Published in final edited form as:

J Mol Biol. 2009 June 19; 389(4): 787–796. doi:10.1016/j.jmb.2009.04.029.

Polymerase-tailored variations in the Water-Mediated and Substrate-Assisted Mechanism for Nucleotidyl Transfer: Insights from a Study of T7 DNA Polymerase

Lihua Wang¹, Suse Broyde^{1,*}, and Yingkai Zhang^{2,*}

¹ Biology Department, New York University, New York, NY 10003, USA

² Chemistry Department, New York University, New York, NY 10003, USA

Abstract

The nucleotidyl transfer reaction catalyzed by DNA polymerases is the critical step governing the accurate transfer of genetic information during DNA replication and its malfunctioning can cause mutations leading to human diseases including cancer. Here utilizing *ab initio* quantum mechanical/molecular mechanical calculations with free energy perturbation, we have carried out an extensive investigation of the nucleotidyl transfer reaction mechanism in the well-characterized high-fidelity replicative DNA polymerase from phage T7. Our defined mechanism entails an initial concerted deprotonation of a conserved crystal water molecule with protonation of the γ -phosphate of the dNTP via a solvent water, and then the proton on the primer 3'-terminus is transferred to the resulting hydroxide ion. Subsequently the nucleophilic attack takes place, with the formation of a metastable pentacovalent phosphorane intermediate. Finally the pyrophosphate leaves, facilitated by the relay of the proton on the γ -phosphate to the α - β bridging oxygen, via solvent water. The computed activation free energy barrier is consistent with kinetic data for the chemistry step with correct nucleotide incorporation in T7 DNA polymerase. This variant of the water-mediated and substrate-assisted (WMSA) mechanism has features tailored to the structure of the T7 DNA polymerase. However, a unifying theme in the WMSA mechanism is the cycling through crystal and solvent waters of the proton originating from the primer 3'-terminus to the α - β bridging oxygen of the dNTP; this neutralizes the evolving negative charge as pyrophosphate leaves and restores the polymerase to its pre-chemistry state. These unifying features are likely requisite elements for nucleotidyl transfer reactions.

Introduction

DNA polymerases are vital enzymes in various key cellular functions. These include not only the critical activity of genome replication per se, but also repair of gaps, bypass of DNA lesions, as well as roles in homologous recombination and maintenance of immune system diversity, among others ¹. Numerous DNA polymerases are now known to participate in these various essential functions ¹, and human diseases result when polymerase performance is flawed ². They manifest a broad range of fidelities in DNA synthesis ^{3; 4} with the polymerases engaged

* Corresponding authors. broyde@nyu.edu, yingkai.zhang@nyu.edu.

Supplementary Material:

Supplementary material associated with this article can be found, in the online version.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

in processively synthesizing long DNA stretches generally having the highest fidelities while polymerases responsible for synthesizing only very short DNA stretches being generally of low fidelity. In recent years, the structural properties of the various polymerases that determine their intrinsic fidelity have evoked great interest and significant numbers of crystal structures containing primer-template DNA and incoming dNTPs have provided insights⁵. Polymerases share an architecture that has been likened to a hand with palm, fingers, and thumb domains and active sites at the intersection of the palm and fingers domains⁶. Here the nucleotidyl transfer reaction takes place: the dNTP that is the Watson-Crick partner to the template base is chemically added to the 3'-OH terminus of the primer strand. It is now understood that the reaction mechanism involves two metal ions, usually Mg^{2+} ^{7; 8}. The catalytic MgA and the nucleotide binding MgB are coordinated with the dNTP, three conserved amino acid residues (usually Asp or Glu), and the 3'-OH terminus of the primer strand. The remainder of the octahedral coordination spheres of the Mg^{2+} are satisfied by liganding to water molecules or other amino acid residues. Of note, polymerase crystal structures from different families in which MgA is present, generally contain a crystal water molecule whose position in the MgA coordination sphere is conserved. Overall this two-metal ion mechanism involves the nucleophilic attack of the 3'-O on the α -phosphate of dNTP with phosphorane intermediate and pyrophosphate leaving, resulting in growth of the primer strand by one nucleotide.

While these features appear to be common to all polymerases, structural studies have revealed that low-fidelity lesion-bypass polymerases of the Y-family differ from high-fidelity polymerases in their structures as they proceed through a replication cycle. In the higher fidelity polymerases, the entry of a dNTP that is the Watson-Crick partner to the template causes the fingers to close on the nascent base pair via an induced-fit mechanism to form a tightly fitting active site that promotes fidelity in the chemical reaction^{9; 10}. However, the low fidelity bypass polymerases do not utilize an open/closing induced-fit mechanism upon dNTP entry¹¹. They have a preformed and spacious active site whose function is to accommodate and transit DNA lesions, and hence they manifest low fidelity on undamaged DNA.

With the recent development of hybrid quantum mechanical and molecular mechanical (QM/MM) methods for elucidating enzyme mechanisms, DNA polymerases have become the focus of considerable interest^{12; 13; 14; 15; 16; 17; 18; 19}, and the possibility that different polymerases utilize mechanisms that differ in detail has arisen. For the low-fidelity Y-family lesion bypass polymerase Dpo4, a water-mediated and substrate-assisted (WMSA) mechanism¹⁶ has been delineated in which the 3'-OH terminus of the primer strand deprotonates by transferring the proton to the α -phosphate via a crystal water, and then to the γ -phosphate through a solvent water molecule; this rate-limiting step is followed by the attack of 3'-O on the α -phosphate, producing a metastable phosphorane intermediate, and then the leaving of pyrophosphate. The essential features of this mechanism have been confirmed in Dpo4 with a 7,8-dihydro-8-oxo-guanine lesion as the templating base¹⁹. It was suggested that the open and solvent-accessible active site region of the Dpo4 bypass polymerase provided a particularly suitable environment utilizing a water-mediated mechanism for chemistry¹⁶. To further explore this idea, we here investigated this mechanism in the high-fidelity replicative polymerase of T7 phage, as well as a number of other mechanisms that had been proposed for this and other polymerases^{12; 13; 14}. Utilizing the pseudobond QM/MM approach with free energy perturbation^{20; 21; 22}, we find that a variant of the WMSA mechanism is also feasible for this case, providing the lowest free energy path among the half dozen we explored.

Results

Based on the crystal structure of a T7 DNA polymerase-DNA-dNTP ternary complex²³, we obtained an enzyme-substrate structure employing molecular modeling/dynamics and subsequently *ab initio* QM/MM minimizations. This structure was selected because it has the

fewest unresolved residues and contains a nearly perfect coordination sphere for the two divalent metal ions required for the nucleotidyl transfer reaction. Details are provided in Supplementary Material. The determined active site for the high-fidelity replicative T7 DNA polymerase complexed with the substrate dCTP and the primer 3'-end is depicted in Figure 1. We note that the Mg^{2+} coordination spheres, Mg^{2+} - Mg^{2+} distance, and $P\alpha$ -O3' distance in this model structure are very similar to those in the X-ray crystal structure of a high resolution Pol β catalytic complex containing a primer terminal 3'-OH²⁴ and to the starting structure we utilized for Dpo4 in our previous study¹⁶. These features provide a near reaction-ready enzyme-substrate structure according to current understanding^{7; 8; 25; 11}.

To characterize the nucleotidyl transfer reaction catalyzed by T7 DNA polymerase, we have explored several plausible reaction schemes with B3LYP(6-31G*) QM/MM calculations. Our calculations suggested that a variant of the WMSA mechanism, as previously proposed for the bypass DNA polymerase Dpo4¹⁶, is feasible for the high-fidelity replicative T7 DNA polymerase. The overall potential energy barrier (ΔE) of this mechanism is at least 10 kcal/mol lower than other pathways that we have investigated. The scheme of the WMSA mechanism is depicted in Figure 2 (left panel), and a movie is provided in the Supplementary Material. Figures 2 (right panel), S1 (Supplementary Material) and 3 provide key structures, energetics, bond distances, and inter- Mg^{2+} distance variations along the path. In this variant of the WMSA mechanism, the initial step of the reaction is the protonation of the γ -phosphate of the incoming dCTP; this is achieved through a concerted proton relay from the conserved water molecule via a second, solvent water, resulting in the formation of a hydroxide ion in the coordination sphere of the catalytic Mg^{2+} ion (I1 in Figure 2). Subsequently, the primer 3'-terminal proton is transferred to the hydroxide ion, producing a negatively charged primer 3'-end (I2 in Figure 2). Then the nucleophilic in-line attack by the primer 3'-terminal O3' on the α -phosphate occurs, forming the leaving pyrophosphate (I3 in Figure 2). A steep decrease in inter- Mg^{2+} distance is observed in the course of nucleophilic attack (Figure 3C), with the minimum inter- Mg^{2+} distance roughly corresponding to a metastable pentacovalent phosphorane intermediate, as suggested by the free energy profile of this reaction path (Figure 3A). This is in agreement 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"²⁵. The nucleophilic attack is accompanied by realignment of the solvent water molecule toward the α - β bridging oxygen of the dCTP. The pyrophosphate finally leaves, facilitated by the transfer of the proton on the γ -phosphate to the β -phosphate through the molecule of solvent water. This mechanism for nucleotidyl transfer is associative with the formation of a metastable pentacovalent phosphorane intermediate preceding the leaving of pyrophosphate. We note a free energy barrier of ~ 14 kcal/mol for the initial proton transfer step, in which the conserved water molecule deprotonates and the γ -phosphate of the dCTP accepts the proton. The primer 3'-end subsequently deprotonates by transferring the proton to the hydroxide ion. Thus the ultimate general base and the general acid in the WMSA mechanism are, respectively, the γ -phosphate of the substrate dCTP and the primer terminal 3'-OH. Mg^{2+} coordination remains intact except for the liganding to the conserved water molecule. This liganding distance goes to ~ 3.4 Å to permit proton transfer from the primer 3'-terminus to the hydroxide.

In order to understand the catalytic roles of surrounding amino acid residues, we also analyzed the contribution of enzyme residues toward stabilizing the transition states (Figure S2, Supplementary Material). For the initial protonation of the γ -phosphate and deprotonation of the primer terminus 3'-OH, the transition states TS1 and TS2 are stabilized by D504, R429, D475 and E655. For the nucleophilic attack, the transition state TS3 is stabilized by R518, K522 and E655. For the pyrophosphate leaving, the transition state TS4 is stabilized by R518, K522 and D654. The initial protonation of the γ -phosphate breaks the hydrogen bonding network between the γ -phosphate and R518; the orphaned R518 is, however, stabilized by the

neighboring D504. The conserved carboxylates D475, D654 and E655 contribute to stabilizing the transition states: they help anchor the two divalent ions in place during catalysis. Most notably, the positively charged R518 and K522 near the dCTP triphosphate tail significantly stabilize the transition states, effectively contributing to neutralizing the locally accumulating negative charge as the P-O bond breaks and the pyrophosphate leaves. Additionally, R429, a residue that is part of the minor groove scanning track where the polymerase tightly binds the nascent duplex^{26; 27; 28; 29; 30}, stabilizes the primer terminal 3'-OH upon its deprotonation.

We investigated extensively several other mechanisms in an effort to consider other plausible pathways. First, we considered the WMSA mechanism delineated for Dpo4 where the α -phosphate of the substrate dCTP serves as the general base with water-mediated proton transfer¹⁶. In this mechanism, the proton on the primer 3'-end is first transferred to the pendant oxygen on the α -phosphate of dCTP, via the crystal water molecule (Figure S3, Supplementary Material). However, a stable intermediate of the α -protonated dCTP could not be obtained: the proton hops back to the 3'-O via the crystal water when the constraint on the reaction coordinate is removed.

We then explored the mechanism proposed for T7 in which the conserved D654 serves as the general base^{12; 13}. Our calculations indicate that the potential energy barrier for the proton transfer to D654 is ~27.5 kcal/mol. This could be attributed to the fact that D654 is involved in the coordination sphere of both Mg^{2+} ions and thus is not a good candidate for the general base. Additionally, the potential energy profile for the subsequent nucleophilic attack and pyrophosphate leaving steps is always uphill, suggesting that protonation of the leaving group may constitute a critical component in the nucleotidyl transfer reaction. Details are given in Figure S4 (Supplementary Material).

Another mechanism we investigated for T7 considered the possibility that E655 acts as the proton acceptor, as in the mechanism for Pol β by Lin et al.¹⁴ where D256 acts in this role. In all crystal structures of the T7 DNA polymerase ternary complex, E655 is not coordinated to MgA and its place is taken by a second crystal water, WAT2 (Figure 1). This is in contrast to Pol β which contains a MgA-coordinated Asp in this location, and in our Dpo4 initial model which has E106 similarly coordinated. It is possible that the absence of the 3'-OH in the crystal structure plays a role in the failure of E655 to engage in the MgA coordination sphere²⁵. Accordingly, in order to investigate E655 as proton acceptor, we remodeled the active site, rotating the E655 side chain torsions N-C α -C β -C γ , C α -C β -C γ -C δ , and C β -C γ -C δ -O ϵ 2 (to 73.3°, -138.3°, and 15.9°, respectively) and removed the WAT2 so that E655 is coordinated to MgA. In this case, the coordination of E655 to MgA was broken during the equilibration for MD, consistent with its instability.

We further investigated E655 as proton acceptor with a mechanism where the O3'-H is transferred directly to the second crystal water (WAT2), and in concert a water proton is transferred to E655. Here the E655, organized as in the crystal, is not coordinated with MgA and thus is a better candidate proton acceptor. While the potential energy barrier for proton transfer is ~12 kcal/mol (Figure S5, Supplementary Material), the energy continuously escalates during the complete course of nucleotidyl transfer if the nucleophilic attack step is forcibly driven through the reaction coordinate.

Yet an additional mechanism involving E655 as proton acceptor was explored. Without releasing the constraint on protonated Glu and following protonation of E655 via the second crystal water, the reaction proceeds with K522 (Figure S6, Supplementary Material) donating a proton to the α - β bridging oxygen of dCTP, neutralizing the evolving charge as the attack of O3' on Pa takes place and pyrophosphate leaves. In this case, also, the energy goes continuously uphill.

With the determined intermediate structure I2 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. As for the case of Dpo4, the resulting reaction energy curve is always uphill (Figure S7, Supplementary Material), suggesting that the water-mediated proton transfer from the γ - to the β -phosphate is essential for completion of the nucleotidyl transfer step in T7 DNA polymerase. 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, Supplementary Material): prior to 3'-O attack on the α -phosphate (I2 in Figure 3), upon formation of the metastable pentacovalent phosphorane intermediate (PPI in Figure 3), and upon completion of pyrophosphate leaving (P in Figure 3). This can be compared with the direct nucleotidyl transfer pathway (Figure S7, Supplementary Material). 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. A low energy product could not be obtained 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, Supplementary Material); hence the dynamic redistribution of negative charge and its neutralization through the water-mediated proton relay determines whether or not the reaction is energetically feasible in T7 DNA polymerase. This feature is similar in our previously investigated WMSA mechanism for Dpo4¹⁶.

Discussion

In the present work, we have investigated the molecular mechanism of the nucleotidyl transfer reaction in the high fidelity replicative T7 DNA polymerase using QM/MM calculations with the pseudobond approach. We utilized the 2.54 Å resolution crystal structure (PDB ID: 1T8E) of this enzyme in a ternary complex with primer-template DNA and incoming dNTP as the basis for this study²³, as it has the fewest unresolved residues and contains a nearly perfect coordination sphere for the two divalent metal ions required for the nucleotidyl transfer reaction. We remodeled the crystal structure by adding a 3'-OH to the primer terminus and incoming dCTP which were dideoxy to prevent reaction, and ran MD simulations to obtain a starting model for the subsequent QM/MM study. To test various mechanisms, in some cases we also remodeled certain amino acid residue side chains in the active site. Our goal was to evaluate the feasibility of the WMSA mechanism for this high fidelity polymerase that we had previously delineated for the low-fidelity lesion-bypass polymerase Dpo4. Among the six mechanisms considered, a variant of the WMSA mechanism was energetically most favorable.

In this mechanism, the initial step is deprotonation of the conserved crystal water molecule and the concerted protonation of the γ -phosphate via a solvent water, producing a hydroxide ion. Then the 3'-OH of the primer terminus is deprotonated with the proton transferred to the hydroxide ion. Subsequently, the 3'-O attacks the P α , producing a metastable pentacovalent phosphorane intermediate; finally the pyrophosphate leaves and the proton on the γ -phosphate is transferred to the α - β bridging oxygen via a solvent water molecule. While this mechanism differs from the WMSA mechanism delineated for Dpo4 in that deprotonation of the crystal water precedes deprotonation of the primer terminus 3'-OH, the final results of the proton transfers converge to production of a deprotonated primer terminus and a γ -protonated dNTP. A comparison of the crystal structures of Dpo4 and T7 DNA polymerases utilized in our computations provides some possible insights on the reasons for this difference (Figure S8, Supplementary Material): the crystal water WAT1 is positioned between the pendant oxygen and the 3'-OH in T7 while it is in a triangular organization with the pendant oxygen and the 3'-OH in Dpo4. In T7, the direct proton relay from O3' to α -phosphate through the crystal water

may be unfavorable because of unfavorable angle involving the lone pair oxygen electrons. However, the variant of the WMSA mechanism delineated above is simply an alternative path that produces a deprotonated primer 3'-end and a γ -protonated dCTP, suggesting that the chain of proton relays may be polymerase-dependant, and may involve one or more crystal and solvent water molecules.

Tsai and Johnson³¹ have determined the rate of the chemistry step for correct nucleotide incorporation by the T7 DNA polymerase to be from 234 s^{-1} to 360 s^{-1} , which corresponds to an activation free energy barrier of $14.1\sim 14.3\text{ kcal/mol}$ based on simple transition state theory. The perfect agreement between our computed overall free energy barrier of 14.2 kcal/mol and the experimental value is fortuitous considering that there are unavoidable approximations in our employed QM method/basis set, MM force field, QM/MM boundary and free energy calculations, but it does suggest that the characterized WMSA mechanism is feasible for T7 DNA polymerase. As shown in Fig. 3A, there is a significant difference between the potential energy curve and the free energy profile, which shows that taking account of protein dynamics in simulating enzyme reactions is important. It should be noted that the free energy profile presented in Fig. 3A does not include the fluctuation of the QM subsystem. Ideally, Born-Oppenheimer *ab initio* QM/MM molecular dynamics simulations should be carried out to take account of dynamics of QM and MM sub-systems on an equal footing; this treatment is currently in progress in our group^{32; 33; 34}. Our mechanism for T7 DNA polymerase is also consistent with experimental studies utilizing D_2O with T7 RNA polymerase and RB69 DNA polymerase which observed two proton transfers in the overall chemical reaction, one of which was suggested to correspond to the deprotonation of the 3'-OH and the other to protonation of the leaving group³⁵.

Meanwhile, we did explore the previously proposed mechanism for T7 in which the conserved D654 serves as the general base^{12; 13} and found that the overall lowest potential energy barrier is greater than 30 kcal/mol (Figure S4, Supplementary Material). This could be attributed to the fact that D654 is involved in the coordination sphere of both Mg^{2+} ions and thus is likely not a good candidate for the general base when thus liganded. We also investigated the crystal structure for other plausible proton acceptor residues. While none were directly feasible in the crystal structure per se, we noted that an indirect proton transfer to E655 via the second crystal water molecule (WAT2), also coordinated to MgA , might be possible, with a potential energy barrier of $\sim 12\text{ kcal/mol}$. Investigating this mechanism we found that a stable intermediate could not be located for the proton transfer step, and the energy continually escalated during the subsequent nucleotidyl transfer step. However, an interesting possibility is suggested from this mechanism: initial proton transfer to E655 through WAT2 could be followed by or occur in concert with uptake of this proton by the γ -phosphate through solvent water molecules, with nucleotidyl transfer then ensuing. This latter part is similar to the WMSA mechanisms presented for T7 in the current work and for Dpo4 in our earlier study¹⁶. A further alternative might allow transfer of the proton on E655 through water to the α - β bridging oxygen, without an intermediate residency on the γ -phosphate. Further work is needed to elucidate these possibilities.

A main feature of the WMSA mechanism¹⁶ is the ultimate transfer of the proton on the γ -phosphate to the α - β bridging oxygen as the pyrophosphate leaves, in order to neutralize the evolving negative charge. This theoretically suggested mechanism was published almost simultaneously with an experimentally characterized two-proton transfer mechanism³⁶, and both are very consistent with each other in emphasizing the importance of both the deprotonation of the 3'-OH nucleophile and the protonation of the pyrophosphate leaving group for the polymerase-catalyzed nucleotidyl transfer. In the two-proton transfer mechanism³⁶, it has been inferred that an active site amino acid residue serves as a general acid to protonate the pyrophosphate leaving group for several polymerases. However, mutation studies indicate

that the assumed general acid is not absolutely essential for the enzyme activity³⁶. For T7 RNA polymerase, whose active site is essentially identical to the T7 DNA polymerase which is studied here, the mutation of the K631 to Leu leads to a decrease in the rate of ~100 fold³⁶ (corresponding to an increase in the activation free energy barrier of ~2.7 kcal/mol).

Examination of the T7 DNA polymerase crystal structure did suggest the possibility for the K522, which is structurally equivalent to the K631 in T7 RNA polymerase, to serve as a general acid to protonate the leaving pyrophosphate. However, the energy was always uphill and a stable intermediate could not be located (Figure S6, Supplementary Material). Meanwhile, in the characterized WMSA mechanism, although K522 does not directly participate in the reaction, it plays a very important role in stabilizing TS3 and TS4 through electrostatic interactions (Figure S2, Supplementary Material). Thus to some degree, our WMSA mechanism, in which the essential pyrophosphate-leaving protonation stems from a relay (Figure 2) that does not necessitate involvement of an active site amino acid residue, would also be able to rationalize the recent experimental results^{35; 37}. Interestingly, the spacious active site of Dpo4 (Figure S8, Supplementary Material) does not contain an amino acid residue positioned comparably to K522 in the T7 DNA polymerase or donors proposed for other polymerases investigated³⁷. Clearly, protonation of leaving pyrophosphate is required to neutralize the evolving negative charge as demonstrated in the WMSA mechanism delineated for Dpo4¹⁶ and in the current study as well as in recent experimental studies^{35; 37}; various pathways may be possible for shuttling the proton from the primer terminus to the leaving group in a complete nucleotidyl transfer cycle, and these may be polymerase dependent. In this connection, careful study of polymerase structures is warranted as possible mechanisms are considered. However, a unifying theme in the WMSA mechanism is the cycling through crystal and solvent waters of the proton originating from the primer 3'-terminus to the α - β bridging oxygen of the dNTP; this neutralizes the evolving negative charge as pyrophosphate leaves and restores the polymerase to its pre-chemistry state. Active site amino acid residues could be involved in the proton shuttling and their protonation states could return to the pre-chemistry ones as part of the proton relay.

Conclusion

Our present study with the high fidelity replicative T7 DNA polymerase indicates that a variant of the WMSA mechanism is an energetically feasible one for this case. A common feature of the WMSA mechanism is the ultimate transfer of the proton on the γ -phosphate to the α - β bridging oxygen as the pyrophosphate leaves, in order to neutralize the evolving negative charge. Further explorations are needed to determine its applicability to other polymerases, to evaluate polymerase structure-dependent variations, as well as to investigate further other possible variants in T7 DNA polymerase.

Computational Methods

We have employed molecular dynamics simulations and *ab initio* QM/MM calculations^{38; 39; 40; 41; 42; 43; 44; 45; 46; 47} to study the nucleotidyl transfer reaction in the T7 DNA polymerase active site, with template guanine and Watson-Crick paired dCTP as the nascent base pair. The QM sub-system is treated by the B3LYP functional with a 6-31G* basis set and contains a total of 75 atoms comprising 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 (9489 atoms) are described by the AMBER99 force field^{48; 49; 50} and the TIP3P water model⁵¹. The QM/MM interface is based on the pseudobond approach^{20; 22}, and an efficient iterative optimization approach and the reaction coordinate driving method²¹ were employed to minimize the reactant structure and map out the minimum energy path. Based on the determined reaction energy path, the free energy changes associated with the QM/MM

interaction were determined by the free energy perturbation (FEP) method²¹. All QM/MM and FEP calculations were carried out with modified versions of the Gaussian03⁵² and TINKER⁵³ programs. The pseudobond *ab initio* QM/MM approach^{20; 21; 22} has been demonstrated to be powerful in its capability for determining reaction pathways consistent with kinetic and mutagenic data for several enzymes^{54; 55; 56; 57; 58; 59}. Full computational details are given in Supplementary Material.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work is supported by NIH grants CA28038 and CA75449 to SB. YZ is grateful for the support from NSF (CHE-CAREER-0448156) and NIH (R01-GM079223). Molecular images were made with PyMOL (DeLano Scientific, LLC.).

References

1. Bebenek K, Kunkel TA. Functions of DNA polymerases. *Adv Protein Chem* 2004;69:137–165. [PubMed: 15588842]
2. Loeb LA, Monnat RJ Jr. DNA polymerases and human disease. *Nat Rev Genet* 2008;9:594–604. [PubMed: 18626473]
3. Kunkel TA. DNA replication fidelity. *J Biol Chem* 2004;279:16895–16898. [PubMed: 14988392]
4. McCulloch SD, Kunkel TA. The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases. *Cell Res* 2008;18:148–161. [PubMed: 18166979]
5. Broyde S, Wang L, Rechko O, Geacintov NE, Patel DJ. Lesion processing: high-fidelity versus lesion-bypass DNA polymerases. *Trends Biochem Sci* 2008;33:209–219. [PubMed: 18407502]
6. Rothwell PJ, Waksman G. Structure and mechanism of DNA polymerases. *Adv Protein Chem* 2005;71:401–440. [PubMed: 16230118]
7. Steitz TA, Steitz JA. A general two-metal-ion mechanism for catalytic RNA. *Proc Natl Acad Sci U S A* 1993;90:6498–6502. [PubMed: 8341661]
8. Steitz TA. A mechanism for all polymerases. *Nature* 1998;391:231–232. [PubMed: 9440683]
9. Beard WA, Wilson SH. Structural insights into DNA polymerase beta fidelity: hold tight if you want it right. *Chem Biol* 1998;5:R7–13. [PubMed: 9479474]
10. Doublié S, Sawaya MR, Ellenberger T. An open and closed case for all polymerases. *Structure* 1999;7:R31–35. [PubMed: 10368292]
11. Yang W, Woodgate R. What a difference a decade makes: insights into translesion DNA synthesis. *Proc Natl Acad Sci U S A* 2007;104:15591–15598. [PubMed: 17898175]
12. Florian J, Goodman MF, Warshel A. Computer simulation of the chemical catalysis of DNA polymerases: discriminating between alternative nucleotide insertion mechanisms for T7 DNA polymerase. *J Am Chem Soc* 2003;125:8163–8177. [PubMed: 12837086]
13. Florian J, Goodman MF, Warshel A. Computer simulations of protein functions: searching for the molecular origin of the replication fidelity of DNA polymerases. *Proc Natl Acad Sci U S A* 2005;102:6819–6824. [PubMed: 15863620]
14. Lin P, Pedersen LC, Batra VK, Beard WA, Wilson SH, Pedersen LG. Energy analysis of chemistry for correct insertion by DNA polymerase β . *Proc Natl Acad Sci U S A* 2006;103:13294–13299. [PubMed: 16938895]
15. Radhakrishnan R, Schlick T. Correct and incorrect nucleotide incorporation pathways in DNA polymerase β . *Biochem Biophys Res Commun* 2006;350:521–529. [PubMed: 17022941]
16. Wang L, Yu X, Hu P, Broyde S, Zhang Y. A water-mediated and substrate-assisted catalytic mechanism for *Sulfolobus solfataricus* DNA polymerase IV. *J Am Chem Soc* 2007;129:4731–4737. [PubMed: 17375926]

17. Cisneros GA, Perera L, Garcia-Diaz M, Bebenek K, Kunkel TA, Pedersen LG. Catalytic mechanism of human DNA polymerase λ with Mg^{2+} and Mn^{2+} from *ab initio* quantum mechanical/molecular mechanical studies. *DNA Repair (Amst)* 2008;7:1824–1834. [PubMed: 18692600]
18. Lin P, Batra VK, Pedersen LC, Beard WA, Wilson SH, Pedersen LG. Incorrect nucleotide insertion at the active site of a G:A mismatch catalyzed by DNA polymerase β . *Proc Natl Acad Sci U S A* 2008;105:5670–5674. [PubMed: 18391201]
19. Wang Y, Schlick T. Quantum mechanics/molecular mechanics investigation of the chemical reaction in Dpo4 reveals water-dependent pathways and requirements for active site reorganization. *J Am Chem Soc* 2008;130:13240–13250. [PubMed: 18785738]
20. Zhang Y, Lee T, Yang W. A pseudobond approach to combining quantum mechanical and molecular mechanical methods. *J Chem Phys* 1999;110:46–54.
21. Zhang Y, Liu H, Yang W. Free energy calculation on enzyme reactions with an efficient iterative procedure to determine minimum energy paths on a combined *ab initio* QM/MM potential energy surface. *J Chem Phys* 2000;112:3483–3492.
22. Zhang Y. Improved pseudobonds for combined *ab initio* quantum mechanical/molecular mechanical methods. *J Chem Phys* 2005;122:024114. [PubMed: 15638579]
23. Brieba LG, Eichman BF, Kokoska RJ, Doubie S, Kunkel TA, Ellenberger T. Structural basis for the dual coding potential of 8-oxoguanosine by a high-fidelity DNA polymerase. *EMBO J* 2004;23:3452–3461. [PubMed: 15297882]
24. Batra VK, Beard WA, Shock DD, Krahn JM, Pedersen LC, Wilson SH. Magnesium-induced assembly of a complete DNA polymerase catalytic complex. *Structure* 2006;14:757–766. [PubMed: 16615916]
25. Yang W, Lee JY, Nowotny M. Making and breaking nucleic acids: two- Mg^{2+} -ion catalysis and substrate specificity. *Mol Cell* 2006;22:5–13. [PubMed: 16600865]
26. Polesky AH, Dahlberg ME, Benkovic SJ, Grindley ND, Joyce CM. Side chains involved in catalysis of the polymerase reaction of DNA polymerase I from *Escherichia coli*. *J Biol Chem* 1992;267:8417–8428. [PubMed: 1569092]
27. Bebenek K, Beard WA, Darden TA, Li L, Prasad R, Luton BA, Gorenstein DG, Wilson SH, Kunkel TA. A minor groove binding track in reverse transcriptase. *Nat Struct Biol* 1997;4:194–197. [PubMed: 9164459]
28. Beard WA, Bebenek K, Darden TA, Li L, Prasad R, Kunkel TA, Wilson SH. Vertical-scanning mutagenesis of a critical tryptophan in the minor groove binding track of HIV-1 reverse transcriptase. Molecular nature of polymerase-nucleic acid interactions. *J Biol Chem* 1998;273:30435–30442. [PubMed: 9804810]
29. Morales JC, Kool ET. Minor groove interactions between polymerase and DNA: More essential to replication than Watson-Crick hydrogen bonds? *J Am Chem Soc* 1999;121:2323–2324.
30. Latham GJ, Forgacs E, Beard WA, Prasad R, Bebenek K, Kunkel TA, Wilson SH, Lloyd RS. Vertical-scanning mutagenesis of a critical tryptophan in the “minor groove binding track” of HIV-1 reverse transcriptase. Major groove DNA adducts identify specific protein interactions in the minor groove. *J Biol Chem* 2000;275:15025–15033. [PubMed: 10747890]
31. Tsai YC, Johnson KA. A new paradigm for DNA polymerase specificity. *Biochemistry* 2006;45:9675–9687. [PubMed: 16893169]
32. Wang S, Hu P, Zhang Y. *Ab initio* quantum mechanical/molecular mechanical molecular dynamics simulation of enzyme catalysis: the case of histone lysine methyltransferase SET7/9. *J Phys Chem B* 2007;111:3758–3764. [PubMed: 17388541]
33. Hu P, Wang S, Zhang Y. Highly dissociative and concerted mechanism for the nicotinamide cleavage reaction in Sir2Tm enzyme suggested by *ab initio* QM/MM molecular dynamics simulations. *J Am Chem Soc* 2008;130:16721–16728. [PubMed: 19049465]
34. Hu P, Wang S, Zhang Y. How do SET-domain protein lysine methyltransferases achieve the methylation state specificity? Revisited by *ab initio* QM/MM molecular dynamics simulations. *J Am Chem Soc* 2008;130:3806–3813. [PubMed: 18311969]
35. Castro C, Smidansky E, Maksimchuk KR, Arnold JJ, Korneeva VS, Gotte M, Konigsberg W, Cameron CE. Two proton transfers in the transition state for nucleotidyl transfer catalyzed by RNA- and DNA-dependent RNA and DNA polymerases. *Proc Natl Acad Sci U S A* 2007;104:4267–4272. [PubMed: 17360513]

36. Yin YW, Steitz TA. The structural mechanism of translocation and helicase activity in T7 RNA polymerase. *Cell* 2004;116:393–404. [PubMed: 15016374]
37. Castro C, Smidansky ED, Arnold JJ, Maksimchuk KR, Moustafa I, Uchida A, Gotte M, Konigsberg W, Cameron CE. Nucleic acid polymerases use a general acid for nucleotidyl transfer. *Nat Struct Mol Biol* 2009;16:212–218. [PubMed: 19151724]
38. Warshel A, Levitt M. Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J Mol Biol* 1976;103:227–249. [PubMed: 985660]
39. Singh UC, Kollman PA. A combined *ab initio* quantum mechanical and molecular mechanical method for carrying out simulations on complex molecular systems: Applications to the $\text{CH}_3\text{Cl} + \text{Cl}^-$ exchange reaction and gas phase protonation of polyethers. *J Comput Chem* 1986;7:718–730.
40. Field MJ, Bash PA, Karplus M. A combined quantum mechanical and molecular mechanical potential for molecular dynamics simulations. *J Comput Chem* 1990;11:700–733.
41. Gao J, Truhlar DG. Quantum mechanical methods for enzyme kinetics. *Annu Rev Phys Chem* 2002;53:467–505. [PubMed: 11972016]
42. Cui Q, Karplus M. Catalysis and specificity in enzymes: a study of triosephosphate isomerase and comparison with methyl glyoxal synthase. *Adv Protein Chem* 2003;66:315–372. [PubMed: 14631822]
43. Warshel A. Computer simulations of enzyme catalysis: methods, progress, and insights. *Annu Rev Biophys Biomol Struct* 2003;32:425–443. [PubMed: 12574064]
44. Friesner RA, Guallar V. *Ab initio* quantum chemical and mixed quantum mechanics/molecular mechanics (QM/MM) methods for studying enzymatic catalysis. *Annu Rev Phys Chem* 2005;56:389–427. [PubMed: 15796706]
45. Zhang Y. Pseudobond *ab initio* QM/MM approach and its applications to enzyme reactions. *Theor Chem Acc* 2006;116:43–50.
46. Hu H, Yang W. Free energies of chemical reactions in solution and in enzymes with *ab initio* quantum mechanics/molecular mechanics methods. *Annu Rev Phys Chem* 2008;59:573–601. [PubMed: 18393679]
47. Senn HM, Thiel W. QM/MM methods for biomolecular systems. *Angew Chem Int Ed* 2009;48:1198–1229.
48. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J Am Chem Soc* 1995;117:5179–5197.
49. Cheatham TE III, Cieplak P, Kollman PA. A modified version of the Cornell *et al.* force field with improved sugar pucker phases and helical repeat. *J Biomol Struct Dyn* 1999;16:845–862. [PubMed: 10217454]
50. Wang J, Cieplak P, Kollman PA. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J Comp Chem* 2000;21:1049–1074.
51. Jorgensen WL, Chandrasekhar J, Madura J, Klein ML. Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 1983;79:926–935.
52. Frisch, MJ.; Trucks, GW.; Schlegel, HB.; Scuseria, GE.; Robb, MA.; Cheeseman, JR.; Montgomery, JA., Jr; Vreven, T.; Kudin, KN.; Burant, JC.; Millam, JM.; Iyengar, SS.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, GA.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, JE.; Hratchian, HP.; Cross, JB.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, RE.; Yazyev, O.; Austin, AJ.; Cammi, R.; Pomelli, C.; Ochterski, JW.; Ayala, PY.; Morokuma, K.; Voth, GA.; Salvador, P.; Dannenberg, JJ.; Zakrzewski, VG.; Dapprich, S.; Daniels, AD.; Strain, MC.; Farkas, O.; Malick, DK.; Rabuck, AD.; Raghavachari, K.; Foresman, JB.; Ortiz, JV.; Cui, Q.; Baboul, AG.; Clifford, S.; Cioslowski, J.; Stefanov, BB.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, RL.; Fox, DJ.; Keith, T.; Al-Laham, MA.; Peng, CY.; Nanayakkara, A.; Challacombe, M.; Gill, PMW.; Johnson, B.; Chen, W.; Wong, MW.; Gonzalez, C.; Pople, JA. Gaussian 03, Revision B.05. Gaussian, Inc.; Wallingford, CT: 2004.
53. Ponder, JW. TINKER. Washington University in St. Louis; St. Louis, WA: 2004.

54. Liu H, Zhang Y, Yang W. How is the active site of enolase organized to catalyze two different reaction steps? *J Am Chem Soc* 2000;122:6560–6570.
55. Zhang Y, Kua J, McCammon JA. Role of the catalytic triad and oxyanion hole in acetylcholinesterase catalysis: an *ab initio* QM/MM study. *J Am Chem Soc* 2002;124:10572–10577. [PubMed: 12197759]
56. Cisneros GA, Liu H, Zhang Y, Yang W. *Ab initio* QM/MM study shows there is no general acid in the reaction catalyzed by 4-oxalocrotonate tautomerase. *J Am Chem Soc* 2003;125:10384–10393. [PubMed: 12926963]
57. Cheng Y, Zhang Y, McCammon JA. How does the cAMP-dependent protein kinase catalyze the phosphorylation reaction: an *ab initio* QM/MM study. *J Am Chem Soc* 2005;127:1553–1562. [PubMed: 15686389]
58. Corminboeuf C, Hu P, Tuckerman ME, Zhang Y. Unexpected deacetylation mechanism suggested by a density functional theory QM/MM study of histone-deacetylase-like protein. *J Am Chem Soc* 2006;128:4530–4531. [PubMed: 16594663]
59. Hu P, Zhang Y. Catalytic mechanism and product specificity of the histone lysine methyltransferase SET7/9: an *ab initio* QM/MM-FE study with multiple initial structures. *J Am Chem Soc* 2006;128:1272–1278. [PubMed: 16433545]

Abbreviations

QM/MM	quantum mechanical and molecular mechanical
WMSA	water-mediated and substrate-assisted

Figure 1.

The determined structure of the T7 ternary complex active site based on molecular modeling/dynamics and subsequent *ab initio* QM/MM minimizations. Dashed lines denote coordination distances (in Å) between the magnesium ions and their ligands. The inter-Mg²⁺ distance is also shown. The values of the corresponding distances in the crystal structure (PDB ID: 1T8E) are given in parentheses. WAT1 and WAT2 are the two crystal water molecules in the coordination sphere of MgA. H₂O is a solvent water molecule. For clarity, the base of the primer 3'-end, the sugar and base of the incoming dCTP are not shown.

Figure 2.

Reaction mechanism (left panel) and critical structures (right panel) determined for the nucleotidyl transfer catalyzed by T7 DNA polymerase. R, reactant (ground state); I1, proton on the crystal water molecule relayed to γ -phosphate via a solvent water molecule; I2, 3'-terminal proton transferred to hydroxide ion; I3, completion of nucleotidyl transfer; P, product. Red arrows with solid heads denote proton transfer; red arrows with hollow heads denote electron rearrangement; d1-d10 denote distances between atoms connected by dashed lines or covalent bonds; their variation along the reaction path is shown in Figure 3B. Shaded areas in the left panel 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. PPI denotes formation of the metastable pentacovalent phosphorane intermediate. See Figure 2 for definitions of d1–d10.

Figure 4.

Group charges of the β - (cyan) and γ - (magenta) phosphates, and their sum (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 proton transfer from 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.