Available online at www.sciencedirect.com
ScienceDirect
Biochemical and Biophysical Research Communications 350 (2006) 521–529
BBRC
www.elsevier.com/locate/ybbrc

# Correct and incorrect nucleotide incorporation pathways in DNA polymerase β

Ravi Radhakrishnan [a,*], Tamar Schlick [b]

[a] Department of Bioengineering, University of Pennsylvania, 240 Skirkanich Hall, 210 S. 33rd Street, Philadelphia, PA, USA
[b] Department of Chemistry and Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, NY 10012, USA

Received 19 August 2006
Available online 25 September 2006

**Abstract**

Tracking the structural and energetic changes in the pathways of DNA replication and repair is central to the understanding of these important processes. Here we report favorable mechanisms of the polymerase-catalyzed phosphoryl transfer reactions corresponding to correct and incorrect nucleotide incorporations in the DNA by using a novel protocol involving energy minimizations, dynamics simulations, quasi-harmonic free energy calculations, and mixed quantum mechanics/molecular mechanics dynamics simulations. Though the pathway proposed may not be unique and invites variations, geometric and energetic arguments support the series of transient intermediates in the phosphoryl transfer pathways uncovered here for both the **G:C** and **G:A** systems involving a Grotthuss hopping mechanism of proton transfer between water molecules and the three conserved aspartate residues in pol β's active-site. In the **G:C** system, the rate-limiting step is the initial proton hop with a free energy of activation of at least 17 kcal/mol, which corresponds closely to measured $k_{pol}$ values. Fidelity discrimination in pol β can be explained by a significant loss of stability of the closed ternary complex of the enzyme in the **G:A** system and much higher activation energy of the initial step of nucleophilic attack, namely deprotonation of terminal DNA primer O3′H group. Thus, subtle differences in the enzyme active-site between matched and mismatched base pairs generate significant differences in catalytic performance.


*Keywords:* Polymerase fidelity; Phosphoryl transfer; Quantum mechanics molecular mechanics simulations

The delicate interplay between DNA damage and repair is crucial to the integrity of the genetic imprint. DNA polymerases are central to these processes because of their role in replication as well as excision repair. Understanding the fidelity of DNA replication, repair, or lesion bypass by polymerases requires tracking of the structural and energetic changes involved, including a characterization of elusive transient intermediates, for nucleotide incorporation at the template/primer DNA junction.

A large body of structural, kinetic, and theoretical studies have established that, for some polymerases, a conformational change in the polymerase triggered by a correct nucleotide binding, followed by a catalytic step that accomplishes phosphoryl transfer, are features of the nucleotide incorporation cycle [1]. Closed and open conformations of polymerases have been solved by X-ray crystallography and are often related by a large subdomain motion of the thumb (for left-handed enzymes like pol β) [2–5] between the DNA-bound polymerase and the substrate, in which the correct substrate (for example, d**C**TP opposite **G**, referred to here as the **G:C** system) leads to a "closed," tightly bound enzyme–substrate complex where catalytic groups are aligned as required for proper synthesis. After synthesis or repair, namely chemical incorporation of the nucleotide in DNA by phosphodiester bond formation [6], the enzyme complex returns to its "open" state, to allow translocation for the next cycle of polymerization. The polymerase conformational change is thought to play a key role in organizing the active-site for

* Corresponding author. Fax: +1 215 573 2071.
*E-mail addresses:* rradhak@seas.upenn.edu (R. Radhakrishnan), schlick@nyu.edu (T. Schlick).

0006-291X/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2006.09.059

DNA synthesis and thereby helping the enzyme discriminate correct from incorrect nucleotides.

Because mechanistic details are difficult to obtain experimentally, modeling and simulation, subject to the usual force field and sampling limitations, can help describe events at the atomic-level and offer insights into rate-limiting conformational and chemical steps [7–20]. Recently, we reported context specific pathways (i.e., for correct versus incorrect nucleotide substrate bound to the polymerase/DNA complex) of the closing conformational change in DNA polymerase β prior to the catalytic step [21,22]. Using transition path sampling simulations and a reaction network model, we uncovered strikingly different transition states in pol β's conformational closing for correct d**C**TP versus incorrect d**A**TP incoming nucleotide opposite a template G [21,22]. These context-specific pathways imply different selection processes: While active-site assembly occurs more rapidly with the correct nucleotide and leads to primer extension, the enzyme remains open longer, has a more transient closed state, and forms product more slowly when an incorrect nucleotide is present. These studies provided an atomic-level delineation of the difference in active-site assembly for the correct and incorrect substrate and revealed the roles of subtle side-chain motions in orchestrating these conformational changes. Here we report corresponding pathways for the rate-limiting catalytic step of DNA extension and analyze marked differences in geometries and energetics for the two systems.

Nucleotide incorporation via the ''two-metal-ion'' catalyzed phosphoryl transfer mechanism is believed to be common to several DNA and RNA polymerases [6]. One divalent metal ($Mg^{2+}$) ion (termed nucleotide binding) coordinates the three phosphates of the incoming dNTP and assists the leaving of the pyrophosphate ($PP_i$). A second metal ($Mg^{2+}$) ion (termed catalytic) positioned between the α-phosphate of the incoming nucleotide and the hydroxyl terminus (O3′H) of the DNA primer terminus facilitates the inline nucleophilic attack on the incoming dNTP α-phosphate ($P_\alpha$). Experimental and theoretical evidence suggests that the resulting attack of the nucleophilic oxygen anion (O3′) on the target phosphorous ($P_\alpha$) proceeds via a conformation resembling a trigonal–bipyramidal transition state formed by a pentacovalent phosphorous intermediate [23–25]. Experimental studies of phosphoryl transfer reactions in other enzymes [24,26,27] indicate that the pentacovalent transition state can either be associative or dissociative. Recently, direct structural evidence based on X-ray crystallography for a *partially associative* transition state in β-phosphoglucomutase was reported, in which the critical distance between nucleophilic oxygen and the target phosphorous was 2.0 Å [24]. A similar geometry was also observed in the crystal structure of a group I intron splicing catalytic RNA [25].

DNA polymerase crystal structures as well as modeling and simulation studies suggest that additional activation barriers must be overcome along the reaction pathway after conformational closing rearrangements [20,28]. Warshel et al. concluded that the primary activation event in the chemical step for nucleotide incorporation in T7 polymerase [14] is deprotonation of the O3′ group of the terminal DNA primer with an activation energy of 12 kcal/mol. A similar conclusion with respect to the activation barrier was reached by Webster [29] for the case of β-phosphoglucomutase [24]. In crystal structures where the catalytic site is less well poised for the reaction because of non-ideal distances or conformations in the reactive group, there is an intriguing possibility that the catalytic step itself could serve as a kinetic checkpoint. Such pre-chemistry changes [20,21,28,30] have been proposed to help interpret varying fidelities of DNA polymerases.

The availability of several high-resolution crystal complexes of pol β makes this study of the chemical reaction possible. To date, pioneering QM simulations on pol β fragments have been reported [19,31]. Both proposed associative reaction pathways; the former group proposed an initial step of proton transfer from 3′OH (primer) to $O(P_\alpha)$, and the latter focused on various protonation possibilities and mentioned that additional protons in the active-site are possible, especially D256.

Our QM/MM simulations here extend these studies by including effects arising from entropic contributions, explicit solvation, long-range electrostatics, and coupling to protein backbone motions. Our earlier QM/MM minimization [21] suggested that two crystallographic water molecules could be involved in proton transfer to D256 from the 3′OH primer. Thus, various mechanisms involving proton transfer from water in the active-site or directly from 3′OH to active-site aspartates are all possible and may be dependent on the initial protonation states of the conserved aspartate triad in pol β's active-site and initial model. (See S1.2 for a discussion on protonation states in supplementary material.)

Here we develop a novel protocol involving energy minimizations, short dynamics QM/MM simulations, and quasi-harmonic free energy analysis that extends beyond typical constrained minimization approaches and suggests feasible pathways involving a water/conserved aspartates mediated proton hopping (Grotthuss) mechanism valid for an initial state with unprotonated D192, and D190. The delineated pathways help explain fidelity discrimination between the **G:C** and **G:A** base-pair systems.

## Methods

An overall summary of our system and methods follows, with details provided in the supplementary material.

(1) System preparation (S1.1 and S1.2). We study two solvated complexes: pol β/DNA/d**C**TP (d**C**TP: deoxyribocytosine 5-triphosphate) or the **G:C** system, and pol β/DNA/d**A**TP or the **G:A** system, both derived and modeled from the 1BPY (closed ternary) crystal structure [4]. The solvated system is shown in Fig. S1 and the active-site is depicted in Fig. S2. Our simulations were initiated from well, characterized initial states (and structures) reported in our previous studies [21,22].

(2) Initial equilibration (S1.3). The initial minimization and 1 ns equilibration were performed using classical force fields.
(3) QM/MM simulations (S1.4). In the QM/MM simulations, the active-site of 86 atoms (Fig. S2, supplementary material) is modeled using density functional theory according to a B3LYP functional and a 6–311 G basis set. The rest of the protein, DNA, solvent, and salt is modeled according to a standard empirical force field. The underlying QM/MM Hamiltonian is defined according to the prescription of Reuter et al. [32], with covalent bridges between the QM region and the MM region treated according to the single-link atom protocol.
(4) Path quenching: Constrained dynamics followed by unconstrained minimizations (S1.5). The phosphoryl transfer reaction was driven to completion by performing constrained molecular dynamics implemented in CHARMM by employing two harmonic restraints with force constants of $K = 300$ kcal/mol/Å$^2$ along the O3′–$P_\alpha$ and O3′–Catalytic $Mg^{2+}$ distances. (The product state was not modeled *a priori*). Starting from 50 different initial configurations uniformly sampled from the constrained trajectories, we performed a series of energy minimizations. Through a structural analysis of the resulting minimized configurations, we identified 4 intermediate states in addition to the reactant and product states.
(5) Quasi-harmonic free energy analysis (S1.6). Born–Oppenheimer MD trajectories for 10 ps each were harvested starting from each of the 6 states in both **G:C** and **G:A** systems (Figs. 1 and 2: a–f) using constant-temperature MD simulations; the states were found to be stable in the 10 ps timeframe. Principal component analysis was performed on the MD trajectories, and the conformational entropies for the intermediate states were calculated using the quasi-harmonic analysis.

Standard limitations of force fields in biomolecular dynamics and electronic structure methods in enzyme catalysis are well documented and appreciated. Of relevance to this work are the neglect of quantum mechanical tunneling, especially in the proton transfer steps, and the omission of polarization functions in the basis set describing the electronic wave functions, especially for the P atom. The savings in CPU time by the omission of polarization functions (i.e., a 6–311G** basis set) are approximately a factor of 2. The potential effect of omitting polarization is destabilization of the perfect trigonal bipyramidal geometry around the central phosphorous, but this was not observed in our simulations: An in-line attack angle of 179° to 180° and the three 120° angles among the valence bonds of the phosphorous and the oxygens were always preserved.

The total CPU time required for the project was about 15 months on a 64-processor, 3.2 GHz Linux system. QM/MM minimization and dynamics runs of 1 ps (1000 steps) required about 18 h on 32 processors of the 3.2 GHz Linux cluster.

## Results

For the **G:C** system, many of the relaxed (energy minimized) intermediate states from the 50 different configurations of the restrained dynamics trajectory resulted in similar active-site geometries. In total, we classified the 50 relaxed configurations into 6 states.

Structural analysis of the intermediate (transient and metastable) states in the **G:C** system revealed a Grotthuss pathway [33] for proton transfer from the attacking

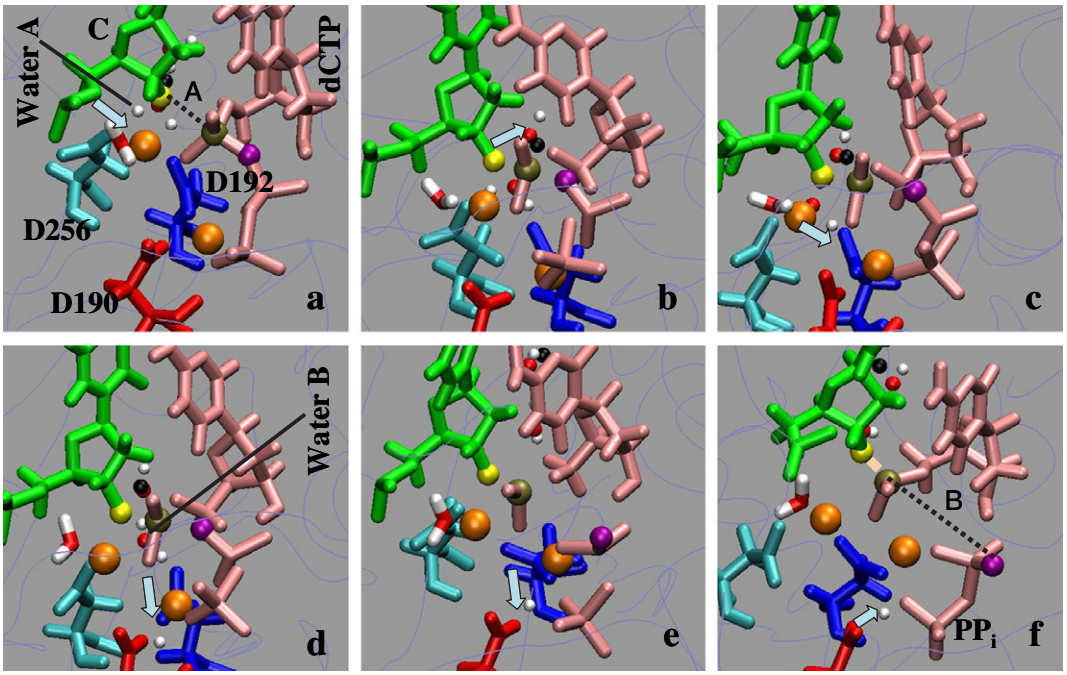



Fig. 1. Captured reaction intermediates for pol β's phosphoryl transfer in the **G:C** system. Key distances A = O3′–$P_\alpha$ and B = $P_\alpha$–O3A are provided: (a) reactant state, A = 2.90 Å, B = 1.7 Å; (b) first intermediate, A = 1.94 Å, B = 1.97 Å; (c) second intermediate, A = 1.73 Å, B = 2.0 Å; (d) third intermediate, A = 1.73 Å, B = 2.41 Å; (e) fourth intermediate, A = 1.70 Å, B = 5.53 Å; (f) product, A = 1.70 Å, B = 5.8 Å. The colors represent: cyan (D256), red (D190), blue (D192), pink (d**C**TP), green (CYT: terminal DNA primer), black (the O3′–H-proton), yellow (the O3′ oxygen, attacking nucleophile), tan (central phosphorous), purple (leaving O3A oxygen), and orange ($Mg^{2+}$). The oxygens and hydrogens of water molecules are in red and white, respectively. The arrows denote the location and direction of proton hop. The water molecules A and B are present in all figure panels and are easily identified by tracking the corresponding oxygen atoms.The hydrogen atoms of these waters are also depicted. However, these constitute the migrating proton. For example, the water A abstracts the O3′H- proton (c), and loses its own proton to water B (d), which in turn loses its proton to $PP_i$ (f). (For interpretation of the references to colour in this figure legend the reader is referred to the web version of this article.)

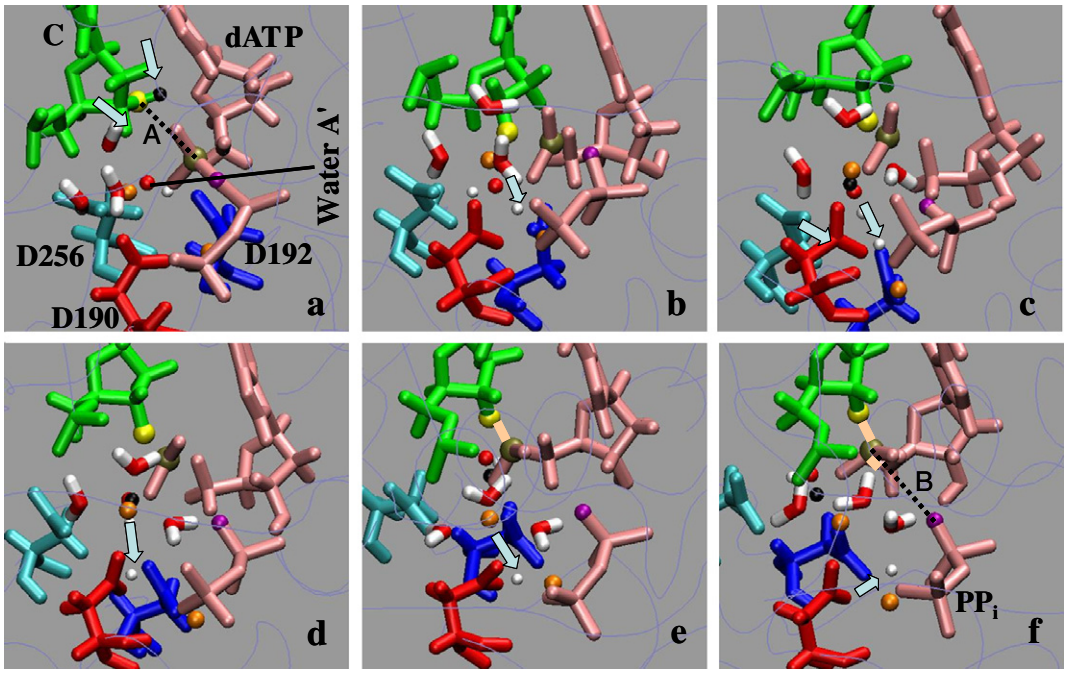
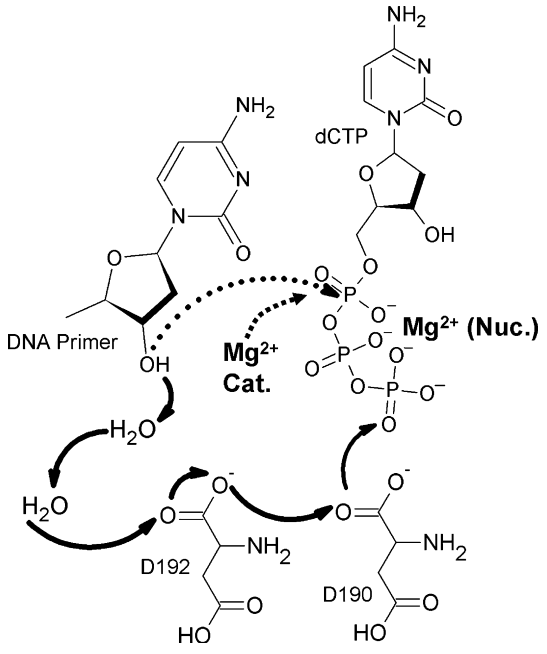



Fig. 2. Captured reaction intermediates for pol β's phosphoryl transfer in the **G:A** system. Key distances A = O3′–$P_\alpha$ and B = $P_\alpha$–O3A are provided: (a) reactant state, A = 6.80 Å, B = 1.69 Å; (b) first intermediate, A = 2.20 Å, B = 1.80 Å; (c) second intermediate, A = 1.73 Å, B = 3.76 Å; (d) third intermediate, A = 1.69 Å, B = 3.92 Å; (e) fourth intermediate, A = 1.68 Å, B = 4.03 Å; (f) product, A = 1.68 Å, B = 4.2 Å. The color scheme is identical to Fig. 1 except the incoming nucleotide (in pink) is d**A**TP. (For interpretation of the references to colour in this figure legend the reader is referred to the web version of this article.)

nucleophilic hydroxyl to the product (Fig. 1 and Scheme 1): each metastable state corresponds to a link in the chain for the transfer of a proton from the initial reactant state (Fig. 1a) to the final product state (i.e., a proton is transferred to the O3G oxygen of the $PP_i$, Fig. 1f). In the initial state (Fig. 1a), the O3′–H proton of the terminal DNA primer is near the O5′ of incoming d**C**TP (2.5–3.0 Å). The activation event for the phosphoryl transfer coincides with a reduction in both the O3′–$P_\alpha$ and O3′–catalytic $Mg^{2+}$ distances, and is accompanied by the transfer of the O3′H-proton to a bridging water molecule (water A), Fig. 1b. This proton transfer is mediated by the stabilization of a hydrogen bond involving the departed proton (now part of the molecule Water A) and the O5′ of d**C**TP. The bridging water molecule (water A) is also hydrogen bonded to a neighboring hydrogen-bonded water molecule (water B), which itself forms two hydrogen bonds to the oxygens of the D256 (OD2) and D192 (OD1), respectively, in a bridging configuration. In Fig. 1b, the O3′–$P_\alpha$ distance and the $P_\alpha$–O3A distances are both ≈1.9 Å and the geometry around the $P_\alpha$ atom is the classic trigonal–bipyramidal configuration with a perfect in-line attack angle of 180° formed by O3′–$P_\alpha$–O3A. The next two intermediate states correspond to the hopping of the proton from water B to the OD1 oxygen of D192 (Fig. 1c) and to the OD2 oxygen of D192 (Fig. 1d). The latter hop coincides with the cleaving of the $P_\alpha$–O3A bond, which results in a bond distance of 2.41 Å in Fig. 1d. (The O3′–$P_\alpha$ distance at this configuration is 1.70 Å.) This sequence of proton hops continues to yield the final two intermediates, first to the OD1 oxygen of D190 (Fig. 1e) and finally to the O1G oxygen of the $PP_i$ (Fig. 1f). In this final configuration, the O3′–$P_\alpha$ and $P_\alpha$–O3A bond distances are 1.70 Å and 5.86 Å, respectively.




Scheme 1. Mechanism of concerted proton-hops during phosphoryl transfer in pol β. Solid arrows indicate the migration path of the proton and the dotted arrows represent the nucleophilic attack.

Note that a key feature of the observed pathway is that a thermal fluctuation that lowers the O3′–$P_\alpha$ distance and simultaneously maintains an O3′–$Mg^{2+}$ distance of about 2 Å is sufficient to drive the O3′–H proton to dissociate. The free energies (see Table 1) point to this as a slow step. After this initial dissociation, the specific context (i.e., positions of water molecules, protonation states of the aspartates) determines the trajectory of the proton-hops and its endpoint. However, except for the O3′–H proton dissociation in the first step, every hop is a Grotthuss-type hop and we expect very low (thermal) activation barrier and fast timescales for the subsequent hops.

Table 1
Energetics of Reaction Intermediates

| State | Energy, $\langle E\rangle$ kcal/mol | Entropy, $S$ J/mol/K | $F = \langle E\rangle - TS$ kcal/mol |
|---|---|---|---|
| ***G:C** System* | | | |
| Reactant<br>A = 2.90 Å, B = 1.7 Å | 0 ± 2.5 | 17.8 ± 2 | −1.2 ± 2.5 |
| Intermediate 1<br>A = 1.94 Å, B = 1.97 Å<br>O3′ deprotonated | 18 ± 2.5 | 27.9 ± 2 | 16 ± 2.5 |
| Intermediate 2<br>A = 1.73 Å, B = 2.0 Å<br>D192 protonated | 11 ± 2.5 | 36.5 ± 2 | 8.3 ± 2.5 |
| Intermediate 3<br>A = 1.73 Å, B = 2.41 Å<br>D192 protonated | 15 ± 2.5 | 34.3 ± 2 | 12.5 ± 2.5 |
| Intermediate 4<br>A = 1.70 Å, B = 5.53 Å<br>D190 protonated | 5 ± 2.5 | 33.1 ± 2 | −2.6 ± 2.5 |
| Product<br>A = 1.70 Å, B = 5.8 Å<br>$PP_i$ protonated | −1 ± 2.5 | 33.2 ± 2 | −3.4 ± 2.5 |
| ***G:A** System* | | | |
| Reactant<br>A = 6.80 Å, B = 1.69 Å | 0 ± 2.5 | 31.8 ± 2 | −2.3 ± 2.5 |
| Intermediate 1<br>A = 2.20 Å, B = 1.80 Å<br>O3′ deprotonated | 21 ± 2.5 | 33.0 ± 2 | 18.6 ± 2.5 |
| Intermediate 2<br>A = 1.73 Å, B = 3.76 Å<br>D192 protonated | 4 ± 2.5 | 33.7 ± 2 | 1.6 ± 2.5 |
| Intermediate 3<br>A = 1.69 Å, B = 3.92 Å<br>D190 protonated | 15 ± 2.5 | 33.7 ± 2 | 12.6 ± 2.5 |
| Intermediate 4<br>A = 1.68 Å, B = 4.03 Å<br>D190 protonated | 3 ± 2.5 | 38.7 ± 2 | 0.2 ± 2.5 |
| Product<br>A = 1.68 Å, B = 4.2 Å<br>$PP_i$ protonated | 0 ± 2.5 | 33.4 ± 2 | −2.4 ± 2.5 |

Here, $\langle E\rangle$ is the average energy, $S$ is the quasi-harmonic entropy, and F is the Helmholtz free energy. The different states correspond to those depicted in a–f of Figs. 1 and 2 for **G:C** and **G:A** systems, respectively; A = O3′–$P_\alpha$ distance and B = $P_\alpha$–O3A distance.

In the reactant state, D192 is relatively far from O3′–H of the terminal DNA primer, so its protonation state may not affect the O3′–H dissociation drastically. The effect of protonated D190 may be more significant. Still, we note that even though D192 and D190 are deprotonated initially, they undergo protonation in the pathway described (by abstracting the O3′–H proton through water molecules).

Based on a similar protocol of restrained dynamics, path quenching, and structural analysis of the intermediates, the pathway for the mismatched **G:A** system reveals a similar sequence of successive proton-hops mediating phosphoryl transfer. However, the geometry of the initial (reactant) state is much less ideal in comparison to the two-metal-ion geometry with O3′–$P_\alpha$ distance at 6.87 Å (Fig. 2a) versus 2.9 Å for **G:C**, Fig. 1a. The next intermediate is defined by proton transfer from the O3′ hydroxyl to a bridging water molecule (water A′) followed by transfer of one of the hydrogens from water A′ to the OD2 oxygen of D192 (Fig. 2b). The corresponding O3′–$P_\alpha$ and $P_\alpha$–O3A distances are 2.2 and 1.8 Å, respectively, with a perfect trigonal–bipyramidal alignment and an in-line attack geometry around the $P_\alpha$ of the incoming d**A**TP. The next intermediate state involves a proton hop to the OD1 oxygen of D192, increasing the $P_\alpha$–O3A distance to 3.76 Å and indicating cleavage of the $PP_i$ group (Fig. 2c). Two successive hops follow—first to OD1 of D190 (Fig. 2d) and then to OD2 of D190 (Fig. 2e)—before the proton is transferred to the O1G oxygen in the $PP_i$, leading to the product state (Fig. 2f).

The six states of the **G:C** system (Fig. 1) were each subjected to a constant-temperature QM/MM dynamics for 10 ps (Section S1.4, supplementary material), before performing quasi-harmonic analysis on the associated trajectories (Section S1.6). A Movie file (Movie 1, S2.1, supplementary information) animates these dynamics trajectories. For each trajectory, the quasi-harmonic frequencies were calculated by diagonalizing the mass-weighted covariance matrix: The principal modes of quasi-harmonic vibrations were obtained by calculating the eigenvectors corresponding to the eigenvalues (see Fig. S3 a,b, supporting information, for corresponding eigenvalue spectra of intermediate states in the **G:C** and **G:A** systems). The trajectories of the systems were projected along these eigenvectors to visualize the principal modes of motion. The animations of the top five dominant modes of motion for the intermediate states of the **G:C** (Movie 2, S2.2) and **G:A** (Movie 3, S2.3) systems allow dissection of key motions. Their relevance to the mechanism is further supported by the fact that reactive distance between O3′ and $P_\alpha$ is coupled (i.e., displays significant fluctuation) with all five dominant modes of the reactant states for both the **G:C** and **G:A** systems.

Specifically, in the **G:C** system, the dominant active-site motion in the reactant state does not perturb the active-site geometry in any significant manner: The relative distances and orientations of the critical atoms are maintained. This demonstrates how the enzyme exercises tight control at the

active-site by possibly suppressing involvement of the incoming and template bases in the usual DNA global breathing modes. In passing to the first intermediate state and subsequent states (including the product state), several cooperative dynamical modes are activated: In intermediate state 1, these involve a wobble motion of the incoming and the template base while, in intermediate state 2 through product, these involve cooperative motions between the two bases analogous to global writhing and twisting of the DNA, as well as translocation of the two bases. The dynamic characteristics unique to the ground state compared to the intermediate states for the **G:C** system are reflected by a sharp increase in the relative conformational entropy (see Table 1).

For the **G:A** system, the modes in intermediate state 2 are similar to those of the **G:C** system. However, the ground (reactant) state also displays the wobble motion of the template and the incoming bases possibly because the mispair association at the active-site is not as strong as in the correct base pair. The subsequent states (state 3 to product) display bending and twisting but not writhing. The absence of writhing may slow down DNA translocation for the subsequent cycle of nucleotide incorporation following the mispair incorporation. In contrast to the trend of entropy increase for **G:C** (activation from initial state to next intermediat1e), no such rise occurs for **G:A** (Table 1).

Table 1 reports the average energies (calculated from the 10 ps Born–Oppenheimer MD simulations), the quasi-harmonic conformational entropies, and the resulting Helmholtz free energies for the reaction intermediates of phosphoryl transfer in pol β's **G:C** and **G:A** incorporation. All values are approximate due to the large fluctuations. The entropy, in particular, approximates the true quasi-harmonic entropy because the principal modes were calculated by including only active-site degrees of freedom. Thus, because we ignore the atoms outside the QM region in our analysis of the covariance matrix, this quantity is closer to an index of active-site plasticity. Therefore, the Helmholtz free energy estimates have a resolution of >1 kcal/mol. A potential of mean force may map the free energy profile more accurately [34], but this entails an order-of-magnitude increase in computational cost.

## Discussion

The concept of coupled proton transfer and enzyme catalysis is common to many mechanisms (for a review, see [35]). This coupling may define a viable pathway for regulating the phosphoryl transfer reaction of polymerase β from our chosen initial state (unprotonated aspartates). Proton transfer from water is a mechanism that complements other possibilities (such as involving initial deprotonations of O3′–H to $OP_\alpha$ or to D256). The captured intermediates represent successive steps in such a proton transfer cascade, and the overall process conforms to the Grotthuss hopping mechanism [33]; proton translocation for each hop is between 0.7 and 1.5 Å but the overall translocation is 4–6 Å. Significantly, the proton hop cascade for both the **G:C** and **G:A** systems occurs simultaneously with the phosphoryl transfer. While true transition states (i.e., saddle points in the free energy surface) could not be identified here, the tight coupling between the key reaction distances of O3′–$P_\alpha$ and $P_\alpha$–O3A distances and the proton jumps support this mechanism.

Specifically, fluctuations in O3′–$P_\alpha$ and O3′–catalytic $Mg^{2+}$ distances prepare for the inline nucleophilic attack and trigger the first proton jump, from O3′–H to $H_3O^+$. The bridging of a water molecule between the OD2 of D256 and OD1 of D192 stabilizes the transition state in a perfect trigonal–bipyramidal, inline attack configuration. The cleaving of the $P_\alpha$–O3A bond is completed with the proton migration to OD1 of D190. Moreover, the average energies and conformational entropies of the intermediates (Table 1) suggest that the rate-limiting step along the phosphoryl transfer pathway involves initial activation that assembles the trigonal–bipyramidal geometry, followed by the step corresponding to the cleaving of the $P_\alpha$–O3A bond.

We emphasize that the path quenching protocol employed here for tracking the transient states does not guarantee a unique pathway, so other pathways connecting the reactants and products (such as by different initial protonation states as discussed above) are possible. Nonetheless, the physical relevance of our pathways for phosphoryl transfer is supported by the following observations.

(1) The rate-limiting hop between the intermediate states in the **G:C** system (see Table 1) corresponds to the initial activation step with a free energy of activation of at least 17 kcal/mol (or 27 $k_BT$). This value agrees with the experimentally determined overall barrier of nucleotide incorporation (26–28 $k_BT$) from $k_{pol}$ measurements [36–38]. The corresponding free energy of initial activation for **G:A** is at least 21 kcal/mol (35 $k_BT$).
(2) Recently, direct structural evidence supporting a *partially associative* transition state in β-phosphoglucomutase was reported, with a critical distance between the nucleophilic oxygen and the target phosphorous of 2.0 Å [24]. Our proposed mechanisms as well as structural intermediates of phosphoryl transfer for both correct and misincorporation in pol β support such a partially associative transition state [27].
(3) The two main barriers along our pathway have major biochemical significance; the first (reactant to intermediate 1) represents nucleophilic attack, and the second (intermediate 2 to intermediate 3) represents product dissociation.
(4) The mechanism that emerges is consistent with a pentacovalent phosphorous intermediate and an inline nucleophilic attack which lead to an inversion of the stereo-conformation about the central

phosphorous. All facts are consistent with the Sawaya structure [4], as well as with other phosphoryl transfer mechanisms [6,39].

(5) Given that this initial hopping step is slow (Table 1), the direction for driving the reaction forward (as inferred from comparing the structures in Fig. 1a and b) appears to be the catalytic $Mg^{2+}$–O3′ distance and the O3′–$P_\alpha$ distance. This result is consistent with the observed fluctuations of the O3′–$P_\alpha$ distance from the principal component analysis (PCA) of the active-site degrees of freedom in the reactant state. (Namely, the first dominant mode involves fluctuations in O3′–$P_\alpha$ and O3′–catalytic $Mg^{2+}$ distances, see Fig. S4). Moreover, applying constraints to these distances triggers the deprotonation of O3′H hydroxyl of the primer terminus and the proton transfer to the bridging water molecule, resulting in the first intermediate in Fig. 1. Thus, the static structures in our pathway and their dynamic attributes are consistently linked. In this respect, a two-dimensional umbrella sampling in a related phosphoryl transfer study involving the self-cleavage reaction of a hammerhead ribozyme has also revealed that this RNA catalyzed phosphoryl transfer reaction can be driven using the distance between the attacking nucleophilic oxygen and the central phosphorous, and the distance between the catalytic metal and the nucleophilic oxygen [Radhakrishnan, unpublished]. Thus, the agreement with an analogous reaction landscape provides additional evidence in support of the pol β's catalytic pathway reported here. Often, the actual reaction coordinate for phosphoryl transfer can be highly complex, involving not only the active-site degrees of freedom, but also long-range interactions and solvent degrees of freedom [40]. A full description of such reaction coordinates requires other approaches, such as transition path sampling [41].

It is also interesting to note that the increase in conformational entropy in the **G:C** case triggered by the initial activation (Table 1) reveals a dynamic activation mechanism as the system translocates from the initial reactant state to the first intermediate state in the pathway. This is evident from the animations provided as supplementary information. This conformational entropic activation likely stabilizes the transition state and is also reminiscent of the broadening in the 2-dimensional NMR spectrum in the experiment of Bose-Basu et al. when dNTP is added [42]. For the **G:A** system, an entropic activation corresponding to the first hop with subsequent entropic stabilization of the transition state is not present because of the large entropy associated with the reactant state, possibly due to the wobbling motion of the base pairs at the active-site.

Thus, a key difference between the matched (**G:C**) and mismatched (**G:A**) systems is the far more distorted active-site in the initial reactant state for the latter. The reaction distances of O3′–$P_\alpha$ and O3′–catalytic $Mg^{2+}$ are 6.87 and 6.5 Å, respectively, for **G:A** compared to 2.90 and 2.23 Å, respectively, for the **G:C** system. This large geometric difference translates to a larger free energy of activation for the first step of deprotonation of O3′H hydroxyl of the DNA primer terminus (Table 1). Moreover, analyses of the dominant modes of motion at the active-site suggest a slower DNA translocation step for the mismatched system, which could further hamper the subsequent nucleotide incorporation. Taken together with our analysis of the closing conformational change pathways for **G:C** [22] and **G:A** [21], the systematic differences between the correct and mismatched systems imply a slowing down of the rate of misincorporation compared to the correct incorporation due to a combination of loss in stability of the closed ternary complex of the enzyme and a higher activation of the initial step of nucleophilic attack. Our findings also suggest further experiments. Mutant systems that will have a profound effect on catalytic activity involve one of the conserved aspartates (D190, D192, or D256). Specifically, the loss of activity in these mutants will reflect three factors.

First, crucial interactions between the aspartates and the $Mg^{2+}$ ions will perturb the stability of the closed complex.

Second, the concerted proton transfer pathway proposed in this present study is expected to be affected. Additional spectroscopy and isotope substitution experiments involving deuterium instead of hydrogen in water could help detect the sensitivity of the proton-hops to the overall reaction mechanism. In this respect, substitution of the oxygen moieties aiding the Grotthuss hops by elements showing altered sensitivity to hydrogen bonding (e.g., thio-elemental substitution of OD2 oxygen in D192) might also help discern the significance of the proton-hops.

Third, crucial differences in the nature of the DNA motion at the active site as revealed by our PCA analyses suggest that force spectroscopy experiments similar to those reported for T7 [43,44] can reveal critical and useful information about the nature of coupling between the DNA and polymerase modes of motion at the (rate-limiting) transition state of the overall nucleotide incorporation reaction.

Finally, we mention that pol β's regulation of nucleotide incorporation by a rate-limiting chemical step for both correct and incorrect systems can be explained by the fact that this catalytic step is tightly regulated via a multi-switch cascade controlled by the active-site geometry, including relative position of the proton being transferred. This tight regulation in the catalytic step of pol β may provide an additional kinetic checkpoint in the nucleotide incorporation cycle.

## Acknowledgments

We thank Dr. P. Sherwood for his help in compiling the CHARMM-GAMESS interface, Dr. M. Karplus for allowing us to use the CHARMM program, and Dr. N. Glykos for his help with the PCA. R. Radhakrishnan

acknowledges a grant from the Whitaker Foundation and computational resources provided by the National Center for Supercomputing Alliance under Grant DAC 1103423 and National Partnership for Advanced Computational Infrastructure under Grant MCB060006. T. Schlick acknowledges support by NSF Grant MCB-0316771, NIH Grants R01 GM55164 and 1R01ES012692, donors of the American Chemical Society Petroleum Research Fund, and Philip Morris International. Computing facilities provided by the N.C.S.A. supercomputing center are highly appreciated.

## Appendix A. Supplementary data

Detailed description of methods and system preparation. Movies capturing the dynamics of the phosphoryl transfer mechanism for the **G:C** system, depicting the top five principal modes of the intermediate states of the **G:C** and **G:A** systems, figures showing the eigenvalue spectrum and a covariance plot from PCA. Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bbrc.2006.09.059.

## References

[1] T.A. Kunkel, K. Bebenek, DNA replication fidelity, Annu. Rev. Biochem. 69 (2000) 497–529.
[2] J.R. Kiefer et al., Visualizing DNA replication in a catalytically active Bacillus DNA polymerase crystal, Nature 391 (6664) (1998) 304–307.
[3] Y. Li, S. Korolev, G. Waksman, Crystal structures of open and closed forms of binary and ternary complexes of the large fragment of *Thermus aquaticus* DNA polymerase I: Structural basis for nucleotide incorporation, EMBO J. 17 (1998) 7514–7525.
[4] M.R. Sawaya et al., Crystal structures of human DNA polymerase beta complexed with gapped and nicked DNA: evidence for an induced fit mechanism, Biochemistry 36 (1997) 11205–11215.
[5] S. Doublie et al., Crystal structure of a bacteriophage T7 DNA replication at 2.2 Å resolution, Nature 391 (6664) (1998) 251–258.
[6] T.A. Steitz et al., A unified polymerase mechanism for nonhomologous DNA and RNA polymerases, Science 266 (1994) 2022–2025.
[7] R. Elber, J. Meller, R. Olender, Stochastic path approach to compute atomically detailed trajectories: application to the folding of C peptide, J. Phys. Chem. B 103 (1999) 899–911.
[8] B. Isralewitz et al., Steered molecular dynamics investigations of protein function, J. Mol. Graph. Model 19 (2001) 13–25.
[9] M.A. Young et al., Dynamic coupling between the SH2 and SH3 domains of c-Src and Hck underlies their inactivation by C-terminal tyrosine phosphorylation, Cell 105 (2001) 115–126.
[10] T. Simonson, G. Archontis, M. Karplus, Free energy simulations come of age: protein–ligand recognition, Acc. Chem. Res. 35 (2002) 430–437.
[11] J. Trylska et al., Ribosome motions modulate electrostatic properties, Biopolymers 74 (2004) 423–431.
[12] N. Rega et al., Hybrid ab initio empirical molecular dynamics: combining the ONIOM scheme with the atom-centered density matrix propagation (ADMP) approach, J. Phys. Chem. B 108 (2004) 4210–4220.
[13] A. Banerjee et al., Structure of a repair enzyme interrogating undamaged DNA elucidates recognition of damaged DNA, Nature 434 (7033) (2005) 612–618.
[14] J. Florian, M.F. Goodman, A. Warshel, Computer simulation of the chemical catalysis of DNA polymerases: discriminating between alternative nucleotide insertion mechanisms for T7 DNA polymerase, J. Am. Chem. Soc. 125 (27) (2003) 8163–8177.
[15] Y. Zhang, J. Kua, J.A. McCammon, Role of the catalytic triad and oxyanion hole in acetylcholinesterase catalysis: an ab initio QM/MM study, J. Am. Chem. Soc. 124 (2002) 10572–10577.
[16] H. Guo et al., Substrate conformational transitions in the active-site of chorismate mutase: their role in the catalytic mechanism, Proc. Natl. Acad. Sci. USA 98 (16) (2001) 9032–9037.
[17] Q. Cui, M. Karplus, Triosephosphate isomerase: a theoretical comparison of alternative pathways, J. Am. Chem. Soc. 123 (10) (2001) 2284–2290.
[18] L. Yang et al., Polymeraseβ simulations suggest that Arg258 rotation is a slow step rather than large subdomain motion per se, J. Mol. Biol. 317 (2002) 651–671.
[19] R.C. Rittenhouse et al., Characterization of the active-site of DNA polymerase beta by molecular dynamics and quantum chemical calculation, Proteins 53 (2003) 667–682.
[20] K. Arora et al., Mismatch-induced conformational distortions in polymeraseβ support an induced-fit mechanism for fidelity, Biochemistry 44 (40) (2005) 13328–13341.
[21] R. Radhakrishnan, T. Schlick, Fidelity discrimination in DNA polymeraseβ: differing closing profiles for a mismatched (**G:A**) versus matched (**G:C**) base pair, J. Am. Chem. Soc. 127 (2005) 13245–13253.
[22] R. Radhakrishnan, T. Schlick, Orchestration of cooperative events in DNA synthesis and repair mechanism unraveled by transition path sampling of DNA polymeraseβ's closing, Proc. Natl. Acad. Sci. USA 101 (16) (2004) 5970–5975.
[23] T.A. Steitz, J.A. Steitz, A general 2-metal-ion mechanism for catalytic RNA, Proc. Natl. Acad. Sci. USA 90 (14) (1993) 6498–6502.
[24] S.D. Lahiri et al., The pentacovalent phosphorus intermediate of a phosphoryl transfer reaction, Science 299 (5615) (2003) 2067–2071.
[25] M.R. Stahley, S.A. Strobel, Structural evidence for a two-metal-ion mechanism of group I intron splicing, Science 309 (5740) (2005) 1587–1590.
[26] K. Kim, P.A. Cole, Kinetic analysis of a protein tyrosine kinase reaction transition state in the forward and reverse directions, J. Am. Chem. Soc. 120 (1998) 6851–6858.
[27] A.S. Mildvan, Mechanisms of signaling and related enzymes, Proteins 29 (1997) 401–416.
[28] Radhakrishnan, R., et al., Regulation of DNA repair fidelity by molecular checkpoints: "Gates" in DNA polymeraseβ's substrate selection. Biochemistry (in press).
[29] C.E. Webster, High-energy intermediate or stable transition state analogue: theoretical perspective of the active-site and mechanism of β-phosphoglucomutase, J. Am. Chem. Soc. 126 (2004) 6840.
[30] C.M. Joyce, S.J. Benkovic, DNA polymerase fidelity: kinetics, structure, and checkpoints, Biochemistry 43 (45) (2004) 14317–14324.
[31] Y.G. Abashkin, J.W. Erickson, S.K. Burt, Quantum chemical investigation of enzymatic activity in DNA polymeraseβ. a mechanistic study, J. Phys. Chem. B 105 (2001) 287–292.
[32] N. Reuter et al., Frontier bonds in QM/MM methods: a comparison of different approaches, J. Phys. Chem. A. 104 (8) (2000) 1720–1735.
[33] N. Agmon, The Grotthuss mechanism, Chem. Phys. Lett. 244 (1995) 456–462.
[34] H. Hu, W.T. Yang, Dual-topology/dual-coordinate free-energy simulation using QM/MM force field, J. Chem. Phys. 123 (4) (2005) 041102.
[35] S.J. Benkovic, S. Hammes-Schiffer, Biochemistry: enzyme motions inside and out, Science 312 (5771) (2006) 208–209.
[36] J. Ahn, B.G. Werneburg, M.-D. Tsai, DNA polymeraseβ: structure-fidelity relationship from pre-steady-state kinetic analyses of all possible correct and incorrect base pairs for wild type and R283A mutant, Biochemistry 36 (1997) 1100–1107.
[37] B.G. Werneburg et al., DNA polymeraseβ: pre-steady-state kinetic analysis and roles of arginine-283 in catalysis and fidelity, Biochemistry 35 (1996) 7041–7050.
[38] J. Ahn et al., DNA polymerase β: effects of gapped DNA substrates on dNTP specificity, fidelity, processivity and conformational changes, Biochem. J. 331 (1998) 79–87.

[39] T.A. Steitz, J.A. Steitz, A general two-metal-ion mechanism for catalytic RNA, Proc. Natl. Acad. Sci. USA 90 (1993) 6498–6502.
[40] M.H. Olsson, A. Warshel, Solute solvent dynamics and energetics in enzyme catalysis: the S(N)2 reaction of dehalogenase as a general benchmark, J. Am. Chem. Soc. 126 (46) (2004) 15167–15179.
[41] P.G. Bolhuis et al., Transition path sampling: throwing ropes over rough mountain passes, in the dark, Annu. Rev. Phys. Chem. 53 (2002) 291–318.
[42] B. Bose-Basu et al., Dynamic characterization of a DNA repair enzyme: NMR studies of methyl-13C methionine-labeled DNA polymeraseβ, Biochemistry 43 (2004) 8911–8922.
[43] C. Bustamante, Z. Bryant, S.B. Smith, Ten years of tension: single-molecule DNA mechanics, Nature 421 (6921) (2003) 423–427.
[44] B. Maier, D. Bensimon, V. Croquette, Replication by a single DNA polymerase of a stretched single-stranded DNA, Proc. Natl. Acad. Sci. USA 97 (22) (2000) 12002–12007.