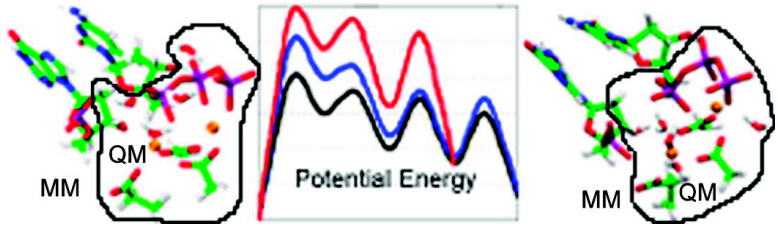
Article

Quantum Mechanics/Molecular Mechanics Investigation of the Chemical Reaction in Dpo4 Reveals Water-Dependent Pathways and Requirements for Active Site Reorganization

Yanli Wang, and Tamar Schlick

J. Am. Chem. Soc., **2008**, 130 (40), 13240-13250 • DOI: 10.1021/ja802215c • Publication Date (Web): 12 September 2008



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES

Published on Web 09/12/2008

Quantum Mechanics/Molecular Mechanics Investigation of the Chemical Reaction in Dpo4 Reveals Water-Dependent Pathways and Requirements for Active Site Reorganization

Yanli Wang and Tamar Schlick*

Department of Chemistry and Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, New York 10012

Received March 25, 2008; E-mail: schlick@nyu.edu

Abstract: The nucleotidyl-transfer reaction coupled with the conformational transitions in DNA polymerases is critical for maintaining the fidelity and efficiency of DNA synthesis. We examine here the possible reaction pathways of a Y-family DNA polymerase, *Sulfolobus solfataricus* DNA polymerase IV (Dpo4), for the correct insertion of dCTP opposite 8-oxoguanine using the quantum mechanics/molecular mechanics (QM/MM) approach, both from a chemistry-competent state and a crystal closed state. The latter examination is important for understanding pre-chemistry barriers to interpret the entire enzyme mechanism, since the crystal closed state is not an ideal state for initiating the chemical reaction. The most favorable reaction path involves initial deprotonation of O3′H via two bridging water molecules to O1A, overcoming an overall potential energy barrier of approximately 20.0 kcal/mol. The proton on O1A−P_α then migrates to the γ-phosphate oxygen of the incoming nucleotide as O3′ attacks P_α, and the P_α−O3A bond breaks. The other possible pathway in which the O3′H proton is transferred directly to O1A on P_α has an overall energy barrier of 25.0 kcal/mol. In both reaction paths, the rate-limiting step is the initial deprotonation, and the trigonal-bipyramidal configuration for P_α occurs during the concerted bond formation (O3′−P_α) and breaking (P_α−O3A), indicating the associative nature of the chemical reaction. In contrast, the Dpo4/DNA complex with an imperfect active-site geometry corresponding to the crystal state must overcome a much higher activation energy barrier (29.0 kcal/mol) to achieve a tightly organized site due to hindered O3′H deprotonation stemming from larger distances and distorted conformation of the proton acceptors. This significant difference demonstrates that the pre-chemistry reorganization in Dpo4 costs approximately 4.0 to 9.0 kcal/mol depending on the primer terminus environment. Compared to the higher fidelity DNA polymerase β from the X-family, Dpo4 has a higher chemical reaction barrier (20.0 vs 15.0 kcal/mol) due to the more solvent-exposed active site.

Introduction

DNA polymerases are fundamental enzymes responsible for DNA replication and repair and thus play a crucial biological role that maintains the continual survival of all living organisms on earth.[1,2] According to primary sequence alignments, DNA polymerases can be grouped into seven different families (A, B, C, D, X, Y, and RT). The recently discovered low-fidelity Y-family DNA polymerases[3] are particularly notable because of their lesion-bypassing capacity; namely, they can transverse bulky lesions at the replication fork whereas replicative DNA polymerases stall there and thus can rescue the replication process in an error-prone or error-free manner.[4−10]

Sulfolobus solfataricus DNA Polymerase IV (Dpo4), a thermophilic archaeal protein capable of translesion synthesis, is one of the Y-family DNA polymerases crystallized and well-studied experimentally. The Dpo4/DNA complexes containing regular as well as lesioned template bases at the active site have been resolved crystallographically.[5−8,11−17] These structures reveal four Dpo4 protein domains, including palm, thumb,

(1) Friedberg, E. C. *Nature* **2003**, *421* (6921), 436–440.
(2) Zhou, B. B.; Elledge, S. J. *Nature* **2000**, *408* (6811), 433–439.
(3) Ohmori, H.; et al. *Mol. Cell* **2001**, *8* (1), 7–8.
(4) Boudsocq, F.; Iwai, S.; Hanaoka, F.; Woodgate, R. *Nucleic Acids Res.* **2001**, *29*, 4607–4616.
(5) Ling, H.; Boudsocq, F.; Woodgate, R.; Yang, W. *Mol. Cell* **2004**, *13*, 751–762.
(6) Ling, H.; Boudsocq, F.; Plosky, B. S.; Woodgate, R.; Yang, W. *Nature* **2003**, *424*, 1083–1087.
(7) Ling, H.; Sayer, J. M.; Plosky, B. S.; Yagi, H.; Boudsocq, F.; Woodgate, R.; Jerina, M.; Yang, W. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 2265–2269.
(8) Ling, H.; Boudsocq, R.; Woodgate, R.; Yang, W. *Cell* **2001**, *107* (1), 91–102.
(9) Perlow-Poehnelt, R. A.; Likhterov, I.; Scicchitano, D. A.; Geacintov, N. E.; Broyde, S. *J. Biol. Chem.* **2004**, *279*, 36951–36961.
(10) Wang, L. H.; Wu, M.; Yan, S. F.; Patel, D. J.; Geacintov, N. E.; Broyde, S. *Chem. Res. Toxicol.* **2005**, *18*, 441–456.
(11) Irimia, A.; Eoff, R. L.; Pallan, P. S.; Guengerich, F. P.; Egli, M. *J. Biol. Chem.* **2007**, *282*, 36421–36433.
(12) Eoff, R. L.; Irimia, A.; Angel, K. C.; Egli, M.; Guengerich, F. P. *J. Biol. Chem.* **2007**, *282*, 19831–19843.
(13) Eoff, R. L.; Irimia, A.; Egli, M.; Guengerich, F. P. *J. Biol. Chem.* **2007**, *282*, 1456–1467.
(14) Vaisman, A.; Ling, H.; Woodgate, R.; Yang, W. *EMBO J.* **2005**, *24*, 2957–2967.
(15) Rechkoblit, O.; Malinina, L.; Cheng, Y.; Kuryavyi, V.; Broyde, S.; Geacintov, N. E.; Patel, D. J. *PLoS Biol.* **2006**, *4* (1), e11.
(16) Zang, H.; Goodenough, A. K.; Choi, J. Y.; Irimia, A.; Loukachevitch, L. V.; Kozekov, I. D.; Angel, K. C.; Rizzo, C. J.; Egli, M.; Guengerich, F. P. *J. Biol. Chem.* **2005**, *280*, 29750–29764.

10.1021/ja802215c CCC: $40.75

fingers, and little finger. Like other DNA polymerases, the active site for nucleotide insertion in Dpo4 is located in the palm domain, where the catalytic triad, Asp7, Asp105, and Glu106, coordinate with two metal ions—the nucleotide-binding and catalytic Mg^{2+}. It is believed that DNA polymerases across different families employ a common "two-metal-ion" mechanism[18−21] during the catalysis for the chemical reaction, in which the catalytic and nucleotide-binding Mg^{2+} ions stabilize the intermediate/transition states by bridging the attacking nucleophilic group O3′ and the α-phosphate of the incoming nucleotide via metal−ligand bonds. Castro et al.[21] recently showed that the two-proton transfer mechanism may even be common for polymerases in general, whether the active site is relatively open or closed. In their study, these two transition steps are deprotonation of O3′H and protonation of the pyrophosphate leaving group. However, they could not rule out water-assisted pathways.

Many studies have been performed to investigate the chemical reaction mechanism in various DNA polymerases, including high-fidelity DNA polymerase T7 from the A-family,[20] polymerase β (pol β) from the X-family,[22−27] and Y-family Dpo4.[28] Employing the Empirical Valence Bond (EVB) method, Florián and co-workers[20] calculated the free energy barriers for the chemical reaction in T7. They examined the possible competing reactions and found that the most favorable reaction pathway in T7 involves proton transfer to Asp654, an active-site catalytic Asp residue coordinating both the catalytic and nucleotide-binding metal ions, with a free energy barrier of 12 kcal/mol. Later, the same authors[29] also applied the linear response approximation and the EVB method to assess the contribution of the substrate binding, p*K*(a) shifts, and chemical steps to T7's overall fidelity.

For DNA pol β, quantum mechanical[22,23,27] and combined quantum mechanics/molecular mechanics (QM/MM) [24−26] studies suggest that different deprotonation routes for O3′H are possible, namely, direct deprotonation to an α-phosphorus oxygen of the incoming nucleotide,[22,27] proton migration to Asp256,[23] and proton transfer to explicit active site water molecules.[24,26] In addition, Warshel and co-workers have performed free energy calculations for the chemical reaction barriers in pol β using the EVB/FEP and the diabatic frozen density functional methods as well as the microscopic linear response approximation (LRA) and the semimacroscopic PDLD/S-LRA methods.[30−33] See also ref 34 on possible pitfalls of QM/MM minimization approaches.

Abashkin and co-workers[22] used a simplified active site model in their QM calculations to map out the reaction potential energy surface. In their model, the Asp256 was represented by an HO^- group, Asp192 was protonated, and no crystal water coordinated the catalytic Mg^{2+}; as a result, this ion was only bound to five ligands. The reaction pathway was found to be associative and initiated by O3′H deprotonation to a phosphate-oxygen on P_α. The rate-limiting step was the nucleophilic attack by O3′ on P_α. A pentacoordinated P_α transition state and a series of other intermediates were captured along the pathway.

Rittenhouse et al.[23] investigated the role of the missing O3′H from the primer and incoming nucleotide by MD and QM simulations of pol β. Results suggest that the missing primer terminus O3′H in the pol β crystal structure might have caused the distortions in the coordination sphere of the catalytic metal ion. They also found an associative pathway for the chemical reaction but not the major intermediates along the reaction path. Several protonation states of coordinating ligands were tested, and an additional protonated Asp256 was deemed likely.

Radhakrishnan and Schlick[24] employed short QM/MM dynamics combined with umbrella sampling to calculate the free energies of intermediates along pol β's reaction path. They proposed a reaction mechanism involving initial O3′H deprotonation to solvent water molecules followed by proton migration to active site Asp residues via a series of Grotthus hopping steps for both G:C and G:A systems. In particular, this work suggests that O3′H deprotonation occurs simultaneously with the nucleophilic attack.

More recently, Bojin and Schlick[27] performed a comprehensive QM investigation of the chemical reaction profile based on several pol β model systems. Several initial proton transfer pathways were examined. The direct route from O3′ to O(P_α) was found energetically the most favorable.

Lin et al.[25] performed QM/MM constrained minimizations using a pol β crystal structure containing a A:T matched nascent base pair, a fully revolved catalytic Mg^{2+}, and primer O3′H.[35] The latter two chemical groups were often missing in other crystal structures and were modeled into the active site by other authors.[22−24,26,27] The Lin et al. work identified a partially associative reaction path starting with O3′H deprotonation to Asp256 followed by the rate-limiting nucleophilic attack of O3′ on P_α. Recently, Lin et al. also investigated pol β's chemical reaction mechanism of incorrect nucleotide incorporation of a G:A mispair using MD, QM, and QM/MM methods.[36] A nonreactive pre-chemistry rearrangement was found to occur before the chemical reaction, while the transition state and rate-

(17) Zang, H.; Irimia, A.; Choi, J. Y.; Angel, K. C.; Loukachevitch, L. V.; Egli, M.; Guengerich, F. P. *J. Biol. Chem.* **2006**, *281*, 2358–2372.
(18) Steitz, T. A. *J. Biol. Chem.* **1999**, *274* (25), 17395–17398.
(19) Steitz, T. A.; Smerdon, S. J.; Jager, J.; Joyce, C. M. *Science* **1994**, *266* (5193), 2022–2025.
(20) Florian, J.; Goodman, M. F.; Warshel, A. *J. Am. Chem. Soc.* **2003**, *125* (27), 8163–8177.
(21) Castro, C.; Smidansky, E.; Maksimchuk, K. R.; Arnold, J. J.; Korneeva, V. S.; Gotte, M.; Konigsberg, W.; Cameron, C. E. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104* (11), 4267–4272.
(22) Abashkin, Y. G.; Erickson, J. W.; Burt, S. K. *J. Phys. Chem. B* **2001**, *105* (1), 287–292.
(23) Rittenhouse, R. C.; Apostoluk, W. K.; Miller, J. H.; Straatsma, T. P. *Proteins* **2003**, *53* (3), 667–682.
(24) Radhakrishnan, R.; Schlick, T. *Biochem. Biophys. Res. Commun.* **2006**, *350* (3), 521–529.
(25) Lin, P.; Pedersen, L. C.; Batra, V. K.; Beard, W. A.; Wilson, S. H.; Pedersen, L. G. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103* (36), 13294–13299.
(26) Alberts, I. L.; Wang, Y.; Schlick, T. *J. Am. Chem. Soc.* **2007**, *129* (36), 11100–11110.
(27) Bojin, M. D.; Schlick, T. *J. Phys. Chem. B* **2007**, *111* (38), 11244–11252.
(28) Wang, L.; Yu, X.; Hu, P.; Broyde, S.; Zhang, Y. *J. Am. Chem. Soc.* **2007**, *129* (15), 4731–4737.
(29) Florian, J.; Goodman, M. F.; Warshel, A. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102* (19), 6819–6824.
(30) Sucato, C. A.; Upton, T. G.; Kashemirov, B. A.; Batra, V. K.; Martínek, V.; Xiang, Y.; Beard, W. A.; Pedersen, L. C.; Wilson, S. H.; McKenna, C. E.; Florian, J.; Warshel, A.; Goodman, M. F. *Biochemistry* **2006**, *46* (2), 461–471.
(31) Xiang, Y.; Oelschlaeger, P.; Florian, J.; Goodman, M. F.; Warshel, A. *Biochemistry* **2006**, *45*, 7036–7048.
(32) Xiang, Y.; Goodman, M. F.; Beard, W. A.; Wilson, S. H.; Warshel, A. *Proteins* **2008**, *70* (1), 231–247.
(33) Xiang, Y.; Warshel, A. *J. Phys. Chem. B* **2008**, *112*, 1007–1015.
(34) Klahn, M.; Braun-Sand, S.; Rosta, E.; Warshel, A. *J. Phys. Chem. B* **2005**, *109*, 15645–15650.
(35) Batra, V. K.; Beard, W. A.; Shock, D. D.; Krahn, J. M.; Pedersen, L. C.; Wilson, S. H. *Structure* **2006**, *14*, 757–766.
(36) Lin, P.; Pedersen, L. C.; Batra, V. K.; Beard, W. A.; Wilson, S. H.; Pedersen, L. G. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105* (15), 5670–5674.

limiting step during chemistry were similar to what has been found earlier in the correct A:T system.

Thus, a range of possibilities for the nucleotidyl-transfer reaction mechanism in pol β was found by QM and QM/MM studies. This might result from both methodological approximations as well as biological versatility in the active site.

Since QM approaches[22−24,27] can only treat the active site region while necessarily neglecting the influence of the protein/DNA and solvent environment on the reaction mechanism, they are less accurate than the QM/MM methods[24−26] in some respects. In addition, different QM methods for describing the reactive core and different initial structural models, such as different protonation states for active-site residues and incoming nucleotide, were used. To address this problem, Alberts et al.[26] recently explored three possible reaction pathways in pol β (i.e., O3′H deprotonation to water molecules, to phosphate oxygen, and to catalytic Asp residues) using QM/MM based on the same simulation model and suggested that the water-assisted deprotonation pathway is more favorable than the other two.

However, besides methodological issues, it is also quite possible that evolution has permitted flexibility and versatility in polymerase mechanisms as well, to accommodate many different biological situations of active site environments, modified substrates, solvation, and salt. Such versatility could help avoid replication or repair malfunction under many varied natural conditions.

For Dpo4, Wang et al.[28] have performed QM/MM calculations implemented in GAUSSIAN03[37] and TINKER[38] and suggested that the nucleotidyl-transfer reaction is initiated by deprotonation of the primer terminus O3′H to a crystal water coordinating the catalytic Mg^{2+} and that the concerted deprotonation of O3′H and proton transfer from the water molecule to α-phosphate oxygen is the rate-limiting step of the entire reaction path, with an energy barrier of approximately 21 kcal/mol (free energy barrier of 13 kcal/mol). Overall, all these prior studies on different DNA polymerases indicate that the chemical reaction is associative in nature with a metastable pentacovalent phosphorane intermediate. The water-assisted deprotonation pathway might be the most favorable pathways for both pol β and Dpo4.

Pre-steady state kinetic experiments have suggested that the conformational transition before the chemical reaction is rate-limiting when Dpo4 incorporates a correct incoming nucleotide to the primer/template junction, while the chemical step is rate-limiting for the incorrect nucleotide insertion.[39] Through long-time dynamics simulations of the Dpo4/DNA complex in a native-like environment, we have captured significant rigid-body motions in the finger and little finger domains that facilitate the DNA translocation.[40] However, compared to the higher-fidelity pol β, these simulation results and crystallographic data[15] suggest that the active-site region of Dpo4 might not undergo a large-magnitude conformational change upon binding of a correct incoming nucleotide as seen in pol β.[41−43] The catalytic site in Dpo4 is also more distorted than in pol β after the conformational changes relative to the ideal reaction-competent conformation,[18,35,44] contributing to the low catalytic efficiency of Dpo4 compared to pol β.

Prior works on pol β and Dpo4 indicate that the polymerases' active sites must undergo subtle reorganization prior to the chemical reaction,[40,44,45] so that the reactive region achieves an "ideal" geometry for the chemical catalysis.[18,35,44] This was termed the "pre-chemistry" avenue[44] to distinguish this phase from conformational rearrangements prior to chemistry and the chemical reaction itself. Of course, one entire free energy profile could also be used to describe the mechanism, but the separate treatment of these three phases helps emphasize their distinct nature. Namely, evidence suggests that pre-chemistry barriers are overcome by stochastic paths that may be distinct from the stepwise path of the chemical reaction itself. Such subtle active-site rearrangements before chemistry in pol β and Dpo4 are further supported by the fact that the ideal geometry is rarely seen, neither in the crystal structures[15,46] nor in the molecular dynamics[40,42] and transition path and free energy sampling[24,43] simulations following the conformational change avenue. This suggests that the DNA polymerase complex can reside in various local minima and that, to reach the reactive state (or the "ideal" geometry) to initiate the chemical reaction, the enzyme must overcome energy barriers in a stochastic manner.[44] The pre-chemistry concept is also supported by the recent work by Lin, et al.,[36] where the authors found that a nonreactive pre-chemistry active-site structure rearrangement step occurs before the nucleotidyl transfer reaction of a G:A mismatch system and forms the major factor in decreasing the rate of incorrect nucleotide incorporation and thus to fidelity enhancement.

Here we simulate the nucleotidyl-transfer reaction in Dpo4 by various models to explore possible reaction pathways, including initial deprotonation of O3′H: to active site water molecules, to α-phosphate oxygen atoms on the incoming nucleotide, and to the catalytic carboxylate side chains on (Asp7, Asp105, and Glu106), using QM/MM simulations implemented in CHARMM/GAMESS-UK. The crystal structure of the Dpo4/DNA/dCTP ternary complex with 2′-deoxy-7,8-dihydro-8-oxoguanosine (8-oxoG) (Figure 1b) at the template position[15] is used to build the simulation systems. We choose this crystal structure for three reasons. First, 8-oxoG is one of the most prevalent lesions found in DNA and thus represents a natural state for Dpo4. Second, because of the changes at N7 and C8, 8-oxoG differs from guanine (Figure 1a) in its capacity to form a Hoogsteen base pair (8-oxoG:A) in addition to an 8-oxoG:C base pair. The chemical reaction mechanism obtained from the 8-oxoG:C system could allow a comparison with that of the G:C template.[28] Third, the active-site region in the chosen structure containing 8-oxoG is more favorable for the chemical reaction than most of the other known Dpo4 crystal structures. This is because of the relatively undisturbed metal ion coordination in the active site due to the small size of the lesion as well as the nicely resolved O3′H group on the incoming nucleotide, a segment usually missing in the other structures.

In comparison, the crystal structure used in ref 28 also contains a reasonably organized active site, where the catalytic

(37) Frisch, M. J.; et al. *Gaussian 03*, revision C.02; Gaussian, Inc.: Wallingford, CT, 2004.
(38) Ponder, J. W. *TINKER*; Washington University in St. Louis: St. Louis, MO, 2004.
(39) Fiala, K. A.; Suo, Z. *Biochemistry* **2004**, *43* (7), 2116–2125.
(40) Wang, Y.; Arora, K.; Schlick, T. *Protein Sci.* **2006**, *15* (1), 135–151.
(41) Yang, L.; Beard, W. A.; Wilson, S. H.; Broyde, S.; Schlick, T. *J. Mol. Biol.* **2002**, *317*, 651–671.
(42) Arora, K.; Schlick, T. *Biophys. J.* **2004**, *87*, 3088–3099.
(43) Radhakrishnan, R.; Schlick, T. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 5970–5975.
(44) Radhakrishnan, R.; Arora, K.; Wang, Y.; Beard, W. A.; Wilson, S. H.; Schlick, T. *Biochemistry* **2006**, *45* (51), 15142–15156.
(45) Arora, K.; Beard, W. A.; Wilson, S. H.; Schlick, T. *Biochemistry* **2005**, *44* (40), 13328–13341.
(46) Sawaya, M. R.; Prasad, R.; Wilson, S. H.; Kraut, J.; Pelletier, H. *Biochemistry* **1997**, *36*, 11205–11215.

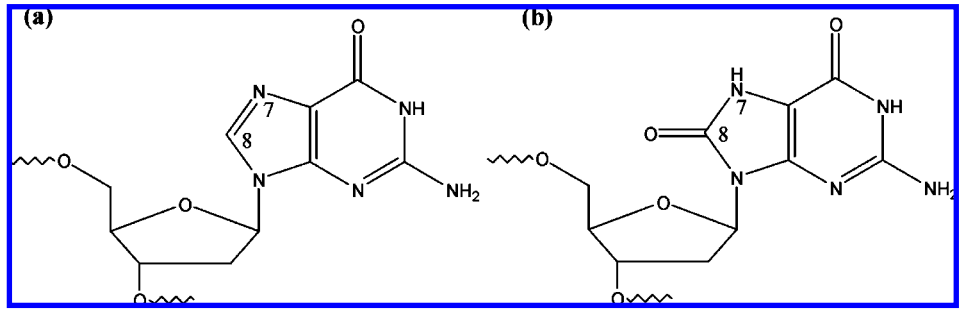


Figure 1. Molecular formulas of (a) guanine and (b) 8-oxoguanine.

metal ion coordinates five ligands (Asp7, Asp105, Glu106, O2A, and a crystal water molecule). However, a bulky lesion (the benzo[a]pyrenyl moiety) was located at the active site of this structure and was replaced by a hydrogen atom before the simulation. This replacement could introduce approximations in the model as well.

We perform two systematic explorations with one model built from the Dpo4/DNA crystal structure,[15] in which the active site region corresponds to the crystal state and not to the "ideal" geometry, and the other model based on an optimized structure with "ideal" ligand binding for the Mg^{2+} ions. Only the latter form was considered in the recent work,[28] which did not consider the pre-chemistry barriers that may be crucial to interpreting polymerase mechanisms.[44] Thus, we aim to describe the possible chemical reaction pathways for Dpo4 and estimate the energy barrier associated with the "pre-chemistry" organization, as well as chemistry.

Methods

Model Preparation. The Dpo4/DNA complex with a correct 8-oxoG:dCTP base pair at the active site was constructed using the available crystal structure[15] (PDB ID: 2ASD). The missing residues 341−352 were modeled and connected to the C-terminus of Dpo4 protein as in our previous work.[40] The crystal waters and ions were kept in the simulation system. The two binding Ca^{2+} ions at the active site of Dpo4 were replaced by Mg^{2+}. Note that replacing Ca^{2+} by Mg^{2+} is a valid approximation. Vaisman et al.[14] have compared the two Ca^{2+} ions in the active site of the Ab-2A structure[5] to the two Mg^{2+} ions in the active site of the T7 DNA polymerase ternary complex[47] and found that Ca^{2+} ions are not the cause for the metal ion displacement observed in the Dpo4/DNA/T:dATP (PBD ID: 2AGQ) structure. Although the catalytic ion in this structure is Mg^{2+} rather than Ca^{2+}, the Mg^{2+} ion does not bind the α-phosphate oxygen on the incoming dATP. Thus, the Ca^{2+} ions in the 2ASD structure can be replaced by Mg^{2+} without causing distortions in the active site. The missing 3′OH group on the primer terminus and the hydrogen atoms were added using CHARMM.[48] (Details for modeling the "ideal" active-site conformation are described in Supporting Information.) The protein/DNA complex was then solvated in a water box using PBCAID.[49] Solvent ions (Na^+ and Cl^-) were added by replacing the water oxygen atoms with the most negative and most positive electrostatic potentials (computed using DELPHI[50,51]), respectively, to neutralize the system and produce an ionic strength of 150 mM. This initial model contains 48801 atoms, including 352 protein residues, 33 nucleotide residues, 13,959 TIP3 water molecules, 3 Mg^{2+}, 47 Na^+, and 28 Cl^- ions as shown in Figure 2a.

The solvated complex was minimized for 10 000 steps using the Steepest Descent (SD) Method followed by 20 000 steps using the Adapted Basis Newton−Raphson (ABNR) procedure.[48,52] This system was then equilibrated for 100 ps at 300 K using the Verlet integrator. These calculations were performed using CHARMM with the all-atom version c31a1 force field.

Since the QM/MM calculations are time-consuming, to reduce the computing cost and maintain the accuracy of the calculations, after equilibrating the initial simulation model, we reduced the model by removing all solvent molecules (water and ions) outside of the 40 Å region of the two binding Mg^{2+} ions at the active site. The resulting complex retains the ionic strength of 150 mM with 25 Na^+, 6 Cl^-, and 7219 water molecules. The protein and DNA were undisturbed (Figure 2b). Note that the solvent ions (Na^+ and Cl^-) were not placed randomly but close to where the most negatively charged and most positively charged regions, respectively. Therefore, they can reduce the polarization effect of the MM part on the QM subsystem compared to the situation where no ions are added to neutralize the charges on DNA and protein residues. In addition, since all the solvent ions are fixed at the initial positions during the QM/MM simulations, the polarization caused by these ions is further minimized.

The protonation states of the catalytic residues at the active site, namely Asp7, Asp105, and Glu106, were predicted using the MCCE program.[53,54] The calculations suggest that their carboxylate sidechains are not protonated before the chemical reaction, each bearing a −1 charge. The triphosphate of the incoming dCTP is dominantly unprotonated under neutral pH according to the pK_a of $H(dCTP)^{3-}$ (approximately 6.52[55]) and carries a −4 charge. All Lys and Arg residues outside of the active site were assigned +1 charge whereas Asp and Glu residues were assigned −1 charge according to the pK_a values of their titratable sidechains.

QM/MM Calculations. We utilize the combined QM/MM approach[53,56,57] implemented in the CHARMM program coupled with GAMESS-UK (an *ab initio* molecular electronic structure program)[58] to describe the chemical reaction and protein environment as in ref 26. The QM methods in GAMESS-UK are used to treat the catalytic core of the Dpo4/DNA complex. This region

(47) Doublie, S.; Tabor, S.; Long, A.; Richardson, C.; Ellenberger, T. *Nature* **1998**, *391*, 251–258.
(48) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187–217.
(49) Qian, X. L.; Strahs, D.; Schlick, T. *J. Comput. Chem.* **2001**, *22*, 1843–1850.
(50) Gilson, M. K.; Sharp, K. A.; Honig, B. H. *J. Comput. Chem.* **1988**, *9*, 327–335.
(51) Klapper, I.; Hagstrom, R.; Fine, R.; Sharp, K.; Honig, B. *Proteins* **1986**, *1*, 47–59.
(52) Schlick, T. Optimization methods in computational chemistry. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; CH Publishers: New York, NY, 1992; pp 1−71.
(53) Georgescu, R. E.; Alexov, E. G.; Gunner, M. R. *Biophys. J.* **2002**, *83* (4), 1731–1748.
(54) Alexov, E. G.; Gunner, M. R. *Biophys. J.* **1997**, *72* (5), 2075–2093.
(55) Sigel, H.; Griesser, R. *Chem. Soc. Rev.* **2005**, *34*, 875–900.
(56) Warshel, A. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 425–443.
(57) Gao, J.; Truhlar, D. G. *Annu. Rev. Phys. Chem.* **2002**, *53*, 437–465.
(58) Guest, M. F.; Bush, I. J.; van Dam, H. J. J.; Sherwood, P.; Thomas, J. M. H.; van Lenthe, J. H.; Havenith, R. W. A.; Kendrick, J. *Mol. Phys.* **2005**, *103*, 719–747. GAMESS-UK is a package of ab initio programs. See: http://www.cfs.dl.ac.uk/gamess-uk/index.shtml.

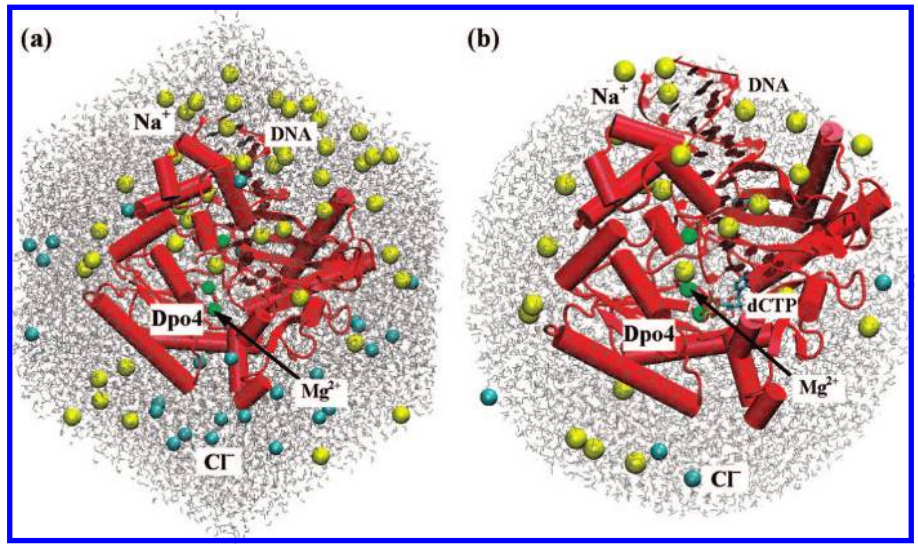


Figure 2. (a) Initial simulation model of the solvated Dpo4/DNA/8-oxoG:dCTP complex. (b) Reduced model of the Dpo4/DNA/8-oxoG:dCTP complex. All atoms within 15 Å of any QM atom are freely optimized during the minimizations, except for the restraints imposed to follow the reaction pathway. Atoms between 15 and 20 Å of any QM atom are semiconstrained and atoms beyond 20 Å are fixed in their original positions. In both models, the Dpo4 protein and its binding DNA are shown in red cartoon representation; the sodium, chloride, and magnesium ions are rendered as yellow, light blue, and green spheres, respectively; the incoming dCTP is shown by bonds, and solvent water molecules are shown by gray lines.

contains 52 atoms, including the three carboxylate sidechains (Asp7, Asp105, and Glu106), the two binding Mg^{2+} ions, the incoming nucleotide dCTP, DNA primer terminus, and four water molecules coordinating the catalytic Mg^{2+} ion and the incoming dCTP. The remaining molecular system is treated with the MM method in CHARMM. Since the boundary between the QM and MM regions cuts through covalent bonds, the single link atom approach, where the link atoms are taken as hydrogen atoms, is used to saturate the valence of the QM atom at the QM/MM interface.[59,60] The interface is positioned at the $C_\alpha-C_\beta$ bond of the amino acids, the C5′−C4′ bond of dCTP, and along the C2′−C1′ and C4′−C3′ bonds of the terminal DNA primer. They were constrained to be colinear along the QM/MM bond vector and placed at a fixed distance from the QM atom.

In the QM/MM calculations, all atoms within 15 Å of any QM atom are unconstrained and allowed to move freely. Atoms 15−20 Å away from any QM atom are semiconstrained with a harmonic force constant of 5.0 kcal/mol/Å^2 and atoms further than 20 Å are fixed in their initial positions throughout the simulations. To identify the favorable reaction pathways and the intermediate structures along each reaction path, we implement several constraints in the energy minimization procedure. Harmonic constraints with force constants K=350.0 kcal/mol/Å^2 are applied along the proton transfer path (O···H···O), while force constants of K = 2000.0 kcal/mol/Å^2 are used to restrain the O3′−P_α and O3A−P_α bond distances (these distances are indicative of the bond formation and bond breaking processes) during the nucleotidyl-transfer reaction. A series of constrained minimizations was conducted in which the forming bond distance was reduced by a step size of 0.1 Å while the breaking bond distance was increased with the same step size. All other internal degrees of freedom were unconstrained during the minimizations. Full geometry optimizations were also performed for the reactant and product states with no constraints.

The Hartree−Fock[61] (HF) method with a relatively small basis set 3−21G was used to identify possible reaction paths using constrained minimization. The optimal reaction pathways were reoptimized using Density Functional Theory (DFT) at the B3LYP/6−31G*[62,63] level. The identified favorable reaction paths were also driven backward to check for convergence. Results show that the forward and backward pathways converge. We expect an error bar of approximately 1.5 kcal/mol due to the limited sampling spaces.

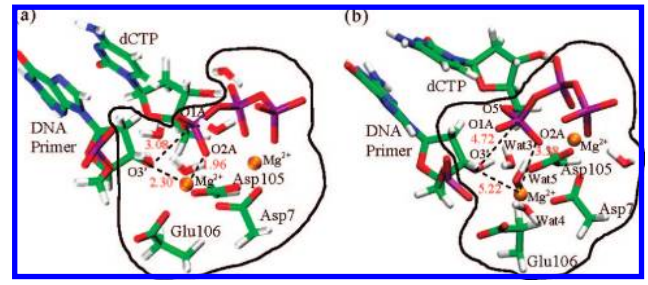


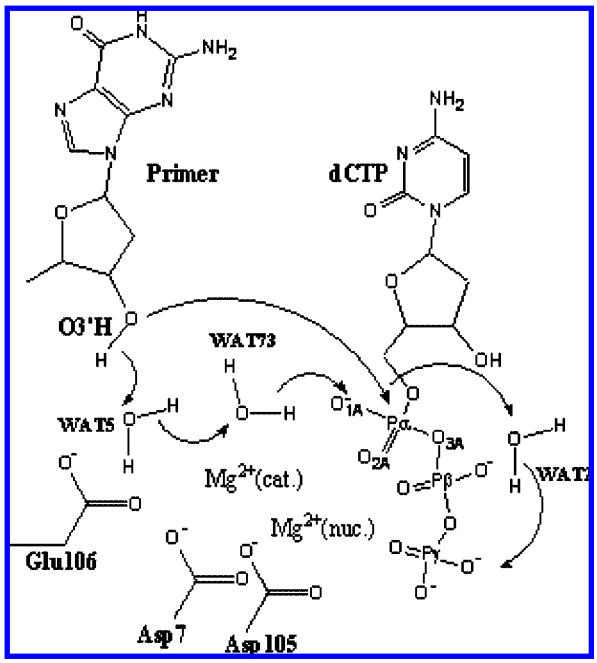
Figure 3. The catalytic sites of Dpo4 in the optimized (a) and the crystal state models (b). QM atoms are enclosed in the black circles. The missing 3′OH of the primer terminus was added using CHARMM in both models. The critical distances, namely O3′−Mg^{2+} (cat.), Mg^{2+} (cat.)−O2A, and P_α−O3′ (in Å) are shown by dashed lines and labeled in red.

Results

In the model constructed according to the "ideal" geometry[18,35,44] (see Supporting Information for details), after minimization and equilibration, the critical distances for the chemical reaction, specifically, O3′−Mg^{2+} (cat.), Mg^{2+} (cat.)−O2A, and P_α−O3′, were reduced to 2.30, 1.96, and 3.08 Å, respectively (Figure 3a). The O3′H proton does not form a hydrogen bond with O5′, O2α, or water molecules, as found in other cases,[26,27] but instead points to the opposite direction of the metal−ligand bond between O3′ and Mg^{2+} (cat.). The catalytic Mg^{2+} coordinates with three carboxylate oxygen atoms on Asp7, Asp105, and Glu106, O1A on P_α of dCTP, primer terminus O3′, and a water oxygen atom. Using the optimized model, we explored all

(59) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11* (6), 700–733.
(60) Das, D.; Eurenius, K. P.; Billings, E.; Sherwood, P.; Chatfield, D. C.; Hodoscek, M.; Brooks, B. R. *J. Chem. Phys.* **2002**, *117*, 10534–10547.
(61) Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; DeFrees, D. J.; Pople, J. A. *J. Chem. Phys.* **1982**, *77* (7), 3654–3665.
(62) Becke, A. D. *J. Chem. Phys.* **1993**, *98* (7), 5648–5652.
(63) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1987**, *37* (2), 785–789.

Scheme 1. Schematic Representation of the Water-Assisted Deprotonation Pathway in the Optimized System[a]



[a] Arrows show the direction of proton transfer from O3′−hydroxyl through the mediating water molecules (WAT5, WAT73, and WAT2) to the leaving pyrophosphate, as well as the direction of nucleophilic attack by O3′ at the α-phosphate.

possible reaction pathways starting with O3′H deprotonation to the formation of the pyrophosphate leaving group. The possible acceptors for O3′H proton considered in our calculations include the phosphate oxygen atom (O1A) on the incoming dCTP, carboxylate groups of the catalytic residues (Asp105 and Glu106), and nearby water molecules coordinating the catalytic Mg^{2+}.

In addition, as mentioned earlier, we also quantitatively estimated the effect of active-site distortions on the chemical reaction pathway and energy barriers, pertaining to the concept of "pre-chemistry" reorganization,[40] by performing QM/MM calculations on another Dpo4/DNA complex model taken from the crystal structure[15] where the active site region was not optimized to an "ideal" configuration.

I. Chemical Reaction Pathways Explored for the Dpo4/DNA Complex with Optimized Active-Site Geometry. Water-Assisted Deprotonation. Applying the constrained minimization protocol, which decreases the distance between the O3′H proton and the WAT5 oxygen while increasing the O3′···H bond distance with a step size of 0.1 Å, we derived the reaction path for proton transfer from O3′ to a nearby water molecule (Scheme 1). Interestingly, the initial O3′H deprotonation triggers a second proton transfer from WAT5 to a hydrogen-bonding water (WAT73). The extra proton on WAT73 then quickly hops to O1A of P_α (Intermediate I in Figures 4 and 5). The proton transfer steps from O3′H to WAT5 and then to WAT73 are concerted and together overcome an overall energy barrier of approximately 20.0 kcal/mol. During proton transfer from O3′H to WAT5 the distance between O3′ and P_α is reduced to 2.64 from 3.09 Å. The second proton transfer from WAT73 to O1A is accompanied by an O3′···P_α distance reduction from 2.64 to 1.96 Å. The proton on O1A then rotates 180° around the P_α−O1A bond with an associated energy barrier of 3 kcal/mol and is stabilized by a hydrogen bond with a free water molecule (WAT2) (Intermediate II). This intermediate is more stable than intermediate I because of the stabilizing hydrogen bond. With O3′ bearing a negative charge, it attacks P_α as the P_α−O3A covalent bond breaks, resulting in a pentacovalent transition state (Intermediate III), where both O3′−P_α and P_α−O3A distances are 1.82 Å. This process costs 7 kcal/mol. The proton on O1A then transfers to an oxygen atom on P_γ via the hydrogen-bonding WAT2 by overcoming an energy barrier of 7 kcal/mol. Finally, the P_α−O3A bond breaks and releases the byproduct pyrophosphate. As the distance between P_α and O3A increases, the negatively charged O3A forms a new hydrogen bond with a free water molecule (WAT100, Figure 4e). This hydrogen bond stabilizes the pyrophosphate and facilitates its release.

In addition to the proton transfer route from O1A to WAT2, we tested another possible pathway, proton transfer from O1A to O3A, which is the bridging oxygen between P_α and P_β and has been found to be a possible proton acceptor in pol β by Abashkin et al.[26] However, we found that the energy barrier for the direct proton transfer from O1A to O3A in Dpo4 was 12 kcal/mol, making it much higher than the water-mediated route. Thus, the proton transfer via a hydrogen-bonding water is a more favorable path.

Phosphate-Oxygen-Assisted Deprotonation. The second reaction pathway we examined for the same initial model starts with deprotonation of O3′H to O1A on P_α (see Figures 6 and 7 and Schemes 2 and 3). This reaction path was optimized by the B3LYP/6−31G* method and the energy barrier for the first proton-transfer step is approximately 25.0 kcal/mol, about 5.0 kcal/mol larger than in the pathway involving explicit water molecules. During formation of Intermediate I, the distance between O3′ and P_α decreases from 3.10 to 2.68 Å. The proton on O1A rotates around the P_α−O1A bond by crossing an energy barrier of 2.5 kcal/mol and forms a hydrogen bond with a free water molecule (WAT2) (Intermediate II). The O3′ group then initiates a nucleophilic attack on the P_α, producing the classic trigonal-bipyramidal configuration of P_α with both O3′−P_α and P_α−O3A bond distances of 1.80 Å (Transition state). The energy barrier for achieving this state is 7 kcal/mol. Finally, the P_α−O3A bond breaks as the proton on O1A transfers to WAT2 and then to O2G on P_γ, resulting in the formation of the pyrophosphate group. The final proton transfer crosses an energy barrier of 8 kcal/mol.

After the bond breaks between P_α and O3A, a free water molecule (WAT100, Figure 6e) forms a hydrogen bond with O3A as in the water-assisted pathway (Figure 4e). This stabilizes the pyrophosphate and speeds up its release together with the nucleotide-binding Mg^{2+}.

Initial Deprotonation via Catalytic Aspartic Acid Residues. The other proton transfer routes we explored involve active-site carboxylate sidechains as the general base. Using constrained minimization, we first tested proton transfer to Asp105. The activation energy associated with the initial deprotonation is approximately 27.3 kcal/mol, 7.3 kcal/mol larger than that in the water-assisted deprotonation pathway. However, the protonation of Asp105 disrupts the active site geometry, particularly the coordination of the magnesium ions. The next proton hop from OD1-Asp105 to OD2-Asp105 leads to a less stable system with 20 kcal/mol higher in energy. In addition, the energy barrier for bond formation and breaking between O3′ and P_α−O3A becomes forbiddingly large (e.g., >40 kcal/mol). Thus, we conclude that the Asp105-assisted deprotonation is not viable.

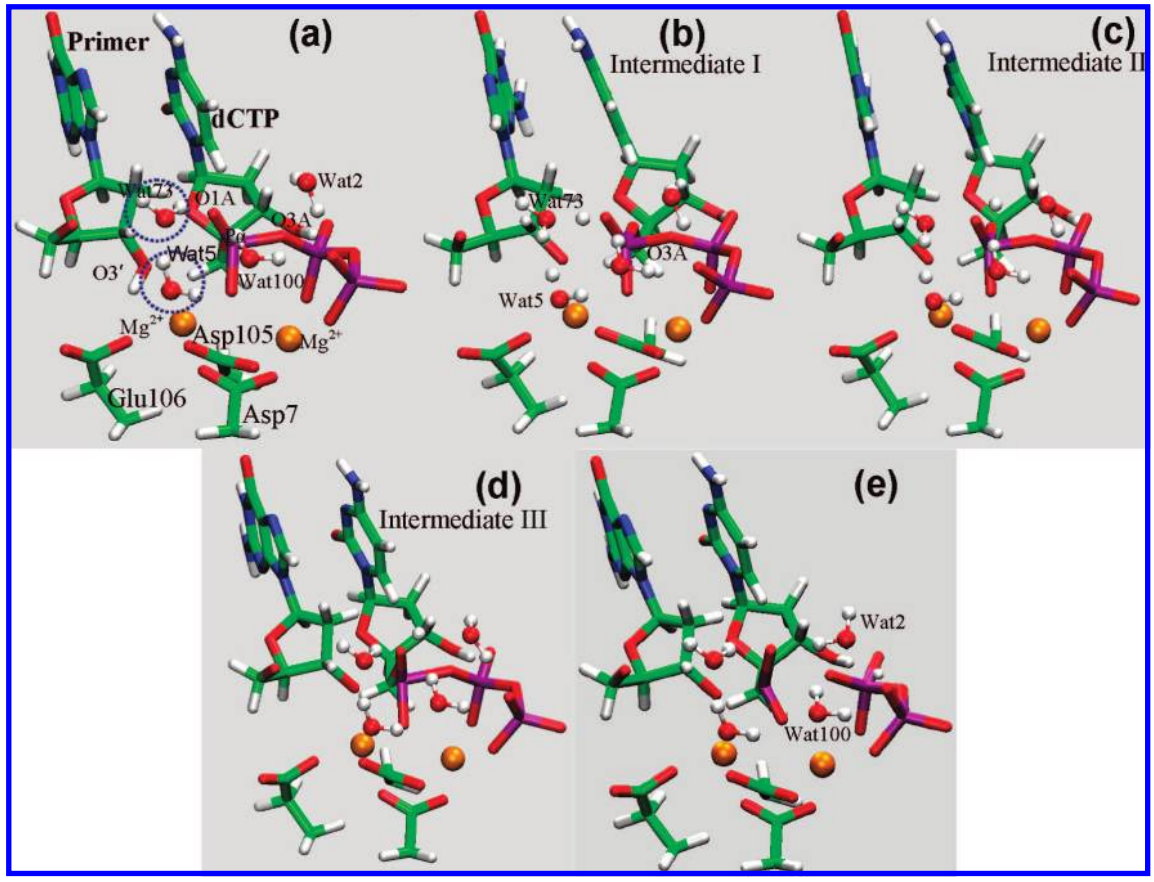


Figure 4. Intermediates along the water-assisted deprotonation pathway, where two explicit water molecules abstract the O3′H proton and transfer it to O1A of the incoming dCTP. (a) Initial reactant state (the two water molecules involved in abstracting the O3′H proton are circled with dashed blue lines); (b) intermediate I, proton on O1A; (c) intermediate II, proton rotation around O1A−P_α; (d) intermediate III, pentacovalent P_α; (e) product state.

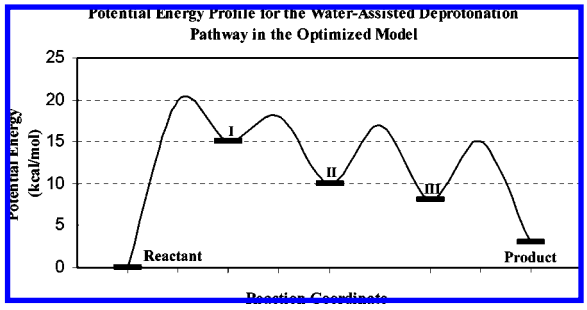


Figure 5. Energy profile for the transient intermediates identified in the water-assisted deprotonation reaction pathway (labeled corresponding to Figure 4).

Constraining O3′H proton to Glu106, which serves as the general base, we found that the energy profile is irregular, and a protonated Glu106 is highly unstable. This suggests that Glu106 is not a favorable proton acceptor. Furthermore, since Asp7 is more than 4 Å away from the 3′-hydroxyl group and the Mg^{2+} (cat.) is in the middle coordinating both of them, a direct proton transfer from O3′H to Asp7 seems unlikely.

II. Chemical Reaction Pathways Explored for the Dpo4/DNA Complex with Crystal State Geometry. Starting with the ternary closed crystal structure of the Dpo4/DNA complex with 8-oxoguanine opposite a correct incoming nucleotide,[15] we built a corresponding starting model by energy relaxation (see Supporting Information for details). We first minimize the entire model system followed by reoptimization and equilibration of the active site region using the QM/MM method. The active site conformation after QM/MM equilibration is shown in Figure 3b. The catalytic magnesium ion coordinates three water molecules, Asp7 and Glu106 carboxylate oxygen atoms, and a phosphate-oxygen on the primer terminus. The nucleotide-binding magnesium coordinates with three phosphate oxygen atoms, a carbonyl oxygen from Phe8, and two carboxylate oxygen atoms from Asp7 and Asp105. There are two extra solvent water molecules binding the Mg^{2+} (cat.) in the crystal state model compared to the ideal active site conformation (Figure 3a). This might arise from the large distance between the primer hydroxyl group and the catalytic Mg^{2+} as well as between O3′H and the triphosphate group of the incoming dCTP. In this optimized structure, the O3′−Mg^{2+} (cat.), Mg^{2+} (cat.)−O2A, and P_α−O3′ distances are 5.22, 3.38, and 5.22 Å, respectively (compared to about 2.30, 1.96, and 3.08 Å in the near-chemistry state).

We explored the possible reaction pathways from this crystal-derived model (Figure 3b) for initial O3′H proton transfer to nearby water molecules, to Asp105 and Glu106, and to O1A of the incoming dCTP using the constrained minimization protocol. However, none of these calculations yielded a favorable reaction path. The resulting energy barriers for the deprotonation and bond making and breaking steps are extremely large (e.g., > 60 kcal/mol). After careful examination of the simulation trajectories, we found that the two water molecules (WAT3 and WAT5) binding Mg^{2+} (cat.) are not in favorable conformation to either accept the O3′H proton or transfer it to other functional groups. More important, their presence hinders the rearrangement of the catalytic site during

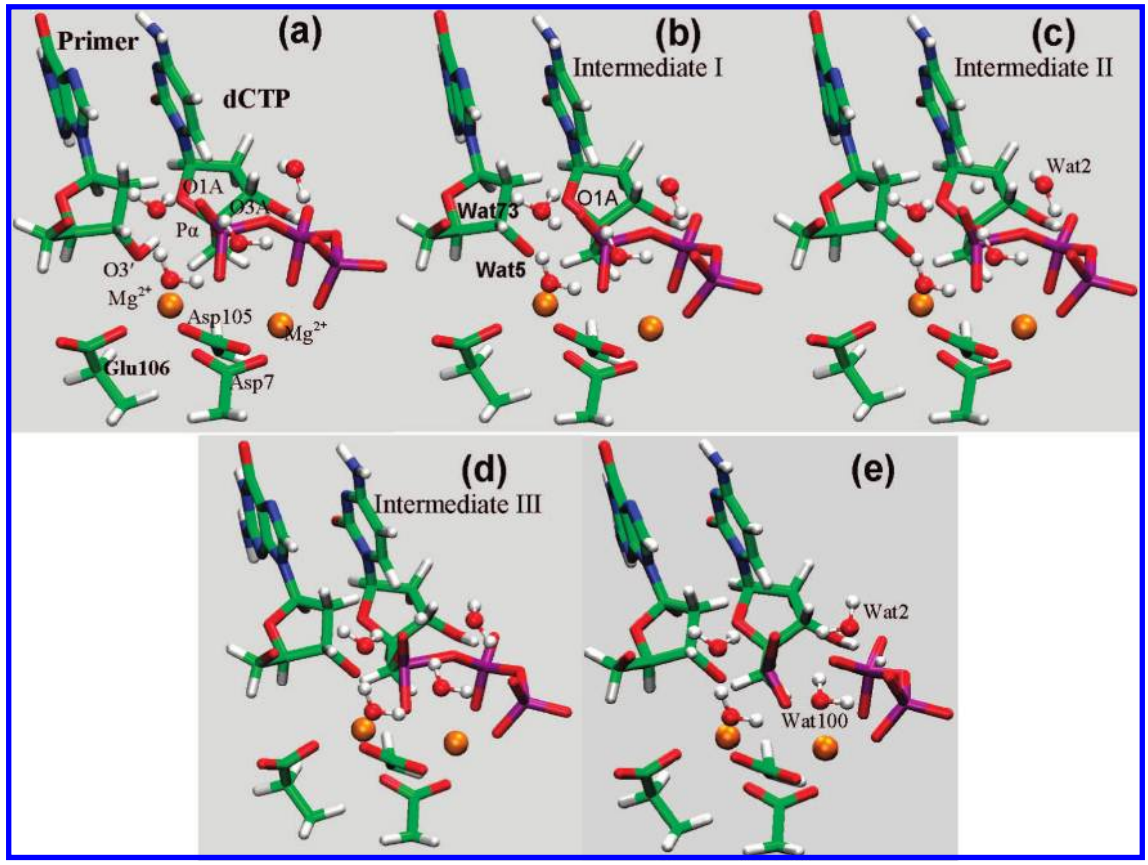


Figure 6. Intermediates along the phosphate-oxygen-assisted reaction pathway, where O1A of the incoming dCTP functions as the general base to abstract the O3′H proton. (a) Reactant state; (b) intermediate I, proton on O1A; (c) intermediate II, proton rotation around O1A−P_α; (d) intermediate III with a pentacovalent P_α; (e) product state.

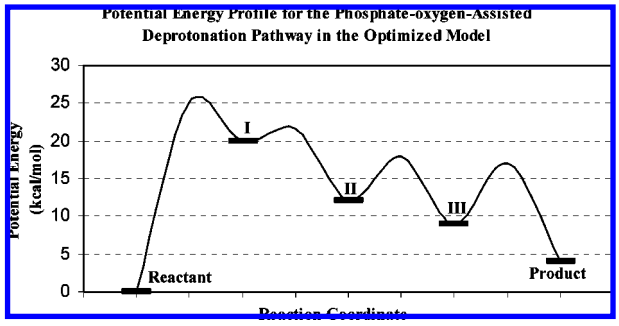
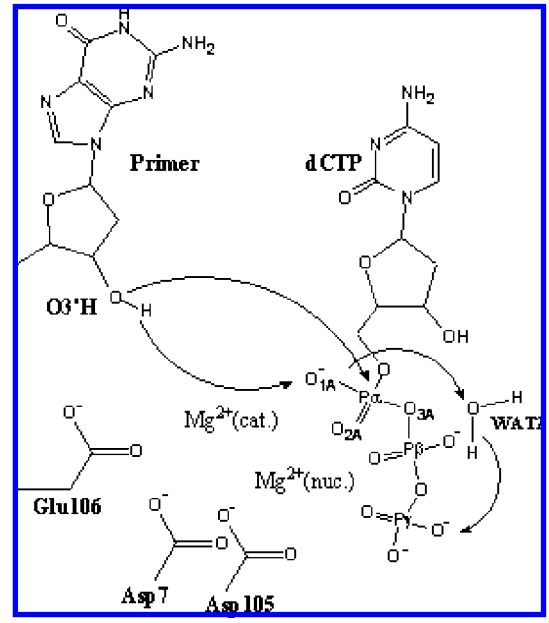


Figure 7. Energy profile for the transient intermediates identified in the phosphate-oxygen-assisted reaction pathway (labeled as in Figure 6).

the nucleophilic attack of O3′ on P_α. Since the coordination of Mg^{2+} (cat.) with O3′ and the α-phosphate moiety from dCTP can bring the two groups together and stabilize the transition state, we regenerated the reaction pathways by removing WAT3 and WAT5, which impede the catalytic magnesium ion's binding of O1A and O3′. In the resulting complex, the Mg^{2+} (cat.) coordinates the remaining water molecule (WAT4), Asp7, Glu106, and the phosphate oxygen on the primer terminus. The subsequent equilibration and constrained minimization for O3′H proton transfer to Asp105, Glu106, and O1A on P_α yielded a favorable mechanism for the proton transfer to O1A.

The energy barrier associated with proton transfer to O1A is approximately 29 kcal/mol (Intermediate I in Figures 8 and 9). Intermediate I, where O1A is protonated, closely resembles Intermediate I in the phosphate-oxygen-assisted deprotonation pathway identified from the optimized system (Figure 6b). Here,

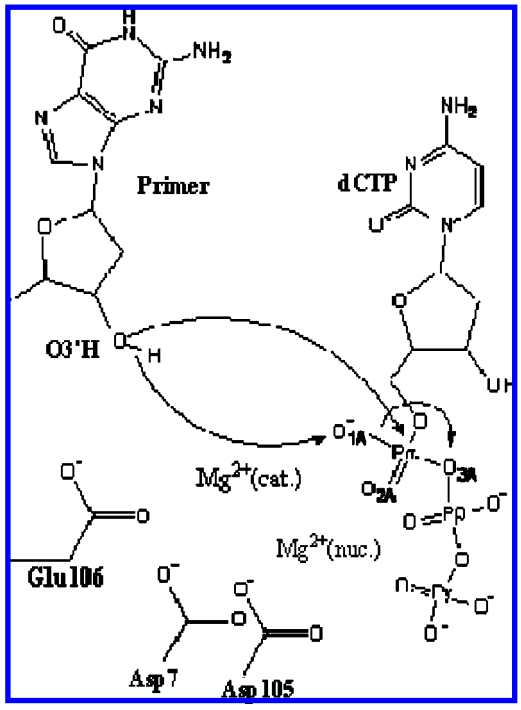
Scheme 2. Schematic Representation of the Phosphate-Oxygen-Assisted Deprotonation Pathway in the Optimized System[a]



[a] Arrows indicate the direction of proton transfer from O3′−hydroxyl directly to O1A and through WAT2 to the leaving pyrophosphate, as well as the direction of nucleophilic attack by O3′ at the α-phosphate.

the catalytic Mg^{2+} coordinates Asp7, Asp105, Glu106, O1A−P_α, and a water molecule (Figure 8b) as found in the

Scheme 3. Overall Phosphate-Oxygen-Assisted Deprotonation Pathway in the Crystal System with Imperfect Active-Site Geometry[a]



[a] Arrows show the direction of proton transfer from O3′-hydroxyl directly to O1A and that of nucleophilic attack by O3′ at the α-phosphate.

optimized system, except that the distance between Mg^{2+} (cat.) and O3′ is 3.09 Å, slightly larger than a metal−ligand bond.

Following the initial deprotonation, the proton on O1A then rotates around the P_α−O1A bond, crossing an energy barrier of 4 kcal/mol (Intermediate II). The system becomes more stable than Intermediate I. Since there remains no free water molecule to form a hydrogen bond with the rotated proton, intermediate II is less stable than that found in the optimized system (Figure 6c). Finally, the negatively charged O3′ interacts with P_α and generates a covalent bond while the P_α−O3A bond breaks, and the proton is transferred from O1A to O3A. The energy barrier associated with the last step is 10 kcal/mol. In contrast to the phosphate-oxygen-assisted deprotonation pathway of the optimized system (Figure 6), there is no free water molecule within hydrogen bonding distance of O3A after the P_α−O3A bond breaks. Thus, the negative charge on the pyrophosphate is less stabilized here than in the optimized model. This difference might also contribute to the slowing release of the pyrophosphate and the higher energy barrier observed in the crystal model.

In addition to the phosphate-oxygen-assisted deprotonation pathway, we also tested proton transfer to Asp105 and Glu106. However, the energy barriers in both cases are much greater than 40 kcal/mol, and we did not pursue these pathways further.

Discussion

Our analysis of possible nucleotidyl-transfer reaction pathways for nucleotide insertion in Dpo4 includes the reaction path initiated by proton transfer from O3′H to explicit water molecules as well as that to oxygen atoms on P_α and to the carboxylate side chains of the catalytic protein residues (Asp105 and Glu106). Our results show that the most favorable reaction path in Dpo4 involves initial deprotonation of O3′H to nearby water molecules, and the associated overall energy barrier is 20.0 kcal/mol. Another less viable reaction path is through deprotonation to the phosphate oxygen on P_α with an overall energy barrier of 25.0 kcal/mol. The difference underscores the mediating effect of the bridging water molecules to lower the energy barrier for proton hopping by approximately 5 kcal/mol. We deem the other reaction pathways involving initial deprotonation to the catalytic carboxylate residues (Asp105 and Glu106) unfavorable.

In the identified reaction pathways of Dpo4, the rate-limiting energy barrier stems from the initial deprotonation rather than the critical bond making and breaking between O3′ and P_α and O3A. A similar rate-limiting step was found for pol β by Radhakrishnan and Schlick and by Alberts et al.,[24,26] while the QM/MM study by Lin et al.[25] suggests that the nucleophilic attack on the P_α by O3′ is rate-limiting for pol β. The initial deprotonation in pol β by explicit water molecules requires less energy than in Dpo4 (15.0 vs 20.0 kcal/mol),[26] and this energy difference (∼5.0 kcal/mol) might be related to the lower efficiency of nucleotide insertion in Dpo4 compared to pol β. Specifically, the pathway in pol β is initiated by O3′H deprotonation to nearby water molecules; the proton then hops to the catalytic side chains of Asp192 and Asp190 as the O3′−P_α bond forms and P_α−O3A breaks, and finally the proton reaches the γ-phosphate of the incoming dCTP.[24,26] In contrast, the water-assisted pathway in Dpo4 first transfers the O3′H proton from two bridging water molecules to O1A, and then the proton hops to the γ-phosphate of the incoming dCTP as the bonds form (O3′−P_α) and break (P_α−O3A).

As known from the crystal structures, Dpo4 has a more open active site than pol β; this additional space leads to more prominent solvation effects on the active site of Dpo4 that can slow down the chemical reaction. This might explain the higher chemical reaction energy barrier in Dpo4 than pol β. Furthermore, the frequent water shuffling within Dpo4's active site can trap the deprotonated O3′H proton among water molecules. This can lead to formation of local energy minima that may reduce the efficiency of the chemical reaction.

Although the water-assisted reaction pathway identified in our work is similar to what was found by Wang et al.,[28] our results show that the initial O3′H deprotonation can be mediated by two water molecules rather than one. Indeed, different numbers of water molecules participating in the chemical reaction are natural in the enzyme environment. Since Dpo4 has a spacious active site compared to other higher-fidelity DNA polymerases, water molecules are expected to constantly migrate in and out of the active site; water-assisted proton transfer might adopt different forms when different numbers of water molecules are present. In simulations, water molecules can migrate to the catalytic site and form a hydrogen bond with the water molecule coordinating the catalytic magnesium ion and O1A on P_α during equilibration. Furthermore, the similarity between the chemical reaction mechanism obtained for the 8-oxoG:C nascent basepair and that of the G:C basepair[28] shows that the presence of an oxidative lesion (8-oxoguanaine) at the active site does not affect the magnitude of the chemical reaction barrier.

With our starting model based on the crystal structure, the catalytic site, especially the coordination of the two magnesium ions, is not in an ideal conformation to start the chemical reaction. This more realistic initial state is important for understanding Dpo4's complete conformational and chemical pathways and hence relevant functional mechanisms.[44] From this geometry, we identified one possible reaction pathway after removing two water molecules binding the Mg^{2+} (cat.) since

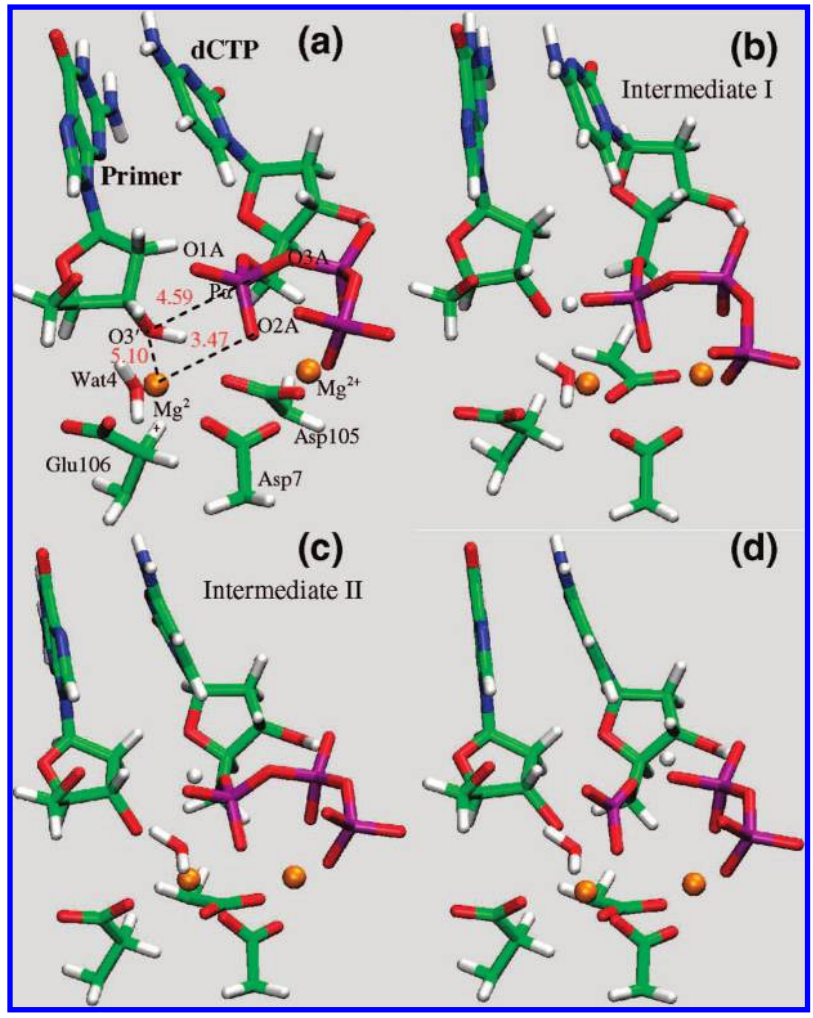


Figure 8. Intermediates along the reaction pathway in the crystal model where proton abstraction occurs directly to O1A of the incoming dCTP. (a) Reactant state after removing the two water molecules from the active site (the critical distances (Å) of Mg^{2+} (cat.)-O2A, Mg^{2+} (cat.)-O3′, and O3′-P_α are shown by dashed lines and labeled in red); (b) intermediate I, proton on O1A; (c) intermediate II, proton rotation around O1A-P_α; (d) product state.

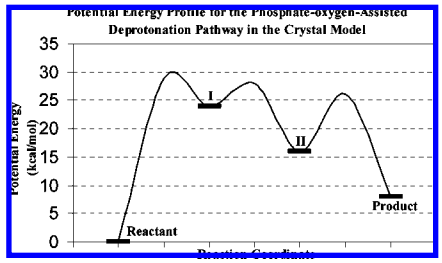


Figure 9. Energy profile for the transient intermediates identified in the phosphate-oxygen-assisted reaction pathway of the crystal system (labeled as in Figure 8).

this allowed the active site to reorganize as needed. In our pathway, the catalytic magnesium ion migrates toward the primer terminus and coordinates the phosphate oxygen (O2A) on P_α of dCTP and Asp105 as found in the optimized system with ideal active-site geometry (Figure 3a) during the deprotonation of O3′H. Then another phosphate oxygen (O1A) on P_α functions as the general base to remove the O3′H proton. The initial deprotonation accompanied by active-site reorganization overcomes an energy barrier of 29 kcal/mol. The O1A proton then migrates to O3A, resulting in the formation of the pyrophosphate moiety.

Compared to the water-assisted and phosphate-oxygen-assisted deprotonation pathways obtained from the model with ideal active-site geometry, the chemical reaction in the crystal state model has an activation energy barrier that is 9.0 and 4.0 kcal/mol higher in magnitude, respectively. The reaction in the crystal system can start by proton transfer from O3′H to O1A (a phosphate-oxygen-assisted deprotonation pathway), followed by proton rotation around O1A and the bond forming/breaking process between O3′, P_α, and O3A. As the O1A proton is transferred to O3A, the pyrophosphate group forms and is released. Since the starting conformations for the two phosphate-oxygen-assisted deprotonation pathways differ in the coordination of the catalytic magnesium ion, the difference in their associated energy barriers suggests that the reorganization of the catalytic site prior to the chemical reaction, namely the "pre-chemistry" avenue,[44] requires approximately 4 kcal/mol. For the chemical reaction to occur through the water-assisted deprotonation pathway, one or two solvent water molecules must enter the active site and favorably bridge the 3′-hydroxyl and O1A−P_α. The water-mediated deprotonation (Figure 5) can lower the overall reaction barrier by 9.0 kcal/mol relative to the pathway started from the crystal state. From the crystal conformation, although there are three water molecules coordinating the catalytic magnesium ion (Figure 3b), because the

distance between 3′-hydroxyl and P_α (4.72 Å) is much greater than the ideal value (3.08 Å), none of these water molecules is in the right conformation to transfer proton from O3′H to O1A.

Note that the obtained energy barriers for the "pre-chemistry" avenue are tentative (see also ref 32) as they are estimated based on the potential energy profiles of two single states (i.e., a crystal state and an "ideal" state) and no sampling was performed to calculate the free energies. However, results agree with extensive other studies using dynamics simulations. Certainly, future computational works delineating the free energy profiles of the "pre-chemistry" and chemistry avenues together with clever experiments are necessary to fully confirm our findings.

In sum, our results suggest that properly oriented water molecules in the active site play critical role in Dpo4's chemical reaction. In fact, they can both act to mediate the reaction pathway (as from the optimized system, where two water molecules function as a bridge to extract the O3′H proton and transfer it to O1A on the α-phosphate of the incoming dCTP), or to hinder active-site reorganization when the water molecules are not in favorable conformation to assist the protein transfer (as we found from simulations near the crystal state). Thus, for Dpo4, the pre-chemistry rearrangements depend on the environment where active-site tightening for the chemical step is crucial. Like the conformational changes prior to chemistry and the chemical reaction itself, the pre-chemistry avenue represents another important kinetic checkpoint in Dpo4's overall catalytic cycle.

Acknowledgment. We thank Dr. Ian L. Alberts for helpful discussions. This work was supported by NSF grant MCB-0316771, NIH grants R01 GM55164 and R01 ES012692, and the donors of the American Chemical Society Petroleum Research Fund to T. Schlick. Research described in this article was supported in part by Philip Morris U.S.A. Inc. and Philip Morris International. Computing facilities provided by the NCSA supercomputer center and the Advanced Biomedical Computing Center at NCI-Frederick are highly appreciated.

Supporting Information Available: Detailed procedure for building the two model systems. Complete refs 3 and 37. This material is available free of charge via the Internet at http://pubs.acs.org.

JA802215C