

A Water-Mediated and Substrate-Assisted Catalytic Mechanism for *Sulfolobus solfataricus* DNA Polymerase IV

Lihua Wang,[†] Xinyun Yu,[†] Po Hu,[‡] Suse Broyde,[†] and Yingkai Zhang^{*‡}

Contribution from the Biology and Chemistry Departments, New York University,
New York, New York 10003

Received December 9, 2006; E-mail: yingkai.zhang@nyu.edu

Abstract: DNA polymerases are enzymes responsible for the synthesis of DNA from nucleotides. Understanding their molecular fundamentals is a prerequisite for elucidating their aberrant activities in diseases such as cancer. Here we have carried out ab initio quantum mechanical/molecular mechanical (QM/MM) studies on the nucleotidyl-transfer reaction catalyzed by the lesion-bypass DNA polymerase IV (Dpo4) from *Sulfolobus solfataricus*, with template guanine and Watson–Crick paired dCTP as the nascent base pair. The results suggested a novel water-mediated and substrate-assisted (WMSA) mechanism: the initial proton transfer to the α -phosphate of the substrate via a bridging crystal water molecule is the rate-limiting step, the nucleotidyl-transfer step is associative with a metastable pentacovalent phosphorane intermediate, and the pyrophosphate leaving is facilitated by a highly coordinated proton relay mechanism through mediation of water which neutralizes the evolving negative charge. The conserved carboxylates, which retain their liganding to the two Mg^{2+} ions during the reaction process, are found to be essential in stabilizing transition states. This WMSA mechanism takes specific advantage of the unique structural features of this low-fidelity lesion-bypass Y-family polymerase, which has a more spacious and solvent-exposed active site than replicative and repair polymerases.

1. Introduction

DNA polymerase enzymes function to catalyze chain elongation, one residue at a time, during polynucleotide synthesis. In this nucleotidyl-transfer reaction, the 3'-OH group in the sugar ring reacts with the α -phosphate of a deoxyribonucleoside triphosphate, pyrophosphate is released, and extension is achieved. A large set of structural and biochemical data has shown that the active sites of all polymerases share certain common features: a pair of metal ions (normally Mg^{2+}) and three universally conserved carboxylates.^{1,2} These elements are proposed to play the key catalytic role: metal ion A lowers the pK_a of the 3'-OH, facilitating the 3'-O attack on the α -phosphate of the incoming nucleotide, while metal ion B assists the leaving of the pyrophosphate. Both metal ions, interacting with the carboxylates and the triphosphate group of the incoming nucleotide, help assemble and organize the enzyme active site. Meanwhile, these enzymes are quite diverse in their overall structure and sequence and have been grouped into different families with shared structural and functional properties.^{3,4} They also have distinct biological roles in cellular transactions involving DNA synthesis in replication and repair.⁴

The precise mechanism of the polymerase-catalyzed nucleotidyl-transfer reaction has been of great interest in recent

years.^{3,5–11} Key issues are the following: (i) What is the general base that accepts the proton from the 3'-OH? (ii) Does the new phosphodiester linkage form before (associative), at the same time (concerted), or after (dissociative) the cleavage of the bond connecting $P\alpha$ with the α,β -bridging oxygen? (iii) What are the reaction intermediates and transition states during this process? (iv) Which step is energetically rate limiting? (v) What are the roles of the conserved carboxylates in the chemical mechanism? (vi) What is the role of water molecules in this process? Understanding these molecular fundamentals of central cellular mechanisms for faithfully synthesizing requisite polynucleotides is necessary for elucidating their malfunctioning in diseases such as cancer.¹²

To address these questions, we have selected a member of the recently characterized Y-family lesion-bypass DNA polymerases^{13–15} for investigation. These polymerases are particu-

[†] Biology Department.
[‡] Chemistry Department.

(1) Steitz, T. A. *Nature* **1998**, *391*, 231–232.
(2) Steitz, T. A. *J. Biol. Chem.* **1999**, *274*, 17395–17398.
(3) Rothwell, P. J.; Waksman, G. *Adv. Protein Chem.* **2005**, *71*, 401–440.
(4) Bebenek, K.; Kunkel, T. A. *Adv. Protein Chem.* **2004**, *69*, 137–165.

(5) 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*, 13294–13299.
(6) Florian, J.; Goodman, M. F.; Warshel, A. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6819–6824.
(7) Showalter, A. K.; Lamarche, B. J.; Bakhtina, M.; Su, M. I.; Tang, K. H.; Tsai, M. D. *Chem. Rev.* **2006**, *106*, 340–360.
(8) Beard, W. A.; Wilson, S. H. *Chem. Rev.* **2006**, *106*, 361–382.
(9) Florian, J.; Goodman, M. F.; Warshel, A. *J. Am. Chem. Soc.* **2003**, *125*, 8163–8177.
(10) Abashkin, Y. G.; Erickson, J. W.; Burt, S. K. *J. Phys. Chem. B* **2001**, *105*, 287–292.
(11) Radhakrishnan, R.; Schlick, T. *Biochem. Biophys. Res. Commun.* **2006**, *350*, 521–529.
(12) Sweasy, J. B.; Lauper, J. M.; Eckert, K. A. *Radiat. Res.* **2006**, *166*, 693–714.
(13) Goodman, M. F. *Annu. Rev. Biochem.* **2002**, *71*, 17–50.
(14) Yang, W. *Curr. Opin. Struct. Biol.* **2003**, *13*, 23–30.
(15) Prakash, S.; Johnson, R. E.; Prakash, L. *Annu. Rev. Biochem.* **2005**, *74*, 317–353.

larly important because they replace replicative polymerases in order to bypass DNA damage.^{4,13,16,17} They synthesize one or a few residues in the lesion vicinity, are usually low fidelity on undamaged DNA, but may transit certain lesions accurately. Understanding the impact of lesions on the reaction mechanism is a long-term goal of this work. Excellent crystallographic characterization of the Y-family polymerase Dpo4 from the crenarchaeon *Sulfolobus solfataricus*¹⁸ has provided the basis for our investigation. Dpo4 is a member of the DinB family of bypass DNA polymerases found in all three kingdoms of life,¹⁹ including the human Pol κ .²⁰ In particular, the ternary complex (enzyme, DNA, and dNTP) we selected contains the critical 3'-OH group, which was absent until very recently²¹ from all other structures to avoid reaction. Dpo4 shares the universal "two-metal ion" organizing center at the active site, as well as the three conserved carboxylates that coordinate the metal ions.²² However, it differs structurally in that its active site is more spacious and solvent-exposed. Importantly, Dpo4, unlike higher-fidelity polymerases such as T7^{23,24} and Pol β ,^{25,26} does not employ an induced-fit mechanism to create a closed, catalytically competent active site where the nascent Watson-Crick base pair in the ternary complex is in van der Waals contact with the polymerase. Dpo4 relies more heavily on Watson-Crick pairing^{27,28} for fidelity than on close steric fit in the ternary complex. In going through a cycle of replication, Dpo4 undergoes a combination of rotational and translational motions²⁹ associated with translocation, rather than the opening-closing of the higher-fidelity enzymes. Interestingly, kinetic data have shown that the nucleotidyl-transfer step is not rate determining in this enzyme,³⁰ in contrast to the case for Pol β .^{31,32}

In the present study, the catalytic mechanism of the lesion-bypass DNA polymerase Dpo4 has been studied on the basis of the pseudobond ab initio quantum mechanical/molecular mechanical (QM/MM) approach,³³⁻³⁵ which has been successfully employed in the study of several enzymes.³⁶⁻⁴¹ Our results

for this lesion-bypass polymerase suggest that the general base is the α -phosphate of the substrate, and the reaction mechanism is associative (which involves a metastable pentacovalent phosphorane intermediate⁴²⁻⁴⁴), has a rate-limiting free energy barrier of 13 kcal/mol for the initial proton transfer to the α -phosphate via a bridging crystal water molecule, and employs a highly coordinated proton-relay mechanism through mediation of two water molecules; the conserved carboxylates retain their liganding to the two Mg²⁺ ions as the reaction progresses. On the basis of a theoretical study, substrate-assisted catalysis was first proposed over a decade ago as a mechanism for the GTP hydrolysis reaction in *Ras* p21;^{45,46} it has recently obtained further support from a crystal structure of a GTP-binding protein complexed with GDP and P_i,⁴⁷ although kinetic isotopic effect studies appear to suggest other possibilities.⁴⁸ A comparison of our results for Dpo4 with results of previous studies of the nucleotidyl-transfer reaction in the high- and medium-fidelity polymerases, T7^{6,9} and Pol β ,⁵ respectively, shows the possibility of unique mechanistic features for the low-fidelity and structurally distinct Y-family Dpo4 enzyme.

2. Computational Methods

We have employed molecular dynamics simulations and ab initio QM/MM calculations⁴⁹⁻⁵⁶ to study the nucleotidyl-transfer reaction in the Dpo4 active site, with template guanine and Watson-Crick paired dCTP as the nascent base pair. The QM sub-system (75 atoms) is treated by the B3LYP functional with a 6-31G* basis set, and it comprises the dCTP, the primer 3'-nucleotide (excluding C5' and phosphate group), the nucleotide-binding and catalytic Mg²⁺ ions, and two mediating water molecules involved in the nucleotidyl-transfer reaction. All other residues and the surrounding water molecules (9843 atoms) are described by the AMBER99 force field⁵⁷ and the TIP3P water model.⁵⁸ The QM/MM interface is based on the pseudobond approach,^{34,35} and an efficient iterative optimization approach and the reaction coordinate driving method³³ have been employed to minimize the reactant structure and map out the minimum energy path. On the basis of the determined reaction energy path, the free energy changes associated with the QM/MM interaction were determined by the free energy perturbation (FEP) method.³³ All QM/MM and FEP calculations were carried out with modified versions of Gaussian03⁵⁹ and TINKER⁶⁰ programs. Our employed pseudobond ab initio QM/MM approach³³⁻³⁵

- (16) Cordonnier, A. M.; Fuchs, R. P. *Mutat. Res.* **1999**, *435*, 111-119.
 (17) Pages, V.; Fuchs, R. P. *Oncogene* **2002**, *21*, 8957-8966.
 (18) Ling, H.; Sayer, J. M.; Plosky, B. S.; Yagi, H.; Boudsocq, F.; Woodgate, R.; Jerina, D. M.; Yang, W. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 2265-2269.
 (19) Boudsocq, F.; Iwai, S.; Hanaoka, F.; Woodgate, R. *Nucleic Acids Res.* **2001**, *29*, 4607-4616.
 (20) Gerlach, V. L.; Aravind, L.; Gotway, G.; Schultz, R. A.; Koonin, E. V.; Friedberg, E. C. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 11922-11927.
 (21) Batra, V. K.; Beard, W. A.; Shock, D. D.; Krahn, J. M.; Pedersen, L. C.; Wilson, S. H. *Structure* **2006**, *14*, 757-766.
 (22) Ling, H.; Boudsocq, F.; Woodgate, R.; Yang, W. *Cell* **2001**, *107*, 91-102.
 (23) Doublé, S.; Ellenberger, T. *Curr. Opin. Struct. Biol.* **1998**, *8*, 704-712.
 (24) Doublé, S.; Tabor, S.; Long, A. M.; Richardson, C. C.; Ellenberger, T. *Nature* **1998**, *391*, 251-258.
 (25) Beard, W. A.; Wilson, S. H. *Chem. Biol.* **1998**, *5*, R7-13.
 (26) Sawaya, M. R.; Prasad, R.; Wilson, S. H.; Kraut, J.; Pelletier, H. *Biochemistry* **1997**, *36*, 11205-11215.
 (27) Kool, E. T.; Sintim, H. O. *Chem. Commun. (Camb)* **2006**, 3665-3675.
 (28) Mizukami, S.; Kim, T. W.; Helquist, S. A.; Kool, E. T. *Biochemistry* **2006**, *45*, 2772-2778.
 (29) Rechkoblit, O.; Malinina, L.; Cheng, Y.; Kuryavyi, V.; Brody, S.; Geacintov, N. E.; Patel, D. J. *PLoS Biol.* **2006**, *4*, e11.
 (30) Fiala, K. A.; Suo, Z. *Biochemistry* **2004**, *43*, 2116-2125.
 (31) Arndt, J. W.; Gong, W.; Zhong, X.; Showalter, A. K.; Liu, J.; Dunlap, C. A.; Lin, Z.; Paxson, C.; Tsai, M. D.; Chan, M. K. *Biochemistry* **2001**, *40*, 5368-5375.
 (32) Showalter, A. K.; Tsai, M. D. *Biochemistry* **2002**, *41*, 10571-10576.
 (33) Zhang, Y.; Liu, H.; Yang, W. *J. Chem. Phys.* **2000**, *112*, 3483-3492.
 (34) Zhang, Y.; Lee, T.; Yang, W. *J. Chem. Phys.* **1999**, *110*, 46-54.
 (35) Zhang, Y. *J. Chem. Phys.* **2005**, *122*, 024114.
 (36) Hu, P.; Zhang, Y. *J. Am. Chem. Soc.* **2006**, *128*, 1272-1278.
 (37) Corminboeuf, C.; Hu, P.; Tuckerman, M. E.; Zhang, Y. *J. Am. Chem. Soc.* **2006**, *128*, 4530-4531.
 (38) Cheng, Y.; Zhang, Y.; McCammon, J. A. *J. Am. Chem. Soc.* **2005**, *127*, 1553-1562.
 (39) Cisneros, G. A.; Liu, H.; Zhang, Y.; Yang, W. *J. Am. Chem. Soc.* **2003**, *125*, 10384-10393.

- (40) Zhang, Y.; Kua, J.; McCammon, J. A. *J. Am. Chem. Soc.* **2002**, *124*, 10572-10577.
 (41) Liu, H.; Zhang, Y.; Yang, W. *J. Am. Chem. Soc.* **2000**, *122*, 6560-6570.
 (42) Lahiri, S. D.; Zhang, G.; Dunaway-Mariano, D.; Allen, K. N. *Science* **2003**, *299*, 2067-2071.
 (43) Lahiri, S. D.; Zhang, G.; Dunaway-Mariano, D.; Allen, K. N. *Biochemistry* **2002**, *41*, 8351-8359.
 (44) Tremblay, L. W.; Zhang, G.; Dai, J.; Dunaway-Mariano, D.; Allen, K. N. *J. Am. Chem. Soc.* **2005**, *127*, 5298-5299.
 (45) Schweins, T.; Geyer, M.; Scheffzek, K.; Warshel, A.; Kalbitzer, H. R.; Wittinghofer, A. *Nat. Struct. Biol.* **1995**, *2*, 36-44.
 (46) Schweins, T.; Langen, R.; Warshel, A. *Nat. Struct. Biol.* **1994**, *1*, 476-484.
 (47) Pasqualato, S.; Cherfils, J. *Structure* **2005**, *13*, 533-540.
 (48) Du, X.; Black, G. E.; Lecchi, P.; Abramson, F. P.; Sprang, S. R. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 8858-8863.
 (49) Cui, Q.; Karplus, M. *Adv. Protein Chem.* **2003**, *66*, 315-372.
 (50) Warshel, A. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 425-443.
 (51) Friesner, R. A.; Guallar, V. *Annu. Rev. Phys. Chem.* **2005**, *56*, 389-427.
 (52) Zhang, Y. *Theor. Chem. Acc.* **2006**, *116*, 43-50.
 (53) Gao, J.; Truhlar, D. G. *Annu. Rev. Phys. Chem.* **2002**, *53*, 467-505.
 (54) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700-733.
 (55) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1986**, *7*, 718-730.
 (56) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227-249.
 (57) Cheatham, T. E., III; Cieplak, P.; Kollman, P. A. *J. Biomol. Struct. Dyn.* **1999**, *16*, 845-862.
 (58) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926-935.

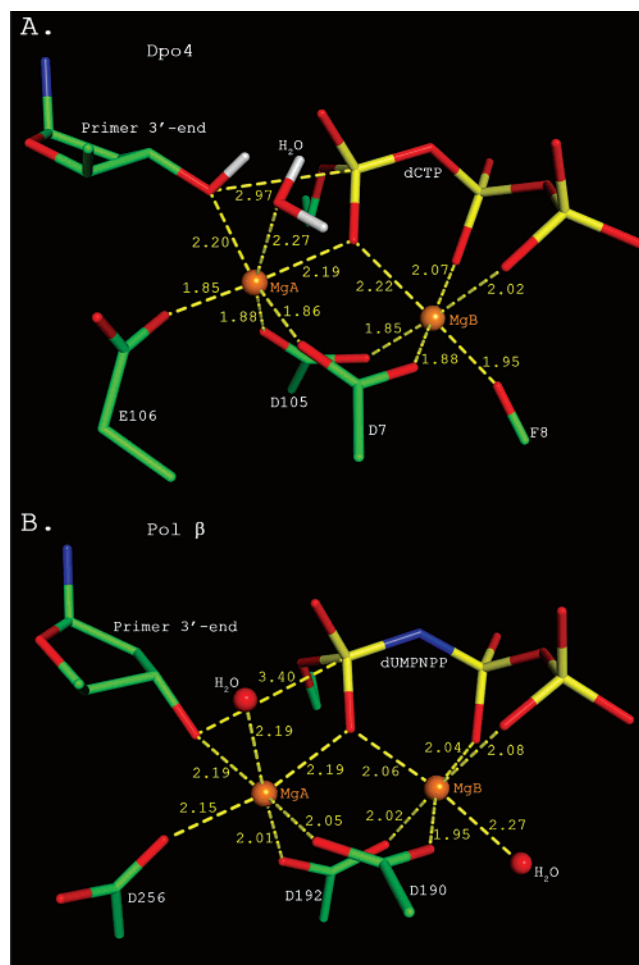


Figure 1. (A) Determined structure of the Dpo4 ternary complex active site based on molecular modeling/dynamics and subsequently *ab initio* QM/MM minimizations. Note that the water was also present in the original crystal structure (PDB ID: 1SOM),¹⁸ which was employed as our initial model. (B) Active site of the Pol β crystal structure (PDB ID: 2FMS).²¹

has been demonstrated to be powerful in the study of several enzymes.^{36–41} Full computational details are given in the Supporting Information.

3. Results

Based on the BP-2 crystal structure of a Dpo4–DNA–dNTP ternary complex,¹⁸ our first task was to employ molecular modeling/dynamics and subsequently *ab initio* QM/MM minimizations to obtain an appropriate enzyme–substrate structure. The determined active site for the lesion-bypass polymerase Dpo4 complexed with the substrate dCTP and the primer 3'-end is depicted in Figure 1A. It is noteworthy that the Mg^{2+} coordination sphere, Mg^{2+} – Mg^{2+} distance, and $\text{P}\alpha$ – $\text{O}3'$ distance in this model structure is very similar to those in a very recently determined X-ray crystal structure of a complete DNA–Pol β catalytic complex²¹ (Figure 1B).

To characterize the nucleotidyl-transfer reaction catalyzed by Dpo4, we have explored several plausible reaction schemes with B3LYP(6-31G*) QM/MM calculations. Our calculations suggested a novel water-mediated and substrate-assisted (WMSA)

mechanism, whose overall potential energy barrier (ΔE) is at least 10 kcal/mol lower than other pathways that we have investigated. A scheme of the WMSA mechanism is depicted in Figure 2 (left panel), and a movie is provided in the Supporting Information. Figures 2 (right panel), S1 (Supporting Information), and 3 provide key structures, energetics, bond distance, and inter- Mg^{2+} distance variations along the path. In this WMSA mechanism, the initial step of the reaction is the transfer of the primer 3'-terminal proton to the α -phosphate via a mediating water molecule present in the crystal structure (PDB ID: 1SOM);¹⁸ next, this proton on the α -phosphate reorients to form a hydrogen bond with a second water molecule and is transferred to the γ -phosphate via this second mediating water. Then the nucleophilic in-line attack of the primer 3'-terminal $\text{O}3'$ on the α -phosphate occurs (I3 in Figure 2), forming a metastable pentacovalent phosphorane intermediate. The nucleophilic attack is accompanied by a synchronized steep decrease in inter- Mg^{2+} distance from ~ 3.62 to 3.52 Å (Figure 3C). This is in line with the hypothesis proposed by Yang et al. that “metal ion A moves toward metal ion B and brings the nucleophile within striking distance for phosphoryl bond formation”.⁶¹ The pyrophosphate finally leaves, facilitated by the transfer of the proton on the γ -phosphate to the β -phosphate through a molecule of water. This mechanism for the phosphate transfer is associative with the formation of a metastable pentacovalent phosphorane intermediate preceding the leaving of pyrophosphate. We note a maximal barrier ($\Delta E \approx 21$ kcal/mol and $\Delta G \approx 13$ kcal/mol) for the initial proton-transfer step, in which the primer 3'-terminal proton is transferred to the α -phosphate via a water. Thus, the general base in the WMSA mechanism is the α -phosphate of the substrate dCTP. Mg^{2+} coordination remains intact except for the liganding to the water molecule. This liganding distance goes to ~ 3.5 Å to permit proton transfer from the primer 3'-terminus to the α -phosphate. For this characterized WMSA mechanism, the calculated energetic results indicate that the making or breaking of the P–O bond is not a rate-limiting step for Dpo4, which is consistent with the available experimental results.³⁰

We explored, in preliminary work, several other mechanisms in an effort to consider other plausible pathways; these include: (i) The conserved E106 is the general base, analogous to the mechanism recently proposed for Pol β ⁵; (ii) the α -phosphate of the substrate dCTP serves as the general base without water-mediated proton transfer, suggested earlier for Pol β ¹⁰ and (iii) the conserved D105 serves as the general base mechanism proposed for T7.^{6,9} Our calculations indicate that each of them had an overall lowest potential energy barrier greater than 30 kcal/mol. Details are given in Figures S2–S4 (Supporting Information).

Of particular interest is the mechanism involving proton transfer to a conserved carboxylate, E106 in Dpo4, followed by nucleophilic attack of the $\text{O}3'$ on the α -phosphate and by pyrophosphate leaving, as for Pol β .⁵ This mechanism remains energetically unfavorable for Dpo4 (Figure S2, Supporting Information), both with an initially γ -protonated dCTP, as in Pol β ,⁵ and an initially unprotonated one, as in our studies. While the potential energy barrier for proton transfer is ~ 26 and ~ 12 kcal/mol for the unprotonated and protonated case, respectively (Figure S2, Supporting Information), the energy continuously

(59) Frisch, M. J. et al. *Gaussian 03*, Revision B.05, 2003; Gaussian, Inc.: Wallingford, CT, 2004.

(60) Ponder, J. W. *TINKER*, 4.2; Washington University in St. Louis: St. Louis, MO, 2004.

(61) Yang, W.; Lee, J. Y.; Nowotny, M. *Mol. Cell* **2006**, *22*, 5–13.

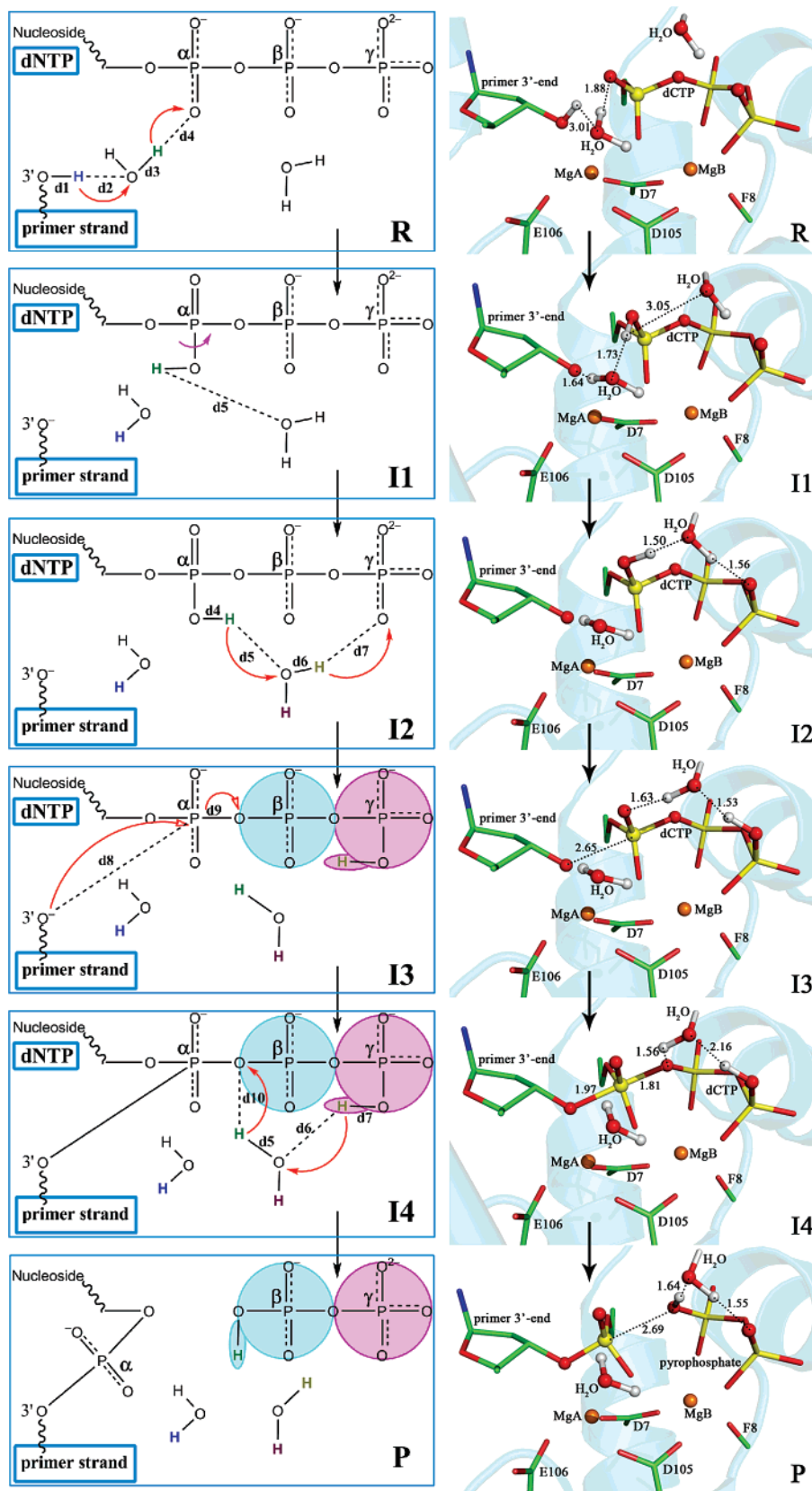


Figure 2. Reaction mechanism (left panel) and critical structures (right panel) determined for nucleotidyl transfer catalyzed by Dpo4. R, reactant (ground state); I1, 3'-terminal proton relayed to α -phosphate via a water; I2, proton on α -phosphate reoriented toward γ -phosphate; I3, proton on α -phosphate relayed to γ -phosphate via a second water; I4, metastable pentacovalent phosphorane intermediate; P, product. Red arrows with solid heads denote proton transfer; red arrows with hollow heads denote electron rearrangement; magenta arrow denotes torsional rotation. $d1$ – $d10$ denote distances between atoms connected by dashed lines or covalent bonds, and their variation along the path is shown in Figure 3B. Shaded areas denote the atoms included to compute group charges of the β - and γ -phosphates (see legend of Figure 4 for details).

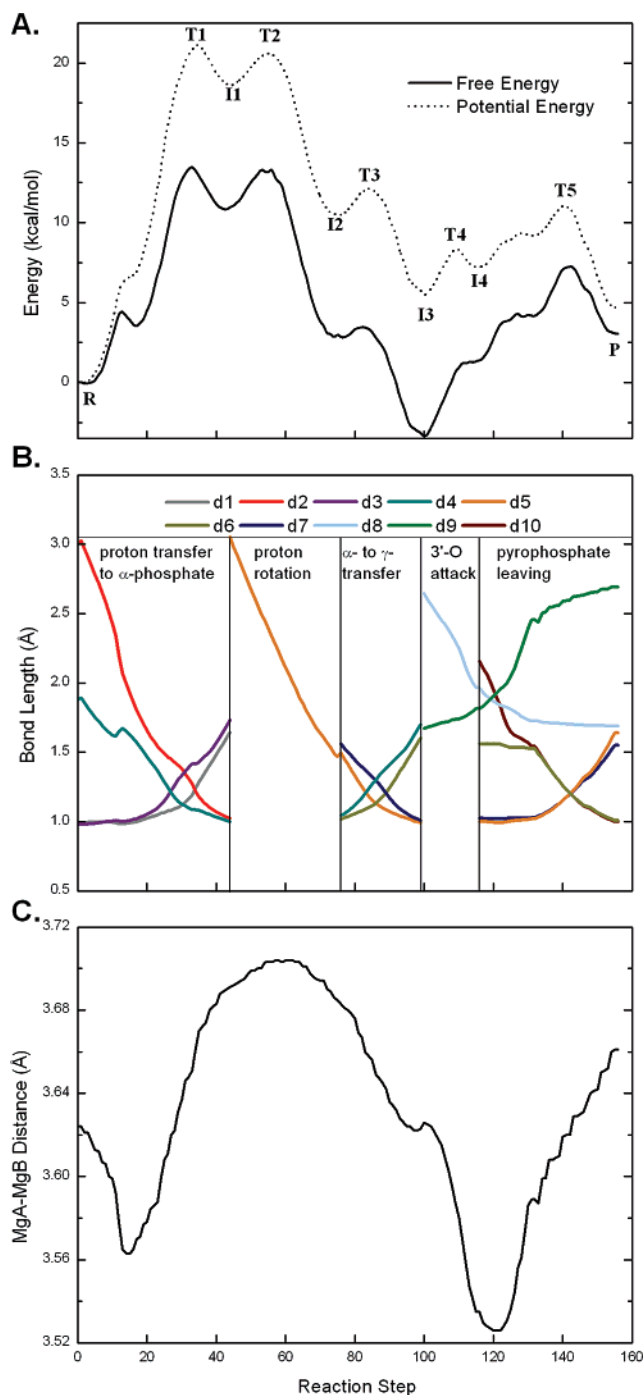


Figure 3. Minimum energy path and free energy reaction profiles (A), critical bond length changes (B), and inter-Mg²⁺ distance changes (C) as a function of reaction step. See Figure 2 for definitions of *d1*–*d10*.

escalates during the complete course of nucleotidyl transfer. The possibility of enzyme-dependent mechanistic differences is thus suggested.

With the determined intermediate structure I3 as shown in Figure 2, we have also explored a direct nucleotidyl-transfer pathway where the proton remains on the γ -phosphate throughout the nucleotidyl transfer. The resulting reaction energy curve is always uphill (Figure S5, Supporting Information). This is consistent with the energy curve in Figure S2 and indicates that the water-mediated proton transfer from the γ - to β -phosphate is essential to facilitate the completion of the nucleotidyl-transfer

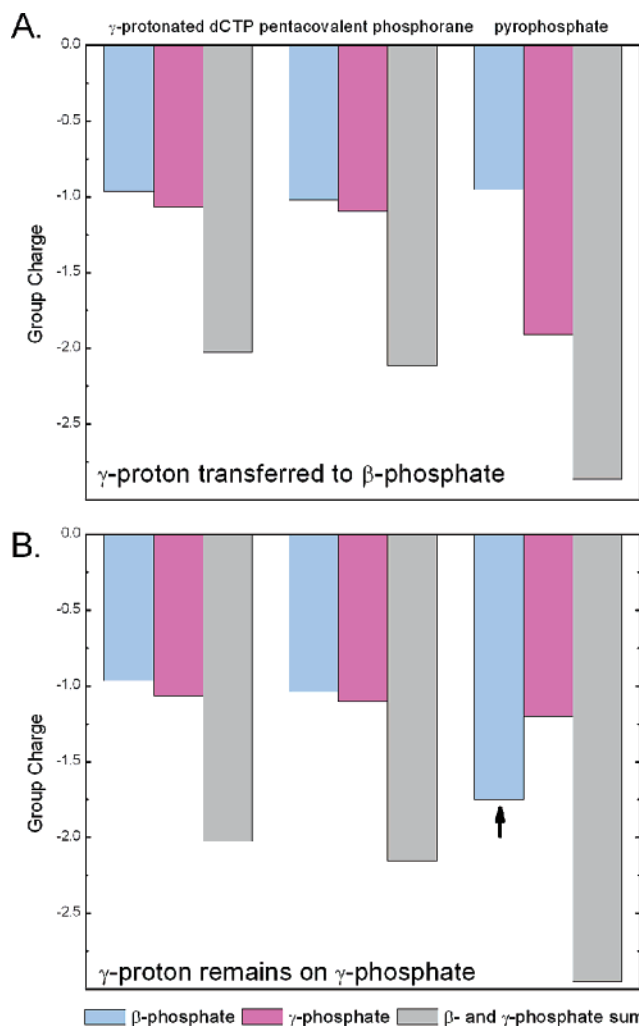


Figure 4. Group charges of the β -phosphate (cyan) and γ -phosphate (magenta) and their sums (gray) in γ -protonated dCTP, pentacovalent phosphorane intermediate, and the leaving pyrophosphate. See shaded areas of Figure 2 (left panel) for atoms included in computing the group charges. The partial charge of the bridging oxygen atoms is considered to be evenly shared between the neighboring phosphate groups. (A) Nucleotidyl-transfer mechanism involving γ - to β -phosphate; (B) nucleotidyl-transfer mechanism with the proton attached to the γ -phosphate throughout the reaction. The arrow highlights the greater accumulation of negative charge on the β -phosphate in B as pyrophosphate leaves.

step in Dpo4. In order to elucidate its origin and the dynamics of the charge evolution during this nucleotidyl-transfer step, we computed the group charges for the β - and γ -phosphate at three stages in the path (Figure 4A and Table S1, Supporting Information): (1) prior to 3'-O attack on the α -phosphate (I3 in Figure 2), (2) upon formation of the pentacovalent phosphorane intermediate (I4 in Figure 2), and (3) upon completion of pyrophosphate leaving (P in Figure 2). This can be directly compared with the direct nucleotidyl-transfer pathway (Figure S5, Supporting Information). As shown in Figure 4B, negative charge accumulates much more significantly on the bridging O3 α during the pyrophosphate-leaving step, in spite of the proton on the γ -phosphate. We were not able to obtain a low-energy product with this mechanism, suggesting that protonation of the bridging oxygen between the α - and β -phosphate is critical in effectively neutralizing the accumulating negative charge and facilitating the pyrophosphate leaving. It should be noted that the sum of the charges on the leaving pyrophosphate

itself is about the same in these two schemes (Figure 4 and Table S1, Supporting Information); hence, the dynamic redistribution of negative charge and its neutralization through the water-mediated proton relay determines whether the reaction is energetically feasible in Dpo4.

4. Discussion

Our characterized WMSA mechanism requires water molecules to mediate proton relay to the γ -phosphate via the α -phosphate and then to the β -phosphate of the dCTP as the reaction proceeds. A crystal water (Figure 1A) is critically situated to transfer the 3'-OH to the α -phosphate, and a second solvent water acts as a relay for protonation on the γ -phosphate which has the most negative partial charge among the three phosphate groups of dCTP.⁶² However, the presence of the crystal water in the catalytic Mg^{2+} ion coordination sphere, which is conserved in many polymerase crystal structures at the same site (e.g., Pol β ²¹ and T7²⁴), as well as that of the solvent water in the vicinity of the leaving group, was not a factor in our selection of starting model, since we did not know the mechanism when we began the work. The subsequent proton transfer to the β -phosphate, accompanied by the pyrophosphate leaving, neutralizes the accumulating negative charge on the β -phosphate (Figure 4A). Thus, this relay process dynamically balances the evolving negative charge on the dCTP as the reaction proceeds; it takes specific advantage of the prevalence of water in the highly solvent-exposed active site of Dpo4. Previous theoretical studies in Pol β ⁵ and T7^{6,9} DNA polymerases suggest that these enzymes, with less solvent-exposed active sites, can utilize other mechanisms. The possibility of enzyme-dependent mechanistic differences for the phosphoryl-transfer reactions in GTP-binding proteins has recently been discussed.⁶³ In our suggested WMSA mechanism for Dpo4, the initial transfer of the primer 3'-terminal proton is to a crystal water¹⁸ and leads then to an intermediate with the α -phosphate protonated; however, the subsequent proton relay to the γ - and then the β -phosphate could involve transfers through the bulk water. Possible tunneling⁶⁴ also remains a consideration.

Other examined paths including ones without water-mediated proton transfer were of much higher energy in Dpo4, including the recent mechanisms described for Pol β ⁵ and T7.^{6,9} This is true whether the dCTP is originally unprotonated with a charge of -4 or protonated with a charge of -3 as employed for Pol β ⁵ (Figure S2, Supporting Information). An interesting possibility suggested from this mechanism, however, could involve initial proton transfer to E106, followed by uptake of this proton by solvent water and further water-mediated transfers to γ - and then β -phosphate, as in our mechanism. This suggestion arises because the proton transfer to E106 has a potential energy barrier of ~ 12 kcal/mol (Figure S2B, Supporting Information), although the overall potential energy for this mechanism continues to escalate if the water-mediated proton transfer is not involved.

The kinetic data of Fiala and Suo³⁰ revealed the nucleotidyl-transfer step is not rate limiting for incorporation of the correct dNTP opposite a normal template base in Dpo4, as found in the present work. The k_{cat} value obtained by Fiala and Suo⁶⁵

corresponds to a ΔG_{act}^0 of 16.4 kcal/mol, based on simple transition state theory:

$$k(T) = (k_B T/h)(c^0)^{1-n} \exp[-\Delta G_{act}^0(T)/RT]$$

According to Fiala and Suo,⁶⁵ the rate-limiting step is conformational change in Dpo4. However, this value does provide an upper limit to the chemical step, which is more rapid according to their work. Thus, this upper limit may be compared with our computed free energy maximal barrier of ~ 13 kcal/mol (Figure 3A), and they are in quite good agreement, considering that there are unavoidable approximations in our employed methods.

Our calculations located a metastable pentacovalent phosphorane intermediate (I4 in Figure 2) as noted both experimentally^{42–44} and theoretically.^{6,9,66,67} Our results for Dpo4 are consistent with the free energy perturbation/empirical valence bond study on the T7 DNA polymerase by Florian et al.,⁹ who also obtained the phosphorane intermediate and found that departure of the pyrophosphate is energetically more costly than the nucleophilic attack.

In order to understand the catalytic roles of surrounding amino acid residues, we have analyzed the contribution of each enzyme residue toward stabilizing the transition states (Figure S6, Supporting Information). For the initial proton transfer from the primer 3'-terminus to the α -phosphate, the transition state TS1 is stabilized by D7 and D105. For the nucleophilic attack, the transition state TS4 is stabilized by R51 and E106. For the pyrophosphate leaving, the transition state TS5 is stabilized by D7, R51, D105, E106, and K159. These results are consistent with experimental studies in Dpo4²² that D105A or E106A mutation leads to enzyme inactivation. Furthermore, the results suggest that it may be worthwhile to carry out mutagenesis experiments involving residues D7, R51, and K159. The conserved carboxylates (D7, D105, and E106 in Dpo4) play the critical role of maintaining active site organization and stabilizing transition states, through their binding with the two Mg^{2+} ions.

In spite of our significant efforts, the number of reaction pathways that have been examined nonetheless remains quite limited. Consequently, our present results are suggestive in nature, in that the characterized WMSA pathway is, at this stage, only one possible mechanism for Dpo4. Further experimental and theoretical investigations for Dpo4 as well as other DNA polymerases are clearly needed.

5. Conclusion

Based on extensive ab initio QM/MM explorations of several possible mechanisms, we have suggested a novel water-mediated and substrate-assisted (WMSA) mechanism for the nucleotidyl-transfer reaction in Dpo4. This mechanism utilizes the substrate itself as the general base for the primer 3'-terminal proton via a bridging crystal water, and employs water-mediated proton transfer to facilitate pyrophosphate leaving by neutralizing the evolving negative charge. A metastable pentacovalent phosphorane intermediate was located which is stabilized by conserved amino acid residues of the Dpo4 active site. The

(62) Perlow, R. A.; Broyde, S. J. *Mol. Biol.* **2002**, *322*, 291–309.

(63) Wittinghofer, A. *Trends Biochem. Sci.* **2006**, *31*, 20–23.

(64) Benkovic, S. J.; Hammes-Schiffer, S. *Science* **2006**, *312*, 208–209.

(65) Fiala, K. A.; Suo, Z. *Biochemistry* **2004**, *43*, 2106–2115.

(66) Uchamaru, T.; Tanabe, K.; Nishikawa, S.; Taira, K. *J. Am. Chem. Soc.* **1991**, *113*, 4351–4353.

(67) Dejaegere, A.; Lim, C.; Karplus, M. *J. Am. Chem. Soc.* **1991**, *113*, 4353–4355.

highly solvent-exposed active site of the lesion-bypass Y-family polymerase Dpo4 may be especially suited for such a mechanism taking specific advantage of the prevalence of the water.

Acknowledgment. This work is supported by NIH Grants CA28038 and CA75449 to SB. Y.Z. is grateful for the support from NSF (CHE-CAREER-0448156), NYSTAR (James D. Watson Young Investigator Award), and NYU (Whitehead Fellowship). Molecular images were made with PyMOL (DeL-

ano Scientific, LLC.). We thank Dr. Wei Yang for her very insightful discussions and suggestions.

Supporting Information Available: Complete ref 59; computational details, supplemental Table S1 and Figures S1–S6; MPG movie file for the characterized WMSA mechanism. This material is available free of charge via the Internet at <http://pubs.acs.org>.

JA068821C