JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
pubs.acs.org/JACS

Optimal and Variant Metal-Ion Routes in DNA Polymerase β's Conformational Pathways

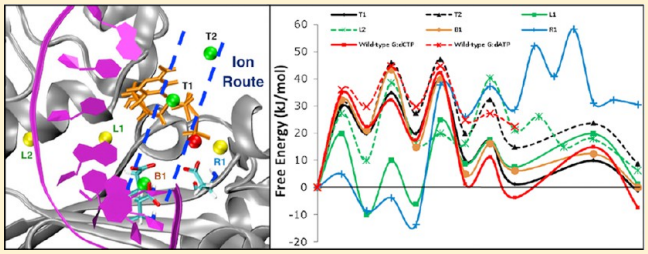
Yunlang Li,[†] Bret D. Freudenthal,[‡] William A. Beard,[‡] Samuel H. Wilson,[‡] and Tamar Schlick*[,†]

[†]Department of Chemistry and Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, New York 10012, United States

[‡]Laboratory of Structural Biology, National Institute of Environmental Health Sciences, National Institutes of Health Research, Triangle Park, North Carolina 27709, United States

Supporting Information

ABSTRACT: To interpret recent structures of the R283K mutant of human DNA repair enzyme DNA polymerase β (pol β) differing in the number of Mg^{2+} ions, we apply transition path sampling (TPS) to assess the effect of differing ion placement on the transition from the open one-metal to the closed two-metal state. We find that the closing pathway depends on the initial ion position, both in terms of the individual transition states and associated energies. The energy barrier of the conformational pathway varies from 25 to 58 kJ/mol, compared to the conformational energy barrier of 42 kJ/mol for the wild-type pol β reported previously. Moreover, we find a preferred ion route located in the center of the enzyme, parallel to the DNA. Within this route, the conformational pathway is similar to that of the overall open to closed transition of pol β, but outside it, especially when the ion starts near active site residues Arg258 and Asp190, the conformational pathway diverges significantly. Our findings should apply generally to pol β, since R283K is relatively far from the active site; further experimental and computational work are required to confirm this. Our studies also underscore the common feature that less active mutants have less stable closed states than their open states, in marked contrast to the wild-type enzyme, where the closed state is significantly more stable than the open form.



INTRODUCTION

DNA polymerase β (pol β) plays a prominent role in DNA repair, specifically base excision repair (BER),[1] which is crucial to the integrity of the genetic imprint. In 30% of human tumors, various altered pol β variants have been observed.[2−6] The polymerase domain of pol β is composed of three functionally distinct subdomains referred to as the D- (DNA-binding, Ile88-Pro151), C- (catalytic, Arg152-Lys262), and N-subdomains (nucleotide-binding, Asp263-Glu335). The amino-terminal 8-kDa lyase domain (Met1-Lys87) contributes an essential deoxyribose phosphate lyase activity necessary for single nucleotide base excision repair.[7] Two global conformational forms of the enzyme have been revealed by X-ray crystallography:[8] an open binary DNA complex and a closed ternary substrate complex. The two forms are related by a significant subdomain repositioning of the N-subdomain (Figure S1, Supporting Information). Kinetic,[9,10] structural,[8,11] and computational studies[12,13] have revealed that the catalytic pathway of pol β follows a general, three-step nucleotide insertion pathway for DNA polymerases: first, following DNA binding, pol β binds a 2′-deoxyribonucleoside 5′-triphosphate (dNTP) to form an open ternary substrate complex, which undergoes a conformational change to align active site residues and form a closed ternary substrate complex; second, the closed pol β complex catalyzes nucleotidyl transfer and forms the closed ternary product complex; third, the product complex undergoes a conformational change back to the open form, allowing the release of pyrophosphate (PP_i). During these conformational steps, subtle side-chain motions of other key residues (e.g., Asp192, Arg258, Tyr271, Phe272, and Arg283) also occur. In addition, two divalent metals, typically magnesium, are essential cofactors that play pivotal roles in nucleotidyl transfer.[14] Accordingly, binding of the ion ligands are expected to influence transitioning of the enzyme to/from the reactive competent state.[13,15,16] Many of these key residues are known to play important roles in pol β's fidelity.[17,18] In mismatched (i.e., not Watson−Crick paired) systems, such key residue motions as well as the related energy barriers are expected to be different.[19−23]

The R283K mutant of pol β has been reported to have lower activity and lower fidelity compared to the wild-type enzyme.[24−27] Recently, a series of structures of the pol β R283K mutant have been reported,[28] including an open structure with no Mg^{2+} ion, an open structure with one Mg^{2+} ion, and a closed structure with two Mg^{2+} ions. It has been hypothesized that the magnesium ions affect the conformational closing of the enzyme and thus can act as factors that

Received: December 20, 2013
Published: February 10, 2014

   dx.doi.org/10.1021/ja412701f | J. Am. Chem. Soc. 2014, 136, 3630−3639

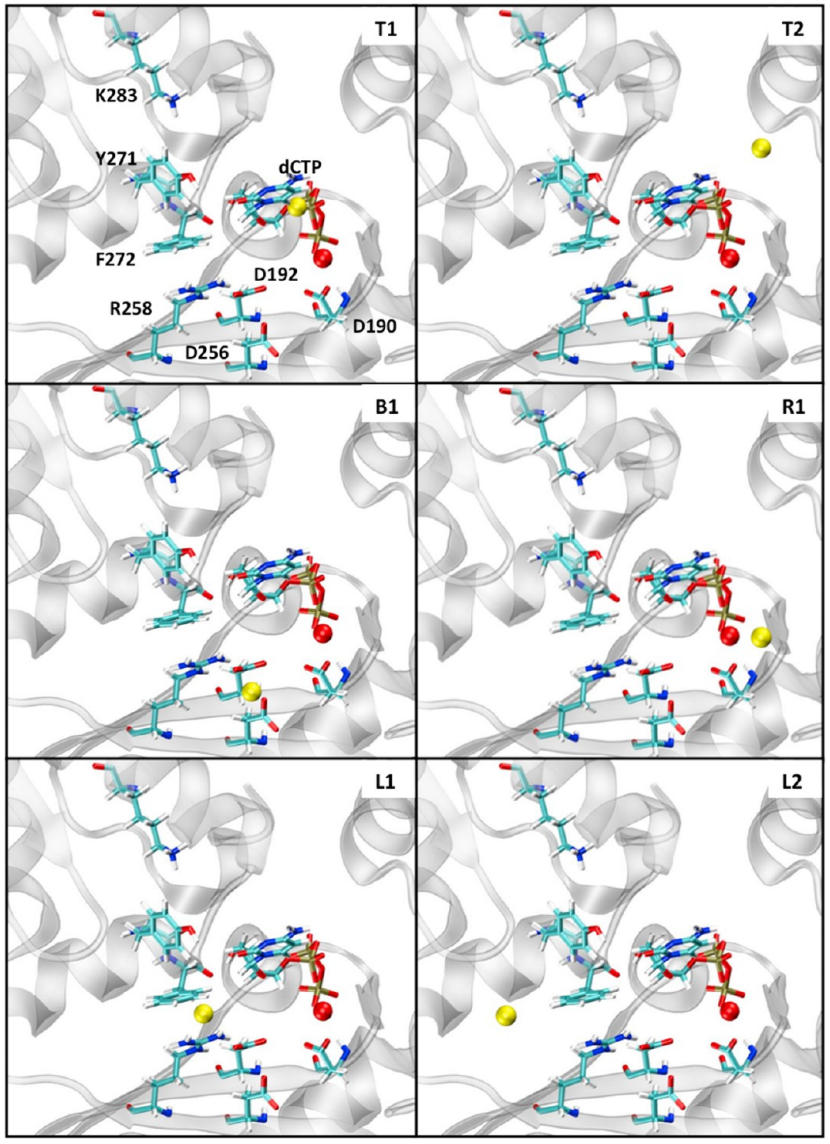


Figure 1. The six models of this study differing in initial position of catalytic Mg^{2+} (yellow). The nucleotide-binding Mg^{2+} is shown in red. Also see Figure S4 (Supporting Information) for stereo views of the starting locations.

regulate activity, among others.[15,29,30] Here we study how differences between the conformational pathways of the pol β R283K mutant and that of wild-type pol β might explain the lower activity and fidelity of R283K, and how Mg^{2+} might affect these pathways. In related studies, it has been shown that the E295K mutation of pol β interferes with the conformational pathway of pol β[31] and distorts its active site; this helps interpret, in part, the low insertion efficiency of the E295K mutant.[32] With regards to the role of Mg^{2+} ion, we hypothesize that the initial ion placement in/relative to the active site influences the conformational pathway and energy barriers. To test this hypothesis, we study the transition between the one-ion (open) and two-ion (closed) states with the transition path sampling (TPS) method we developed for biomolecules.[16,20] We base our starting models on the crystal structure of the open one-ion state (PDB entry 4F5O).[28] In each starting model, the catalytic Mg^{2+} was added at a different position. From these models, we have performed regular molecular dynamics and TPS simulations to examine the different possible conformational pathways from these one-ion models to the closed two-ion model based on the closed crystal structure (PDB entry 4F5Q).[28]

Our work supports the hypothesis that initial ion placement influences the conformational pathway and energy barriers. We suggest that a preferred "Mg^{2+} ion route" acts to steer the ion to a low-energy pathway. Whether the initial catalytic Mg^{2+} falls in the preferred route or not influences the conformational pathway significantly. The existence of such a preferred "Mg^{2+} ion route" and its relation to the conformational pathway may also hold for wild-type pol β and possibly other polymerases. In addition, we show that the partially closed state of R283K is less stable than the corresponding closed state of wild-type pol β, and this may further affect the chemical step. This energetic difference agrees with the lower activity and fidelity of R283K observed in experiments.[25,33]

METHODS

Simulation Setup. Our TPS simulations aim to link the one-ion (nucleotide-binding only) to two-ion crystal structures; to accomplish this, we explore various initial settings for the catalytic ion with respect to the active site, and the ion route in particular. Our starting models

dx.doi.org/10.1021/ja412701f | J. Am. Chem. Soc. 2014, 136, 3630−3639

are based on the R283K mutant pol β/DNA substrate complexes from the one-ion state (PDB entry 4F5O) crystal structure. All missing atoms are added to the models using the CHARMM program (version c35b2).[34] The catalytic Mg^{2+} is placed manually at various initial positions to build six starting models, as follows (Figure 1): In sets T1 and T2 (T for Top), the catalytic ion (yellow atom) is placed on the incoming nucleotide side; in sets B1 (B for Bottom), the catalytic ion is placed near the catalytic aspartates; in sets L1 and L2 (L for Left), the catalytic ion is near the upstream primer; and in set R1 (R for Right), the catalytic ion is positioned on the side of downstream primer. In sets T1, B1, L1, and R1, the initial catalytic Mg^{2+} is placed ~5 Å from the catalytic metal site in the two-ion crystal structure (approaching the distance between the catalytic Mg^{2+} and nucleotide-binding Mg^{2+} in their final positions, 3.6 Å); in sets T2 and R2, the initial catalytic Mg^{2+} is placed ~10 Å from its final position. Further minimization and equilibration were applied to optimize the system (see below).

The system is solvated with the explicit TIP3P water model in a water box via the VMD program.[35] The smallest image distance between the solute and the faces of the periodic cubic cell is set to 8 Å. The total number of water molecules is 14 179. To obtain a neutral system at an ionic strength of 150 mM, 52 Na^+ and 30 Cl^- ions are added to the system. All of the Na^+ and Cl^- ions are placed at least 8 Å away from both the protein and DNA atoms and from each other. A model of the ending structure based on the two-ion state (PDB entry: 4F5Q) crystal structure is built in a similar way.

Figure S2 (Supporting Information) compares the active site conformation of our starting (one-ion) and ending (two-ion) models. In the one-ion model, key residues such as Asp192, Arg258, and Phe272 are in their binary (open) states, and change to their ternary (closed) conformations in the two-ion model. The conformations of Asp190 and Asp256 also change between the one-ion model and two-ion model.

Minimization and Equilibration. Initial energy minimizations and equilibration simulations are performed using CHARMM. The system is minimized with fixed positions for all heavy atoms of protein or nucleotides, using steepest descent (SD) for 10 000 steps followed by the adopted basis Newton–Raphson (ABNR) method for 25 000 steps. The equilibration process is started with a 200 ps simulation at 300 K using the single-time step Langevin dynamics, while keeping all the heavy atoms of protein or nucleotides fixed. The SHAKE algorithm is then employed to constrain the bonds involving hydrogen atoms. This is followed by unconstrained minimization consisting of 10 000 steps of SD and 20 000 steps of ABNR. The system is then transferred to NAMD[36] and equilibrated for 500 ps at constant pressure and temperature. Pressure is maintained at 1 atm using the Langevin piston method with a piston period of 100 fs, a damping time constant of 50 fs, and a piston temperature of 300 K. The temperature is maintained at 300 K using weakly coupled Langevin dynamics[37] of nonhydrogen atoms with a damping coefficient of 10 ps^{-1}. Bonds to all hydrogen atoms are kept rigid using SHAKE, producing good stability with a time step of 2 fs. The system is simulated in periodic boundary conditions with full electrostatics computed using the particle mesh Ewald method[38] with grid spacing on the order of ≤1 Å. Short-range nonbonded terms are evaluated at every step using a 12 Å cutoff for van der Waals interactions and a smooth switching function. The final dimensions of the system are 80.1 Å × 79.2 Å × 79.7 Å.

Transition Path Sampling Simulations. The TPS method explores sequences of states constituting dynamical trajectories[39,40] through random walks using standard Monte Carlo (MC) procedures. Starting from an initial trajectory generated here by targeted molecular dynamics (TMD) that captures a barrier crossing, TPS uses the Metropolis MC method to sample the trajectory space by performing a random walk with the shooting algorithm;[41] the random walk is biased so that the most important regions of the trajectory space are adequately sampled.[40] The frequency of a trajectory region being visited is determined by its probability; thus, even when a random walk is initiated far from a representative transition pathway, the bias can drive the system to important transition regions after sampling. Therefore, despite the unphysical nature of the initial sampling trajectory obtained by TMD, TPS can lead the system to the most important transition regions and yield physically meaningful trajectories.

To obtain the initial trajectories that connect the two states during the transition, we apply TMD simulations to connect our modeled open and closed forms of the pol β R283K mutant complexes. The most challenging part of TPS is to describe the order parameters representing the transitions. To choose appropriate order parameters for TPS simulations, we use the crystallographic data,[28] molecular dynamics,[12,13] and prior TPS studies[16,19,22] on wild-type pol β as reference. Since these works have shown that key active-site residues (Asp190, Asp192, Arg258, Tyr 271, and Phe272), α-helix N of the N-subdomain, and the Mg^{2+} motion serve as measures of pol β's closing pathway, we start testing values associated with these residues and ions as well as the RMSD value of α-helix N atoms (residue 275 to 295). The complete set of order parameters is listed in Table 1.

Table 1. Transition States Properties for the Closing Conformational Profile of the R283K Pol β Mutant

Event	χ-order parameter	χ_{max} state A	χ_{min} state B
Partial N-subdomain motion	RMSD of residues 275–295 with respect to closed form	3.3 Å	2.5 Å
Asp192 flip	Dihedral angle C_γ–C_β–C_α–C	150°	180°
Arg258 rotation	Dihedral angle C_γ–C_δ–N_ε–C_ζ	100°	170°
Phe272 flip	Dihedral angle C_α–C_β–C_γ–$C_{\delta 2}$	110°	40°
Shift of Tyr271	Distance of Tyr271: OH – Lys283: N_ζ	5.4 Å	8.3 Å
Asp190 flip	Dihedral angle N–C_α–C_β–C_γ	60°	−70°
Asp190 flip back	Dihedral angle N–C_α–C_β–C_γ	−70°	60°
Ion motion[a]	Distance of catalytic Mg^{2+} to OD2 of Asp256	4.3–5.5 Å[b]	1.7 Å

[a]In set R1, the transition state is characterized by the rotation of Asp256, but the same distance value is used as the order parameter. [b]The initial distance between Mg^{2+} and Asp256 varies in different sets.

We use the TMD code implemented in NAMD to generate the initial constrained trajectories. An energy restraint based on the RMSD of the system relative to the final form is applied to force the open pol β complexes to close. From the TMD trajectory, we select frames that bracket the transition regions and perform unconstrained dynamics simulations. We perturb the atomic momenta of the frames and integrate the equations of motion forward and backward over short trajectories of order 10–100 ps (see below) to generate new physical, unbiased trajectories to connect the open and closed states. On the basis of these unconstrained simulations, we determine the adequate length of sampling trajectories for all the transition states. Specifically, for the mutant complex, the trajectories for Asp190 flipping back in set R1 are simulated for 25 ps, and those trajectories for other key residues and Mg^{2+} motions are run for 10 ps. To capture the transition states of N-subdomain closing in the two complexes, the sampling trajectories have to be propagated for 100 ps.

Using one of the newly generated physical trajectories as the starting trajectory, we perform path sampling for each individual conformational change with the shooting and shifting algorithm and a Monte Carlo protocol. The entire process is performed by using a PERL script that interfaces with NAMD. The velocity Verlet integrator in NAMD with a time step of 1 fs is used to generate the individual MD trajectories in TPS. All other parameters are the same as those in the equilibration process. To obtain an acceptance rate of 30–45%, the momentum perturbation magnitude (dP) of each transition state are varied from 0.001 to 0.005. To identify the transition states, 200 accepted trajectories for each transition state are collected.

The convergence of the harvested sampling trajectories is verified by computing the autocorrelation function associated with order

dx.doi.org/10.1021/ja412701f | J. Am. Chem. Soc. 2014, 136, 3630–3639

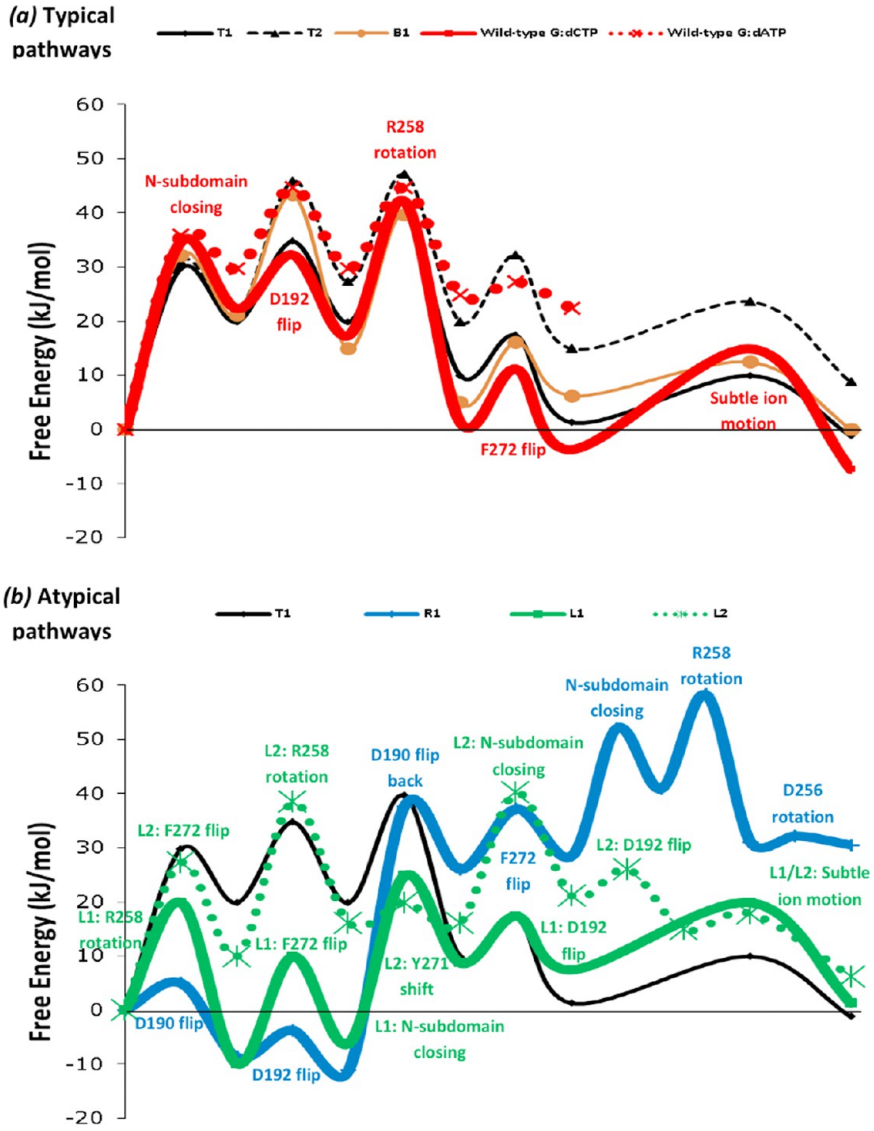


Figure 2. Free-energy pathways for the six models examined. The six pathways are grouped as (a) typical pathways as in sets T1, T2, and B1 (wild-type pol β as reference), and (b) atypical pathways as in sets R1, L1, and L2 (T1 as reference). The sequences of events in typical pathways are the same as that in wild-type pol β. Reference plot for wild-type pol β is from ref 16. Also see Table S1 (Supporting Information).

parameters to check for decorrelation of paths. The new trajectories are essentially decorrelated if the autocorrelation function shows a gradual transition between $\langle\chi_A\rangle^2$ and $\langle\chi_A\rangle\langle\chi_B\rangle$.

Free Energy Barrier and Rate Constant Calculations. The free energy barriers for transition states are evaluated using the "BOLAS" protocol,[42] an efficient procedure for getting free energies with relatively low error bars using the TPS trajectory harvesting idea. To calculate the free energy, 2000 trajectories for each transition state (10 overlapping windows, 200 trajectories per window) are collected. The potential of mean force plots obtained on each window are combined by adding/subtracting a constant to match the free energy values of the overlapping region. The free energy barriers for the conformational transitions are then calculated from the over free energy plots. The error bar for the free energy calculations is determined by repeating umbrella sampling on one window of a transition for 10 times with the same initial trajectory but different starting pseudorandom numbers. The standard deviation for each barrier (2–4 kJ/mol) is used as the error bar.

RESULTS AND DISCUSSION

Kinetic studies have indicated that the order of metal binding during active site assembly consists of, first, the nucleotide-binding ion and, second, the catalytic ion.[9] Generally, this sequential binding is reflective of the induced-fit hypothesis for selection of the correct nucleotide, where the dNTP binds first and then the enzyme samples for base-pair complementarity. For dNTP binding there is the need to neutralize the triphosphate oxygens of the incoming dNTP as it binds to active site aspartate residues; a nonbridging oxygen on Pα (pro-R_P) also provides a coordinating ligand for the nucleotide metal. We have observed multiple structures of pol β where ternary complexes contain the incoming nucleotide and associated nucleotide metal, but lack the catalytic metal.[28,30] These structures are consistent with nucleotide metal-containing ternary complexes poised for binding the catalytic metal. Binding of the catalytic metal induces a 3′-endo sugar pucker at the primer terminus aligning the O3′ atom for an in-line attack on Pα, and the bound catalytic metal also serves to activate the nucleotidyl transferase reaction.[43]

The closing conformational pathways of all six model sets are displayed in Figure 2. Also see Table S1 (Supporting Information) for the sequence of events in each pathway. On

dx.doi.org/10.1021/ja412701f | J. Am. Chem. Soc. 2014, 136, 3630–3639

Table 2. Free Energy Barrier of Sets T1, L1, and R1 Estimated by Transition-State Theory

	Asp190 flip	N-subdomain closing	Asp192 flip	Asp190 flip back	Arg258 rotation	Phe272 flip	ion motions	total[a]
				R283K mutant set T1				
$\beta\Delta F_{AB}^{\text{barrier}}$[b]	–	30 ± 4	15 ± 2	–	20 ± 3	7 ± 1	9 ± 1	40 ± 6
$\beta\Delta F_{AB}^{\text{barrier}}$	–	10 ± 2	15 ± 2	–	29 ± 3	16 ± 2	11 ± 2	40 ± 4
				R283K mutant set L1				
$\beta\Delta F_{AB}^{\text{barrier}}$	–	(31 ± 4)[c]	(9 ± 1)	–	(20 ± 3)	(20 ± 2)	12 ± 1	25 ± 7
$\beta\Delta F_{AB}^{\text{barrier}}$	–	(16 ± 2)	(10 ± 1)	–	(30 ± 4)	(16 ± 2)	18 ± 2	23 ± 3
				R283K mutant set R1				
$\beta\Delta F_{AB}^{\text{barrier}}$	5 ± 1	(24 ± 3)	5 ± 1	49 ± 4	(17 ± 3)	(11 ± 1)	1 ± 1	58 ± 8
$\beta\Delta F_{AB}^{\text{barrier}}$	14 ± 2	(11 ± 2)	7 ± 1	12 ± 3	(27 ± 4)	(9 ± 1)	2 ± 1	28 ± 4
				Wild-type matched system (G:dCTP)				
$\beta\Delta F_{AB}^{\text{barrier}}$	–	35 ± 5	10 ± 4	–	25 ± 4	10 ± 3	19 ± 5	42 ± 8
$\beta\Delta F_{AB}^{\text{barrier}}$	–	12 ± 3	15 ± 4	–	40 ± 6	15 ± 4	22 ± 4	49 ± 8

[a]The total energy barrier for conformational pathway before chemistry from open to closed state. [b]$\beta\Delta F_{AB}^{\text{barrier}}$ is the free energy of the transition state region between basin A and B relative to basin A, in kJ/mol. [c]To make the energy values comparable, the orders of events in the table for sets L1 and R1 have been rearranged as reflected by values in parentheses. See Table S1 (Supporting Information) for the sequence of transition states in each system.

the basis of the similarity of the conformational pathway to that of wild-type pol β, we categorize the conformational pathways for the six sets of simulations into "typical" (similar to that of wild-type pol β), and "atypical" (significantly different from that of wild-type pol β) pathways. The free energy values of representative sets T1, L1, and R1 as described in the Methods are listed in Table 2.

Typical Pathway. The conformational pathways of sets T1, T2, and B1 (starting from the incoming nucleotide side and near the catalytic aspartates) are similar to the conformational pathway of wild-type pol β. The representative set for T1 in Figure 3 shows that the closing of the N-subdomain (TS1) occurs first as the catalytic ion moves into the active site. Following N-subdomain closing, Asp192 flips (TS2). Unlike in wild-type pol β, Asp192 does not directly bind to the Mg^{2+} after the flip, only indirectly through a water molecule. However, this flip breaks the hydrogen bonds with Arg258, and the rotation of Arg258 follows (TS3). The next step is the rearrangement of Phe272 (TS4). After the Phe272 flip, Asp192 binds to the catalytic Mg^{2+} and helps the Mg^{2+} transit toward its final position, which is the last step of the conformational pathway (TS5). The altered residue Lys283 does not directly interact with Asp192, Arg258, or Phe272. However, after all the residue motions, the N-subdomain does not fully close as it does in wild-type pol β (the final RMSD of the N-subdomain compared to the closed state is ~2.2 Å, while the final RMSD of the N-subdomain in wild-type pol β is ~1.5 Å,[16] at a significance level of 90%). This may explain why the final closed state of R283K has a higher or similar energy level as its open state, while the final closed state of the wild-type pol β has an energy level lower than the open state (~7 kJ/mol) (see below).

Sets T2 and B1 have the same sequence of events in the prechemistry conformational pathway as T1. Though the three sets T1, T2, B1 share a similar conformational pathway, the energy barriers for each transition state differ. For example, after the flip of Asp192, the energy level of set T2 is higher than that of T1 and B1 (7 and 12 kJ/mol, respectively). This difference may arise because in set T2, the catalytic Mg^{2+} is relatively far away from the active site after the flip of Asp192, and thus cannot interact with Asp192. Consequently, the active site is less stable. Overall, these results illustrate the importance of the catalytic Mg^{2+} ion to the conformational pathway.

Atypical Pathway I. In sets L1 and L2, where the Mg^{2+} is initially close to key residues Arg258 and Phe272 side, the conformational pathway is significantly different from that of wild-type pol β. Figure 4 shows the sequence of events for L1. Arg258 first flips away from the Mg^{2+} (TS1). Since the positively charged Arg258 is farther away from the Mg^{2+}, the energy level of the system significantly decreases to approximately −10 kJ/mol (at a significance level of 95%). Second, the flip of Phe272 follows (TS2). However, because of the steric hindrance of the catalytic Mg^{2+}, Phe272 does not fully move to its closed state, between Arg258 and Asp192, as it does in wild-type pol β or in sets T1, T2, and B1. As a result, after the flip of Phe272, the energy level does not decrease as it does in wild-type pol β. The movement of Phe272 to its closed state happens during the closing of the N-subdomain closing (TS3) while the catalytic Mg^{2+} further moves into the active site. Following the N-subdomain closing, the flip of Asp192 (TS4) occurs, and finally the catalytic Mg^{2+} moves to its final position (TS5). As in sets T1, T2, and B1, the N-subdomain does not fully close in the final state (RMSD of the N-subdomain compared to the closed state is ~2.3 Å); subsequently, the final closed state also has a slightly higher energy level than its open state. The energy barrier of the whole conformational pathway is lower (~25 kJ/mol) than that in wild-type pol β (~42 kJ/mol) (at a significance level of 80%), mainly because of the energetic decrease associated with the Arg258 rotation. The lower energy barrier in L1 is mainly due to the starting catalytic ion position disrupting the Asp192/Arg258 interaction. From previous experimental and computational studies,[17,44,45] we know that the Asp192/Arg258 interaction plays an important role in pol β's open-to-closed conformational pathway. Therefore, it is not likely that the catalytic ion should initially disrupt this interaction, and the L1 pathway is unfavorable.

Set L2 has a similar sequence of events as L1 (both L1 and L2 have initial positions near the upstream primer) except that Phe272 flips prior to that of Arg258. A transition state for Tyr271 shifting before the N-subdomain closes also occurs, but the energy barrier is relatively small (~3 kJ/mol) and within the simulation error bar. However, because the catalytic Mg^{2+} is farther away from Arg258 when Arg258 rotates, the electrostatic effect is less significant and the rotation of Arg258 in L2 results in an energetic increase. Thus, the energy barrier of the

dx.doi.org/10.1021/ja412701f | J. Am. Chem. Soc. 2014, 136, 3630−3639

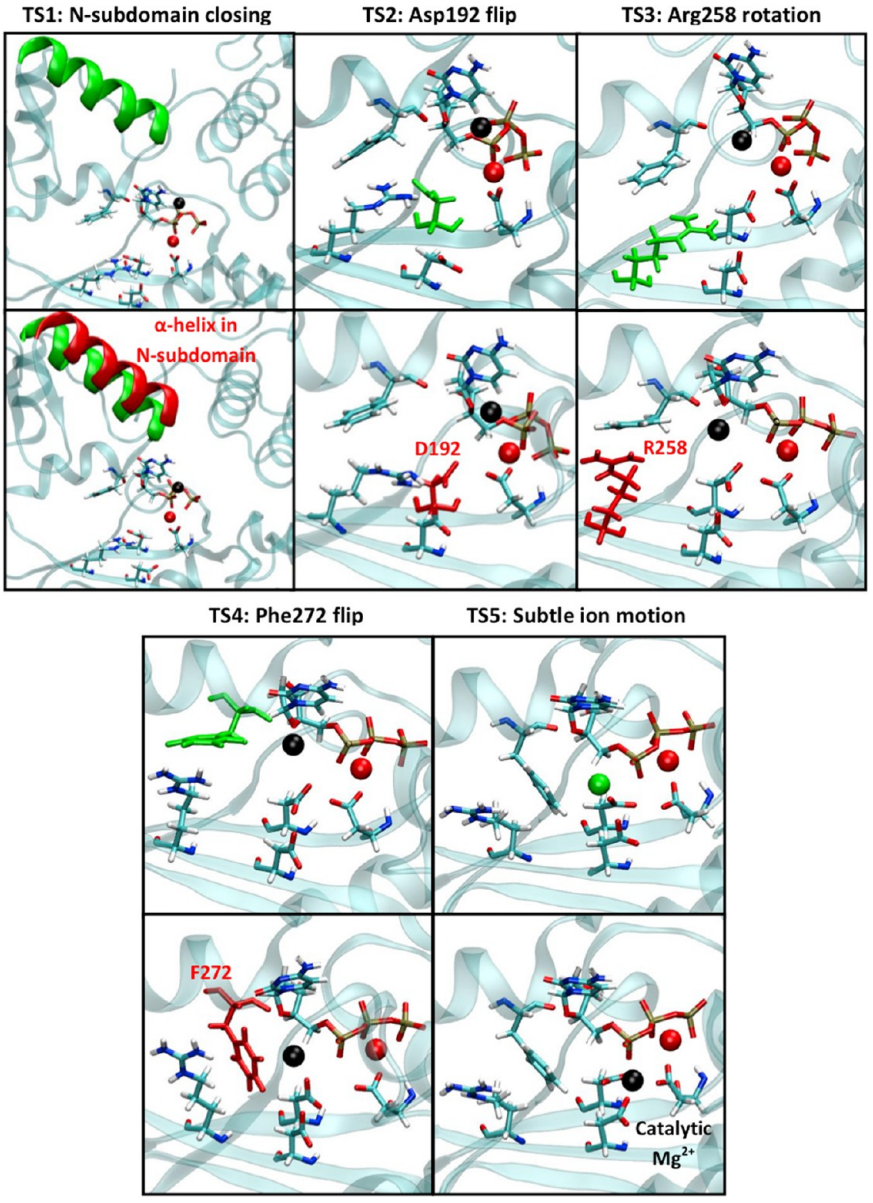


Figure 3. Transition states identified in set T1. The residue/region before the transition is shown in green (upper) and after transition in red (lower). In the last step (subtle ion motion), the catalytic Mg^{2+} after transition is shown in black instead of red. Also see Figure S5 (Supporting Information) for normalized probability distribution of the order parameters for the transition states revealed (TS1 to TS5), and Figure S6 (Supporting Information) for stereo views of each transition state.

whole conformational pathway (~40 kJ/mol) is similar to that of wild-type pol β.

Atypical Pathway II. In set R1, where Mg^{2+} is initially close to Asp190, we also discover a significantly different conformational pathway, as shown in Figure 5. As the catalytic Mg^{2+} approaches the active site, Asp190 flips toward it (TS1). Asp190 and dCTP coordinate with the catalytic Mg^{2+} (Figure S3a, Supporting Information), so the system is relatively stable and the energy level decreases to approximately −9 kJ/mol. After the flip of Asp190, the rearrangement of Asp192 follows (TS2). Interestingly, in set R1, the flipped Asp192 does not bind to the catalytic Mg^{2+} either directly or through water molecules. Instead, after Asp190 flips to coordinate with the catalytic Mg^{2+}, the nucleotide-binding Mg^{2+}, which is originally coordinated with Asp190, shifts toward Asp192. Coordination between Asp192 and the nucleotide-binding Mg^{2+} through water is occasionally observed. After the flip of Asp192, the energy level of the system further decreases to approximately −11 kJ/mol. The next step is the Asp190 flip back to its original position, which involves a large energy barrier of ~48 kJ/mol, since the coordination between Asp190/dCTP and the catalytic Mg^{2+} is relatively stable and difficult to disturb. We observe that before the flipping back of Asp190, the catalytic Mg^{2+} shifts slightly toward the dCTP to weaken its binding with Asp190 (Figure S3b, Supporting Information). After this flipping back of Asp190, Asp190 binds to the nucleotide-binding Mg^{2+} again (TS3). The next residue to rearrange is Phe272 (TS4). Partial N-subdomain closing (TS5) occurs after the flip of Phe272, followed by the rotation of Arg258 (TS6). Unlike in wild-type pol β or other R283K sets, in set R1 the catalytic Mg^{2+} is almost at its final position following the Arg258 rotation. However, we observe a final transition state of Asp256 rotation (TS7) toward the catalytic Mg^{2+} to seal rearrangements of the active site prior to the chemical step. To be consistent, we use the distance

dx.doi.org/10.1021/ja412701f | J. Am. Chem. Soc. 2014, 136, 3630−3639

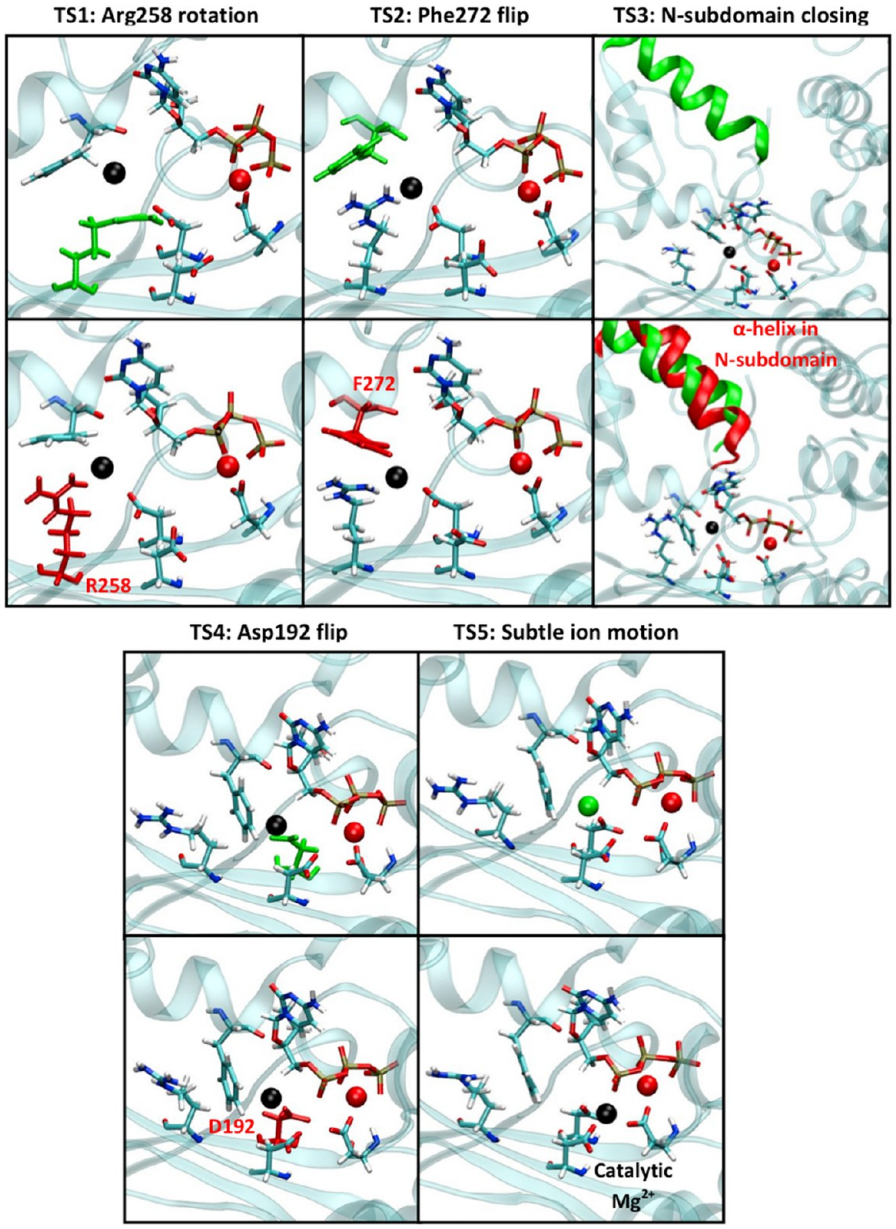


Figure 4. Transition states identified in set L1. The residue/region before the transition is shown in green (upper) and after transition in red (lower). In the last step (subtle ion motion), the catalytic Mg^{2+} after transition is shown in black instead of red. Also see Figure S7 (Supporting Information) for normalized probability distribution of the order parameters for the transition states revealed (TS1 to TS5), and Figure S8 (Supporting Information) for stereo views of each transition state.

between OD2 on Asp256 and the catalytic Mg^{2+} as the reaction coordinate in set R1.

The energy barrier of the entire conformational pathway (~58 kJ/mol) is higher than that of wild-type pol β (at a significance level of 80%), mainly because of the costly step associated with the flipping back of Asp190. The final closed state is also less stable than the open state, in terms of both higher energy level (~30 kJ/mol, at a significance level of 95%) and the partially closed N-subdomain (RMSD of 2.5 Å compared to the closed state). However, key distances within the active site (e.g., catalytic Mg^{2+} to O3′ on the primer terminus) do not deviate much from those of matched wild-type pol β system, compared to G:dATP mismatched or E295K mutant systems. Therefore, the active site is less deformed overall. Since R283K is farther away from the active site than E295K, which we studied in detail,[32] or the incorrect incoming nucleotide, and the change from arginine to lysine is more conservative (compared to glutamate to lysine in E295K), the mutation of R283K affects the active site less overall.

Preferred Mg^{2+} Ion Route. We show all initial catalytic Mg^{2+} positions in Figure 6. Our studies indicate that a preferred "Mg^{2+} ion route", which is approximately parallel to the DNA, affects overall enzyme motions. The initial Mg^{2+} positions in sets T1, T2, and B1 are located within this route, while the initial Mg^{2+} positions in set L1, L2, and R1 are located outside of this route. When the initial Mg^{2+} position does not fall in this preferred route, it is more facile to impact motions of key residues such as Arg258 or Asp190 to generate atypical conformational pathways. Since the R283K mutation does not contact the metal ion, this preferred ion route and related typical/atypical conformational pathways should also hold for the wild-type pol β. Further work on wild-type pol β to test the

dx.doi.org/10.1021/ja412701f | J. Am. Chem. Soc. 2014, 136, 3630−3639

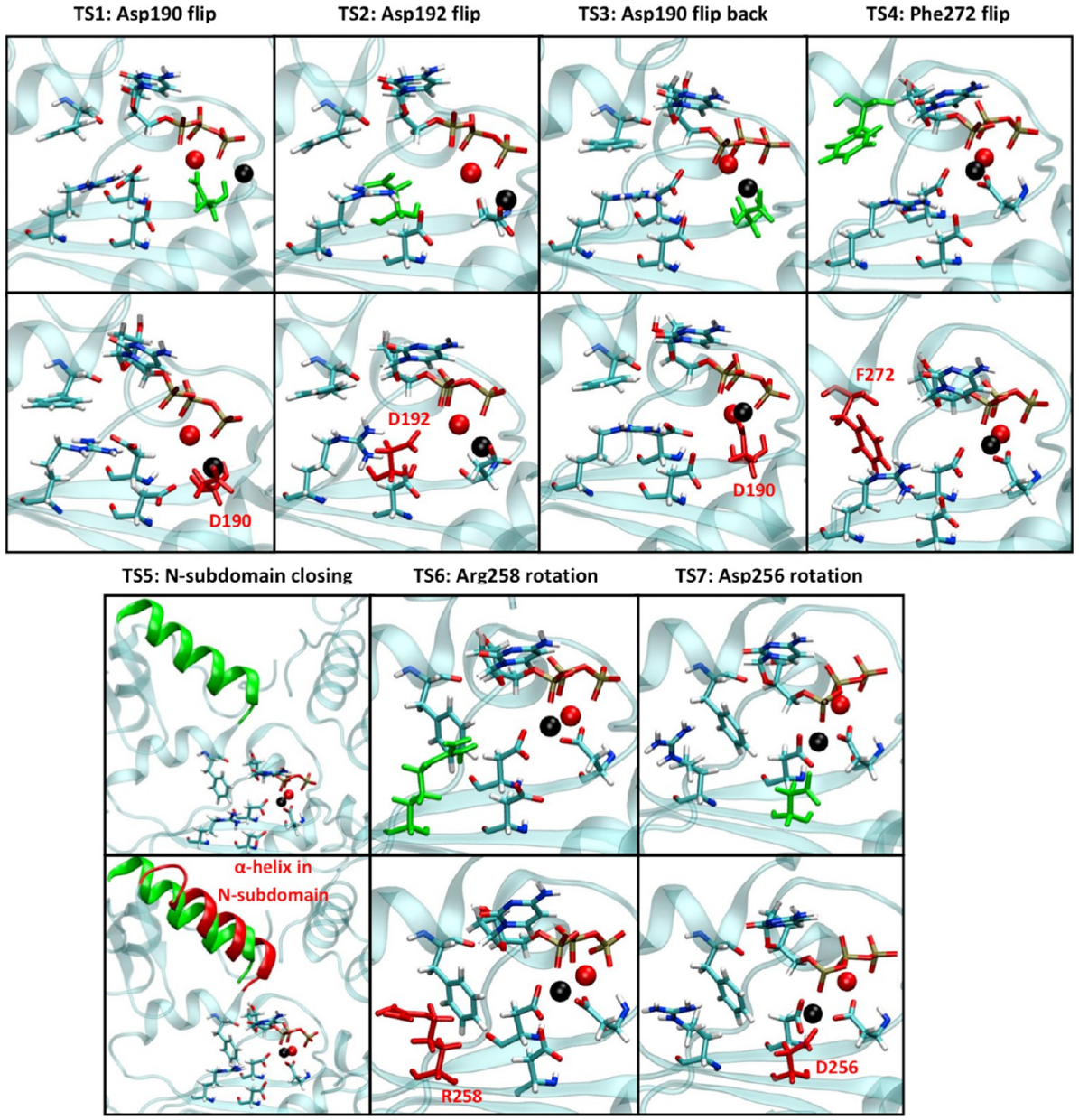


Figure 5. Transition states identified in set R1. The residue/region before the transition is shown in green (upper) and after transition in red (lower). Also see Figure S9 (Supporting Information) for normalized probability distribution of the order parameters for the transition states revealed (TS1 to TS7), and Figure S10 (Supporting Information) for stereo views of each transition state.

existence of such a preferred ion route forms a natural extension of this work. We further hypothesize that in vivo it is more probable that the catalytic Mg^{2+} moves into the active site through the preferred route to generate a typical conformational pathway, as reported in wild-type pol β. Though in set L1, the energy barrier of the whole conformational pathway is lower than that of the typical pathway, the catalytic Mg^{2+} disrupts the Arg258/Asp192 interactions, and the sequence of events in L1 differs significantly from the typical pathway and also from that of wild-type pol β.

Recent studies have revealed that during the conformational pathway after chemistry, a third Mg^{2+} appears in the active site.[11,46] The reported position of the third Mg^{2+} is close to the initial Mg^{2+} position in set T1, which falls within the "Mg^{2+} ion route". Therefore, it is likely that the third Mg^{2+} also moves into the active site through the route, which may validate our hypothesis on the function of route.

Unstable Closed State in R283K. Though Lys283 does not directly interrupt the key residue motions in the active site, we observe that in the R283K mutant, the N-subdomain does not fully close as it does in wild-type pol β (RMSD of 2.2−2.5 Å compared to the closed state, compared to ∼1.5 Å in wild-type pol β). This partially closed N-subdomain represents a less stable closed state compared to wild-type pol β (see Figure 2), and the distorted final state during prechemistry conformational pathway in turn may be unfavorable for the following chemical reaction. This significant difference from the active, wild-type enzyme may help explain experimental observations concerning lower activity and fidelity of the R283K mutant. This theme of lower stability of the closed state in certain mutants and mismatched systems was first discussed in our

dx.doi.org/10.1021/ja412701f | J. Am. Chem. Soc. 2014, 136, 3630−3639

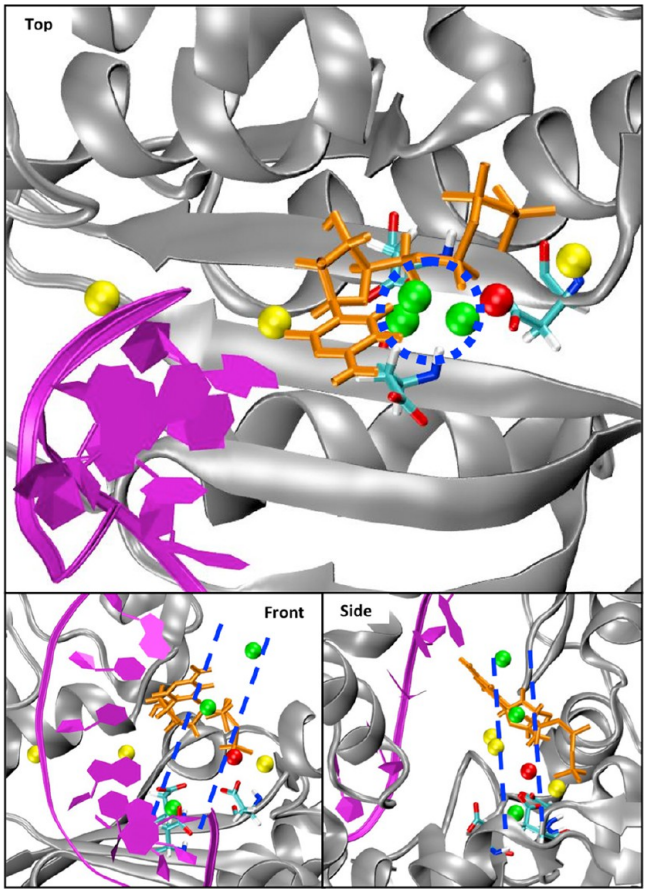


Figure 6. Three views of the Mg^{2+} ion route in dotted blue line. Initial catalytic Mg^{2+} positions in sets T1, T2, and B1 are shown in green, while those in sets L1, L2, and R1 are shown in yellow. The nucleotide-binding Mg^{2+} is shown in red. For clarity, the DNA primer strand and lyase domain are not shown.

work for pol β,[19,20,22,32,47] and later shown for other polymerases.[48]

CONCLUSIONS

We have used transition path sampling simulations to investigate the conformational transition pathways before chemistry for the pol β R283K mutant with varying initial positions of the catalytic Mg^{2+} ion. On the basis of the crystal structure with only the nucleotide-binding Mg^{2+}, we have built several models to study how different initial positions of catalytic Mg^{2+} influence the subsequent closing conformational pathway. Our analyses reveal that the conformational pathway depends significantly on the initial position of catalytic Mg^{2+}, regarding both the sequence of events and individual energy values. The combined energy barrier ranges from 25 to 58 kJ/mol. Though the active site in the final closed form of the mutant is not distorted, the N-subdomain of the mutant does not fully close after the transition. On the basis of the similarity to the conformational pathway of wild-type pol β, we have categorized the conformational pathway of R283K mutant as "typical" and "atypical". We also highlight the importance of a preferred "Mg^{2+} ion route", located roughly parallel to the DNA. When the initial catalytic Mg^{2+} falls on this route, the conformational pathway is typical; when the initial catalytic Mg^{2+} occurs outside the route, the ion may interact with key residues such as Asp190 and Arg258 and skew the conformational pathway. Since residue R283 is relatively distant from the active site, we also suggest that the observed dependence of conformational closing pathway on the initial Mg^{2+} position may hold generally to pol β and allow robustness and variability in enzyme dynamics and thus activity. This may be examined by future experimental and computational works. Finally, our finding that the R283K mutant has a less stable partially closed state than wild-type pol β agrees with the lower activity and fidelity of the mutant and underscores a pattern we revealed previously. Further experimental and computational studies are required to examine the significance of surrogate and variable number of cations near the polymerase active site and their effect on enzyme dynamics and catalysis.

ASSOCIATED CONTENT

Supporting Information

Choices of protonation states, shooting algorithm and test of convergence of TPS, Tables S1 and S2, and Figures S1−S10. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

schlick@nyu.edu

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Dr. Ravi Radhakrishnan for providing the initial scripts for transition path sampling simulations. This work was supported in part by the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences (project number Z01-ES050158) and was in association with NIH grant 1U19CA105010. The computations in this study were conducted using the resources of the CCNI supported by the New York State Foundation for Science, Technology and Innovation (NYSTAR), and the Dell computer cluster by New York University Information Technology Services (NYU ITS). Molecular images were generated using the VMD[35] programs.

REFERENCES

(1) Beard, W. A.; Wilson, S. H. *Chem. Rev. (Washington, DC, U. S.)* **2006**, *106*, 361.
(2) Dalal, S.; Kosa, J. L.; Sweasy, J. B. *J. Biol. Chem.* **2004**, *279*, 577.
(3) Starcevic, D.; Dalal, S.; Sweasy, J. B. *Cell Cycle* **2004**, *3*, 998.
(4) Dalal, S.; Hile, S.; Eckert, K. A.; Sun, K. W.; Starcevic, D.; Sweasy, J. B. *Biochemistry* **2005**, *44*, 15664.
(5) Lang, T. M.; Dalal, S.; Chikova, A.; DiMaio, D.; Sweasy, J. B. *Mol. Cell. Biol.* **2007**, *27*, 5587.
(6) Dalal, S.; Chikova, A.; Jaeger, J.; Sweasy, J. B. *Nucleic Acids Res.* **2008**, *36*, 411.
(7) Joyce, C. M.; Steitz, T. A. *Annu. Rev. Biochem.* **1994**, *63*, 777.
(8) Sawaya, M. R.; Pelletier, H.; Kumar, A.; Wilson, S. H.; Kraut, J. *Science* **1994**, *264*, 1930.
(9) Balbo, P. B.; Wang, E. C.; Tsai, M. D. *Biochemistry* **2011**, *50*, 9865.
(10) Tanabe, K.; Bohn, E. W.; Wilson, S. H. *Biochemistry* **1979**, *18*, 3401.
(11) Freudenthal, B. D.; Beard, W. A.; Shock, D. D.; Wilson, S. H. *Cell* **2013**, *154*, 157.
(12) Arora, K.; Schlick, T. *Biophys. J.* **2004**, *87*, 3088.
(13) Arora, K.; Schlick, T. *J. Phys. Chem. B* **2005**, *109*, 5358.
(14) Beese, L. S.; Steitz, T. A. *EMBO J.* **1991**, *10*, 25.

(15) Yang, L.; Arora, K.; Beard, W. A.; Wilson, S. H.; Schlick, T. *J. Am. Chem. Soc.* **2004**, *126*, 8441.
(16) Radhakrishnan, R.; Schlick, T. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 5970.
(17) Radhakrishnan, R.; Arora, K.; Wang, Y.; Beard, W. A.; Wilson, S. H.; Schlick, T. *Biochemistry* **2006**, *45*, 15142.
(18) Schlick, T.; Arora, K.; Beard, W.; Wilson, S. *Theor. Chem. Acc.* **2012**, *131*, 1.
(19) Radhakrishnan, R.; Schlick, T. *J. Am. Chem. Soc.* **2005**, *127*, 13245.
(20) Radhakrishnan, R.; Schlick, T. *Biochem. Biophys. Res. Commun.* **2006**, *350*, 521.
(21) Wang, Y. L.; Reddy, S.; Beard, W. A.; Wilson, S. H.; Schlick, T. *Biophys. J.* **2007**, *92*, 3063.
(22) Wang, Y. L.; Schlick, T. *BMC Struct. Biol.* **2007**, 7.
(23) Berlow, R. B.; Swain, M.; Dalal, S.; Sweasy, J. B.; Loria, J. P. *J. Mol. Biol.* **2012**, *419*, 171.
(24) Beard, W. A.; Shock, D. D.; Yang, X. P.; DeLauder, S. F.; Wilson, S. H. *J. Biol. Chem.* **2002**, *277*, 8235.
(25) Werneburg, B. G.; Ahn, J.; Zhong, X. J.; Hondal, R. J.; Kraynov, V. S.; Tsai, M. D. *Biochemistry* **1996**, *35*, 7041.
(26) Osheroff, W. P.; Beard, W. A.; Yin, S.; Wilson, S. H.; Kunkel, T. A. *J. Biol. Chem.* **2000**, *275*, 28033.
(27) Osheroff, W. P.; Beard, W. A.; Wilson, S. H.; Kunkel, T. A. *J. Biol. Chem.* **1999**, *274*, 20749.
(28) Freudenthal, B. D.; Beard, W. A.; Wilson, S. H. *Structure* **2012**, *20*, 1829.
(29) Cisneros, G. A.; Perera, L.; Garcia-Diaz, M.; Bebenek, K.; Kunkel, T. A.; Pedersen, L. G. *DNA Repair* **2008**, *7*, 1824.
(30) Batra, V. K.; Beard, W. A.; Shock, D. D.; Krahn, J. M.; Pedersen, L. C.; Wilson, S. H. *Structure* **2006**, *14*, 757.
(31) Kirby, T. W.; Derose, E. F.; Cavanaugh, N. A.; Beard, W. A.; Shock, D. D.; Mueller, G. A.; Wilson, S. H.; London, R. E. *Nucleic Acids Res.* **2012**, *40*, 2974.
(32) Li, Y.; Gridley, C. L.; Jaeger, J.; Sweasy, J. B.; Schlick, T. *J. Am. Chem. Soc.* **2012**, *134*, 9999.
(33) Beard, W. A.; Osheroff, W. P.; Prasad, R.; Sawaya, M. R.; Jaju, M.; Wood, T. G.; Kraut, J.; Kunkel, T. A.; Wilson, S. H. *J. Biol. Chem.* **1996**, *271*, 12141.
(34) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.
(35) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics Modell.* **1996**, *14*, 33.
(36) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781.
(37) Feller, S. E.; Zhang, Y. H.; Pastor, R. W.; Brooks, B. R. *J. Chem. Phys.* **1995**, *103*, 4613.
(38) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.
(39) Pratt, L. R. *J. Chem. Phys.* **1986**, *85*, 5045.
(40) Bolhuis, P. G.; Chandler, D.; Dellago, C.; Geissler, P. L. *Annu. Rev. Phys. Chem.* **2002**, *53*, 291.
(41) Bolhuis, P. G.; Dellago, C.; Chandler, D. *Faraday Discuss.* **1998**, *110*, 421.
(42) Radhakrishnan, R.; Schlick, T. *J. Chem. Phys.* **2004**, *121*, 2436.
(43) Batra, V. K.; Perera, L.; Lin, P.; Shock, D. D.; Beard, W. A.; Pedersen, L. C.; Pedersen, L. G.; Wilson, S. H. *J. Am. Chem. Soc.* **2013**, *135*, 8078.
(44) Menge, K. L.; Hostomsky, Z.; Nodes, B. R.; Hudson, G. O.; Rahmati, S.; Moomaw, E. W.; Almassy, R. J.; Hostomska, Z. *Biochemistry* **1995**, *34*, 15934.
(45) Batra, V. K.; Beard, W. A.; Shock, D. D.; Pedersen, L. C.; Wilson, S. H. *Mol. Cell* **2008**, *30*, 315.
(46) Nakamura, T.; Zhao, Y.; Yamagata, Y.; Hua, Y. J.; Yang, W. *Nature* **2012**, *487*, 196.
(47) Arora, K.; Beard, W. A.; Wilson, S. H.; Schlick, T. *Biochemistry* **2005**, *44*, 13328.
(48) Kirmizialtin, S.; Nguyen, V.; Johnson Kenneth, A.; Elber, R. *Structure* **2012**, *20*, 618.