

∗To whom correspondence should be addressed (e-mail:schlick@nyu.edu; fax: 212-995-4152; Tel: 212-998-3116)

July 18, 2005

Abbreviations: Dpo4, DNA polymerase IV; 8-oxoG, 8-oxoguanine; pol β, DNA polymerase β; dCTP, 2′-deoxyribocytidine 5′-triphosphate; dNTP, 2′-deoxyribonucleoside 5′-triphosphate; ddATP, 2′,3′-dideoxyriboadenosine 5′-triphosphate; LF, little finger; PPi, pyrophosphate; PCA, principal component analysis.
Summary

The possible conformational changes of DNA polymerase IV (Dpo4) before and after the nucleotidyl-transfer reaction are investigated at atomic level by dynamics simulations to gain insight into the mechanism of low fidelity polymerases. The absence of significant conformational changes in Dpo4 before chemistry when the incoming nucleotide is removed suggests that the “induced-fit” mechanism employed to interpret fidelity in some DNA polymerases may 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 changes may define subtle conformational adjustments 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 variable conformational-mobility potential as required for bypassing different DNA adducts by low-fidelity polymerases.
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 DNA damage 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 DNA polymerase IV (Dpo4) from the thermophilic archael *Sulfolobus solfataricus* P2. 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 dimers (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; Ling et al., 2003; Ling et al., 2004a; Ling et al., 2004b). Kinetic and *in vitro* experimental works have measured Dpo4 fidelity and proposed possible mechanisms for Dpo4’s enzymatic function (Fiala and Suo, 2004a; Fiala and Suo, 2004b; Boudsocq et al., 2004).

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 (Rechkoblit and Patel, a) (Figures 1a and b). 8-OxoG (Figure 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 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 (Shibutani and Grollman, 1994; Cheng et al., 1992; Moriya, 1993).

The superimposition of the ternary and binary crystal structures of the Dpo4/DNA complex with 8-oxoG according to their palm domains (Figure 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 and towards the finger domain by 12° before and after the reaction, respectively, so that 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 basepair’s distance in both the pre- and post-chemistry processes so that the DNA moves forward by a basepair after one complete reaction cycle.

The catalytic core of Dpo4 is shaped like a right hand comprised of thumb, palm, and fingers 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 1a and compared to their analogues in human DNA polymerase β (pol β). 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 analogues in pol β (Table 1a). Tyr10, Thr45, Tyr48, Arg51, and Lys159 form hydrogen bonds with 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; Ling et al., 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-N6-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 of primer dG, and three water oxygen atoms, while the nucleotide-binding ion B coordinates with Asp7, Asp105, mainchain carbonyl (C=O) group of Phe8, and three phosphate oxygen atoms on dCTP.

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 DNA polymerase $\beta$’s (pol $\beta$’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 that affect the DNA synthesis fidelity.

While recent pre-steady state kinetic data (Fiala and Suo, 2004b) 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 Schär, 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 archael Y-family polymerases) were swapped; it was found 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 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.

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 $\beta$ (Yang et al., 2002; 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 (Rechkoblit and Patel, a) to construct initial models. For root-mean-square deviation (RMSD) analysis, we use the analogous binary crystal complexes.
Results

Some moderate and high fidelity polymerases like human DNA pol β, T7 DNA polymerase, and HIV-1 RT undergo large-scale subdomain conformational transition from an “open” to “closed” state upon binding the correct dNTP. This “induced-fit” mechanism has been deduced on the basis of several structural and gained support from computational studies (Sawaya et al., 1997; Krahn et al., 2004; Arora and Schlick, 2004). In contrast, clear evidence of such conformational changes is lacking for Y-family polymerases including Dpo4, except that the little finger (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). To explore the nature of conformational transitions involved in Dpo4 during DNA replication, we performed two dynamics simulations before the nucleotidyl-transfer reaction and two simulations after the reaction.

Results 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 ns 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 expect that, if the “induced-fit” mechanism exists, 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 display more fluctuations than in the first trajectory with the incoming nucleotide. This can be seen from the root-mean-square deviation (RMSD) plots of the finger and LF domains with respect to the binary and ternary crystal complexes for both trajectories in Figure 2a 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 notable deformations occur over longer timescales than
23 ns 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; Ling et al., 2003; Ling et al., 2004a; Ling et al., 2004b), 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 coordinating ligands in the trajectory are listed in Table 1b. 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 of the primer terminus rather than the α-phosphate oxygen 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 Pα of the incoming nucleotide and the O3’ of the primer terminus is about 4.6 Å on average, larger than the ideal Pα–O3’ distance (3.3 Å) for the phosphoryl transferase reaction to proceed via a dissociative mechanism (Mildvan, 1997).

The active-site residues of the Dpo4/DNA complex undergo 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 (Figure 3b). Notably, Glu106 rotates toward Tyr108 and is stabilized by forming a hydrogen bond with Tyr108. Asp7 forms hydrogen bond with Lys159 and Asp105 has 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 sidechains (Arg258, Phe272, and Asp192) as well as the thumb subdomain motion (Yang et al., 2002; Radhakrishnan and Schlick, 2004; Arora and Schlick, 2004).

Results After Chemistry

Domain and DNA Motions. In our post-chemistry simulation where the bound ions and PPi are retained in the active site of the Dpo4/DNA complex, we capture the LF domain rotation back to the binary conformation within 13 ns. The simulated structure (average over the last 600 ps) 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 towards 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 (Figure 2e) indicates that the conformational transition starts at 6.8 ns and has a magnitude of rotation of approximately 3 Å. The finger and thumb domains, however, show no significant transitions, and their final conformations are similar to those in the ternary crystal structure (Figures 2f, 4 and Supplementary Figure 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 thumb domains when superimposing the palm domains of the ternary before chemistry and binary after chemistry crystal structures (Figure 1).

Motivated by the critical effects of the metal ions (Mg2+) on pol β’s thumb conformational change (Yang et al., 2004), we performed the second post-chemistry simulation where both metal ions and PPi were removed from the active site. In this trajectory, we found the finger and LF domains to exhibit large rigid-body movements. The RMSD evolution of the LF domain in Figure 2g shows that the LF domain starts changing conformation at the beginning of the simulation and reaches the maximum deviation from the ternary crystal structure by 12 ns. The thumb domain does not change conformation as dramatically as the LF domain, although it displays motion (Supplementary Figure S1c). The shape similarity in the RMSD plots of the LF and finger domains (Figure 2g and h) suggests a correlation of these conformational changes for the two domains.
During the period from 4.3 ns to 14 ns in the second post-chemistry simulation, the finger and LF domains deviate most significantly from the ternary structure. The averaged structure of the trajectory from 12 ns to 12.5 ns indicates that the LF domain moves forward (direction shown by a black arrow in Figure 5a) by as much as 3.5 Å and rotates around the DNA major groove (Figure 5c) by approximately 12°, as measured by the rotation of α-helices L and M on the LF domain (labeled in Figure 1b). Similarly, the finger domain also translates in the same direction as the LF domain by about 3.5 Å. The LF and 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 basepair’s rise (Figure 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 the ternary crystal structure and final 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 towards its conformation in Dbh (see Figure 5c), though the magnitude of the rotation is small (~12°) compared to the apo-structure of Dbh (which would require a 77° rotation and 2 Å translation (Ling et al., 2004a)). This suggests that the LF domain of Dpo4 might lose 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.

**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 largest two 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 towards the finger domain in the largest principal mode (Figure 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 (Figure 6b), 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 modes, the palm and DNA primer strand that reside in 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 (Supplementary Figure S2b) (the two principal modes above correspond to 56% of the total motion). In the largest eigenvalue-eigenvector pair (Figure 6c), both the finger and 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 primer strands display large magnitude motions in both the DNA duplex axis direction and LF rotation direction away from the palm domain. The thumb domain also moves slightly along the DNA and towards 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 basepair as well as the DNA duplex away from the active site.

In the second mode (Figure 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 (Figure 7a). Asp7 interacts with Lys159 after the simulation, while Asp105 rotates towards Ser103 and forms a weak hydrogen bond with Ser103. Glu106 is anchored by Lys152 and the metal ions in the ternary crystal structure. By the end of the simulation, Glu106’s sidechain flips 120° so
that it forms an additional hydrogen bond with Tyr108.

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 (Rechkoblit and Patel, a). 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 ns and interacts with residues Asp7 and Phe8 as well as O2P atom of the nascent primer nucleotide (Supplementary Figure 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.

Though the sugar pseudorotation angle (Altona and Sundaralingam, 1972; Arora and Schlick, 2003) of dCTP remains at C3'-endo before chemistry, this angle changes from C3'-endo (~18°) to C2'-endo (~160°) around 14 ns (Figure 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, the 3'-OH group on the sugar ring of dCTP forms a hydrogen bond with one non-bridging oxygen atom on β-phosphate (Figure 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 basepair.

**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 deter-
mining Dpo4’s enzymatic properties, such as processivity, fidelity, and lesion-bypassing. The authors 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\textsubscript{i} released, we analyze the sidechain profile of each charged or aromatic residue using dials and windows (see Supplementary Figure S4.1 to 4.5), for these residues are expected to have greater contributions to the LF’s surface electrostatic potential than other noncharged residues. We find that the Arg336 sidechain (C\textsuperscript{α}-C\textsuperscript{β}-C\textsuperscript{γ}-C\textsuperscript{δ}) rotates by 120\textdegree (Figure 8a and e) at 8.1 ns, 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 Figure 1a). One of these hydrogen bonds remains and strengthens after Arg336 rearranges, while the other weakens and is replaced by a hydrogen bond between Arg336 and O5' atom of T9 (Figure 8d). The shortened sidechain of Arg336 consequently pulls the template strand towards the LF domain, and triggers the sugar repuckering of T8 (Thymine) from C2'-endo to C3'-endo (Figure 8b and e). Arg336 and T8 remain at the new conformations for another 10 ns until the end of the simulation (Figure 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 (Figure 8c and f). The first hydrogen bond forms when the distance between the Arg298 and P8 decreases significantly at 10 ns. The second hydrogen bond starts to form at 10 ns and forms completely by 16.7 ns.

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 sidechain 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 sidechain conformational changes of active-site residues were noted with both binding metal ions and PP\textsubscript{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 before chemistry of the ternary crystal structure of the Dpo4/DNA complex with and without the incoming nucleotide at the active site capture no large conformational changes, suggesting 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 binary states. The absence of conformational changes in Dpo4 is different from observations in pol β simulations, where the thumb subdomain closes or opens depending on the presence or 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 (\(\sim 12^\circ\)) 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 (Figure 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° towards the finger domain, the LF rotation observed here when Dpo4 bypasses 8-oxoG is small, perhaps due to the modest distortion introduced to the active site by the 8-oxo group.

In contrast, our 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 pucker transition of the nascent primer nucleotide), in the LF and finger domains, and in the DNA. This suggests that the release of the metal ions may trigger a more systematic and correlated-among-the-domain Dpo4 conformational changes and 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 (Rechkoblit and Patel, a) — 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 (Figure 5a and 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 basepairs in the active site of Dpo4 slide as much as one basepair’s distance relative to the palm domain, with the template strand translocating by ~3.4 Å and the primer strand by ~2 Å (Figure 5b).

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.

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 (Figures 3a and 7a) resembles the flip of Asp192 in pol β after chemistry (Sawaya et al., 1997) (Table 1a). In addition, Asp7 and Asp105 are stabilized by forming hydrogen bonds with surrounding residues after simulation. These sidechain 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.

The sugar repuckering in the newly incorporated nucleotide from C3′-endo (A-DNA-like) to C2′-endo (B-DNA-like) (Figure 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. (Marquez et al., 2004) suggests that cellular polymerases prefer exclusively nucleotide substrates with sugar puckers in the C3′-endo conformation before the chemical reaction.

Studies have indicated that the LF domain may play a critical role in DNA translocation (Ling et al., 2001; Ling et al., 2004a; Boudsocq et al., 2004). In our post-chemistry
simulation without the metal ions, we note 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 sidechain 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.

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 (Figure 1b) (Rechkoblit and Patel, a); however, such thumb domain movements were not captured in our simulations. Perhaps DNA translocation requires milliseconds or longer in vivo. Nevertheless, the motions of the LF domain observed here demonstrate how they may facilitate DNA translocation; 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 may translocate both DNA strands by one basepair; 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.

The fact that we do not observe the finger domain to move away from the nascent basepair in all our simulations of Dpo4 as in pol β (Sawaya et al., 1997) suggests 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. The 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.

This idea is sketched in Figure 9. Thus, 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.

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 (Frisch et al., 1998) 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 (Case et al., 2002).

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 (Rechkoblit and Patel, a) 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 were kept intact as in the ternary crystal structure; these
units 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-Huckel distribution using SOLVATE (Grubmüller, 1996). 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 about 70,000 atoms (341 protein residues, 32 DNA basepairs, 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 ps 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 fs, a damping time constant of 50 fs, and piston temperature of 300 K. Temperature coupling was enforced by velocity reassignment every 2 ps. Then, the production dynamics was performed at constant temperature and volume. Constant temperature was maintained at 300K using weakly coupled Langevin dynamics of nonhydrogen atoms with a damping coefficient $\gamma$ of 5 ps$^{-1}$; pressure was maintained at 1 atm using a Langevin piston Nosé-Hoover barostat with an oscillation period of 200 fs and a decay time of 100 fs. Bonds to all hydrogen atoms were kept rigid using SHAKE (Ryckaert et al., 1977), permitting a time step of 2 fs. 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 and 23.2 ns, respectively.

System Preparation and Dynamics Simulations After Chemistry
To build the after-chemistry structure of the Dpo4/DNA complex, the nucleotidyl-transfer reaction was performed manually by reacting the $\alpha$-phosphate of the dCTP with the 3'-hydroxyl of the DNA primer strand in the ternary crystal structure of the Dpo4/DNA complex
with 8-oxoG and the incoming dCTP (Rechkoblit and Patel, a). The resulting pyrophosphate (PP$_i$) group and the two specific binding ions 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 using the crystal structure of polymerase IV from E. 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., 1987). All ions added were placed at least 8 Å away from the protein or DNA or from each other. The systems contain 48798 and 48804 atoms in the presence and 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 non-bonded 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 ps at 300K 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; finally, the system was equilibrated with the stochastic LN approach (Schlick et al., 1997; Barth and Schlick, 1998a; Barth and Schlick, 1998b; Schlick, 2001) for 150 ps with all the atoms released.

For the production runs, the multiple-timestep Langevin integrator, LN was used with timesteps $\Delta \tau / \Delta t_m / \Delta t$ of 1/2/150 fs 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 reliability of LN for biomolecular systems in terms of thermodynamic, structural and dynamic properties are examined thoroughly in Ref. (Yang et al., 2002) compared to single-timestep Langevin as well as Newtonian (Velocity Verlet) propagators). The damping constant $\gamma$ was set as 10 ps$^{-1}$ to couple the system to a 300°C heat bath. The first trajectory was simulated for 13 ns with both divalent metal ions present in the active site and the second was simulated for 18 ns without the presence of the metal ions and PP$_i$. SHAKE was applied to all bonds containing hydrogen atoms. Coordinates were saved every 3 ps. The 13 and 18 ns trajectories took about 86 and 129 days on eight 600MHz 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 hours are required to complete a one-nanosecond run for our Dpo4 system using 64 CPUs in parallel on NCSA IA32 LINUX supercluster, compared to 158 CPU hours per nanosecond on 8 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 ps) of the Dpo4 system, as follows.

A covariance matrix $\mathbf{C}$ 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 \( X_k \) 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 \( \Lambda \) from the decomposition

\[ V^T C V = \Lambda \] (1)

or

\[ C V_n = \lambda_n V_n, \quad n = 1, 2, \ldots, 3N, \] (2)

where \( \Lambda \) is the diagonal matrix with the eigenvalues

\[ \lambda_i = diag(\lambda_1, \lambda_2, \ldots, \lambda_{3N}). \] (3)

Each eigenvector \( V_n \) 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 \( (\lambda_n \sum_{n=1,3N} \lambda_n) \) indicates the relative percentage of the trajectory motions along eigenvector \( V_n \). Thus, an arbitrary structure \( Y \) can be generated from the average structure by a displacement \( D \) along the linear combination of all eigenvectors \( V_n \) with \( 3N \) scalars \( \alpha_n \), where

\[ Y = \langle X \rangle + D = \langle X \rangle + \sum_{n=1,3N} \alpha_n V_n. \] (4)

**Supporting Information**: 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 sidechain dihedral angles of the charged and aromatic residues on 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 stimulating 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. Schlick. 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 utilized to generate figures in this article.
References


Grubmüller, H. Solvate 1.0.


Rechkoblit, O. and Patel, D. J. Personal communication.


**Figures Legends**

**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 (Rechkoblit and Patel, a). 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 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 basepair, 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. Parts (d) and (e) display the molecular formulas of 8-oxoG and guanine, respectively.

**Figure 2.** Evolution of the RMSD for the Cα atoms in the LF and finger domains of Dpo4 in the simulations before and after chemistry relative to the crystal binary (green) and ternary complexes (red). (a) to (d) are 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) to (h) are those in the simulations after chemistry with and without metal ions in the active site. Superimpositioning is performed according to the palm domains.

**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 1b. (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.

**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.

**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.

**Figure 6.** Dpo4 domain motions revealed by PCA from the two trajectories after chemistry. (a) and (b) represent the two largest principal modes extracted from the simulation with the metal ions and PP$_i$ present in the active site. (c) and (d) are 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 green, thumb is red, finger is purple, and LF is yellow. The DNA backbone is traced, with P atoms in brown, O atoms in red, and C atoms in cyan.

**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 Sundaralingam, 1972; Arora and Schlick, 2003) of the incoming nucleotide (dC) in this trajectory.

**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) sidechain 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
(Cα-Cβ-Cγ-Cδ) and T8 sugar pucker pseudorotation angle P. (f) Time evolution of atom-atom distances between Arg298 and O2P atom of P8.

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

<table>
<thead>
<tr>
<th>Residues in Dpo4</th>
<th>Y12, V32, A42, A44, G58</th>
<th>Y10, T45, Y48, R51, K159</th>
<th>D7, D105, E106</th>
<th>Y108</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subdomain</td>
<td>Finger</td>
<td>Finger &amp; Palm</td>
<td>Palm</td>
<td>Palm</td>
</tr>
<tr>
<td>Analogues in Pol β</td>
<td>Y271, F272, T273, G274, S275, D276, R283</td>
<td>S180, R183, G189</td>
<td>D190, D192, D256</td>
<td>R258</td>
</tr>
<tr>
<td>Subdomain</td>
<td>Thumb</td>
<td>Palm</td>
<td>Palm</td>
<td>Palm</td>
</tr>
</tbody>
</table>

| Function | Interact with the template and incoming nucleotide bases | Interact with the triphosphate of the incoming dNTP | Conserved catalytic triad; D7, D105, and E106 reorients after the metal ions are released (discussed later) | Y108 stabilize E106 when the catalytic ion is released (discussed later) |

Table 1b. Magnesium ions coordination distances (in Å) in the simulated Dpo4 ternary complex and its crystal form.

<table>
<thead>
<tr>
<th>Distance</th>
<th>Dpo4/DNA/dCTP (Simulated 14.5 ns)</th>
<th>Dpo4* (X-ray)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mg$^2+$ (A)–Asp7:O$^{δ1}$</td>
<td>1.83</td>
<td>2.47</td>
</tr>
<tr>
<td>Mg$^2+$ (A)–Glu106:O$^{δ2}$</td>
<td>1.83</td>
<td>2.31</td>
</tr>
<tr>
<td>Mg$^2+$ (A)–354:OIP</td>
<td>1.89</td>
<td>2.32</td>
</tr>
<tr>
<td>Mg$^2+$ (A)–WAT1</td>
<td>1.99</td>
<td>n/a</td>
</tr>
<tr>
<td>Mg$^2+$ (A)–WAT2</td>
<td>2.01</td>
<td>n/a</td>
</tr>
<tr>
<td>Mg$^2+$ (A)–WAT3</td>
<td>1.98</td>
<td>n/a</td>
</tr>
<tr>
<td>Mg$^2+$ (B)–Asp7:O$^{δ2}$</td>
<td>1.89</td>
<td>2.30</td>
</tr>
<tr>
<td>Mg$^2+$ (B)–Phe8:O</td>
<td>1.89</td>
<td>2.30</td>
</tr>
<tr>
<td>Mg$^2+$ (B)–Asp105:O$^{δ2}$</td>
<td>1.84</td>
<td>2.30</td>
</tr>
<tr>
<td>Mg$^2+$ (B)–dCTP:O2$^{α}$</td>
<td>1.92</td>
<td>2.48</td>
</tr>
<tr>
<td>Mg$^2+$ (B)–dCTP:O1$^{β}$</td>
<td>1.90</td>
<td>2.28</td>
</tr>
<tr>
<td>Mg$^2+$ (B)–dCTP:O1$^{γ}$</td>
<td>1.82</td>
<td>2.31</td>
</tr>
</tbody>
</table>

Mg$^2+$ (A), catalytic magnesium; Mg$^2+$ (B), nucleotide binding magnesium; dCTP, 2’–deoxyadenosine 5’–triphosphate; 354, primer terminus; n/a, absent in the crystal structure

* In the crystal complex metal ions were resolved to be Calcium
### Binary complex (Before chemistry)

$$5'\text{CTAACGCTACATCCACC}3' \quad 5'\text{GATGGTAGTTGG}5'$$

### Ternary complex (Before chemistry)

$$5'\text{CTAACGCTACATCCACC}3' \quad 5'\text{GATGGTAGTTGG}5'$$

### Binary Complex (After chemistry)

$$5'\text{CTAACGCTACATCCACC}3' \quad 5'\text{CGATGGTAGTTGG}5'$$

**Figure 1**
Figure 2
Figure 3
Figure 5

(a) Fingers
(b) OxoG C
(c) Little finger
Figure 6
Figure 7
Figure 8
Figure 9