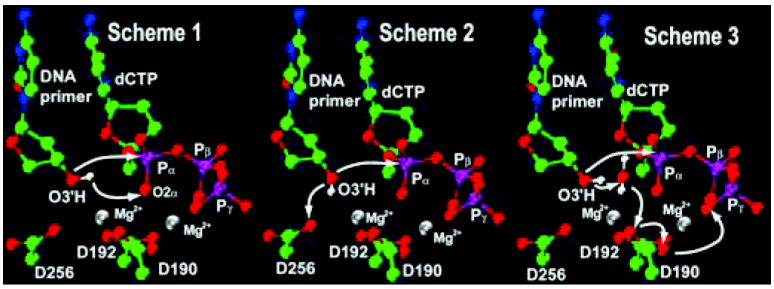
Article

DNA Polymerase β Catalysis: Are Different Mechanisms Possible?

Ian L. Alberts, Yanli Wang, and Tamar Schlick

J. Am. Chem. Soc., **2007**, 129 (36), 11100-11110• DOI: 10.1021/ja071533b • Publication Date (Web): 16 August 2007



More About This Article

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

- Supporting Information
- Links to the 5 articles that cite this article, as of the time of this article download
- 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 08/16/2007

DNA Polymerase β Catalysis: Are Different Mechanisms Possible?

Ian L. Alberts,† Yanli Wang, and Tamar Schlick*

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

Received March 16, 2007; E-mail: schlick@nyu.edu

Abstract: DNA polymerases are crucial constituents of the complex cellular machinery for replicating and repairing DNA. Discerning mechanistic pathways of DNA polymerase on the atomic level is important for revealing the origin of fidelity discrimination. Mammalian DNA polymerase β (pol β), a small (39 kDa) member of the X-family, represents an excellent model system to investigate polymerase mechanisms. Here, we explore several feasible low-energy pathways of the nucleotide transfer reaction of pol β for correct (according to Watson−Crick hydrogen bonding) G:C basepairing versus the incorrect G:G case within a consistent theoretical framework. We use mixed quantum mechanics/molecular mechanics (QM/MM) techniques in a constrained energy minimization protocol to effectively model not only the reactive core but also the influence of the rest of the enzymatic environment and explicit solvent on the reaction. The postulated pathways involve initial proton abstraction from the terminal DNA primer O3′H group, nucleophilic attack that extends the DNA primer chain, and elimination of pyrophosphate. In particular, we analyze several possible routes for the initial deprotonation step: (i) direct transfer to a phosphate oxygen O(P_α) of the incoming nucleotide, (ii) direct transfer to an active site Asp group, and (iii) transfer to explicit water molecules. We find that the most probable initial step corresponds to step (iii), involving initial deprotonation to water, which is followed by proton migration to active site Asp residues, and finally to the leaving pyrophosphate group, with an activation energy of about 15 kcal/mol. We argue that initial deprotonation steps (i) and (ii) are less likely as they are at least 7 and 11 kcal/mol, respectively, higher in energy. Overall, the rate-determining step for both the correct and the incorrect nucleotide cases is the initial deprotonation in concert with nucleophilic attack at the phosphate center; however, the activation energy we obtain for the mismatched G:G case is 5 kcal/mol higher than that of the matched G:C complex, due to active site structural distortions. Taken together, our results support other reported mechanisms and help define a framework for interpreting nucleotide specificity differences across polymerase families, in terms of the concept of active site preorganization or the so-called “pre-chemistry avenue”.

Introduction

Genomic DNA replication and repair is crucial for the survival and maintenance of living organisms.[1,2] Several cooperatively acting proteins are known to be intimately involved in these processes, including DNA polymerases, which play a central role in the cell replication and repair machinery.[3] Experimental and computational investigations have recognized that the DNA polymerase-catalyzed nucleotide incorporation process entails a conformational rearrangement of the enzyme, upon binding the incoming nucleotide substrate, followed by the phosphoryl transfer reaction that extends the DNA duplex by one base pair.[4−6] Of interest is to rationalize the fidelity behavior of these enzymatic systems by discussing characteristics that apply to matched (according to Watson−Crick hydrogen bonding) versus mismatched nucleotide systems at the primer/template junction.[7] Understanding the mechanism of action of these systems at the atomic level has immense importance for elucidating the origin of polymerase fidelity and specificity.

The eukaryotic enzyme DNA polymerase β (pol β), a member of the X-family, is involved in single nucleotide gap-filling in the base excision repair (BER) process.[8,9] As one of the smallest (39 kDa) known DNA polymerases, pol β represents an excellent model system for exploring the nucleotide insertion process. Pol β consists of a 31 kDa C-terminal segment comprising thumb, palm, and finger subdomains, all responsible for the enzyme’s role in BER, and an 8 kDa N-terminal domain displaying deoxyribose phosphate lyase activity.[10] Several high-quality X-ray crystal structures have been resolved for pol β. These structures reflect open and closed forms of the enzyme complex system and, therefore, suggest a large subdomain

† Current address: Schrödinger, Inc., 120 West 45th Street, 29th Floor, New York, New York 10036.

(1) Friedberg, E. C. *Nature* **2003**, *421*, 436−440.
(2) Zhou, B.-B. S.; Elledge, S. J. *Nature* **2000**, *408*, 433−439.
(3) Alberts, B. *Nature* **2003**, *421*, 431−435.
(4) Beard, W. A.; Wilson, S. H. *Chem. Rev.* **2006**, *106*, 361−82.
(5) Kunkel, T. A.; Bebenek, K. *Annu. Rev. Biochem.* **2000**, *69*, 497−529.
(6) Mildvan, A. S. *Proteins: Struct., Funct., Genet.* **1997**, *29*, 401−416.
(7) Beard, W. A.; Shock, D. D.; Vande Berg, B. J.; Wilson, S. H. *J. Biol. Chem.* **2002**, *277*, 47393−47398.
(8) Matsumoto, Y.; Kim, K. *Science* **1995**, *269*, 699.
(9) Piersen, C. E.; Prasad, R.; Wilson, S. H.; Llord, R. S. *J. Biol. Chem.* **1996**, *271*, 17811.
(10) Sawaya, M. R.; Prasad, R.; Wilson, S. H.; Kraut, J.; Pelletier, H. *Biochemistry* **1997**, *36*, 11205.

10.1021/ja071533b CCC: $37.00

motion of the thumb component induced by substrate binding.[10−13] Several recent computational studies using molecular dynamics (MD) simulations support this induced-fit mechanism. Thus, a correct incoming nucleotide matching the complementary base in the template DNA strand leads to a "closed" complex structure with catalytic groups appropriately assembled for the subsequent chemical reaction,[14−16] and, at the same time, insertion of an incorrect nucleotide promotes a transition to a more "loosely" associated active site that is significantly distorted from the reaction-competent state.[17,18] Intriguingly, according to transition path sampling techniques, the overall conformational closing activation barrier was found to be similar for correct and incorrect incoming nucleotides; however, the "closed" enzyme structure was deemed unstable relative to the "open" form for the mismatch.[19,20] This implies a potential basis for fidelity discrimination against the mismatch in terms of context specific organization of the pol β active site. In a comprehensive and fascinating recent review, a so-called "pre-chemistry avenue" was defined, which advocates that active site preorganization, in the spirit of Warshel and co-workers,[21,22] prior to the chemical reaction but after the conformational closing, is key to pol β fidelity and may provide a unifying framework across different polymerases.[17,23]

Complementary studies of the pol β rate-limiting nucleotidyl transfer chemical step have been reported that identify transient intermediate structures and associated energetics along the reaction pathway. The reaction involves P−O bond making and bond breaking, leading to DNA primer extension and elimination of pyrophosphate. Recent evidence from experimental and computational studies supports the concept of a two-metal ion assisted mechanism for enzyme-catalyzed nucleotide incorporation processes.[14,24−26] In this scheme, two metal ions coordinated to three conserved acidic residues (typically two aspartates and either another aspartate or glutamate)[27] appear crucial both for the catalytic activity of the enzyme as well as for maintaining the structural integrity of the active site. It has been postulated that the phosphoryl transfer reaction proceeds through an associative or partially associative transition state or intermediate, rather than a corresponding dissociative reaction mechanism, involving nucleophilic attack of the terminal DNA primer hydroxyl oxygen (O3′) on the α-phosphorus (P_α) of the incoming nucleotide substrate.[6,28] The nucleophilic character of the hydroxyl oxygen is enhanced by deprotonation of the hydroxyl hydrogen (O3′H), with the resultant oxyanion stabilized by coordination to the catalytic Mg^{2+} ion.

As the nucleotidyl insertion reaction involves bond breaking and bond making, quantum mechanical (QM)-based methods are necessary to correctly model the corresponding changes in electron density. Several excellent previous theoretical studies have suggested different possible pathways for the crucial initial deprotonation of the terminal primer O3′H group. Using a QM approach applied to a small model of the pol β active site region, Abashkin and workers suggest direct deprotonation to an α-phosphorus oxygen O(P_α) of the incoming nucleotide substrate,[28] whereas Rittenhouse et al. propose proton migration toward conserved active site D256.[29] The latter route was also advocated by Florian et al. for the activation of the chemical step for DNA polymerase T7 using Empirical Valence Bond (EVB) methods.[30] Recently, Bojin and Schlick reported a more comprehensive QM investigation of the chemical reaction profile for several model pol β systems, in which several possible initial proton-transfer pathways were investigated.[31] Again, the direct route from O3′ to O(P_α) was found energetically most favorable.

QM approaches necessarily neglect the influence of the protein/template primer DNA and solvent environment on the reaction mechanism. To rectify these limitations, two recent works combined quantum mechanics with molecular mechanics procedures (QM/MM) to simulate the entire solvated pol β system. Lin et al. reported initial proton abstraction of O3′H to D256 by a series of QM/MM constrained minimizations along a partially associative nucleophilic substitution reaction pathway involving incoming dTTP opposite template A.[32] Radhakrishnan and Schlick developed and applied a novel procedure employing short 10 ps QM/MM dynamics in conjunction with umbrella sampling to estimate free energies of intermediates along the phosphoryl transfer reaction coordinate.[33] Their dominant pathway involves O3′H deprotonation to multiple water molecules followed by migration to active site Asp residues in a series of Grotthus hopping steps[34] for both G:C and G:A systems. Furthermore, by comparing the reaction energetics in a substrate-specific context, Radhakrishnan and Schlick showed that the mismatched G:A basepair system results in a significantly higher free energy barrier than the G:C matched basepair case (by about 7 kcal/mol). Their study also advocated that the initial deprotonation step occurs concurrently with the nucleophilic attack of O3′ on P_α; Lin et al. instead proposed that the proton migration takes place before nucleophilic attack.

Although all prior explorations of the pol β chemical reaction pathway favor an associative (or partially associative) mechanism, a diversity exists concerning the key activating step: initial deprotonation of the terminal DNA primer hydroxyl proton. This likely stems from the different theoretical methods as well as pol β structural models employed. For example, the two QM/MM studies discussed above[32,33] differed in initial pol β crystal

(11) Beard, W. A.; Wilson, S. H. *Structure* **2003**, *11*, 489.
(12) Batra, V. K.; Beard, W. A.; Shock, D. D.; Pedersen, L. C.; Wilson, S. H. *Structure* **2005**, *13*, 1225−1233.
(13) Krahn, J. M.; Beard, W. A.; Wilson, S. H. *Structure* **2004**, *12*, 1823−1832.
(14) Yang, L.; Arora, K.; Beard, W. A.; Wilson, S. H.; Schlick, T. *J. Am. Chem. Soc.* **2004**, *126*, 8441−8453.
(15) Arora, K.; Schlick, T. *J. Phys. Chem. B* **2005**, *109*, 5358.
(16) Arora, K.; Schlick, T. *Biophys. J.* **2004**, *87*, 1−12.
(17) Arora, K.; Beard, W. A.; Wilson, S. H.; Schlick, T. *Biochemistry* **2005**, *44*, 13328−13341.
(18) Yang, L. J.; Schlick, T. *J. Mol. Biol.* **2002**, *321*, 458−487.
(19) Radhakrishnan, R.; Schlick, T. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 5970−5975.
(20) Radhakrishnan, R.; Schlick, T. *J. Am. Chem. Soc.* **2005**, *127*, 13245−13252.
(21) Warshel, A. *J. Biol. Chem.* **1998**, *273*, 27035−27038.
(22) Florian, J.; Goodman, M. F.; Warshel, A. *J. Phys. Chem. B* **2002**, *106*, 5739−5753.
(23) Radhakrishnan, R.; Arora, K.; Wang, Y.; Beard, W. A.; Wilson, S. H.; Schlick, T. *Biochemistry* **2006**, *45*, 15142−15156.
(24) Steitz, T. A. *J. Biol. Chem.* **1999**, *274*, 17395−17398.
(25) Steitz, T. A.; Smerdon, S. J.; Jager, J.; Joyce, C. M. *Science* **1994**, *266*, 2022−2025.
(26) Florian, J.; Goodman, M. F.; Warshel, A. *J. Am. Chem. Soc.* **2003**, *125*, 8163−8177.
(27) Wang, J.; Sattar, A. K. M. A.; Wang, C. C.; Karam, J. D.; Konigsberg, W. H.; Steitz, T. A. *Cell* **1997**, *89*, 1087−1099.
(28) Abashkin, Y. G.; Erickson, J. W.; Burt, S. K. *J. Phys. Chem. B* **2001**, *105*, 287−292.
(29) Rittenhouse, R. C. *Proteins* **2003**, *53*, 667.
(30) Florian, J.; Goodman, M. F.; Warshel, A. *J. Am. Chem. Soc.* **2003**, *125*, 8163−8177.
(31) Bojin, M.; Schlick, T. *J. Phys. Chem. B*, in press.
(32) 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.
(33) Radhakrishnan, R.; Schlick, T. *Biochem. Biophys. Res. Commun.* **2006**, *350*, 521−529.
(34) Agmon, A. *Chem. Phys. Lett.* **1995**, *244*, 456−462.

models, including variable protonation states of key active site groups and different incoming and template nucleotides, and the use of differing QM methods to describe the reactive core. Crucially, each study generally explored in detail only a single viable reaction coordinate and rationally argued its validity, often on the basis of structural features of the initial model. This makes it difficult to quantitatively assess alternative reaction pathways, especially concerning the initial activating proton abstraction step. Our goal, thus, was to examine multiple feasible pol β-catalyzed reaction pathways within a consistent theoretical framework and compare the energetics involved.

With this directive, we report here results of our QM/MM study producing several potentially low-energy mechanisms for the phosphoryl transfer reaction within the pol β active site. By exploring relative energies for the matched G:C case as compared to the G:G mismatch, we can also address pol β's fidelity mechanism. Our calculations adhere to a partially associative nucleophilic attack of the terminal DNA primer O3′ atom on P_α of the substrate and explore several possible initial O3′H deprotonation routes: directly to $O2_\alpha$ of the incoming nucleotide, directly to active site Asp residues, or to nearby solvent molecules, in a similar manner to Warshel and co-workers in their EVB studies of T7 and DNA Polymerase I catalysis.[26,35] Our results support the notion that the dominant energetic pathway involves initial proton abstraction to water followed by proton transfer to active site Asp residues and finally to the leaving pyrophosphate group. The initial deprotonation can occur in conjunction with nucleophilic attack at the P_α center and represents the highest point on the reaction surface. The higher activation barrier for an incorrect nucleotide insertion (misinsertion) by about 5 kcal/mol than the correct basepair is a consequence of the greater deviation of the initial active site structure from the ideal reaction-competent state. This difference allows rationalization of fidelity discrimination for pol β in terms of the "pre-chemistry avenue"[17,23] arising from active site preorganization concepts.[21,22]

Methods

Pol β Model Preparation. Two pol β model systems were investigated: pol β/DNA/dCTP, where dCTP denotes 2′-deoxyribocytosine 5′-triphosphate, the G:C system; and pol β/DNA/dGTP, the G:G system. The G:C model was constructed from the high-resolution crystal structure of the ternary complex.[36] Hydrogen atoms were added at pH 7 using InsightII,[37] and the missing hydroxyl group at the 3′ terminus of the DNA primer strand, as well as missing residues 1−9 of pol β, were added. Further, the Na^+ ion located at the position of the catalytic metal ion in the pol β active site was replaced by Mg^{2+}. The system was flooded with explicit water molecules in cubic periodic domains via PBCAID.[38] Using Delphi to calculate electrostatic potentials, the system was neutralized to an ionic strength of 150 mM by replacing water molecules with minimal electrostatic potential by Na^+ ions and those with maximal electrostatic potential by Cl^- ions. These Na^+ and Cl^- ions were positioned over 8 Å away from any atoms of the protein or DNA system and from each other. The final G:C system comprised 40 238 atoms, including 11 249 water molecules.

According to the pK_a values of titratable side-chains, all pol β His, Lys, and Arg residues were assigned +1 charged states and Asp and Glu residues −1 charged states, including key active site residues D190, D192, and D256. The initial choice of protonation states for the three active site Asp residues is not definitive, and different protonations may influence ensuing reaction pathways. These issues are explored in further detail by Bojin and Schlick.[31] This setup leads to a charge of +7 for pol β, −29 for the primer DNA, −4 for the dCTP, and +4 for the two Mg^{2+} ions. In conjunction with 42 Na^+ and 20 Cl^- ions, these assignments result in an overall neutral G:C system. The assembly was minimized for 10 000 steps using the Steepest Descent Method followed by 20 000 steps using the Adapted Basis Newton Raphson procedure.[39,40] This was followed by system equilibration for 30 ps at 300 K using the Verlet integrator. These simulations were conducted using CHARMM (Chemistry Department, Harvard University, Cambridge MA)[41] with the all-atom version c31a1 force field.

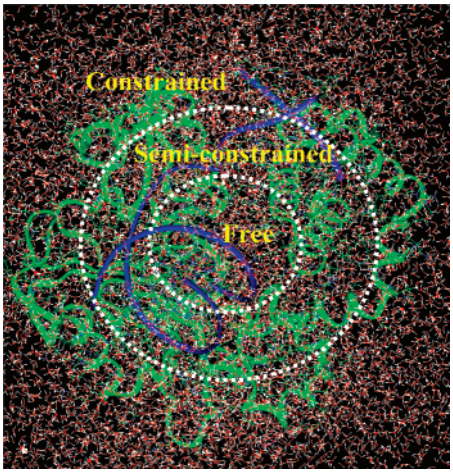


Figure 1. Initial model of solvated pol β/DNA/dGTP complex comprising over 40 000 atoms. All atoms within 15 Å of any QM atom are freely optimized during the minimizations, except for the restraints imposed to follow the reaction pathway. Atoms between 15 and 25 Å of any QM atom are semi-constrained, and atoms beyond 25 Å are fixed in their original positions.

The G:G model was taken from the endpoint of MD simulations conducted to analyze pol β conformational changes in the presence of dGTP.[17] These simulations were initiated from a solvated intermediate pol β model, built as an average of the crystal structures of the open, binary complex (PDB[42] code 1BPX) and the closed ternary complex (PDB code 1BPY). In this model, the incoming G and the template G were arranged in an anti conformation in accordance with pol β crystal structures comprising mispairs.[13] This solvated system was subject to minimization, equilibration, and dynamics protocols as described in detail by Arora et al.,[17] and the endpoint of the simulations involving a distorted active site arrangement was chosen as the starting point for our study of the subsequent chemical reaction mechanism. The solvated G:G system (Figure 1) comprised the same number of atoms, with all titratable side-chains assigned the same protonation states as the corresponding G:C system; similarly, the G:G system was overall neutral.

QM/MM Calculations. Our QM/MM approach involves the combination of CHARMM and GAMESS-UK,[43] the latter an ab initio

(35) Florian, J.; Goodman, M. F.; Warshel, A. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6819.
(36) Batra, V. K.; Beard, W. A.; Shock, D. D.; Krahn, J. M.; Pedersen, L. C.; Wilson, S. H. *Structure* **2006**, *14*, 757−766.
(37) InsightII, Accelrys Inc.
(38) Qian, X.; Strahs, D.; Schlick, T. *J. Comput. Chem.* **2001**, *22*, 1843−1850.
(39) Brooks, B. R.; Bruccoleri, E. R.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.
(40) Schlick, T. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; CH Publishers: New York, 1992; Vol. III, pp 1−71.
(41) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.
(42) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235−242.
(43) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. J.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347−1363.

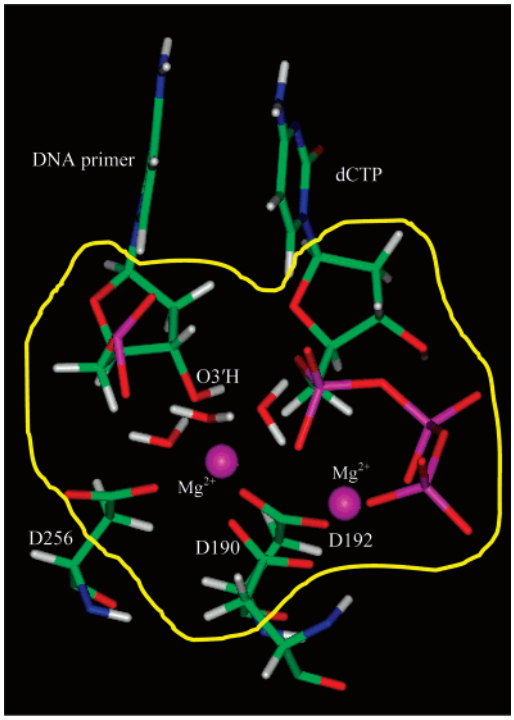


Figure 2. The reactive species of pol β defined as the QM region. QM atoms are within the yellow circle.

molecular electronic structure program. Further details of the QM/MM methodology adopted can be found in the Appendix A of the Supporting Information. In our application of the QM/MM approach, the reactive species within the active site of pol β were modeled by QM methods and the rest of the enzyme, DNA, incoming nucleotide, NA^+, Cl^-, and solvent molecules were designated as MM atoms (see Figure 2). The QM region consists of two Mg^{2+} ions, D190, D192, D256, terminal DNA primer, incoming nucleotide, and explicit water molecules within hydrogen-bonding distance of any QM atoms. The single link atom approach was used to satisfy the valence requirements of the QM atoms in the interface of the QM and MM regions. The link atoms are positioned along the $C_\alpha-C_\beta$ bond of the amino acids, on the C5′−C4′ bond of dNTP, and along the C2′−C1′ and C4′−C3′ bonds of the terminal DNA primer. The link atoms are taken as hydrogen atoms. They are constrained to be collinear along the frontier QM−MM bond vector and placed at a fixed distance from the frontier QM atom (in this respect, frontier atoms refer to those QM and MM atoms at the QM/MM interface that are joined by the severed bond). This leads to a total of 65 atoms in the QM region, including link atoms.

Several levels of theory were used to describe the QM atoms in the QM/MM simulations. The Hartree−Fock (HF) method with the minimal STO-3G basis set, denoted HF/STO-3G, was used to identify preliminary geometries of reactant, transient intermediates, and product along several possible nucleotide transfer reaction pathways. These structures were subsequently used as the starting points for geometry optimizations at higher levels of theory, including HF/6-31G* (valence double-ζ with polarization functions on all heavy atoms)[44] and B3LYP/6-31G*. The B3LYP approach[45,46] within the Density Functional Theory (DFT) formalism has been shown to be successful for locating intermediates on potential energy surfaces. Finally, single point energy calculations were conducted with Møller−Plesset perturbation theory,[47] MP2, with a 6-311+G(d,p) basis set (valence triple-ζ with polarization functions on all atoms and diffuse functions on heavy atoms) at the B3LYP/6-31G* optimized geometries for the QM region to obtain the most accurate energies.

QM/MM Constrained Minimizations. Geometry optimizations were performed using the QM/MM Hamiltonian. In these simulations, all atoms within 15 Å of any QM atom were unconstrained and allowed to freely optimize subject to the restraints considered below. Atoms at a distance of 15−25 Å from any QM atom were semi-constrained harmonically with a force constant of 5.0 kcal/mol/Å^2 and atoms further than 25 Å were fixed in their initial positions. To follow the catalytic reaction mechanism of pol β and identify relevant intermediate structures along the pathway, we implemented several constraints in the energy minimization procedure. Harmonic constraints with force constants, $K = 1000.0$ kcal/mol/Å^2 along the $O3'-P_\alpha$ and $O3_\alpha-P_\alpha$ distances, were used. These distances correspond to the bond making and bond breaking processes during the phosphoryl transfer reaction. A series of constrained minimizations were conducted in which the $O3'-P_\alpha$ distance was reduced from 3.8 to 1.7 Å in 10 equal steps and the $O3_\alpha-P_\alpha$ distance was increased from 1.6 to 4.0 Å in 10 equal steps. Other constraints were also used involving specific O−H distances to describe proton-transfer steps as explicated in the Results section below. In CHARMM, these restraints are incorporated by augmenting the energy with the harmonic term $0.5K(R-R_0)^2$, where $R-R_0$ is the difference between the actual distance and the target distance for each constraint. 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.

Results

Reaction Pathways Explored. We systematically investigated several possible mechanisms for the pol β catalytic reaction in the context of both the correct G:C and the incorrect G:G cases within a consistent theoretical framework. All pathways involve nucleophilic attack of the terminal primer O3′ on the phosphate P_α atom leading to the generation of a partially associative intermediate, comprising a pentacoordinated P_α. This was followed by complete formation of the $O3'-P_\alpha$ bond that extends the DNA primer chain and cleavage of the $P_\alpha-O3_\alpha$ bond that triggers pyrophosphate elimination. These pathways differ in the vital initial deprotonation of the terminal 3′-hydroxyl group that activates the nucleophilic O3′ atom. The following deprotonation steps are explored in detail: (i) direct deprotonation of the terminal primer O3′H proton to the $O2_\alpha$ phosphate oxygen atom of the incoming nucleotide, (ii) deprotonation of the O3′H proton to either D190, D192, or D256, and (iii) deprotonation of the O3′H proton to nearby hydrogen-bonded water molecules. Note that the first pathway was suggested in works of Abashkin et al.[28] and Bojin and Schlick,[31] the second pathway suggested by the Pedersen group,[32] and the third by Radhakrishnan and Schlick[33] as well as by Bojin and Schlick.[31] Identification of the rate-limiting step in these pathways for both G:C and G:G cases provides insights into the nucleotide fidelity and selectivity processes of pol β.

We did not consider another possible mechanism involving the formation of a metaphosphate-like PO_3^- intermediate or transition state by cleavage of the $P_\alpha-O3_\alpha$ bond in a dissociative nucleophilic reaction, followed by attack of the terminal primer O3′H group; Abashkin et al. found the dissociative pathway to have a barrier about 14 kcal/mol above that of the associative route.[28]

All relative energies considered in this section refer to QM/MM calculations using MP2 with a 6-311+G(d,p) basis set to describe the QM region (with the standard CHARMM force

(44) Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; DeFrees, D. J.; Pople, J. A.; Gordon, M. S. *J. Chem. Phys.* **1982**, *77*, 3654.
(45) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(46) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785−789.
(47) Head-Gordon, M.; Pople, J. A.; Frisch, M. J. *Chem. Phys. Lett.* **1988**, *153*, 503.

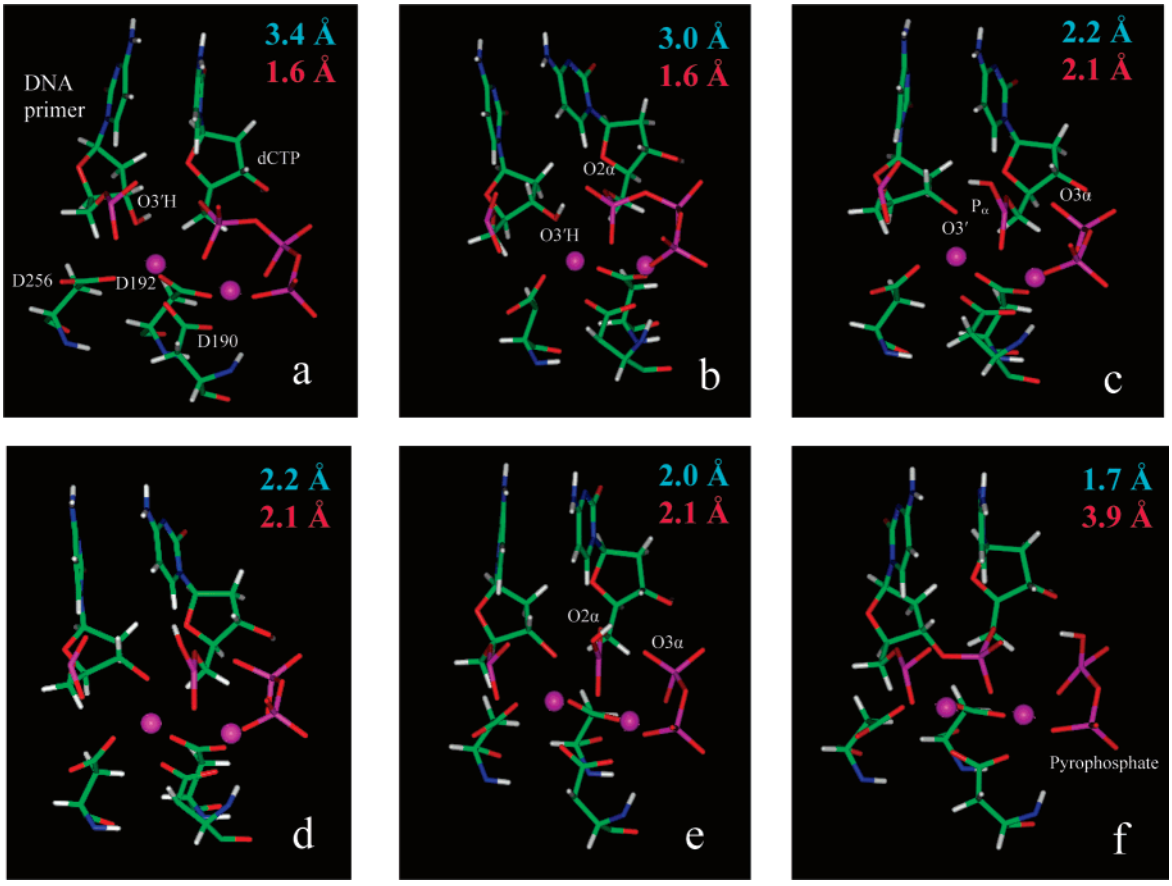


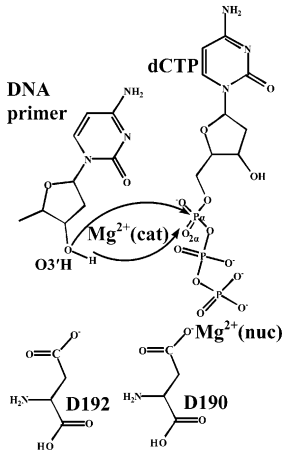
Figure 3. Transient intermediates identified along reaction pathway 1 involving direct deprotonation of the terminal primer O3′H to $O2_\alpha$ of the incoming dCTP in conjunction with nucleophilic attack at the P_α center. (a) Initial reactant state; (b) modified reactant (intermediate 1); (c) intermediate 2; (d) transition state for rotation about the $P_\alpha-O2_\alpha$ bond; (e) intermediate 3; and (f) product state with eliminated pyrophosphate. Key distances O3′−P_α and $P_\alpha-O3_\alpha$ are shown for each structure.

field for the MM atoms). The effect of different levels of theory for the QM region is reported in Appendix B in the Supporting Information. Moreover, the identified reaction pathways (1 to 3) were also driven backward to check for convergence. Our inverse backward pathways in each case reproduced the forward pathways and so confirmed the convergence and findings of each of the routes studied.

To examine perturbations caused by the relatively large step size (0.2 Å), we also performed the same calculations with a step size of 0.1 Å (as used in ref 32) in the vicinity of the initial transition state-like structure (highest energy structure) located for each pathway. The new structures and energies obtained with the smaller step size match those reported here (to within 1%). This suggests that the approach used in this work is satisfactory and does not lead to significant perturbation errors.

Pathway 1. Nucleophilic Attack Initiated by Direct Deprotonation of the Terminal DNA Primer O3′H Proton to the $O2_\alpha$ Phosphate Oxygen Atom of the Incoming Nucleotide. We obtained a series of transient intermediates (Figure 3 and Scheme 1) using constrained minimization from the initial reactant to product state, involving transfer of the primer O3′H proton to $O2_\alpha$ and eventually to $O3_\alpha$ of the leaving pyrophosphate group. In the reactant state of the G:C system (minimized initial structure, Figure 3a), the O3′H terminal DNA primer proton forms a hydrogen bond with the bridging O5′ oxygen atom of the incoming nucleotide, as has been recognized in earlier studies.[20,29,32,33] This organization does not afford a feasible pathway for the crucial abstraction of the O3′H proton.

Scheme 1. Sketch of Reaction Pathways[a]



[a] Arrows show direct deprotonation of the terminal primer O3′H to $O2_\alpha$ of the incoming dCTP, in conjunction with nucleophilic attack at the P_α center.

To facilitate this initial deprotonation, an exploration of the QM/MM potential energy surface close to the reactant state revealed a second local minimum structure, reached by a small rotation of the system around the catalytic magnesium ion center, in which the O3′H proton is hydrogen bonded to the $O2_\alpha$ phosphate oxygen (Figure 3b). This configuration (intermediate 1 in Table 1) is about 6 kcal/mol higher in energy than the original G:C reactant structure (Figure 4) and suggests a proton migration from the terminal primer hydroxyl group to the $O2_\alpha$ atom of

Table 1. Energetics of Structures Located during Pathway 1 from QM/MM Calculations Using MP2/6-311+G(d,p) for the QM Region

structure	energy [kcal/mol] G:C	G:G
reactant	0.0	0.0
intermediate 1 (modified reactant)	6.2	N/A
intermediate 2 (O3′ deprotonation to $O2_\alpha$ and associative intermediate)	21.0	29.4
transition state ($P_\alpha-O2_\alpha$ bond rotation)	22.4	31.2
intermediate 3 (H-bond to $O3_\alpha$)	18.9	27.1
product ($O3_\alpha$ of pyrophosphate protonated)	−2.9	−2.1

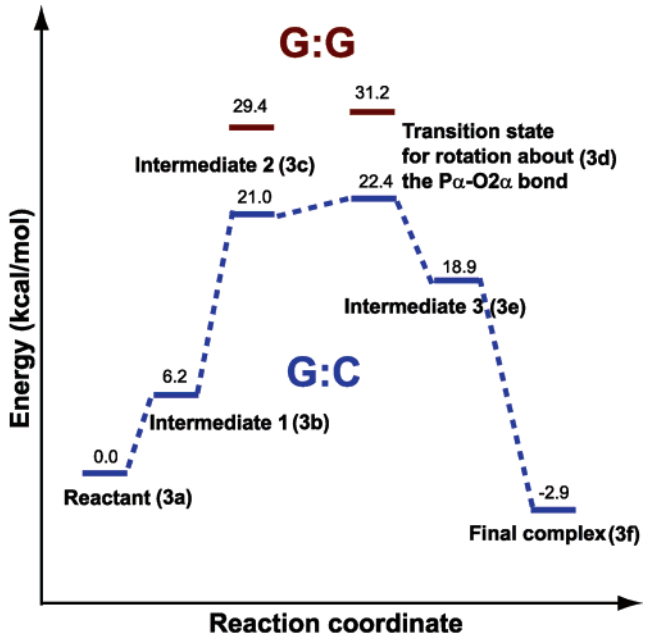


Figure 4. The energy profile for the endpoints and transient intermediates (labeled 3a−f corresponding to Figure 3) identified in reaction pathway 1. The blue and red bars represent the relative energies of the intermediates for the G:C and G:G system, respectively.

the incoming dCTP. Such an initial rearrangement was also found by Bojin and Schlick.[31]

The series of constrained minimizations then defined a pathway, in which the $O3'-P_\alpha$ distance was decreased to promote the forward reaction. A second set of constraints involving appropriate O···H···O distances was also imposed to drive the transfer of the terminal primer O3′H proton to the $O2_\alpha$ atom of dCTP. This leads to intermediate 2, about 15 kcal/mol above the modified reactant structure (intermediate 1) and about 21 kcal/mol above the initial reactant state (see Table 1 and Figure 4). Intermediate 2 involves a pentacoordinated P_α atom surrounded by a trigonal bipyramidal arrangement of coordinating groups with almost equal $O3'-P_\alpha$ and $P_\alpha-O3_\alpha$ bond distances of 2.1−2.2 Å. This structure corresponds to the partially associative intermediate for the phosphoryl transfer reaction (Figure 3c).

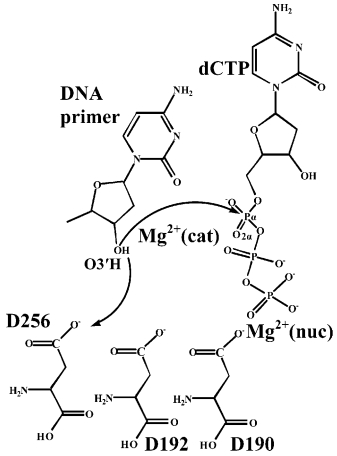
In intermediate 2, the transferred proton still points back to the terminal primer O3′ oxygen atom, and, in a totally unconstrained energy minimization from this configuration, the proton hops back to the primer oxygen and the structure reverts back to the modified reactant (intermediate 1). For this reason, the next stage of the reaction corresponds to rotation about the $P_\alpha-O2_\alpha$ bond so that the proton is directed away from the primer oxygen and toward phosphate oxygen $O3_\alpha$ (Figure 3d and e). This rotation costs about 2 kcal/mol as the O3′···H hydrogen bond is broken (Figure 3d) to produce "transition state" in Table 1. This energy is recovered by formation of another hydrogen bond between the proton and $O3_\alpha$ (Figure 3e) to yield intermediate 3 in Table 1. The critical $O3'-P_\alpha$ and $P_\alpha-O3_\alpha$ bond lengths are not significantly modified during this rotation.

The reaction coordinate then proceeds toward the product. The migrating proton is transferred to $O3_\alpha$ of the leaving pyrophosphate group, and the $O3'-P_\alpha$ bond length decreases to 1.7 Å and the $P_\alpha-O3_\alpha$ distance increases to 3.9 Å. In the final product state, which is about 3 kcal/mol below the initial unmodified reactant state (Table 1 and Figure 4), the $O3'-P_\alpha$ bond has been completely formed leading to extension of the DNA primer chain with the $P_\alpha-O3_\alpha$ bond fully cleaved (promoting elimination of the pyrophosphate group) (Figure 3f). Although the highest point on the potential energy surface for this pathway is the transition state for rotation about the $P_\alpha-O2_\alpha$ bond, the key activation stage corresponds to the initial terminal primer O3′H deprotonation step in concert with nucleophilic attack of the O3′ atom at the P_α center of the incoming dCTP. This reaction pathway resembles that described by Abashkin et al.[28] and Bojin and Schlick,[31] although the highest energy structure on the route identified here, 22 kcal/mol above the reactant, is higher than that identified in the earlier works (∼15 kcal/mol above the reactant). This is primarily a consequence of the initial structural rearrangement of the reactant state, which costs about 6 kcal/mol according to our QM/MM calculations.

Because crystal structures are not available for the mismatched systems, we explored pathway 1 for the G:G system from the endpoint of MD simulations as discussed in the Methods section. QM/MM constrained minimizations were conducted following a similar series of events as described above for the corresponding G:C matched case. The only exception was that, because no hydrogen bond was present for G:G between the terminal primer O3′H atom and O5′ of the incoming dGTP, there was no reactant structural rearrangement and the first step of the reaction corresponds to direct O3′H proton transfer to the $O2_\alpha$ phosphate oxygen atom. The significant distortions of the initial active site structure in the mismatched system imply that the system must overcome larger energy barriers to reduce the critical distances (Mg^{2+}(cat)−OD1−D256, Mg^{2+}(cat)−O3′, Mg^{2+}(cat)−O1A, $O3'-P_\alpha$)[6] to nearly ideal values for the chemical reaction. Consequently, the partially associative intermediate for nucleophilic attack is about 8 kcal/mol above that for the G:C system (Table 1). The highest point energetically on this reaction pathway again corresponds to the $P_\alpha-O2_\alpha$ bond rotation transition state, about 29 kcal/mol above the initial reactant state (Table 1).

Pathway 2. Nucleophilic Attack Initiated by Direct Deprotonation of the Terminal DNA Primer O3′H Proton to Conserved Active Site Asp Residues. QM/MM constrained minimizations were then performed to force the terminal primer O3′H proton to migrate directly to one of the three conserved

Scheme 2. Sketch of Reaction Pathways[a]



[a] Arrows show direct deprotonation of the terminal primer O3′H to active site Asp residues, in conjunction with nucleophilic attack at the P_α center.

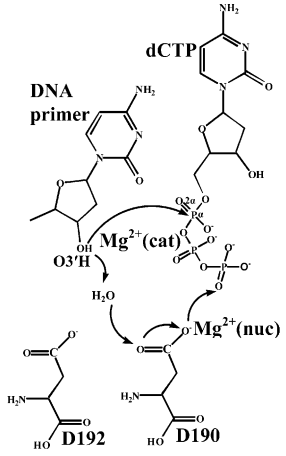
Table 2. Energetics of Structures Located during Pathway 2 from QM/MM Calculations Using MP2/6-311+G(d,p) for the QM Region

	energy [kcal/mol]	
structure	G:C	G:G
reactant	0.0	0.0
intermediate (D256 protonated)	26.3	33.9
intermediate (D192 protonated)	35.1	
intermediate (D190 protonated)	36.4	

active site Asp residues, D190, D192, or D256, to drive the nucleophilic attack at the phosphate center of the incoming nucleotide (Scheme 2). The energetics for these possible pathways is shown in Table 2. For the correct dCTP incoming nucleotide (G:C), initial deprotonation of the terminal primer O3′H proton directly to D256 leads to an intermediate state that is about 26 kcal/mol above the reactant state. Initial proton transfer to either D190 or D192 leads to even higher energy intermediates, about 35 and 36 kcal/mol, respectively, above the reactant structure. The activation energy for pathway 2 is, therefore, significantly higher than that for phosphoryl transfer mediated by initial deprotonation to active site water molecules (pathway 3) as well as by direct proton transfer to the $O2_\alpha$ phosphate oxygen atom (pathway 1). Initial proton transfer to either D190 or D192 also causes structural degradation of the active site in the ensuing intermediate, such that the protonated Asp residue loses coordination to the active site metal ions. Furthermore, in unconstrained minimizations starting from the intermediate located via proton transfer to D256, the proton hops back to the terminal primer O3′ oxygen and the minimized structure resembles that of the original reactant. The corresponding energetics involved for the incorrect dGTP incoming nucleotide (G:G case) for pathway 2 are even higher. For these reasons, we did not investigate such unfavorable, high-energy phosphoryl transfer reaction pathways further. However, as discussed in the next section, indirect proton transfer to the active site Asp residues via mediating water molecules represents a feasible initiating step for the reaction.

Pathway 3. Nucleophilic Attack Initiated by Deprotonation of the Terminal DNA Primer O3′H Proton to Explicit Water Molecules. In this pathway, QM/MM constrained minimizations were conducted to examine the influence of explicit water molecules on the terminal primer O3′H deprotonation pathway. The QM active site region includes four explicit water molecules (Figure 2) that interact with active site atoms, and one of them is also within hydrogen-bonding distance of the terminal primer O3′H group. Proton transfer from the initial reactant to this water molecule, followed by a series of further proton migrations, directs the phosphoryl transfer reaction toward the product state in this pathway (Scheme 3 and Figure 5).

Scheme 3. Sketch of Reaction Pathways[a]



[a] Arrows show proton-transfer steps from the terminal primer O3′H to pyrophosphate, via active site water and D190, in conjunction with nucleophilic attack at the P_α center.

From the original reactant state for the G:C case (Figure 5a), the constrained minimizations involve reduction in the O3′−P_α distance in conjunction with additional constraints to steer the terminal primer O3′H proton transfer toward its hydrogen-bonded water partner. This procedure leads to the detection of intermediate 1, about 15 kcal/mol above the initial reactant state (see Table 2 and Figure 6). Intermediate 1 is consistent with a partially associative nucleophilic attack at the P_α center and involves a trigonal bipyramidal coordination arrangement around P_α with O3′−P_α and P_α−$O3_\alpha$ bond distances both close to 2.0 Å (Figure 5b). The protonated water molecule forms a hydrogen bond to atom OD2 of D192, and the next two steps of the reaction involve successive proton migrations from the water to OD2 of D192 (intermediate 2 in Table 3, Figures 5c and 6) and then subsequently to OD1 of D192 (intermediate 3 in Table 3, Figures 5d and 6). After these proton hops, the P_α−$O3_\alpha$ distance is 2.7 Å, corresponding to a cleaved P_α−$O3_\alpha$ bond, and the O3′−P_α bond distance is 1.9 Å. Finally, the proton is transferred to the $O1_\gamma$ oxygen atom of the leaving pyrophosphate group and the product state is reached, with corresponding O3′−P_α and P_α−$O3_\alpha$ bond distances of 1.7 and 4.0 Å, respectively (Figure 5e). The final state comprises a fully formed P_α−$O3_\alpha$ bond, which translates to extension of the DNA primer chain, and lies about 4 kcal/mol below the initial reactant state (Figure 6). The critical activation step for this pathway (with a barrier of about 15 kcal/mol) corresponds to the initial proton-transfer step from the terminal primer O3′H group to a hydrogen-bonded water molecule, which is driven by reduction in the O3′−P_α distance. This reaction pathway is similar to that reported by Radhakrishnan and Schlick[33] although with a different sequence of proton hopping steps.

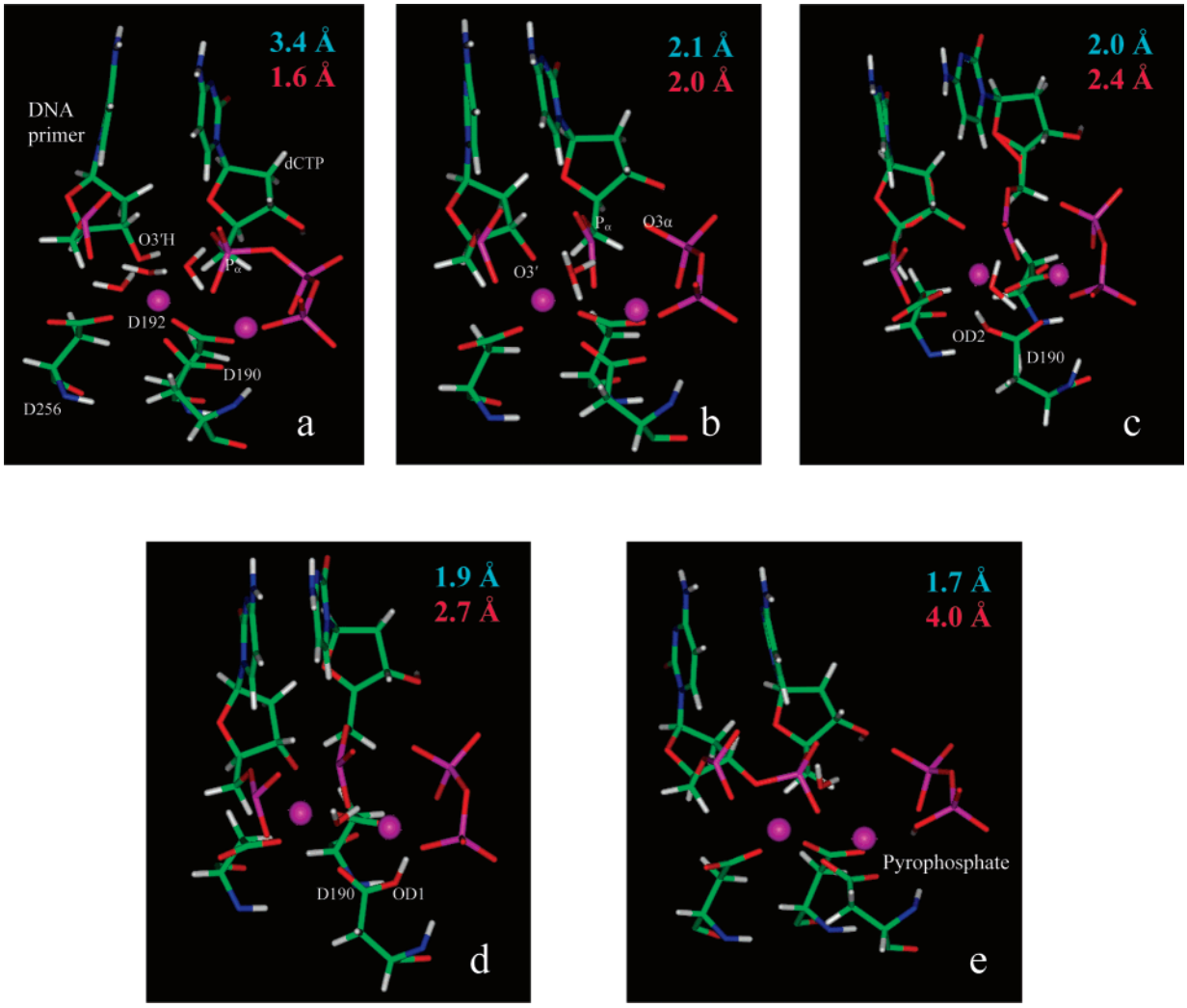


Figure 5. Transient intermediates identified along reaction pathway 3 involving proton transfer from the terminal primer O3′H to $O1_\gamma$ of pyrophosphate, via water and D190. (a) Initial reactant state; (b) intermediate 1; (c) intermediate 2; (d) intermediate 3; and (e) product state with eliminated pyrophosphate. Key distances O3′−P_α and P_α−$O3_\alpha$ are shown for each structure.

Phosphoryl transfer reaction pathway 3 for the incorrect G:G system, initiated from the endpoint of MD simulations, using the QM/MM constrained minimization protocol followed a similar sequence of proton migration steps. As a consequence of the significant structural deviations of the initial active site structure for the G:G mismatch from an ideal reaction-competent state for the two-metal ion process, the activation barrier is almost 7 kcal/mol above that for the corresponding G:C matched case (Table 3 and Figure 6). The highest energy stage of this reaction pathway again corresponds to nucleophilic attack at the P_α center, which prompts the initial deprotonation of the O3′H group to an explicit hydrogen-bonded water molecule.

The succession of proton shifts involving active site waters described in pathway 3 is not unique. Distinct sequences of proton-transfer events to D256 and D192 via active site water molecules are undoubtedly feasible (making the connection to the work of Lin et al.[32] and Radhakrishnan and Schlick[33]), although, in the former case, it is not evident how subsequent hops will direct the proton toward the pyrophosphate leaving group.

Discussion

Our work explored the phosphoryl transfer chemical reaction catalyzed by pol β that leads to extension of the DNA primer chain in a systematic way using QM/MM methods that accounts for both the active site and the enzyme environment. The three possible reaction pathways we explored for both the correct C and the incorrect G incoming nucleotide opposite a template G by QM/MM constrained minimizations all involve an associative nucleophilic substitution mechanism, but differ in the initial deprotonation step that activates the nucleophilic character of the terminal primer O3′ oxygen. Although previous works examined in detail a single possible pathway, the possibilities scrutinized here permit a quantitative energetic comparison of discrete pathways and allow identification of the most favorable possibilities.

Emerging as the energetically most viable for both the G:C and the G:G systems is the pathway involving a series of proton-transfer steps that are intimately coupled with the phosphoryl transfer catalytic process. The initial deprotonation of the O3′H proton to water is the rate-limiting activating step. This migration occurs concomitantly with nucleophilic attack of the terminal primer O3′ oxygen at the P_α center of the incoming nucleotide, leading to the formation of a partially associative intermediate structure with a trigonal bipyramidal arrangement around P_α. This intermediate is stabilized by the two Mg^{2+} ions, in

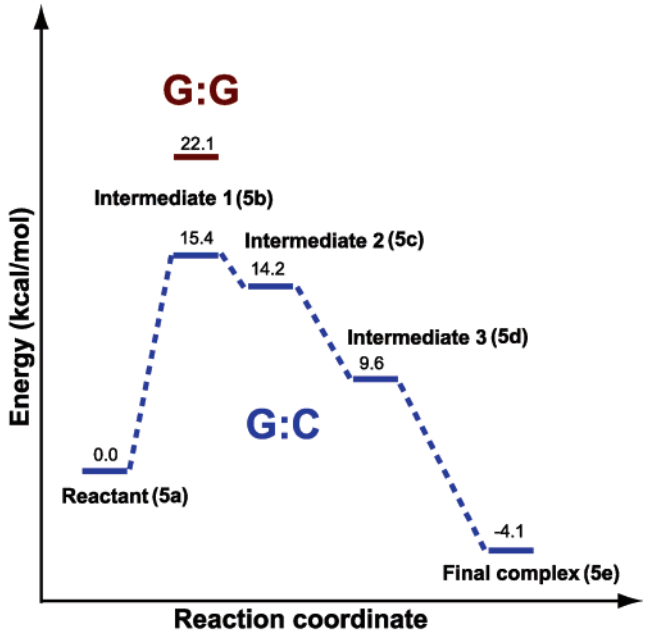


Figure 6. The energy profile for the endpoints and transient intermediates (labeled 5a−e corresponding to Figure 5) identified in reaction pathway 3. The blue and red bars represent the relative energies of the intermediates for the G:C and G:G system, respectively.

Table 3. Energetics of Structures Located during Pathway 3 from QM/MM Calculations Using MP2/6-311+G(d,p) for the QM Region

	energy [kcal/mol]	
structure	G:C	G:G
reactant	0.0	0.0
intermediate 1 (O3′ deprotonation to water and associative intermediate)	15.4	22.1
intermediate 2 (OD2 of D190 protonated)	14.2	21.6
intermediate 3 (OD1 of D190 protonated)	9.6	18.5
product ($O1_\gamma$ of pyrophosphate protonated)	−4.1	−2.5

accordance with the two-metal ion assisted mechanism, as well as by the protonated water that also interacts with D190. As the system evolves toward the product state, further proton hopping events occur, accompanied by full formation of the O3′−P_α bond, leading to extension of the primer chain, and complete cleavage of the P_α−$O3_\alpha$ bond causing elimination of pyrophosphate. A similar series of events were identified for both the matched G:C and the mismatched G:G cases.

The initial rate-limiting step for the G:C case yields an activation energy of about 15 kcal/mol, as compared to about 22 kcal/mol for the G:G mismatch. These values are in close agreement with activation barriers for the overall nucleotide insertion process derived from kinetic measurements, which translate to about 16 kcal/mol for the G:C system and 21 kcal/mol for G:G.[48,49]

The alternative pathways we explored for different proton-transfer events, initiated by either direct deprotonation of O3′H to $O2_\alpha$ of the phosphate group of dNTP or to active site Asp residues, prompts significantly higher activation energies. Interestingly, in the former case (initial deprotonation from O3′H to $O2_\alpha$), this higher energy is probably due to an essential structural rearrangement of the active site prior to the proton abstraction event. In this respect, our work generally supports the notion of critical subtle active site rearrangements (pre-chemistry avenue) prior to the chemical reaction.[23]

Given that the activation barriers for the pol β closing conformational rearrangements prior to chemistry are about 11 kcal/mol for both the G:C case and a related misincorporated G:A system,[20] the rate-limiting step in the overall nucleotide insertion process of pol β for both correct and incorrect incoming nucleotides corresponds to the phosphoryl transfer chemical reaction. The instability of the closed ternary complex relative to the open form for the mismatched case[20] in conjunction with the higher activation barrier for the chemical step (as shown here and in ref 33) translates to a reduction in the rate of incorrect nucleotide insertion as compared to correct insertion.

The QM/MM protocol presented here for probing the pol β catalytic reaction mechanism does not preclude the possibility that other routes for the reaction may be feasible. The reaction energetics will invariably be influenced by the initial structural model chosen, such as initial protonation states assigned for D190, D192, D256, and the terminal phosphate oxygen atoms of the incoming dNTP, as well as the theoretical methodology utilized (MM force field, size of QM region, link atom approach, QM method, etc.). Furthermore, the choice of initial crystal structure or MD snapshot can have profound consequences. For these reasons, the energetically most likely pathway identified here may not be unique, although the agreement of the calculated activation barriers with experiment and agreement with prior works provide a convincing supporting argument.

Our higher predicted activation barrier for the mismatched G:G system as compared to the matched G:C case can be attributed to the more extensively deformed active site structure in the initial reactant state for the mismatch. Basepair stacking rather than formation of Watson−Crick hydrogen bonds is the major cause of such severe active site distortions in the presence of the incorrect incoming nucleotide.[17] As Figure 7 shows, the coordination spheres for both Mg^{2+} ions are severely distorted in the initial G:G active site, as compared to the G:C case, and several key reaction distances also exhibit major disparities. For example, the O3′−P_α, O3′−Mg^{2+}(cat), and Mg^{2+}(cat)−O1A distances are 4.8, 4.3, and 4.3 Å, respectively, for the G:G system versus 3.4, 2.1, and 2.0 Å, respectively, for G:C. These values can be compared to so-called ideal distances for the phosphoryl transfer reaction of <3.0, ∼2.0, and ∼2.0 Å for O3′−P_α, O3′−Mg^{2+}(cat), and Mg^{2+}(cat)−O1A, respectively.[6] The significant deviation of the initial active site geometry from the ideal structure for the chemical reaction in the presence of the mismatched incoming nucleotide discourages incorrect nucleotide insertion and translates to a higher activation barrier for the initial O3′H deprotonation step in the G:G case.

The above argument for rationalizing the reduced rate of nucleotide misincorporation relates to the concept of enzyme active site preorganization,[21,22,50] the pre-chemistry avenue.[17,23] This emerging notion advocates that additional structural rearrangements are critical after the closing conformational

(48) Ahn, J.; Werneburg, B. G.; Tsai, M.-D. *Biochemistry* **1997**, *36*, 1100−1107.
(49) Ahn, J.; Kraynov, V. S.; Zhong, X.; Werneburg, B. G.; Tsai, M.-D. *Biochem. J.* **1998**, *331*, 79−87.
(50) Cannon, W. R.; Benkovic, S. J. *J. Biol. Chem.* **1998**, *273*, 26257−26260.

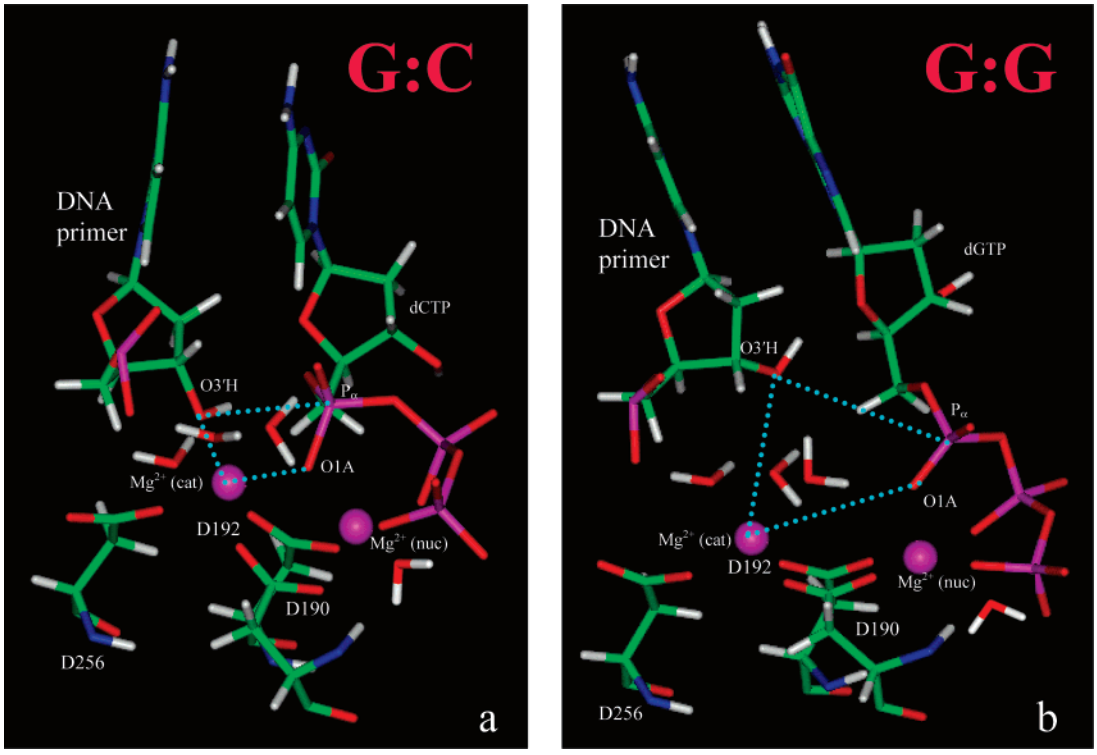


Figure 7. Active site structures in the initial reactant state. (a) G:C system, O3′−P_α, 3.4 Å; O3′−Mg (cat), 2.1 Å; Mg (cat)−O1A, 2.0 Å; (b) G:G system, O3′−P_α, 4.8 Å; O3′−Mg (cat), 4.3 Å; Mg (cat)−O1A, 4.3 Å. Key reaction distances specified above are indicated by blue dotted lines.

change yet prior to the chemical reaction. These active site modifications evolve the system toward a reaction-competent state that is congruent with the subsequent two-metal ion assisted chemical reaction step.[51] The pre-chemistry avenue is postulated to involve a series of subtle stochastic alterations to the metal ion and phosphate coordination distances, as well as to provide further kinetic checkpoints prior to the nucleotide incorporation reaction.[23] For polymerases in which the phosphoryl transfer reaction corresponds to the rate-limiting step and the closing conformational change is relatively fast, fidelity behavior may be interpreted on the basis of the higher energy barrier to drive the misincorporated system toward the ideal geometry for chemistry as compared to that for the correct nucleotide. For the matched case, the enzyme, therefore, provides a preoriented environment (e.g., stable closed state) that stabilizes the onset of the transition state for the chemical reaction.[21,22] For the mismatched systems, the pre-chemistry stochastic rearrangements may need to evolve a partially closed/open state closer to a reaction-competent state.

Specifically, the smaller energy barriers in the matched G:C case as compared to those of the mismatched G:G system are a direct consequence of a more suitable active site preorganization in the former; the better organization translates to a relatively low barrier for the chemical reaction. In G:C, the deprotonation to water occurs in conjunction with nucleophilic attack at the P_α center, and all of the relevant atoms are aligned so that the process involves a relatively low energetic cost. For example, the transfer of the proton from O3′ to a hydrogen-bonded water is facilitated by favorable interactions between the proton and O5′. This water molecule is also metal-bound, which in turn emphasizes how the organization of the active site promotes the reaction, leading to the formation of the pentacovalent P_α intermediate.

For the mismatched G:G case, however, the active site is more open, and the active site atoms are not well predisposed toward the chemical reaction. In particular, the water molecules coordinating the catalytic Mg^{2+} are not hydrogen-bonded to O3′, which is much further away from P_α. This arrangement implies no stabilizing interaction with O5′ during the initial proton transfer to water. Thus, the active site organization in the G:G system does not promote the deprotonation step, which again takes place in concert with nucleophilic attack at P_α. This translates to a higher barrier for the initial rate-limiting step.

This work, which uses high-level QM methods for the reactive core, provides accurate values of the enthalpy (internal energy) of the systems, as well as the transition state-like structures along the reaction coordinates. The main error of the energy calculations stems from neglect of protein reorganization along each pathway, which may influence free energies. Although QM/MM sampling, via dynamics for example, has been conducted recently,[33] such simulations are very limited because of large computational requirements. We expect errors to be about ±3 kcal/mol,[33] without any qualitative altering of our results.

In summary, we have identified several feasible relatively low-energy pathways for the nucleotide incorporation reaction catalyzed by DNA polymerase β using a combination of quantum mechanics and molecular mechanics simulations. In combination with earlier studies concerning the associated conformational changes leading to the "active" closed polymerase configuration prior to the chemical reaction, our work provides key insights into polymerase fidelity mechanisms by identifying critical kinetic gates and checkpoints in the overall nucleotide insertion process. Furthermore, the idea of active site

(51) Lahiri, D. S.; Zhang, G.; Dunaway-Mariano, D.; Allen, K. N. *Science* **2003**, *299*, 2067−2071.

preorganization may offer a unifying concept that can rationalize nucleotide specificity across distinct polymerase families.

Acknowledgment. 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.S. Research described in this Article was supported in part by Philip Morris USA 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: QM/MM formalism (Appendix A) and relative energies of structures determined along reaction pathways 1, 2, and 3 (Table S1, S2, and S3, respectively) using different QM methods (Appendix B). This material is available free of charge via the Internet at http://pubs.acs.org.

JA071533B